J. Fluid Mech. (1995), *vol.* 300, *pp.* 1–41

Geometric focusing of internal waves

By LEO R. M. MAAS AND FRANS-PETER A. LAM
Netherlands Institute for Sea Research, PO Box 59, 1790 AB Texel, The Netherlands

(Received 2 January 1995)

The spatial structure of the streamfunction field of free, linear internal waves in a two-dimensional basin is governed by the canonical, second-order, hyperbolic equation on a closed domain. Its solution can be determined explicitly for some simple shapes of the basin. It consists of an algorithm by which 'webs' of uniquely related characteristics can be constructed and the prescription of one (independent) value of a field variable, related to the streamfunction, on each of these webs. The geometric construction of the webs can be viewed as an alternative version of a billiard game in which the angle of reflection equals that of incidence with respect to the *vertical* (rather than to the normal). Typically, internal waves are observed to be globally attracted ('focused') to a limiting set of characteristics. This attracting set can be classified by the number of reflections it has with the surface (its *period* in the terminology of dynamical systems). This period of the attractor is a fractal function of the normalized period of the internal waves: large regions of smooth, low-period attractors are seeded with regions with high-period attractors. Occasionally, all internal wave rays fold exactly back upon themselves, a 'resonance': focusing is absent and a smooth pattern, familiar from the cellular pattern in a rectangular domain, is obtained. These correspond to the well-known seiching modes of a basin. An analytic set of seiching modes has also been found for a semi-elliptic basin. A necessary condition for seiching to occur is formulated.

1. Introduction

Study of the canonical hyperbolic equation (the wave equation) is usually performed on *half-open* domains only. This is because in those cases one of the independent variables is *time* and no future behaviour of the solution is normally imposed. In the present study the wave equation governs the *spatial* structure (of the streamfunction) of linear, monochromatic internal waves in a stratified basin. It should thus be solved on a *closed* domain on which boundary the streamfunction vanishes. Magaard (1962, 1968) showed that this equation is solved by a functional relation that can be rewritten as a mapping between successive surface intersections (reflections) of characteristics. The interval between two successive surface intersections is referred to as a *fundamental interval*. Once the field variable is prescribed in a fundamental interval the complete solution can be determined in two steps. First, from the specified value of the field variable at the surface in that fundamental interval this field variable can be constructed over the whole surface domain. Second, the streamfunction at any point of the interior domain is obtained as the difference of the value of this field variable that is carried invariantly along the characteristics intersecting at that point. Magaard basically restricts his study to propagation of internal waves that have a frequency for which the basin bottom is *sub*critical (characteristic steeper than bottom) which constitute two detached *monotonic* maps: one for rightward and one for leftward propagation. In this study we extend this to a consideration of internal waves with frequencies for which the bottom is *super*critical: internal waves bounce back and forth between the sides of the basin. The right- and leftward modes of propagation get connected and present a *bi-modal* map. For certain simple cross-sectional profiles this map can be obtained explicitly. The characteristics fold back over and over again to form what will be referred to as a *web*. The construction of webs of characteristics and the prescription of the field variable in a unique (fundamental) interval form the two independent parts of the solution of this problem. It is the former, geometric aspect that is most influential however. Irrespective of what the field variable may be it predicts the possible existence of certain limiting characteristics to which the solution is attracted.

In §2 the equations governing internal gravity waves in a two-dimensional stratified basin are derived and the functional relation of Magaard (1962, 1968) is reviewed. In §3 this is applied to a non-trivial, one-parameter topography – the parabolic basin – for which the bi-modal map can be derived. Webs of characteristics and their asymptotic states are constructed with this map. Several geometrical aspects are pointed out for these attractors, giving rise to a conjecture on nested maps. In §4 an example of a solution of the complete problem is given for a special choice of the field variable in the fundamental intervals. Standing versus propagating modes of internal gravity waves are discussed.

One would like to view a boundary value problem like the one presented here as an eigenvalue problem. Solutions of such a problem are usually obtained as a (finite or infinite) set of discrete eigenfrequencies separated by compact regions where no such frequencies reside. The solutions of the hyperbolic equation in the only geometry for which analytical solutions are presently available, the rectangle, however signal that there are some unusual facets to this kind of eigenvalue problem (Münnich 1994). First, the eigenfrequencies (eigenperiods) are degenerate: for any eigenperiod there is an infinite number of spatial structures corresponding to it (spatial multiples of the horizontal and vertical structure of the basic state). Second, the eigenfrequencies are dense: *every* rational frequency is an eigenfrequency, much like for inertial motion on a torus. Thus the 'eigen-ness' of the eigenfrequencies is becoming dubious terminology. In §5 the solution for the rectangle, obtained by separation of variables, will be compared with that using the method of characteristics, employed here. It is argued that for irrational frequencies the characteristics are plane-filling and thus the width of the fundamental interval over which the field variable can be independently specified shrinks to zero (a single point). Hence the streamfunction at any point – being the difference of two sampled values of the field variable – vanishes and no free solution for such frequencies exists. In contrast to what is found in a parabolic basin, for rational frequencies, each characteristic exactly folds back upon itself: a *resonance*. In this section, finally, it is pointed out that analytical solutions, bearing a one-to-one relationship to those found for the rectangle, can be obtained in a (semi) elliptic basin.

For other non-rectangular geometries there also appear to be frequencies for which stationary internal wave patterns do not exist. The notion of the existence of certain discrete 'eigenfrequencies', however, regresses even further, since, in one sense, these now constitute *compact* domains.

Section 6 discusses the same features for some other simple basin shapes. Section 7, finally, discusses the relevance of the present study for oceanic and lake applications. It also summarizes the main results and limitations of this approach.

2. Internal-wave equation and solution by functional relation

2.1. *Internal-wave equation*

Internal waves in a uniformly stratified, inviscid, linear, hydrostatic, non-rotating, two-dimensional Boussinesq fluid are governed by the momentum equations, conservation of density and continuity equation (e.g. Turner 1973):

$$\frac{\partial u}{\partial t} = -\frac{1}{\rho_*}\frac{\partial p}{\partial x}, \tag{2.1a}$$

$$\frac{1}{\rho_*}\frac{\partial p}{\partial z} = b \equiv -g\frac{\rho}{\rho_*}, \tag{2.1b}$$

$$\frac{\partial b}{\partial t} + wN^2 = 0, \tag{2.1c}$$

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0. \tag{2.1d}$$

Here t is time and u and w are the velocity components in the horizontal (x) and vertical (z) directions in a Cartesian frame of reference whose origin is located at the surface on the basin centreline. The positive z-direction is antiparallel to gravity. Gravitational acceleration is denoted as g. Perturbation density and pressure fields, ρ and p, are expanded about a density field $\rho_* + \rho_0(z)$ and a hydrostatically-related pressure field, where $\rho_* \gg \rho_0(z) \gg \rho(x,z,t), \forall\{x,z,t\}$. Buoyancy b is defined in (2.1b) and N is the Brunt-Väisälä frequency defined through $N^2(z) = -(g/\rho_*)(\mathrm{d}\rho_0/\mathrm{d}z)$, which acts as the upper bound of internal wave frequencies (Groen 1948).

Elimination of the buoyancy b between (2.1b) and (2.1c) yields

$$\frac{1}{\rho_*}\frac{\partial^2 p}{\partial z \partial t} = -wN^2, \tag{2.2a}$$

while, with (2.1a), subsequent elimination of p/ρ_* gives

$$\frac{\partial^3 u}{\partial z \partial t^2} = \frac{\partial w}{\partial x}N^2. \tag{2.2b}$$

Equation (2.1d) suggests the use of a streamfunction $\Psi(x,z,t)$ related to the velocities by $u = -\partial\Psi/\partial z,\ \ w = \partial\Psi/\partial x$, with which (2.2$b$) becomes

$$\frac{\partial^4 \Psi}{\partial z^2 \partial t^2} + N^2\frac{\partial^2 \Psi}{\partial x^2} = 0. \tag{2.3}$$

For monochromatic waves of frequency ω

$$\Psi(x,z,t) = \psi(x,z)\mathrm{e}^{-\mathrm{i}\omega t},$$

this reduces to

$$\frac{\partial^2 \psi}{\partial x^2} - \frac{\omega^2}{N^2}\frac{\partial^2 \psi}{\partial z^2} = 0. \tag{2.4}$$

It will be assumed that the stratification is uniform, so that N is a constant. In an infinite medium (2.3) is satisfied by planar waves (with horizontal and vertical wavenumbers k and m), which obey the dispersion relation

$$\omega = \pm N\frac{k}{m}.$$

The frequency is therefore just a function of the angle that the wave vector makes with the vertical. From the dispersion relation the perpendicular nature of internal wave propagation – group velocity vector normal to phase velocity vector – can be inferred and is such that the vertical components of these two vectors are always in opposition (Lighthill 1978). Demonstrations of this type of internal wave propagation have been given in the laboratory studies of Görtler (1943), Mowbray & Rarity (1967) and Thorpe (1968), while ray-like propagation of internal waves was also observed in the ocean by deWitt *et al.* (1986) and Pingree & New (1991).

By scaling x with the basin half-width L and z with $\omega L/N$, (2.4) obtains the canonical form of a second-order, hyperbolic equation:

$$\frac{\partial^2 \psi}{\partial x^2} - \frac{\partial^2 \psi}{\partial z^2} = 0. \tag{2.5a}$$

This is the 'wave equation' in *spatial* coordinates only. For a wide class of non-uniform, ocean-like stratification profiles a coordinate and variable transformation exists for which (2.4) can also be reduced to standard form (Magaard 1962; Baines 1973). This transformation, however, also affects the description of the form of the boundary. For the sake of simplicity, therefore, we will stick to the assumption of uniform stratification (N = constant).

The boundary condition on the basin wall is one of vanishing streamfunction such that the flow is parallel to it:

$$\psi = 0 \quad \text{at} \quad z = 0, z = -\tau h(x). \tag{2.5b}$$

Here

$$\tau \equiv \frac{ND}{\omega L}, \tag{2.6a}$$

which can be interpreted as the scaled period of the monochromatic internal wave (N/ω). Alternatively τ can be viewed as the scaled aspect ratio (depth divided by half-width, D/L) of the basin. Non-dimensional topography is given by $h(x), |x| \leqslant 1$, with $h(\pm 1) = 0$, and, for symmetric topographies, $h(0) = 1$. A scaling like this may seem inconvenient, as for fixed geometry (D, L, N and $h(x)$) the 'depth', τ, changes with changing frequency of the wave and one cannot draw rays of waves having different frequencies in one and the same diagram. This is offset, however, by the advantage that for each frequency wave rays make one and the same angle of 45° with respect to the vertical, which allows quick visual assessment of diagrams. Note that this angle also applies after reflection off sloping boundaries. This makes internal-wave reflection unusual when compared to, for instance, the coastal reflection of obliquely incident surface gravity waves, which obey the specular law of reflection in which the angle of incidence, measured with respect to the coast's normal, equals the angle of reflection. The peculiar nature of reflection of internal gravity waves implies that there exists a critical slope of the topography, as when it equals that of the characteristics. With the non-dimensionalization employed here these slopes are $\pm 45°$, below and above which the waves reflect along or against the original x-direction. For these and other basic aspects of internal wave propagation see e.g. Turner (1973) and Lighthill (1978).

For values of ω as low as $10^{-4}\mathrm{s}^{-1}$ – typical for semi-diurnal tides – the Coriolis frequency f (twice the angular velocity of the Earth multiplied by sine of latitude) can no longer be neglected. Also, for high-frequency waves, non-hydrostatic effects are no longer negligible. Both effects merely lead to a slight change in the definition

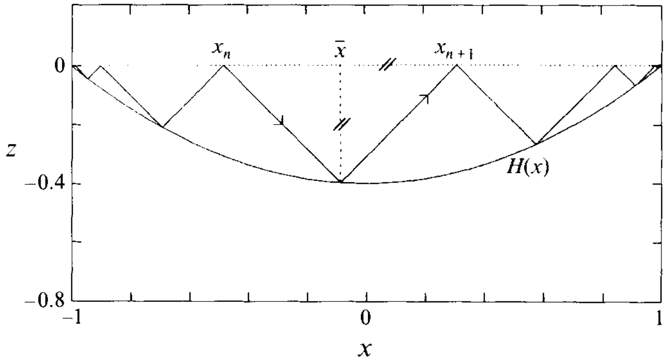

FIGURE 1. Sketch of a uniformly stratified, parabolic basin with subcritical bottom slope, showing the approach of the characteristics towards the corners of the basin for $\tau = 0.4$.

of τ, (Baines 1973):

$$\tau = \left(\frac{N^2-\omega^2}{\omega^2-f^2}\right)^{1/2}\frac{D}{L}, \tag{2.6b}$$

which provides a mapping of the internal wave band ($f < \omega < N$) onto the positive real axis of the scaled period τ. Lakes and oceans are characterized by values of $\tau \approx (0.1\text{–}1)$, based on $f = 5 \times 10^{-5}\mathrm{s}^{-1}$, $\omega = 10^{-4}\text{–}10^{-3}\mathrm{s}^{-1}$, $N = 10^{-2}\mathrm{s}^{-1}$, $D = 10^2\text{–}5\times10^3\mathrm{m}$, $L = 10^4\text{–}2.5\times10^6\mathrm{m}$.

2.2. *Solution with functional relation*

Magaard's (1962, 1968) work is succinctly presented in Sandstrom (1976) which will be followed here. Equation (2.5*a*) is solved by arbitrary complex functions $f_-(x-z)$ and $f_+(x+z)$ of the real characteristic variables $x-z$ and $x+z$:

$$\psi(x,z) = f_-(x-z) + f_+(x+z). \tag{2.7}$$

Applying the surface boundary condition (2.5*b*) shows that the functional forms of $f_\pm$ are related,

$$f_+(x) = -f_-(x) \equiv -f(x),$$

on dropping the subscript. Hence,

$$\psi(x,z) = f(x-z) - f(x+z). \tag{2.8}$$

Let us denote the bottom as $H(x) = \tau h(x)$. Then, application of the boundary condition (2.5*b*) at the bottom $z = -H(x)$ of the basin yields

$$f(x+H(x)) = f(x-H(x)), \tag{2.9}$$

a functional relation for $f(x)$. If successive surface intersections are denoted as $x_n, x_{n+1}, \ldots$ (where n runs over all positive and negative integers), then, from figure 1, it is obvious that

$$\frac{x_{n+1}-x_n}{2} = sH(\bar{x}), \tag{2.10a}$$

where

$$\bar{x} = \frac{x_{n+1}+x_n}{2} \tag{2.10b}$$

and where sign $s = +1, -1$ determines the two modes of the map for rightward and leftward moving characteristics respectively. Equation (2.9), applied at $\overline{x}$, therefore, can be interpreted as

$$f(x_{n+1}) = f(x_n), \tag{2.11}$$

which states that the 'field variable' f is invariant under map (2.10a). In fact f is unchanged along the entire trajectory of reflecting characteristics, such that at any point in the interior, the streamfunction value can be readily obtained as the difference of the values of the field variable on the two characteristics that go through it, see (2.8). Its validity is obvious for any topography that is entirely subcritical (a special case of which is shown in figure 1), since neighbouring characteristics retain their ordering (i.e. for $y_0 > x_0$ one has $y_1 > x_1$). Equation (2.11) equally applies for topographies that are partially supercritical, at least, when f is real, (see §3), which is less obvious because the ordering is destroyed due to back-reflection. The region between two successive surface intersections will be referred to as a *fundamental* interval, since, when we prescribe the field variable $f(x)$ at the surface for $x \in [x_n, x_{n+1})$ then, because of (2.11), $f(x)$ is uniquely determined for all $x \in [-1, 1]$. This definition applies to the subcritical case discussed above. Identification of the fundamental interval for supercritical cases will be addressed in §4.

The 'solution' thus consists of two parts that will be discussed separately in the next two sections: (i) a geometric aspect, that may be captured in the set of successive surface intersections, $S(x_0) = \{..., x_{-2}, x_{-1}, x_0, x_1, x_2, ..\}$ and (ii) the prescription of the field variable $f(x)$ in a fundamental interval $x \in [x_n, x_{n+1})$. The set $S(x_0)$, together with connecting characteristics will be referred to as the *web* belonging to x_0 – a name that is more readily appreciated for the supercritical topographies, discussed in the next section. The largest fundamental interval will be called the *primary* interval. For subcritical topographies, like the one in figure 1, the limiting points of $S(x_0)$ are the corners of the basin:

$$\lim_{n\to\pm\infty} x_n = \pm 1;$$

these constitute the *attractor* of the rightward and leftward 'moving' characteristic respectively. More complicated attractors will be obtained in the next section. Note that in spite of the terminology no real movement towards the attractor can be meant here, since time has been removed from the hyperbolic equation.

The generation of the 'web' is merely a part of constructing the spatial structure of $\psi(x, z)$. Nevertheless, Wunsch (1969), considering internal waves in a subcritical wedge, concludes that the corner of a subcritical topography *does* act as a physical attractor of the internal wave field. He does so on considering some laboratory experiments which led him to re-interpret his earlier theoretical analysis of the problem in which he obtained a standing internal wave pattern (Wunsch 1968). He concludes that only the incoming solution should be physically acceptable, since all of the energy of the internal wave field will be absorbed, because of the intensification of the wave field and subsequent breaking and mixing – and breakdown of the linear theory – that accompany the approach of the corner. Similar results will occur for supercritical basins: a standing wave pattern can in principle be constructed, but, again in view of the intensification of the internal wave field, the energy of the incoming wave will be deposited near the physical location of the attractor. Sandstrom (1976) obtained an explicit solution like that of Wunsch (1968) for a particular closed, subcritical and symmetrical basin. Manton & Mysak (1971) pointed out that the functional relation (2.9) can be used to construct internal wave solutions for arbitrary topographies, a viewpoint that we share. The approach they take, however, is different from that in the following and results should thus be considered to be complementary.

3. Explicit bi-modal map for a parabolic basin

For certain simple basin shapes, $H(x)$, the implicit map, given by (2.10*a*) and (2.10*b*) can be made explicit. In the following we will consider several such topographies taken as (piecewise) linear or quadratic polynomials. They can be classified according to the number of parameters needed to specify them. One parameter (τ) is related to the product of the ratio of the buoyancy frequency and wave frequency and the aspect ratio, see (2.6*a*). Other parameters are sometimes needed to specify piecewise-defined topographies. Apart from the rectangle and the ellipse (that will be discussed in §5) only one other one-parameter topography will be considered: the *parabolic basin*. This will be our main example. To some extent the results obtained in that case are representative for those found for multi-parameter topographies, like the *bucket*. This is a piecewise, linear topography having sloping side walls and a flat bottom in between. However, since some new features arise for these cases they will be given separate attention in §6.

3.1. *Explicit bi-modal map*

When the basin shape is parabolic, $h(x) = 1 - x^2$,

$$H(x) = \tau(1 - x^2), \tag{3.1}$$

map (2.10*a*)–(2.10*b*), with $x \equiv x_n$ and $x_{r,l} \equiv x_{n+1}$ for rightward ($s = +1$) and leftward ($s = -1$) moving characteristics, becomes

$$x_r = -x - \frac{1}{\tau} + \left(\frac{4x}{\tau} + 4 + \frac{1}{\tau^2}\right)^{1/2} \equiv X(x), \tag{3.2a}$$

$$x_l = -x + \frac{1}{\tau} - \left(\frac{-4x}{\tau} + 4 + \frac{1}{\tau^2}\right)^{1/2} = -X(-x), \tag{3.2b}$$

where signs, in front of the radicals, have been chosen such that $x_r > x$ and $x_l < x$. It can be verified that $x_r(x_l(x)) = x$ and vice versa: the right- and leftward maps are each other's inverse, $x_r^{-1}(x) = x_l(x)$. Also, for a symmetrically shaped topography, $x_l(x) = -x_r(-x)$, and hence $x_r(-x_r(-x)) = x$. Because of this we will also denote $x_r(x)$ just by $X(x)$ and $x_l(x)$ by $-X(-x)$.

The topography has maximum slope at its corners, $x = \pm 1$, where it is $\pm 2\tau$. It is therefore everywhere subcritical (i.e. makes an angle with the horizontal which is less than 45°), when $\tau < 1/2$, and these two modes are detached. Rightward moving characteristics end up in the right corner and vice versa (see figures 1 and 2*a*). These are the fixed points (attractors) of the map.

When the topography is supercritical however, the two modes get connected as the corners no longer act as fixed points. The naive map (3.2*a*)–(3.2*b*) formally computes, for some range of x-values, a new surface intersection which lies outside the basin domain, $-1 \leqslant x \leqslant 1$, see figure 2(*b*). For a rightward 'moving' characteristic this happens for $x > x_s$, with

$$x_s \equiv \frac{2}{\tau} - 3$$

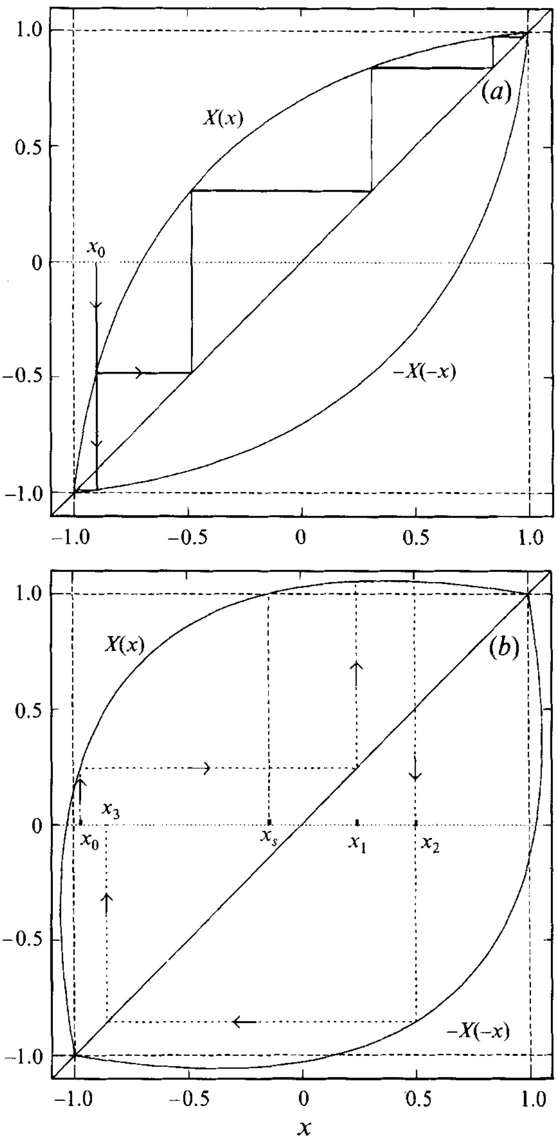

FIGURE 2. (*a*) Subcritical ($\tau < 1/2$) and (*b*) supercritical ($\tau > 1/2$) maps of successive surface intersections of characteristics for rightward (upper curve, $X(x)$) and leftward (lower curve, $-X(-x)$) 'moving' characteristics. Construction of successive surface intersections is shown for one particular value of x_0. The diagonal line is drawn for convenience.

being the point that is mapped onto the right corner (that can be obtained from $x_l(1)$), see figure 3. For values $x < x_s$ the simple forward map applies. For $x > x_s$, however, the new virtual value, X, not only lies outside the basin domain, but also has two pre-images, x and Y say. The latter is in fact the true image of x (see figure 3). Neglecting the virtual points that appear, the sequence $\{x_n\}$ can be constructed graphically as in figure 2(*b*). One often wants the explicit functional dependence, however. This can be obtained as follows. The leftward map of X (the two roots of the quadratic that is obtained from (2.10*a*) with $H(x)$ given by (3.1)), gives the two pre-images

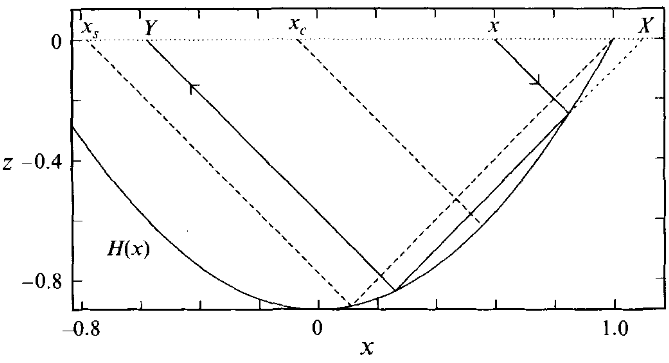

FIGURE 3. Sketch showing the construction of successive surface intersections of characteristics for a supercritically reflecting bottom. The critical characteristic (surface intersection x_c) and characteristic going through the right-hand corner (intersecting at x_s) are also shown.

x and Y:

$$x = \frac{1}{\tau} - X + \left(\frac{-4X}{\tau} + 4 + \frac{1}{\tau^2}\right)^{1/2}, \tag{3.3a}$$

$$Y = \frac{1}{\tau} - X - \left(\frac{-4X}{\tau} + 4 + \frac{1}{\tau^2}\right)^{1/2}. \tag{3.3b}$$

Adding these yields

$$Y = \frac{2}{\tau} - x - 2X(x), \tag{3.4}$$

where $X(x)$, the inverse of (3.3*a*), is given by (3.2*a*). In figure 3, two regions in the interval $x \in [x_s, 1]$ can be recognized which determine whether leftward reflection occurs for a characteristic coming from below, or from above. The dividing line is the *critical characteristic* (which intersects the bottom at the point where the bottom is critical). Its intersection with the surface is at

$$x_c \equiv \frac{3}{4\tau} - \tau.$$

Physically, internal waves propagating along that critical characteristic tend to be mainly dissipated (due to breaking resulting from strong amplification), see Cacchione & Wunsch (1974) and Ivey & Nokes (1989). The mathematical approach pursued in this section, however, merely aims to construct webs of characteristics without implying anything about the physical fields carried along them. It considers construction of the critical characteristic as a limiting process. For characteristics approaching the critical characteristic the reflected ray resides just at the other side of it. Thus reflection on the critical characteristic itself should result in complete back-reflection along that same ray, from which x_c is obtained as fixed point of map $Y(x)$, i.e. x_c satisfies $Y(x_c) = x_c$.

Leftward reflection of an initially rightward moving characteristic should be accompanied by a sign change of s, indicating that one should shift to the leftward map. For a supercritical topography then, the complete bi-modal map, $\boldsymbol{T}(x,s) \equiv (T_1(x,s), T_2(x,s))$, is specified by two parameters giving the new surface intersection, $T_1(x,s)$, as well as

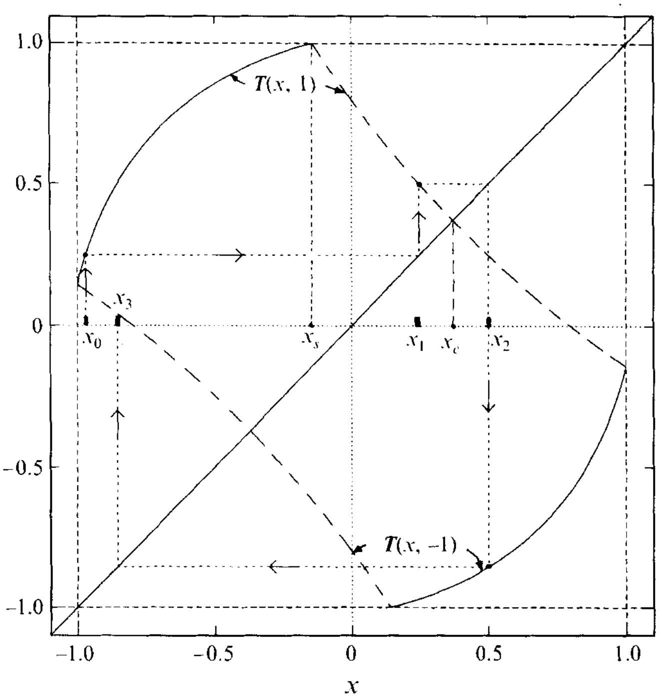

FIGURE 4. Bi-modal map for $\tau = 0.7$ with successive surface intersections x_n, $n \in \{0,1,2,3\}$. The rightward (leftward) map is given by the upper (lower) curve; the solid (dashed) part of it indicates that the sign is unchanged (changed). Short-dashed lines give graphical construction of successive surface intersections.

the new sign, $T_2(x,s)$, where $T_2 \in \{-1, 1\}$:

$$\boldsymbol{T}(x,s) = \begin{cases} (X(x), s) & \text{if } s = +1, \quad -1 \leqslant x \leqslant x_s \\ (Y(x), -s) & \text{if } s = +1, \quad x_s \leqslant x \leqslant 1 \\ (-X(-x), s) & \text{if } s = -1, \quad -x_s \leqslant x \leqslant 1 \\ (-Y(-x), -s) & \text{if } s = -1, \quad -1 \leqslant x \leqslant -x_s. \end{cases} \tag{3.5}$$

Alternatively, the map can be written as $(x_n, s_n) = \boldsymbol{T}^{(n)}(x_0, s_0)$, where x_0 and s_0 indicate the initial position and direction of the ray and $n > 0$ (< 0) relates to $s_0 = +1$ (-1). The map is plotted for a particular value of τ in figure 4. The graphical construction of successive surface intersections is a slight variation of the usual procedure in iterated maps (e.g. Schuster 1984) owing to the bi-modality of the map. For a given x_0 one might read off x_1 from the graph and then read off x_2 etc. a process that is reduced by reflection in the diagonal. When an $x \in [x_s, 1)$ is obtained for initially rightward motion (upper curve) one should shift to the leftward mode (lower curve), and vice versa when $x \in (-1, -x_s]$. This corresponds to the dashed parts of the map in figure 4 and indicates that sign changes occur. The solid parts indicate that no sign change occurs. In the remainder of this paper dashing of branches on which the map changes sign will be suspended, on the understanding that sign changes will still occur according to the definition above.

In this paper τ-values will be restricted by the arbitrary, additional requirement that there is at least one point that is mapped simply forward, $x_s \geqslant -1$, or $\tau \leqslant 1$. This restriction is made just for the sake of simplicity, since now characteristics reflect

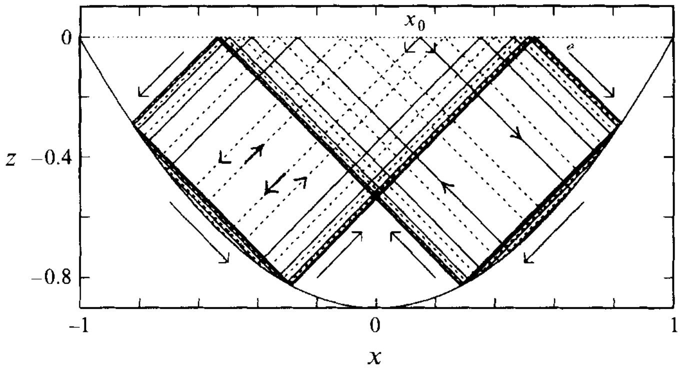

FIGURE 5. Construction of web for $\tau = 0.9$ and $x_0 = 0.15$ by iterated mapping. Right- and leftward 'moving' characteristics are drawn as solid and dashed lines respectively. The final sense in which the limit cycle is traversed has been indicated by arrows.

from the bottom at most twice prior to reaching the surface. Construction of the map for larger values of τ can be done along the lines indicated in the Appendix and is made explicit there for $1 \leqslant \tau \leqslant 3/2$.

3.2. *Construction of web by iteration of the map*

Given a single position x_0, the complete web, $S(x_0)$, can be constructed by forward and backward iteration of the map, following the characteristics passing through that point both in rightward and leftward directions. In this way, for a particular value of τ, the web shown in figure 5 is constructed. It is observed that the rays are rapidly attracted towards a limit cycle, that can be characterized by the number of surface intersections it has. This number is referred to as the *period* of the attractor in accordance with the usage in dynamical systems. There should be no confusion with the period of the wave (which, in scaled form, appears here as the central parameter τ) in (2.6*a*). Thus, for this particular example, the period of the attractor is two. Surprisingly, for this value of τ, this limit cycle is the only one present. Irrespective of the value x_0 the same limit cycle is reached. This applies both for characteristics 'initially' moving to the right as well as for those moving to the left. This insensitivity to initial position and direction is a consequence of the symmetry of the final attractor.

For odd-period attractors there are two separate limit cycles (which are each other's images when mirrored in the line $x = 0$). The limit cycle that will be reached depends on starting position, x_0, as well as on direction, s_0. For the 3-cycle this relation of initial values (x_0, s_0) to the 'final state' of the characteristics is illustrated in figure 6. An arbitrary (but typical) value of the scaled period ($\tau = 0.72$) within the period-3 interval is sketched in figure 6(*a*). The attractor with two negative and one positive surface intersection will be called the positive attractor (solid line), since the product of these three values is positive. Correspondingly its mirror image is called the negative attractor (dashed line). The bars in the upper part of figure 6(*a*) show whether the positive (black) or negative (white) attractor is reached for different starting values x_0. The upper (lower) bar corresponds with rightward (leftward) starting characteristic, $s_0 = +1$ (-1).

Owing to the combination of starting value and rightward/leftward direction, there are four possible final states:

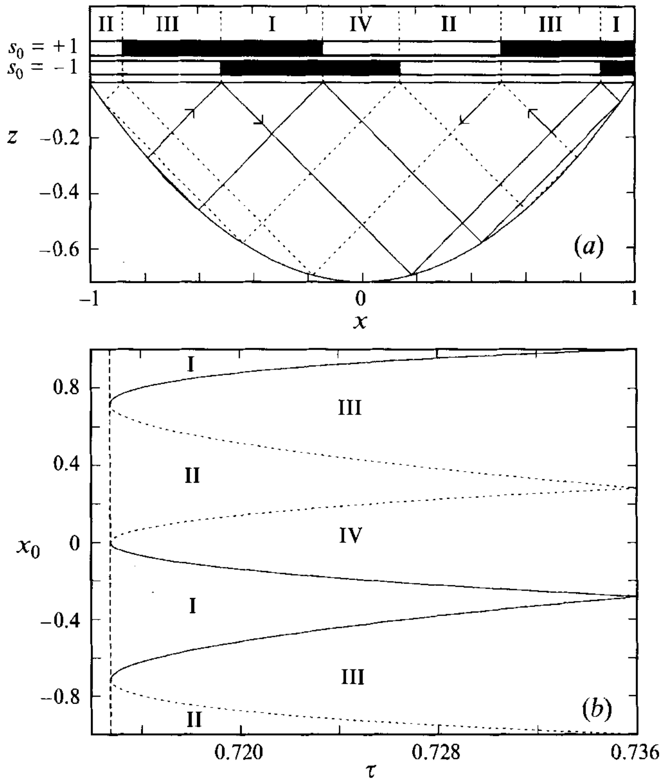

FIGURE 6. The two possible attractors are shown (a) for a typical period ('depth') τ in the 3-cycle interval ($\tau = 0.72$). The horizontal bars on top of the figure show which final state is reached for all possible initial values x_0 and directions s_0. The upper (lower) bar corresponds to rightward (leftward) initial direction, $s_0 = +1$ (-1). Black (white) bars denote the solid (dashed) attractor as final state. The four possible combinations of the attractors reached (both bars black/white and two combinations) are denoted as regions I–IV, and are explained in detail in the text. In (b) the locations of these regions are given for the whole 3-cycle interval $0.715\ldots < \tau < \sqrt{5} - 3/2 = 0.736\ldots$.

(I) the positive attractor (solid line) is reached for both starting directions (both bars black),

(II) the negative attractor (dashed line) is reached for both starting directions (both bars white),

(III) the positive attractor is reached for a rightward ($s_0 = +1$) start, while the negative attractor is reached for a leftward start (upper bar black, lower bar white),

(IV) the reverse of case III: upper (lower) bar is white (black).

In figure 6(b) the four defined regions are given for the whole 3-cycle interval.

The approach of the limit cycle can also be appreciated from successive iterations directly in a graph of the map. Figures 7(a) and 7(b) give examples of a 2- and 3-cycle respectively. In the latter figure the initial position is such that two different attractors are reached for rightward (solid) and leftward (dashed) moving characteristics (state IV).

It has been mentioned that the web is to be considered a spatial structure and that, in spite of the terminology used, the iterative procedure, by which it is constructed, should not be viewed as a temporal process. The global convergence of all webs

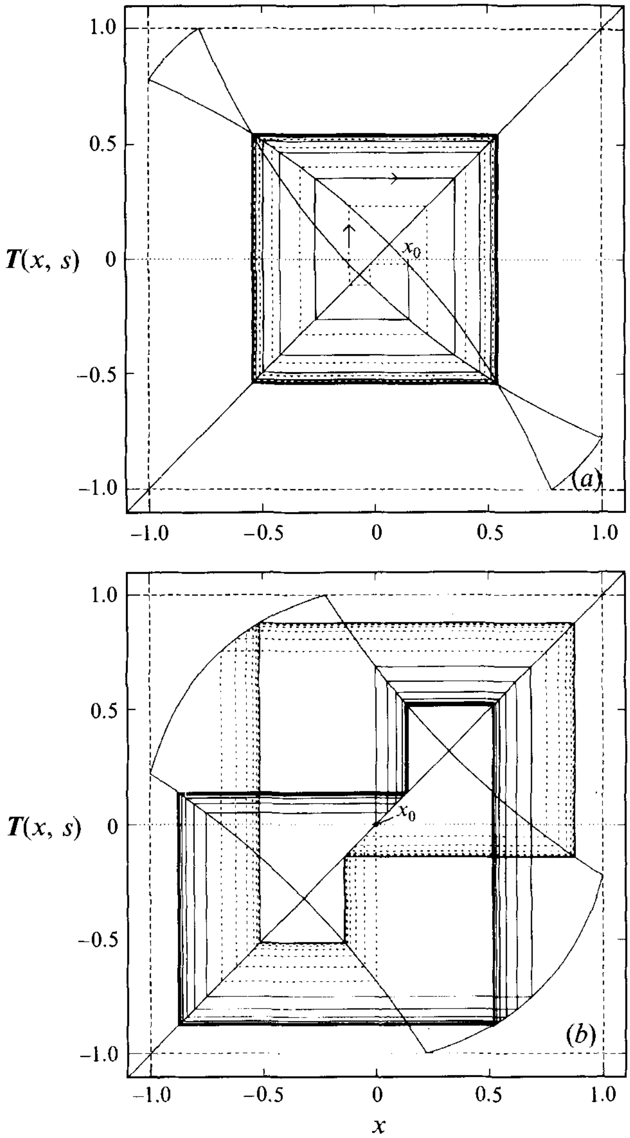

FIGURE 7. Successive mappings in the case of (*a*) a 2-cycle, $x_0 = 0.15$ for $\tau = 0.9$ and (*b*) a 3-cycle, $x_0 = 0$, $\tau = 0.72$. Solid (dashed) lines are used for initially rightward (leftward) moving characteristics.

towards a limit cycle for sufficiently often iterated maps suggests, however, that when the field variable that is 'advected' along the characteristic is complex, this may nevertheless be interpreted as propagation along the characteristic. In particular this then implies that distributed fields tend to get focused along the limit cycle. This focusing process appears to be generic, and happens irrespective of the precise value of the field variable itself. This geometric effect, therefore, seems to be the most important factor determining the complete solution.

3.3. *Asymptotic state(s) as a function of* τ

The limiting characteristics (limit cycles) can succinctly be summarized by their surface intersections (limit points): a Poincaré section. Recall that for $\tau < 1/2$ the two corners

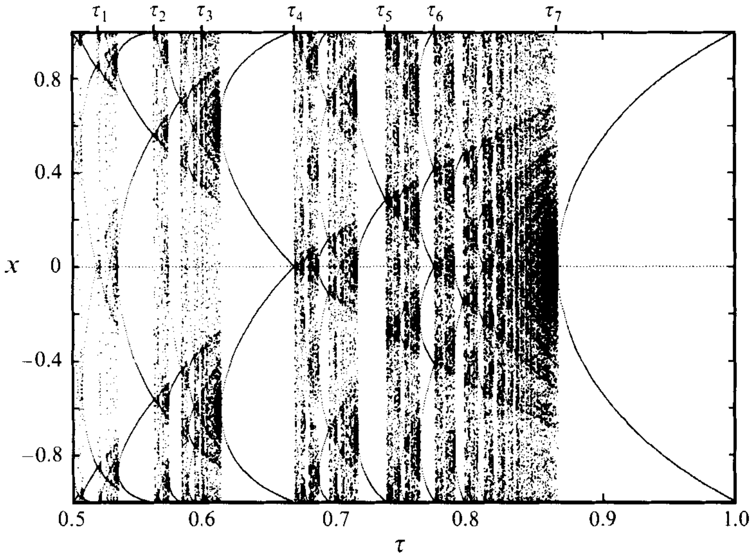

FIGURE 8. Poincaré plot of $x_{900}-x_{1100}$ of map (3.5) for $x_0 = 0.123456789$ and $s_0 = +1$ in the interval $1/2 \leqslant \tau \leqslant 1$ where τ is incremented with 1/1600 of this interval. Indicated at the top are some special values of the map parameter that can be computed algebraically from the lines in figure 10: $\tau_1 = (9-\sqrt{41})/5$, $\tau_2 = (\sqrt{17}-3)/2$, $\tau_3 = \sqrt{\frac{3}{8}}$, $\tau_4 = 2/3$, $\tau_5 = \sqrt{5}-3/2$, $\tau_6 = 2(\sqrt{137}-9)/7$, $\tau_7 = \sqrt{3}/2$.

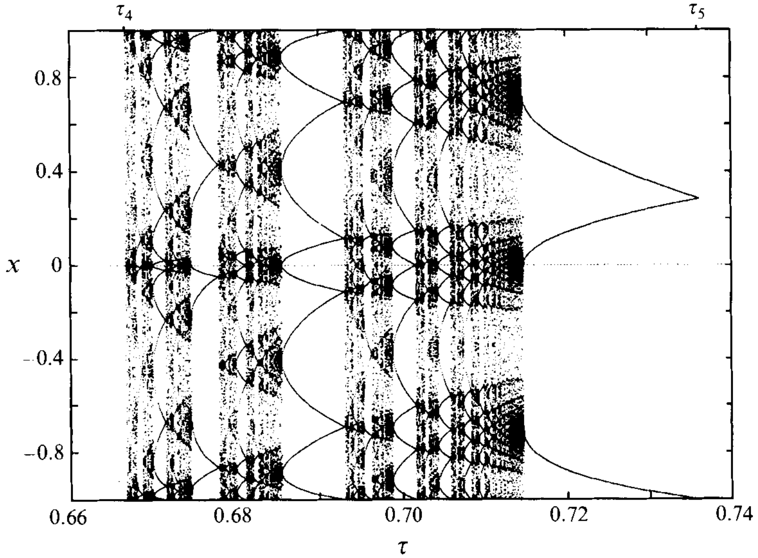

FIGURE 9. Expansion of figure 8 by employing 1600 points to cover the interval $2/3 \leqslant \tau \leqslant \sqrt{5}-3/2$ using the same initial value x_0.

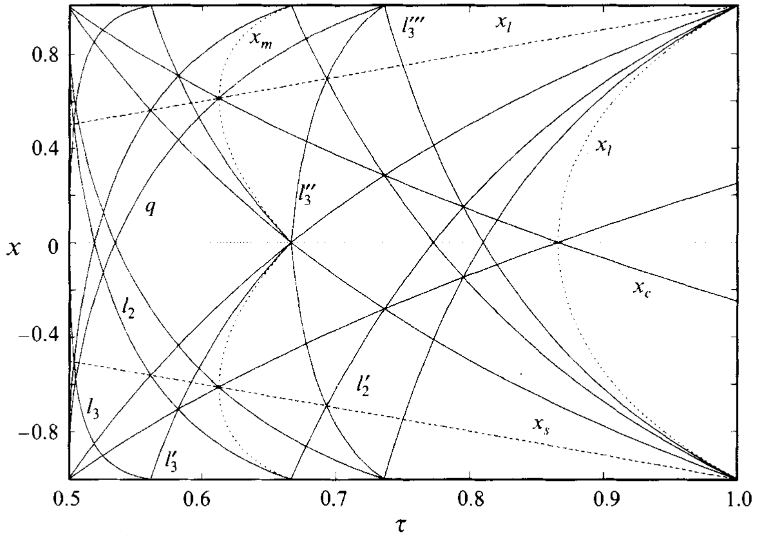

FIGURE 10. 'Skeleton' of figure 8. The labels on the lines refer to those in table 1.

$x = \pm 1$ are the two limit points. In figure 8 Poincaré sections have been plotted for a sequence of τ-values by taking one particular x_0, iterating that along the initially rightward direction, $s_0 = +1$, a large number of times (here 1100) and plotting the last few hundred (here 200) iterates. This plot will be referred to as a Poincaré plot. This figure shows that there is a complicated dependence of the period of the attractor on the map parameter τ – the scaled period of the internal wave. Regular windows, in which the attractor period stays constant and the limit points gradually move out, are interrupted by high-period regions. These high-period windows, in turn, appear to have, at a finer scale, a similar fractal-like division in high- and relatively low-period windows (figure 9). None of these windows contains chaos, however, as will become clear when one considers Lyapunov exponents (see below). For increasing values of τ each of these windows undergoes a kind of bifurcation towards a point where the period increases indefinitely. No regular, period-doubling bifurcation is obtained in this case though. Because only the rightward direction has been traced here, asymmetric structures appear for odd-period attractors, their mirrored parts being obtained for other initial values and/or directions.

3.3.1. *Skeleton of the Poincaré plot*

Some of the 'lines' that can be discerned in figure 8 can, in fact, be related to the two 'special points', $x_c(\tau)$ and $x_s(\tau)$, defined previously. The latter one is the leftward image of the corner point $x = 1$, i.e. $-X(-1)$. Likewise, some of the other lines in figure 8 consist of points that are pre-images of the corner points. This is shown in the *skeleton* of the Poincaré plot, figure 10. From the intersections of these lines, the τ-values that specify the borders of some of the windows can be calculated algebraically (see caption of figure 8). The functions appearing in figure 10 are given in table 1, along with some other frequently used functions.

The distances between the successive windows, converging at $\sqrt{3}/2$, do not seem to converge at the Feigenbaum rate (Schuster 1984), as might be expected at first. This has not been further elaborated yet.

Function	Definition	Expression	τ-interval
x_s	$-X(-1)$	$\frac{2}{\tau}-3$	$\left(\frac{1}{2},1\right)$
x_c	$Y(x_c)-x_c=0$	$\frac{3}{4\tau}-\tau$	$\left(\frac{1}{2},1\right)$
x_l	$Y(-Y(-x_l))-x_l=0$	$\pm\left(4-\frac{3}{\tau^2}\right)^{1/2}$	$\left(\frac{1}{2}\sqrt{3},1\right)$
x_m	$Y(X(x_m))+x_m=0$	$\pm\left[\frac{3}{5}\left(\tau-\frac{1}{\tau}\right)\pm\left(\frac{8}{5}-\frac{3}{5\tau^2}\right)^{1/2}\right]$	$\left(\frac{1}{2}\sqrt{\frac{3}{2}},\frac{2}{3}\right)$
q	$X(-x_c(\tau))$	$\left(8-\frac{2}{\tau^2}\right)^{1/2}-\frac{1}{4\tau}-\tau$	$\left(\frac{1}{2},\sqrt{5}-\frac{3}{2}\right)$
l_2	$-X(-x_s)$	$3-\frac{1}{\tau}-\left(4-\frac{7}{\tau^2}+\frac{12}{\tau}\right)^{1/2}$	$\left(\frac{1}{2},\frac{2}{3}\right)$
l_2'	$-Y(-x_s)$	$-3-\frac{2}{\tau}+2\left(4-\frac{7}{\tau^2}+\frac{12}{\tau}\right)^{1/2}$	$\left(\frac{2}{3},1\right)$
l_3	$-X(-l_2)$	$-3+\frac{2}{\tau}+\left(4-\frac{7}{\tau^2}+\frac{12}{\tau}\right)^{1/2}+$ $-\left(4+\frac{5}{\tau^2}-\frac{12}{\tau}+4\frac{\left(4-\frac{7}{\tau^2}+\frac{12}{\tau}\right)^{1/2}}{\tau}\right)^{1/2}$	$\left(\frac{1}{2},\frac{1}{2}\left(\sqrt{17}-3\right)\right)$
l_3'	$-Y(-l_2)$	…	$\left(\frac{1}{2}\left(\sqrt{17}-3\right),\frac{2}{3}\right)$
l_3''	$X(l_2')$	…	$\left(\frac{2}{3},\sqrt{5}-\frac{3}{2}\right)$
l_3'''	$Y(l_2')$	…	$\left(\sqrt{5}-\frac{3}{2},1\right)$
x_t	$X(x_t)+x_t=0$	τ	$\left(\frac{1}{2},1\right)$
$X(x)$	$(X-x)/2=\tau(1-(X+x)^2/4)$	$-\frac{1}{\tau}-x+\left(4+\frac{1}{\tau^2}+\frac{4x}{\tau}\right)^{1/2}$	
$Y(x)$	$\frac{2}{\tau}-x-2X(x)$		

TABLE 1. Definition and expressions for lines indicated in figure 10.

3.3.2. *Lyapunov exponents*

Even though the windows in figure 8 are reminiscent of the chaotic regions in the logistic map (Schuster 1984), they are nevertheless very different. Chaos is associated with divergence of nearby trajectories characterized by positive Lyapunov exponents. Lyapunov exponents for the bi-modal map in a parabolic basin, however, are always less then zero (within numerical precision).

In figure 11 the convergence rate with which the limiting characteristics are approached has been quantified by calculating Lyapunov exponents. The Lyapunov exponent, λ_+, is defined as

$$\lambda_+ = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \ln\left|\frac{\mathrm{d}T_1(x_n, s_n)}{\mathrm{d}x}\right|,$$

where $T_1(x,s)$ is the first part of map (3.5) and x_n and s_n denote the nth iterates, starting from x_0 and $s_0 = +1$. The associated Lyapunov exponent, λ_-, can be obtained by starting with $s_0 = -1$. The Lyapunov exponent is, in principle, a function of the starting position x_0. But, typically, the same λ_+ is obtained for almost any x_0. Also, the degree to which the true Lyapunov exponent is approximated depends on N in a *non-monotonic* (typically oscillatory) way. Therefore some averaging should, in principle, be performed, although for N large this can be safely ignored.

The Lyapunov exponent measures the total convergence or divergence along a characteristic. It is observed that figure 11(a) mimics certain aspects of the Poincaré plot (figure 8). The latter figure is summarized by calculating the period of the attractor, P say, by determining the number of iterations for which asymptotically (N large) a certain x_N recurs with sufficient accuracy, ϵ, i.e. the smallest integer $P \in \mathbb{N}$ for which $|x_{N+P} - x_N| < \epsilon$. Here $N = 900$ and $\epsilon = 10^{-7}$ have been used. For visual similarity with figure 11(a),

$$\nu \equiv -1/P$$

is plotted, see figure 12. The graph of the Lyapunov exponent shows that the map is strongly attracting for regular (low-period) regions of the Poincaré plot like in the regions with period 2, 3 and 4. In between, the curve is less negative and in particular seems to reach zero at some discrete set of points. A blow-up (figure 11b,c) demonstrates that self-similarity also appears in the Lyapunov exponent, a self-similarity that can be discerned in the (inverse) period, figure 12(b,c), too.

Vanishing of the Lyapunov exponent means that all points retain their mutual distances. This can either happen when neighbouring points are all situated at distinct limit-cycles of the same finite period (a situation encountered in the rectangle, considered in §5), or when they migrate in unison. In the latter case, however, it implies that the attractor has infinite period. This is the situation occurring in the parabolic basin.

Note that overall negativeness of Lyapunov exponents implies that the bi-modal map is, in the terminology of dynamical systems, *dissipative*. This happens despite the fact that the physical model is inviscid.

3.4. *Integral quantification of webs*

Figures 8 and 9 are unsatisfactory as a classification of entire webs as they concentrate on just the asymptotic part of them, corresponding to the limit cycle.

Also, in figures 11 and 12 the Lyapunov exponent as well as the period of the web have been given for one particular starting value x_0, i.e. for one single web. To do more 'justice' to each complete web one needs integral measures to characterize them (as a function of x_0). One such measure, the sum of iterates

$$\mu_N(x_0) = \sum_{n=-N}^{N} x_n, \tag{3.6}$$

has been employed here, see figure 15. This is similar to the Poincaré plot, figure 8, except that it contains 'information' about the whole parameter plane. For most values in the (x_0, τ)-plane this quantity is independent of N (for large values of it)

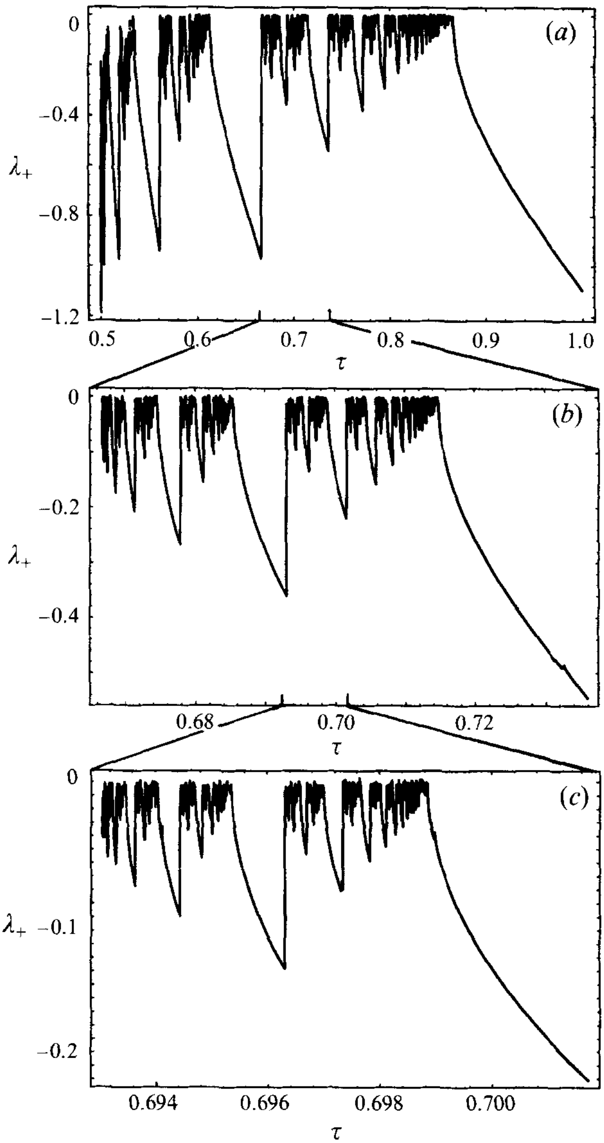

FIGURE 11. (*a*) Lyapunov exponent, λ_+, as a function of τ, with (*b*, *c*) two successive enlargements. For each graph 1200 τ-values have been used.

owing to the existence of symmetric limit cycles. Particularly if, for n large, $x_{-n} = -x_n$, this sum of iterates stabilizes. For other values of τ and x_0 this antisymmetry of the iterates does not exist and a stable value of the sum in (3.6) would be obtained only after averaging over the period, M say, of this cycle. Such an averaging has not been done in figure 15, though. Similarly, for odd-period attractors that have their x_0-values in intervals for which forward and backward iteration leads to different (mirrored) limit cycles (figures 6 and 7*b*), cancellation occurs between terms with index n and $-n$. The only contribution to the sum in (3.6) thus comes from small (absolute) values of the index n and stays approximately in the range $(-1, 1)$, see figure 15. Only for τ-intervals with odd-period attractors, which have x_0-regions that reach the *same* asymmetric attractor for rightward and leftward iteration, there is a

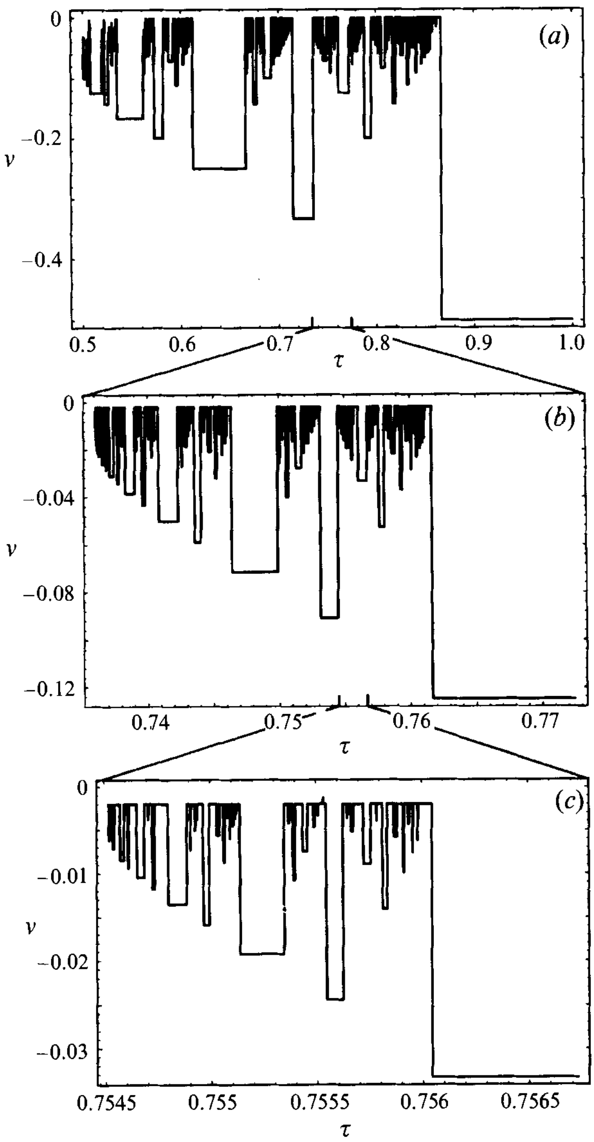

FIGURE 12. $v = -1/P$, related to the period of the attractor, P, for $N = 1200$ arbitrarily truncated when the period is in excess of 400, as a function of τ with (*b*, *c*) two successive enlargements; (*c*) suffers from numerical convergence problems, though.

net 'drift' (N-dependence) of μ_N (the dark regions in figure 15). The sum of iterates is an odd function of x_0 owing to the symmetry of the topography.

3.4.1. *Conjecture: two-parameter independence of infinite sum of iterates*

There exist regions where μ_N is constant to within round-off error. These are particularly visible as the white regions in figure 15 (where $\mu_N \approx 0$). This allows us to conjecture the following, surprising two-parameter independence of an 'infinite' sum

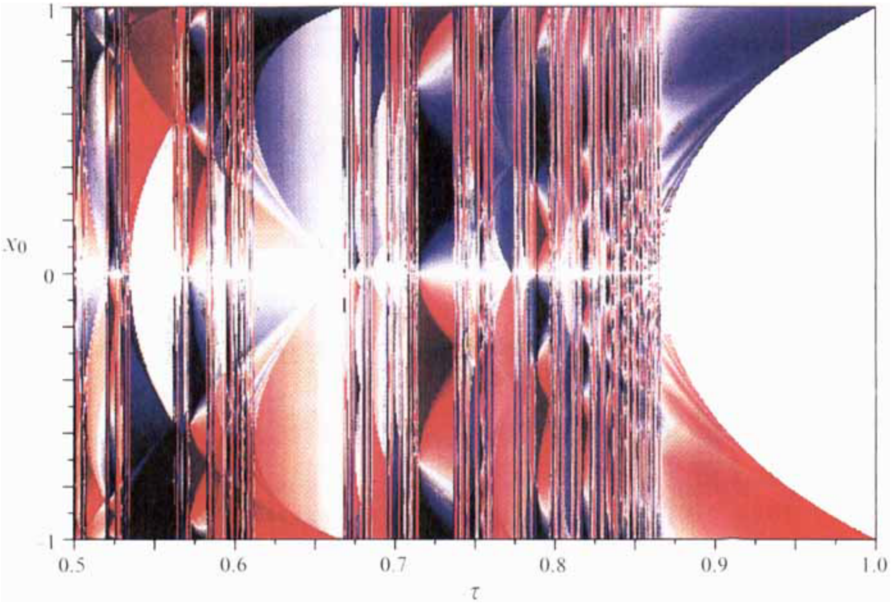

FIGURE 13. For caption see facing page.

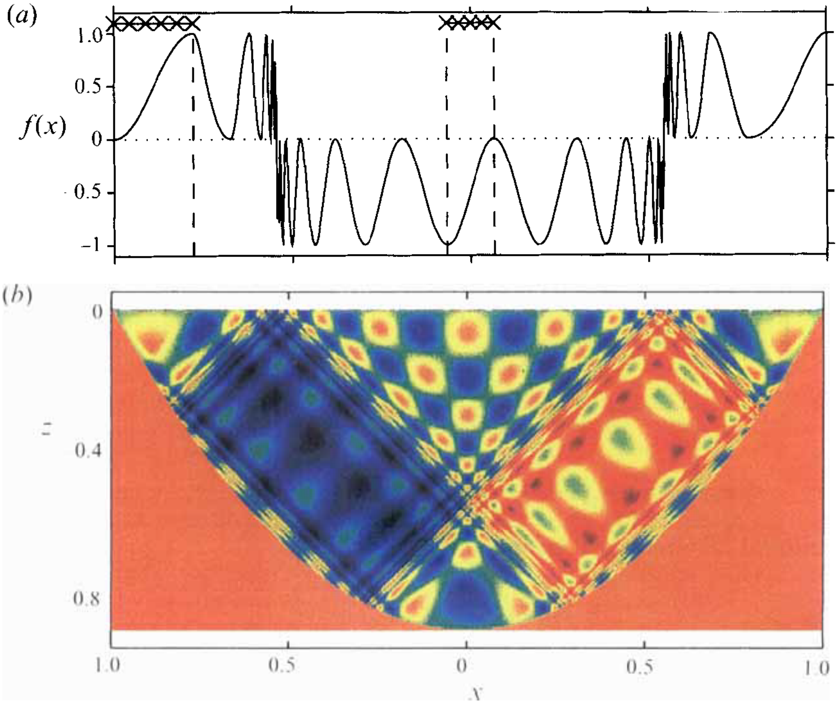

FIGURE 15. For caption see facing page.

of iterates:

$$\sum_{n=-\infty}^{\infty} x_n(x_0,\tau) = 0 \quad \left\{\forall x_0,\tau \mid |x_0| \leqslant \left(4-\frac{3}{\tau^2}\right)^{1/2}, \tau \in \left(\frac{\surd 3}{2},1\right)\right\}.$$

Since, for the region indicated, sign changes occur at *every* reflection from the bottom (see figure 7), just the second and fourth alternatives of (3.5) apply. Hence this can also be rewritten as

$$\lim_{N\to\infty} T_N(x,\tau) = 0, \quad \left\{\forall x,\tau \mid |x| \leqslant \left(4-\frac{3}{\tau^2}\right)^{1/2}, \tau \in \left(\frac{\surd 3}{2},1\right)\right\},$$

where

$$T_N(x,\tau) = x + \sum_{n=1}^{N}(-1)^n(g^{(n)}(x,\tau) - g^{(n)}(-x,\tau)),$$

with a recursively defined $g^{(n)}(x,\tau)$:

$$g^{(n+1)}(x,\tau) = g(g^{(n)}(x,\tau),\tau),$$

and

$$g^{(1)}(x,\tau) = g(x,\tau) \equiv -\frac{4}{\tau} - x + 2\left(4+\frac{1}{\tau^2}+\frac{4x}{\tau}\right)^{1/2}.$$

No rigorous proof for the validity of this conjecture has been obtained as yet, however. Figure 15 suggests that similar algorithms should exist in other regions of the parameter plane where μ_∞ approaches a constant. No formulation of these has been attempted, though.

3.5. *Relation to a billiard*

The construction of the web of reflecting characteristics can be viewed as an alternative to the classical billiard problem (Berry 1981). A 'billiard' is defined as a closed region of the plane for which the trajectory of a point particle is studied. The particle reflects elastically according to the law that the angle of reflection equals that of incidence with respect to the normal to the boundary at the point of incidence. Successive bounces label the orbit of the particle and can be described by the distance along the boundary and the angle of incidence. This constitutes a mapping of a (related) two-dimensional parameter space onto itself. Three types of behaviour are encountered: first, periodic motion when an orbit closes onto itself; second, motion in parameter space along an invariant curve; and third, chaotic motion in which part of the parameter plane is traversed. The actual behaviour depends on the shape of the boundary and the particular aspect ratio it has.

For acoustic waves, the spatial structure is determined by an (elliptic) Helmholtz

---

FIGURE 13. Sum of iterates, μ_N for map (3.5) as a function of τ (601 points) and initial position of web, x_0 (601 points). Number of iterates is $2N+1 = 399$. About 95% of the values of μ_N reside in the -1 (bright blue) to $+1$ (bright red) range. White indicates a zero value of μ_N.

FIGURE 15. (*a*) Function $f(x)$, specified in the two primary fundamental intervals (hatched parts of x-axis indicated at the top, corresponding to those indicated on the right-hand dashed line of figure 14*a*), and subsequently calculated values of $f(x)$ in remaining parts of domain for $\tau = 0.9$. (*b*) Spatial structure of streamfunction field, $\psi(x,z)$, obtained from $f(x)$ with (2.8). Zero value of the streamfunction field is indicated with green. Values range from -2 (dark blue) to $+2$ (dark red).

equation. For a monochromatic wave in the WKB-approximation, there exists a one-to-one correspondence with the billiard problem (Abdullaev 1993), in which the wave rays obey the specular law of reflection. In particular this implies the possibility of chaos in ray dynamics. The internal wave rays, determined by the map in (3.5), however, are more constrained, since a different reflection law operates: motion can occur only in two directions (labelled by s). Thus we are dealing with a reduced parameter space. Apparently, as a consequence, a different type of behaviour – focusing of trajectories – is observed. Negative Lyapunov exponents have not been encountered in the standard billiard. Conversely, in the internal wave problem, no *positive* Lyapunov exponents – the hall mark of chaos – have been observed. The two types of billiard thus seem to have complementary features.

4. Standing internal-wave patterns

Having obtained the geometrical structure of the rays we are now in a position to specify the field variable $f(x)$ on one or more independent fundamental intervals at the surface. Having specified this the function can be determined completely at each point of the surface domain. With this specification, according to (2.8), the value of the streamfunction field at any point within the basin can be readily obtained as the difference of the f-values carried along the characteristics which intersect at this point.

The first question we have to address is whether we can specify non-overlapping fundamental intervals. This question has here been solved by inspection in two simple cases: in τ-intervals with asymptotic cycles of period 2 and 4. Consider in figure 14(a) the vertical line at the right, in the region with asymptotic 2-cycles. The two hatched parts of that line indicate the two independent intervals on which $f(x)$ can be arbitrarily specified. Figures 14(b) and 14(c) show that rays coming from these regions are mutually exclusive and plane filling. These constitute two separate domains of attraction even though the attractor itself is the same. The 'inner' domain, figure 14(b), is seen to be affected by just a particular part of the bottom and it is independent of any deviations that the bottom might exhibit in the outer (or central) region, like a flattening of the bottom, characteristic of near-shore shoaling. Similarly, the 'outer' domain, figure 14(c), is unaffected by the shape of the bottom in the intermediate regions (that is, as long as the bottom does not intersect any of the wave rays above it). Notice that there are regions in which the orthogonal rays come from both domains. Conversely there are complementary regions for which one ray comes from the inner and one from the outer domain. It is also remarkable that the critical characteristics, situated in the inner domain, act as *repellors*, thus downplaying the relevance of the failure of linear, inviscid theory in that case (Cacchione & Wunsch 1974; Ivey & Nokes 1989). Consistent with existing theory, however, it is observed that downward reflection from the supercritical part of the sloping bottom always leads to convergence of wave rays (focusing of characteristics), a focusing which is partly offset by subsequent reflection from the bottom leading to divergence of wave rays. The net effect of focusing and defocusing is, owing to a larger 'scattering cross-section' – the interval-size over which focusing/defocusing extends – necessarily dominated by the former, so that net convergence of wave-rays is the rule. This will be more explicit for the 'bucket'-topography, considered in §6.

In figure 15(a) an example is given in which $f(x)$ is specified to be a sine with an offset in the two fundamental intervals. The offset has been chosen of a different sign

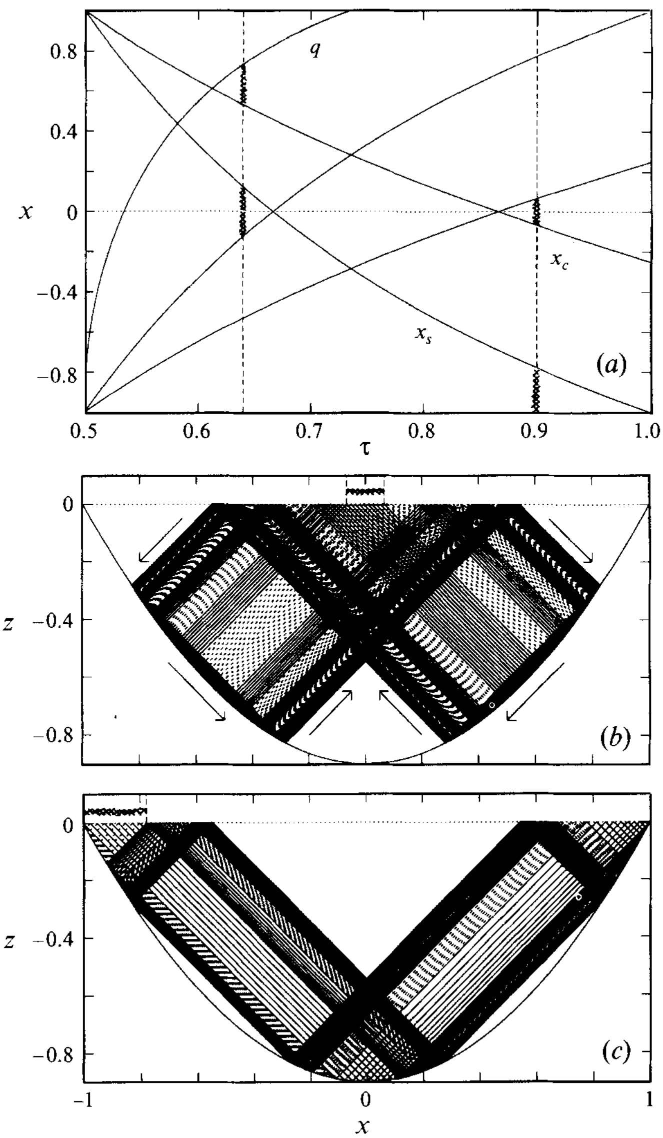

FIGURE 14. (*a*) Selection of the 'skeleton', figure 10, of the Poincaré plot for the parabolic basin with two vertical lines in the regions where 2- and 4-cycles exist. Hatched parts of these lines indicate primary (fundamental) intervals for this τ, where function $f(x)$ can be arbitrarily specified. (*b*) Rays coming from the inner and (*c*) outer primary fundamental intervals in the specific 2-cycle case. The location of the primary intervals has been indicated at the top of these last two figures. In the latter case this is ambiguous as the mirrored interval might also be adopted as the primary interval.

in the two regions. Based on its prescribed value in the primary fundamental intervals $f(x)$ has been determined for all $x \in (-1, 1)$ with the aid of (3.5), see figure 14(*b*, *c*). From this graph the standing wave pattern $\psi(x, z)$ has been obtained, Figure 15(*b*).

On the vertical line at the left in figure 14(*a*) the regions have been indicated where primary fundamental intervals reside in the case of a 4-cycle. Again, just two independent intervals arise, suggesting this to be true for each even-period attractor.

In contrast, preliminary analysis shows that for odd-period attractors three such regions exist.

Because of the fact that the fundamental intervals constitute finite-sized domains an arbitrary function $f(x)$ can be specified on these intervals by a Fourier series, which, depending on the symmetry or antisymmetry of this function, is given by a cosine or sine series. The two classes of 'solution' constructed in this way are similar and comparable to (though not as explicit as) those obtained by Wunsch (1968) for a wedge. That is, they represent 'blinking', standing internal-wave patterns. Wunsch (1969) argued, however, that laboratory observations, as well as field experiments, showed that internal wave energy (for the subcritical wedge) only approaches the wedge and does not return to form a standing pattern. This is more adequately described by a field of waves whose phase and group velocity have a component in the direction of the corner, such as obtained by a linear combination of the two standing wave solutions. In a similar vein, here too it is unlikely that standing wave patterns are obtained as there will be no 'reflection' from the attractor. It is therefore necessary too in this case to construct a propagating wave pattern that has phase and group velocity which have components approaching the attractor. Propagating wave solutions have not been addressed in this study though.

The solutions discussed above are in a sense 'free' or unforced solutions. Forcing of internal waves can be due to a variety of mechanisms, see Krauss (1973). In this study we will, for the sake of simplicity, restrict ourselves to forcing by pressure variations at the surface, because a specification of the pressure means a direct specification of $f(x)$. From the description of u by means of a streamfunction we find, non-dimensionally, $u(x,0,t) = 2f'(x)\exp(-it)$. Vanishing of the horizontal velocity field in the corners therefore requires $f'(\pm 1) = 0$. With (2.1a), applied at the surface, $z = 0$, we find that $f(x)$ is directly related to the pressure which is assumed to be given by the air pressure, $p_a(x)$:

$$f(x) = -\frac{\mathrm{i}}{2}p_a(x). \tag{4.1}$$

Note that the imaginary unit implies an out-of-phase relationship of pressure and f. This, however, demonstrates the paradoxical nature of this kind of specification, because, apparently, is one free to specify surface pressure only in one, or two disconnected fundamental intervals. Specifying the pressure, proportional to $f(x)$, over the whole interval must inevitably lead to inconsistencies. The paradox is resolved by concluding that only when such inconsistencies do not arise is one able to construct stationary solution-patterns for the frequency under consideration, but that otherwise, one necessarily has to employ internal-wave solutions that propagate away from the forcing area. This still assumes the forcing to be stationary. The solution of a true initial value problem for closed basins is further complicated by the fact that the wave field has to satisfy a radiation condition to guarantee causality (Baines 1971a), a problem which indeed has, to our knowledge, not even been solved for the rectangle.

5. Explicit solutions: the rectangle and semi-ellipse

Having obtained a solution of the spatial hyperbolic equation (2.5a), with boundary conditions (2.5b), in the parabolic domain by solving the functional equation (2.10a)–(2.10b) makes one wonder what the ray method yields in a geometry for which solutions can also be obtained by another method. Two such geometries, the rectangle and the semi-ellipse will be discussed now. Solutions for the first are well-known from literature, those for the second geometry are derived below.

5.1. *The rectangular basin*

It is well-known that in the case of a rectangle (2.5*a*)–(2.5*b*) can also be solved by separation of variables and yields

$$\psi(x,z) \equiv a_1 \sin m\pi x' \sin n\pi z', \tag{5.1}$$

where $x' \equiv (x+1)/2$ and $z' = z/\tau$, provided non-dimensional depth (τ) is a rational number: $\tau = 2n/m$, with mutual prime numbers $m,n \in \mathbb{N}$. Here a_1 is an undetermined amplitude of the mode (suppressing the dependence on the mode numbers). This (m,n)-mode is not unique for the frequency ω (map parameter τ, see (2.6*a*)) under consideration, since, as Münnich (1994) remarks, any multiple – a (jm,jn)-mode, with $j \in \mathbb{N}$ – equally satisfies the hyperbolic equation while vanishing at the boundary. In this terminology, used in Münnich (1994), an (m,n)-mode describes a cellular pattern with m horizontal and n vertical cells. This non-uniqueness can be employed to directly solve the forced problem, when the forcing is by pressure variations imposed at the top in a fundamental interval. A fundamental interval at the surface is, in this case, recognized below as an interval in between two zeros of the gravest-mode streamfunction field (see figure 16). The response to an arbitrarily shaped, oscillatory pressure field in that interval is directly obtained as the sum over the Fourier modes of that function, which act as the amplitudes of the gravest and higher-order streamfunction modes.

5.1.1. *The map*

The characteristic theory applied to the rectangle has been discussed in Magaard (1968). He also included the effect of sheared mean currents, which gives an asymmetry in the leftward and rightward sloping characteristics. In the absence of sheared currents, map $\boldsymbol{T}(x,s)$, for a rectangle of depth τ, has the same form as (3.5), where, in this case,

$$x_s = 1 - 2\tau \tag{5.2}$$

and

$$X(x) = x + 2\tau, \tag{5.3a}$$

$$Y(x) = 2(1-\tau) - x = 2 - X(x). \tag{5.3b}$$

The behaviour is strikingly different for rational and irrational values of τ. In the former case the map has a periodic structure such that each x_0, after a fixed, finite number of iterations, turns exactly back to its starting value. In stark contrast, for irrational values of the map parameter (which, recall, may alternatively be interpreted as the period of the wave, or the depth of the basin) the trajectory of each x_0 comes arbitrarily close to any point in the domain. Remarkably, this difference in behaviour is not signalled by the Lyapunov exponent which, being the sum of the logarithm of the absolute value of the derivative of the map, identically vanishes for all values of τ, as inspection of (5.3*a*)–(5.3*b*) tells us. The map thus is neutrally stable, but may either have an infinite set of closed orbits (on which the field variable, $f(x)$, can be freely specified), or just one single orbit, such that only one value of $f(x)$ can be specified, which leads, according to (2.8), to an everywhere-vanishing streamfunction field. The artificial restriction used in the discussion of the parabola to values of τ for which there is at least one x for which just the forward map, $X(x)$, applies ($\tau \leqslant 1$)

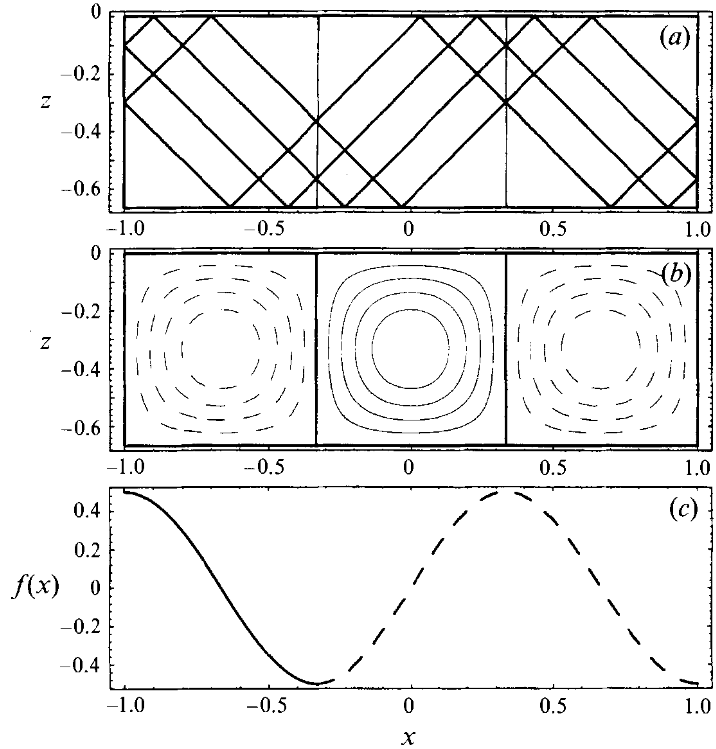

FIGURE 16. (*a*) Definition sketch for the (3,1)-mode ($\tau = 2/3$) of the rectangle. There are three fundamental intervals at the surface, in between successive zeros of the streamfunction field. Specifying the surface pressure in one of these intervals allows one to determine the amplitudes of the higher harmonic structures – $(3j, j)$-mode, $j \in \mathbb{N}$ – as the Fourier amplitudes of the specified spatial pressure structure in that interval. Two closed rays have been drawn. (*b*) Streamfunction field $\psi = \sin 3\pi x' \sin \pi z'$ for $\tau = 2/3$, corresponding to (*c*) $f(x) = (\cos 3\pi x')/2$. In (*c*) the solid part of $f(x)$ has been specified, while the dashed part has been inferred.

limits our solutions to those which have at least two fundamental intervals in the x-direction. This restriction can, of course, easily be eliminated. In particular the square domain ($\tau = 2$), also discussed by Magaard (1968), has a single fundamental interval stretching out over the entire x-domain, for which each surface-point lies on a period-1 characteristic. The rays in the basin corresponding to the example in figure 16 illustrate that the fundamental interval in this case is 1/3 of the size of the x-domain and also that each x_0 lies on a 3-cycle. That is, $x_3 = x_0, \forall x_0 \in (-1, 1)$. Compilation of a figure with the asymptotic state of the trajectories, like figure 8, shows a heavily dotted, structureless set due to the fact that the adopted discretized values of τ, employed in the construction of that figure, correspond to very high-period cycles, because the numerator of the rational number, going with those τ-values, is very large.

5.1.2. *Determination of streamfunction field by characteristic method*

For $f(x)$, given in a fundamental interval, this function can be determined over the whole x-range (figure 16*c*). From this, the streamfunction field is directly determined by (2.8) for $z + 1 > |x|$. Near the vertical boundaries rays are reflected and the complete description of the streamfunction is given by

$$\psi(x,z)=\begin{cases} f(x-z)-f(x+z) & \text{for } z+1>|x| \\ f(2-x+z)-f(x+z) & \text{for } z+1<x \\ f(x-z)-f(-2-x-z) & \text{for } z+1<-x. \end{cases}$$

Although $f(x)$ can be specified at will in any of the fundamental intervals it is now clear that the same streamfunction field can also be obtained from the modal solutions (5.1) as

$$\psi(x,z)=\sum_{j=1}^{\infty} a_j \sin jm\pi x' \sin jn\pi z', \tag{5.4a}$$

where

$$a_j=\int_{-1}^{-1+2/m} f(x)\sin jm\pi x' \mathrm{d}x \tag{5.4b}$$

are the Fourier components of $f(x)$ on the first fundamental interval $x \in (-1,-1+2/m)$. The two methods thus yield the same results and the indeterminacy is in both cases resolved by specifying the 'pressure', $f(x)$, in one fundamental interval only. The characteristic method, however, is more direct and enjoys a slight preference over the modal method as no Fourier decomposition is needed. The superiority of the characteristic method becomes more evident, however, in the case that 'non-trivial' topographies are taken into consideration (§§4 and 6).

5.2. *The semi-elliptic basin*

There is at least one other class of 'bottom' profiles for which the solution of the hyperbolic equation can be obtained in terms of modes: the ellipse. Since this equation is to be solved in cases where there exists a flat surface this is here further restricted to the semi-ellipse. As for the rectangle the trajectories are either periodic, or of infinite period. No previous derivation of this set of solutions exists as far as the authors are aware, but it can be readily derived by a variation of complex function theory.

It is well known that the Laplace equation,

$$\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial y^2}=0,$$

is solved by the real and imaginary parts of arbitrary functions $F(x+\mathrm{i}y)$. In particular, the polynomials $(x+\mathrm{i}y)^p$, with $p \in \mathbb{N}$, have simple expressions. For instance, for a cubic ($p=3$) one finds that both x^3-3xy^2, and y^3-3yx^2 satisfy the Laplace equation. By simply replacing $y=\mathrm{i}z$ in both the Laplace equation and its solutions, functions satisfying the hyperbolic equation (2.5*a*) are obtained. From the example of the cubic we find, by adding a linear term, that we thus obtain an explicit solution also satisfying the boundary condition (2.5*b*) at $z=-\tau(1-x^2)^{1/2}$:

$$\psi=z(z^2+3(x^2-1)),$$

provided the 'depth' $\tau=\sqrt{3}$. Since this consists of just one circulation cell we might consider this as the (1,1)-mode of a semi-ellipse. It appears that more complicated modes are obtained by considering higher-degree polynomials (in general, a particular combination of either even or odd polynomials). In this way the first few cellular modes, listed in table 2, have been obtained (see also figure 17). For the rectangular basin it was argued that any frequency that is a rational number is an eigenfrequency

p	(m,n)	τ	$\psi(x,z)$
3	(1,1)	$\sqrt{3}$	$z(z^2+3x^2-3)$
4	(2,1)	1	$xz(x^2+z^2-1)$
5	(3,1)	$(5-2\sqrt{5})^{1/2}$	$z(z^2+(5-2\sqrt{5})(x^2-1))(z^2+(5+2\sqrt{5})x^2+(\sqrt{5}-5)/2)$
5	(1,2)	$(5+2\sqrt{5})^{1/2}$	$z(z^2+(5+2\sqrt{5})(x^2-1))(z^2+(5-2\sqrt{5})x^2-(\sqrt{5}+5)/2)$
6	(4,1)	$1/\sqrt{3}$	$xz(z^2+3x^2-1)(x^2+3z^2-1)$
6	(2,2)	$\sqrt{3}$	$xz(z^2+3x^2-3)(x^2+3z^2-3)$

TABLE 2. Expressions for the streamfunction field satisfying the hyperbolic equation and vanishing at the surface, $z = 0$, and semi-ellipse, $z = -\tau(1-x^2)^{1/2}$. Columns give respectively: power of polynomial p, modal structure (m,n) denoting the number of horizontal and vertical cells respectively, τ related to the eigenfrequency, and streamfunction field $\psi(x,z)$.

and, moreover, that each frequency is infinitely degenerate, because of the existence of multiples of the modal structure: modes (jm, jn), $j \in \mathbb{N}$. For the (semi) ellipse this happens too, except that the eigenfrequencies are now no longer rational, but rather a subset of the real numbers. The infinite set of eigenfrequencies for the ellipse can be ordered (and be made denumerable) by the modal structure of the related streamfunction field. Also, by association of the modal structure with the corresponding structure in the rectangle one finds a one-to-one correspondence between both infinite sets of accompanying eigenfrequencies. The enumeration of successively more complicated modes proceeds on the (m,n)-lattice along lines $m+2n=p$, where $p \in \mathbb{N}$ denotes the highest power of the polynomial under consideration. Thus for $p=3$ and 4 we find just the $(1,1)$- and $(2,1)$-modes, respectively. For $p=5$, however, there are two modes satisfying this constraint, which provides us both with the $(3,1)$- and $(1,2)$-modes. Algebraic computation of the eigenfrequencies gets increasingly more complicated, although one should not have difficulty finding a good numerical approximation to the eigenfrequencies along these lines.

Not only do we find a denumerable, infinite set of eigenfrequencies, but also each eigenfrequency is degenerate. For the same frequency, multiples of fundamental modes exist (for which m and n do not have a common divisor), as the first and last rows of table 2 indicate. The last mode ($p=6$) has the same eigenperiod as the first mode ($p=3$), but has twice the number of cells in the horizontal and vertical direction.

As for the rectangle it can be verified that $x_m = x_0$ for each x_0 in the interval $x_0 \in (-1,1)$, for some $m \in \mathbb{N}$. There are thus infinitely many, closed m-periodic orbits. In contrast, for the parabolic basin, periodicity of the map was obtained only asymptotically and then approaches just one or two closed orbits, which are reached irrespective of the starting value of x_0. There thus appear two types of internal-wave solutions, which strongly depend upon the shape of the boundary.

In the case of the semi-ellipse the streamfunction field can also (and more readily) be constructed by means of the characteristic method along the lines indicated above. This has not been elaborated here. For completeness we give the bi-modal map, $\boldsymbol{T}(x,s)$, which is defined by (3.5), with

$$x_s = (\mu-3)/(\mu+1) \tag{5.5}$$

with $\mu \equiv 1/\tau^2$ (assumed to be larger than one, to restrict oneself again to at most two

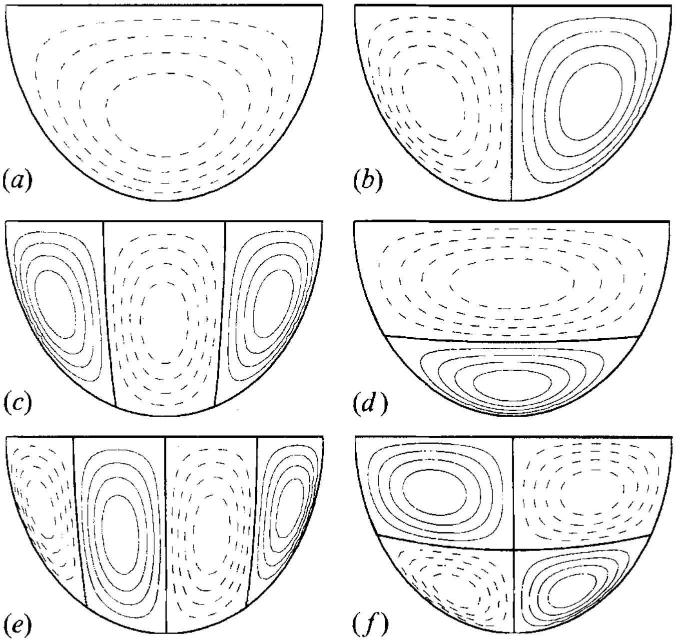

FIGURE 17. Nodal lines ($\psi = 0$, thick solid) and streamlines (positive: thin solid, and negative: dashed) for the exact modal solutions of the hyperbolic equation, within a semi-ellipse of 'depth' τ equal to (*a*) $\sqrt{3}$, (*b*) 1, (*c*) $(5-2\sqrt{5})^{1/2}$, (*d*) $(5+2\sqrt{5})^{1/2}$, (*e*) $1/\sqrt{3}$, (*f*) $\sqrt{3}$. For visualization purposes the z-coordinate has been rescaled such that the basin acquires one and the same semi-elliptical shape which implies that the characteristics in this representation have a direction related to the period of the wave (as given by the above values of τ).

reflections at the bottom) and

$$X(x) = \frac{(\mu-1)x + 2(1+\mu(1-x^2))^{1/2}}{1+\mu}, \tag{5.6a}$$

$$Y(x) = -x + \frac{2(\mu-1)X(x)}{\mu+1}. \tag{5.6b}$$

As for the rectangle the Poincaré plot of successive reflections of a ray yields a featureless dotted figure, while the Lyapunov exponents are also zero to within the numerical accuracy of its computation.

6. The 'bucket' and other basin shapes

The rectangular and semi-elliptic basins have nice, cellular patterns of the streamfunction field as solutions of the hyperbolic equation. Since these modal structures are to be multiplied with a sinusoidal temporal variation these have traditionally been interpreted as the internal, seiching modes of the basin (Defant 1941). It appears, however, that this behaviour is not generic, as the example of the parabolic basin, §§3 and 4, has shown. On the contrary, from a number of examples considered we get the impression that these two cases are exceptional. Instead, in general, seiching modes are either absent and focusing of the internal wave field to a well-defined attractor appears to be the rule, or, as we will try to demonstrate with the following example, hybrid situations may arise that exhibit the existence of both regular (neutrally stable) modes, as well as focusing (with, as its limiting case, infinite-period, plane-filling) orbits.

6.1. *Bucket-shaped basin*

Consider a bucket-shaped basin given by

$$H(x)=\begin{cases}\mu(1-x), & x\in(d,1)\\ \mu(1-d), & x\in(-d,d)\\ \mu(1+x), & x\in(-1,-d).\end{cases} \tag{6.1}$$

This 'bucket' is a two-parameter topography, with d the relative size of the interval where the bottom is flat compared to the width of the basin. This geometric quantity is normally regarded fixed. The second parameter, μ, the tangent of the angle of the sloping sidewall, is, by our convention to put the frequency of the internal wave into the apparent depth, a variable quantity. The dimensionless depth is given by $\tau=\mu(1-d)$. The bi-modal map corresponding to this case is given by (3.5) with

$$x_s=1-2\mu(1-d) \tag{6.2a}$$

and

$$X(x)=x+2\mu(1-d), \tag{6.2b}$$

$$Y(x)=\begin{cases}\dfrac{1+\mu}{1-\mu}x+2\mu\dfrac{d-\mu+d\mu}{\mu-1}, & x\in(x_s,d(1+\mu)-\mu)\\ \dfrac{1-\mu}{1+\mu}x-2\mu(1-d)+2\dfrac{\mu}{1+\mu}, & x\in(d(1+\mu)-\mu,1).\end{cases} \tag{6.2c}$$

For $d=4/5$ a Poincaré plot of the asymptotic values of the surface intersections is given in figure 18(a). The upper bound of $\mu=1/(1-d)$ has again been chosen to restrict the mapping to a regime where bottom reflections occur at most twice. It demonstrates that there are a number of windows with low-period attractors accompanied by other windows of very high-period attractors. This figure looks like an incomplete version of figure 8. In fact it has features in between those of the parabola and those displayed in the rectangle to which it approaches for $d\rightarrow 1$. In the latter limit the compact, low-period windows vanish. If d drops below a half ($d\leqslant 1/2$), however, the 'noisy' windows disappear and we are just left with a period-2 window for all frequencies in this band! (The latter configuration is presumably the easiest case for testing the occurrence of focusing in a laboratory experiment).

6.1.1. *Lyapunov exponents*

It is worthwhile to notice that the Lyapunov exponent in the case of the bucket is particularly easy to determine. If we look at map (6.2a)–(6.2c) it is obvious that the only contribution comes from parts where the slope is not equal to one. The slopes of the map for the remaining two cases, see (6.2c), however, are reciprocal and the logarithm of its value, $\ln[(\mu-1)/(\mu+1)]$, can thus be factored out. The remaining determination of the Lyapunov exponent, then, reduces to simple book-keeping of the number of times for which the 'divergent' (steeply sloping), 'neutral' (slope equal to one) and 'convergent' (weakly sloping) parts of the map are reached by a particular ray. This factor can be recognized in the generally increasing form of the Lyapunov exponent. In fact, the dashed curves in figure 18(b), related to this function, match exactly for the low-period attractors. This is because, e.g. for the 2-cycle, for the particular starting position adopted, points successively sample only the convergent branches of the forward and backward map (which have the same slope). Indeed,

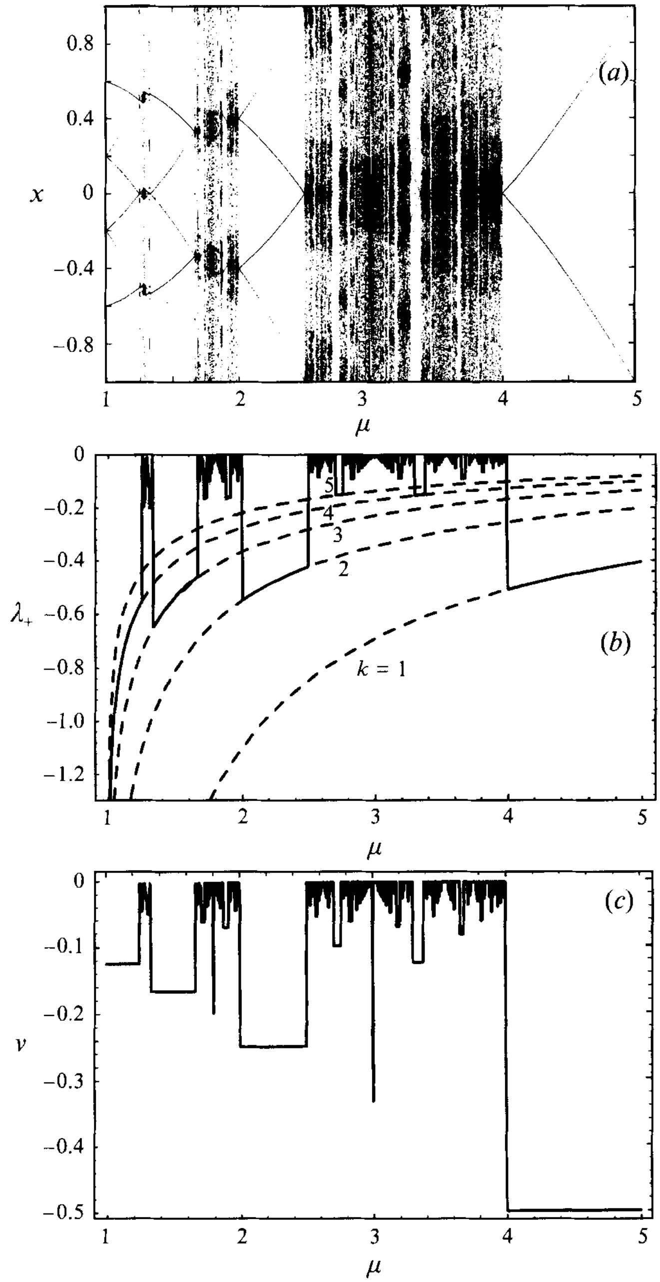

FIGURE 18. (*a*) Poincaré plot, (*b*) Lyapunov exponents, λ_+ and (*c*) $\nu = -1/P$, where P is the period of the attractor for a bucket with $d = 4/5$ as a function of μ. The dashed lines in (*b*), referred to in the text, are given by $\ln[(\mu - 1)/(\mu + 1)]/k$, for given k.

normalizing the Lyapunov exponent with this logarithmic curve would enable us to calculate (twice) the period, figure 18(c), rather than determining the latter (or, its related value, v) numerically.

6.1.2. *Resonance*

In figure 18(a) we clearly recognize the 2-, 4-, 6- and 8-cycles. Suspiciously lacking, however, are the odd-period attractors, with, most notably, the period-3 attractor. Closer inspection of this figure, however, shows that this interval has not vanished altogether, but rather has shrunken to the size of a point (located at $\mu = 3$ for this value of d). Surprisingly, as a phoenix rising from its ashes of almost-zero (infinite-period) Lyapunov exponents, we recover a neutrally stable, period-3 mode (see figure 18b,c)! Each initial value returns to that same value after three mappings. A global attractor no longer exists. Other stable, periodic modes are obtained in the Lyapunov diagram whenever the Lyapunov exponent approaches zero both for μ coming from above as well as from below the point where it exactly vanishes. For instance, at $\mu = 9/5$ we find a period-5 mode. These periods have also been captured in figure 18(c), and appear there as spikes. Higher-period stable modes have not been resolved, however, and can only be detected by the above-formulated rule-of-thumb.

The stable 3-mode appears to be present for every value of $d > 1/2$. Its μ-location in figure 18(a) can be obtained from the observation in figure 19(a) that it occurs when the ray stemming from the upper corner reaches the opposite corner at the bottom (that now contains the critical characteristic):

$$x_s + x_c = 0.$$

Figure 19(a) shows a sketch demonstrating the configuration of these rays for this (3, 1)-mode. Generalizing this to arbitrary d one may obtain this mode at a value of μ determined by the requirement that $3\mu(1-d) + 1 - d = 2$, or $\mu = (1+d)/3(1-d)$, as inspection of this figure shows. Likewise, for mode $(m, 1)$, $\mu = (1+d)/m(1-d)$ yields the 'eigenperiods' $\tau = (1+d)/m$ for which neutrally stable modes exist. Since we require $\mu > 1$, in order for the bucket topography to be supercritical, this implies that a finite number m_μ exists, where m_μ equals the largest integer smaller than $(1+d)/(1-d)$, such that $m \leqslant m_\mu$.

Of course, there also exists a (1, 1)-mode for which a ray emanating from the corner directly intersects the opposing bottom corner, but that has been excluded from consideration by our artificial requirement that reflection at the bottom should occur at most twice prior to the surfacing of the ray. This led us to require $\mu < 1/(1-d)$, or $\tau < 1$. The above criterion, applied to $n = 1$, would yield $\mu = 9$ for $d = 4/5$ and is therefore formally outside the range of figure 18. Extension of the algorithm that computes successive surface intersections when the rays make multiple reflections at the sidewalls does not cause any difficulties in principle though. In particular the (1, 1)-mode is computed below, albeit for a different value of d.

6.1.3. *Analytical solution*

An analytical solution for the (1, 1)-mode for $d = 1/2$ and $\mu = 3$ has been reported on by Cushman-Roisin, Tverberg & Pavia (1989) modelling a fjord environment. They considered a continuously stratified fluid within a bucket-shaped trench at the bottom of a broader, otherwise flat basin. The (oscillatory) flow in the main basin being prescribed (having a spatial part of the horizontal velocity field $u = x - 1$, corresponding to a stagnation flow) a solution in the bucket is sought that has no slip at the interface (the top of the bucket). With this prescribed velocity field at

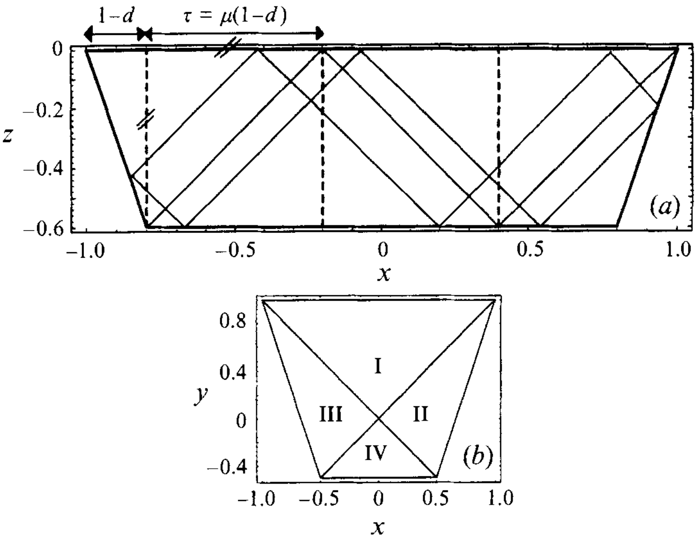

FIGURE 19. (*a*) Ray configuration in the case $d = 4/5$ and $\mu = 3$. Also indicated are intervals at the surface, used in the text, relating to these two parameters. (*b*) Definition of regions for a bucket with $d = 1/2$ and $\mu = 3$. Here $y = z + 1$.

the top and the subsequent requirement that the streamfunction field vanishes at all the boundaries they obtained, apparently by inspection, an exact solution of the hyperbolic equation describing a stationary wave pattern of the internal-wave field in the trench. With a displaced vertical coordinate, $y = z + 1$, it is given, in terms of the streamfunction, by

$$\psi = \begin{cases} (y-1)(x-1), & x, y \in \mathrm{I} \\ (y-x)^2 - (y+x-2)^2/4, & x, y \in \mathrm{II} \\ (x-y)^2/4 - (1+x+y)^2, & x, y \in \mathrm{III} \\ -(1+2x)(1+2y), & x, y \in \mathrm{IV}, \end{cases} \tag{6.3}$$

where the four regions have been indicated in figure 19(*b*). This solution for the streamfunction is depicted in the lower panel of figure 20(*a*). It is characterized by the existence of a vortex sheet (along $y = -x$) due to the fact that the prescribed horizontal velocity field does not vanish in the left-hand corner, $f'(-1) \neq 0$. The occurrence of a vortex sheet is typical in the generation of internal tides as observed in theoretical (Wunsch 1968, 1969; Robinson 1969; Larsen 1969) and experimental (Sandstrom 1969; Baines & Fang 1985) models, as well as, to some extent, in numerical models and nature (deWitt *et al.* 1986). Cushman-Roisin *et al.* (1989) used this solution to validate their numerical model. They remark that "notwithstanding the authors' effort no other nontrivial analytical solution has been found". However, we may recognize their solution as a particular case of the general solution obtained with the characteristic method, which would read:

$$\psi = \begin{cases} f(x-y+1) - f(x+y-1), & x, y \in \mathrm{I} \\ f(1-2x+2y) - f(x+y-1), & x, y \in \mathrm{II} \\ f(x-y+1) - f(-1-2(x+y)), & x, y \in \mathrm{III} \\ f(1-2x+2y) - f(-1-2(x+y)), & x, y \in \mathrm{IV}, \end{cases} \tag{6.4}$$

where $f(x)$ can be specified at will along the surface domain (the entire domain now being a fundamental interval). This function, $f(x)$, has the same meaning as before, being related to the surface pressure, equation (4.1), while its derivative relates to the horizontal, surface velocity field. The streamfunction field (6.3) is obtained with $f(x) = (x-1)^2/4$. This function has been displayed in the upper panel of figure 20(*a*). A few other choices of $f(x)$ and their corresponding streamfunction fields, obtained from (6.4), are shown in the other panels of that figure. In particular, we note the absence of free shear layers when the derivative of the prescribed function $f(x)$ – the horizontal component of the velocity field – vanishes at the corners, $f'(\pm 1) = 0$, (figure 20*b*).

The $(1,1)$- and $(3,1)$-modes, determined above, are characterized by the fact that a ray connects the surface corner with the opposing bottom corner. It is natural to inquire what happens when corner points are connected differently. Thus when a surface-corner point is connected with an adjacent bottom corner we obtain for instance the $(1,2)$- and $(3,2)$-modes, the latter occurring when $\mu = [1 + d + 2(1 - d + d^2)^{1/2}]/3(1-d)$. Connecting two surface, or two bottom corner points (for the situation depicted in figure 18 occurring at $\mu = 5$ and 4 respectively), however, does not yield the missing even horizontal modes, since these are clearly observed to have negative Lyapunov exponents and thus to consist of focusing modes. Physically this is obvious, since, following a ray, in this case the sloping sidewalls are always approached from above, which leads to convergence of wave rays. In contrast, for odd modes (in the case that there is a ray connecting the surface and bottom corner) the left and right sloping walls are approached successively from above and from below, so that convergence is exactly offset by subsequent divergence.

Resuming, the ‘bucket’ is a truly hybrid geometry, showing both focusing and neutrally stable (‘seiching’) modes. Of the latter, the modes with an even number of cells in the (x) direction are entirely absent, while also only a finite number of odd modes appear. Although it is nice to have explicit solutions in those cases where the rays are strictly periodic (neutrally stable – folding back upon themselves), it is necessary to emphasize that these form a very restricted class amongst all possible solutions, see figure 18(*b*,*c*). It is just for the particular frequencies corresponding to these cases that such a solution (a ‘resonance’ in the terminology of Cushman-Roisin *et al.* 1989) exists. (Curiously, these authors fail to identify their example of the (1,1)-mode, discussed above, in terms of a resonance, even though it is the prototype example in which each characteristic returns to its original position in one iteration: $x_1 = x_0$.) For all other frequencies rays are attracted to a particular limit cycle and focusing should thus be considered as the generic behaviour.

6.2. *Other basin geometries and artificial maps*

A basin with a flat bottom, τ, for $x \in [-d, d]$ with convex sidewalls which are segments of a hyperbola $\tau[1 - ((x^2 - d^2)/(1-d^2))^{1/2}]$, for $|x| \in (d, 1]$ has also been analysed. Here central depth is defined as $\tau = b(d^{-2} - 1)^{1/2}$, with $b \in (d/(1-d^2)^{1/2}, d(1-d^2)^{1/2})$, in order for the sidewalls to be supercritical and restricting the number of bottom reflections to at most two. For given value of length scale d, the Poincaré plot, the Lyapunov exponent and period of the asymptotic state have been computed as a function of b. The results are qualitatively similar to those obtained for the parabolically shaped basin. Upon cursory inspection no neutrally stable mode is obtained; all asymptotic states are globally attracting. The result does, of course, depend upon the particular value of d, but this does not invoke a qualitative change.

Each of the maps considered so far have two lines of symmetry, namely the lines

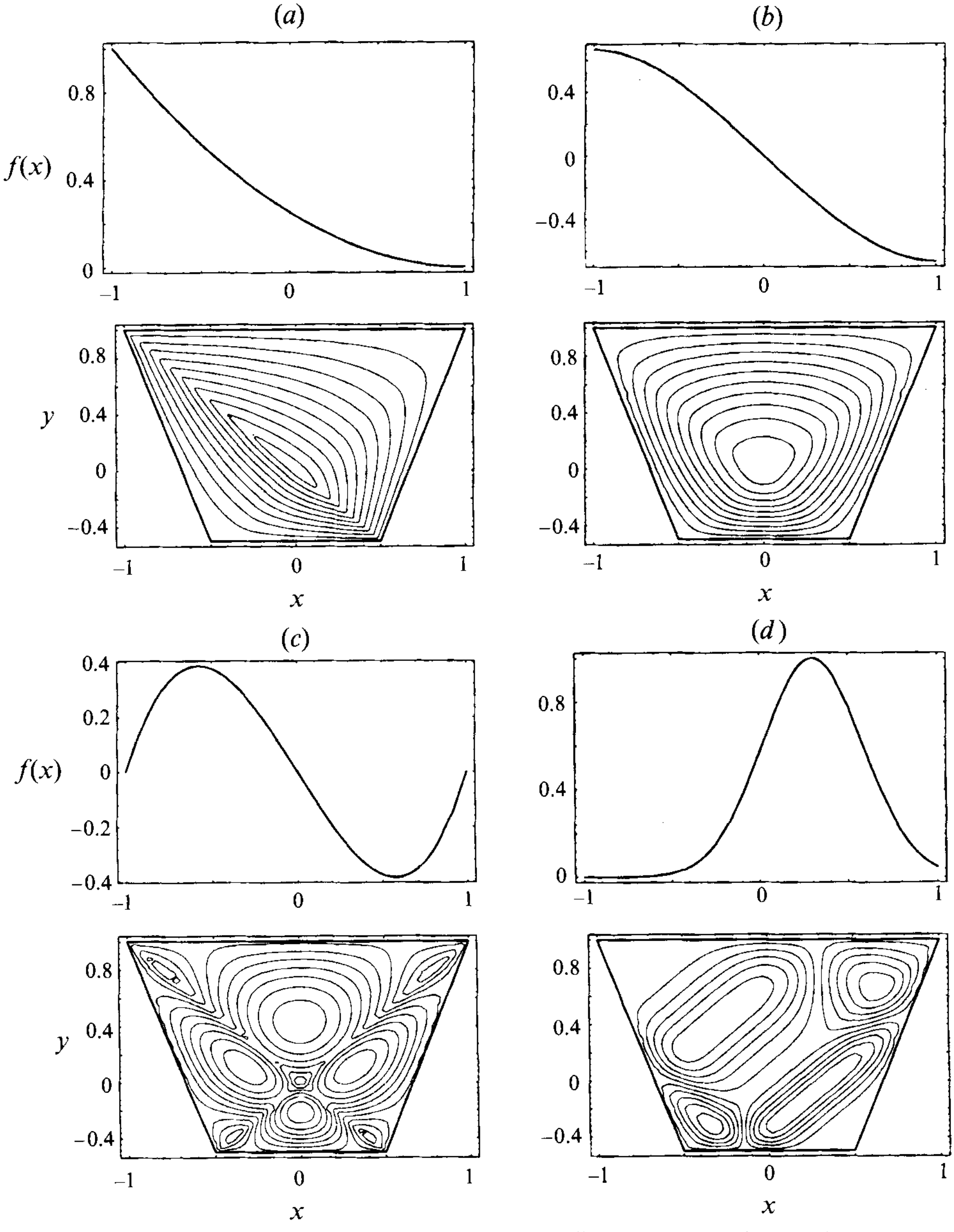

FIGURE 20. Plots of $f(x)$ (upper part) and corresponding streamfunction field (lower part) for a bucket with 'depth' 3/2 ($d = 1/2$, $\mu = 3$) for $f(x)$ given by (*a*) $(x-1)^2/4$, (*b*) $x^3/3 - x$, (*c*) $x^3 - x$, (*d*) $\exp[-(5x/2 - 3/4)^2]$.

$y = \pm x$ (if we designate the successor of x momentarily as y), see figure 4. To examine the effect of asymmetries in these maps various other topographies and also artificial bi-modal maps have been considered. For instance, a skew parabolic basin, which matches two parabolas at a trough which is located off-centre (see Münnich 1994), has a map for which the symmetry in the line $y = -x$ is broken. As a result the corresponding Poincaré plot is also no longer symmetric, owing to an asymmetry in the attractors. However, qualitatively, nothing happens: the same alternation of low- and high-period attracting regions, with changing values of τ, is obtained as for the symmetric case. Artificially breaking the second symmetry, by adding linear terms to the map under the restriction that the position and height of the map's maximum stay fixed, also does not yield very different results.

Finally, an abstraction of the bi-modal map has been considered. Each of the forward parts (upper curve) of the 'realistic' maps considered (like in figure 4) has a single maximum, to the left (right) of which the sign is invariant (changes). Also, the backward map (lower curve) is a point-mirrored version of the forward map (upper curve). Adopting a simple parabola, $1-(x-b)^2/(1+b^2)$, with $b \in (-1, 1)$ as forward map, a completely artificial bi-modal map is examined. It turns out to have features in common with the 'bucket': firstly, it has both low- and high-period windows, secondly it has Lyapunov exponents which are always less than or equal to zero, and thirdly, it has particular values of b for which resonances exist (exact vanishing of the Lyapunov exponent and, as a consequence, each point lies on its own stable orbit). It is thus expected that these features are quite common and should be expected to occur in more realistic circumstances too. Further examination of these and other cases is necessary, though.

7. Discussion and conclusions

7.1. *Focusing*

It is well known that internal waves are fundamentally different from surface waves. This is because in the former case the phase propagates in a direction *perpendicular* to the energy-propagation direction (as given by the group velocity vector), rather than along it. Nevertheless, one would like to interpret stationary, internal-wave patterns in terms of seiching modes – a resonance, familiar from oscillating surface-gravity waves in an enclosed basin (Münnich 1993). It is shown in this study that such an interpretation is only occasionally justified. In general, internal waves are focused towards a limiting attractor, while increasing their amplitude and reducing their wavelength and group velocity. Its position depends on parameters characterizing the geometry. For the simple topographies (and stratification) considered here this attractor consists of one or two fixed sets of lines in the basin, whose locations depend on only one parameter, τ: the product of the buoyancy and wave frequency ratio and the aspect ratio of the basin. The attractor can be classified by the number and location of reflections of the asymptotic ray with the surface. This is a fractal function of the scaled period of the internal-wave field, τ: it can change very rapidly in certain intervals, while remaining qualitatively similar (characterized by the same period of the attractor) in other intervals.

The attractor is the limiting trajectory of a ray's orbit. The ray path itself is constructed, following Magaard (1962, 1968), by means of an iterated map. This map can be made explicit for piecewise linear, or quadratic shapes of the topography. The map consists of a rightward and leftward mode that get coupled for two-sided, supercritically sloping basins. This bi-modal map has the property that Lyapunov exponents are less than or equal to zero, corresponding to focusing or neutrally stable modes (resonances), respectively. The solution of the canonical, hyperbolic equation which the streamfunction satisfies, is completed by finding regions (the fundamental intervals) over which rays can be uniquely identified. By specifying the value of a field variable, related to the surface pressure, within these fundamental intervals, the streamfunction in the entire basin can be computed. With this procedure, formally, only standing wave patterns have been obtained. Therefore, it is assumed that internal waves manage to 'bounce back' from the attractor. This is unlikely to happen in reality, though, since amplification of the internal-wave amplitude will inevitably lead to viscous decay (neglected so far) and 'deposition of energy' near the attractor. A proper description of a stationary (and modulated) propagating internal-wave field is currently being studied.

7.2. *Resonance*

In some cases a resonance does occur. Here we use the word resonance in a slightly broader sense than that introduced by Cushman-Roisin *et al.* (1989). They refer to each ray that returns to itself as a separate resonance. This is perhaps appropriate in their finite difference approach where this number, because of the discretization employed, is necessarily denumerable. Neglecting the attractors in the focusing cases (which in their terminology would, formally, also be regarded a resonance), we define a resonance as a situation when on one interval each point taken from that interval acts as starting point of a ray that eventually returns to its position after crisscrossing the basin a finite number of times. Perhaps more appropriately this should be referred to as a seiching mode. A necessary condition for seiching to occur is apparently that the wave ray stemming from the surface corner ends up in the critical characteristic. This condition is fulfilled by the classical, separable solutions in the rectangle, as well as by the resonant cases found for a bucket-shaped basin. We observed that it also applies to the separable solutions constructed for a (semi) elliptic basin, but not for a parabolic basin (see below). Physically, the relevance of resonant modes comes from the fact that, in principle, they lack the presence of vortex sheets, along which incoming internal wave energy will be degraded by viscous effects. These modes are therefore able to store more energy and will stand out globally. In practice, vortex sheets may still occur in these resonant cases too, as when there is a non-vanishing, surface pressure gradient in the corners of the basin, but they are of a different nature than those related to the attractors.

Cushman-Roisin *et al.* (1989) argue that the existence of closed ray paths is supposedly the rule rather than the exception. We feel that this may in part be due to their approximation of true topography by horizontal and vertical line segments, which precludes the possibility of obtaining focusing with the numerical algorithm. Besides this there is a semantic difference. Our statement that closed ray paths are exceptional means that in the focusing (resonant) cases there is a set of x_0-values of measure zero (one) for which ray paths are closed. Hence, since focusing is more common, this allows us to use this phrase. Conversely, Cushman-Roisin *et al.* (1989) stress that both in focusing and resonant cases there exists at least one closed ray and contrast this with the (exceptional) ergodic cases for which no ray ever closes on itself. This allows them to state that ray closing is generic. There is thus no contradiction between these two viewpoints.

Now, Münnich (1993, 1994) also obtains resonances (seiches) for the parabolic basin, a geometry where no such solutions exist according to the characteristic method. In particular he found the (1,1)-mode (consisting of one vertical and one horizontal cell) for aspect ratio $\tau = \sqrt{2}$. We believe that this mode was artificially obtained because the numerical procedure was formulated such that it defined a (1,1)-mode as one that minimizes the basin-averaged shear of the 'solution'. In this way, we think that this numerical procedure trades off accuracy in favour of finding a minimally-sheared, modal solution. This value of τ is interesting though, because it is the only value for which a period-1 attractor exists for $1 \leqslant \tau \leqslant 3/2$. This attractor consists of the ray starting at the surface from the centre of the basin ($x = 0$, $z = 0$), which exactly returns to itself after traversing a square-shaped ray path in the interior. All other starting positions lead to focusing towards this attractor, though. One may check in particular that rays starting in the corner do not coalesce with the critical characteristic (see Appendix), thus denying the existence of a resonance.

From a numerical point of view too there are some advantages in using the characteristic method. First, the solution is exact along the characteristics and second, in regions of high shear, a higher accuracy is automatically obtained, because more rays are present there. This is not surprising, because discretization along characteristics (which includes information on the topography used; e.g. figure 14*b*, *c*) is better suited to the problem than discretization on a 'random' grid.

Notwithstanding the difference in interpretation of 'resonance', one of the reasons for Cushman-Roisin *et al.* (1989) to introduce this concept was that it illuminates the fact that a complete specification of the internal-wave field is not always given in terms of its prescription at a particular part of the boundary. In their case it was a connecting shelf on which the internal-wave field was prescribed. As in the laboratory studies of Robinson (1969) and Sandstrom (1969) and in the field studies of Cushman-Roisin & Svendsen (1983) and deWitt *et al.* (1986), this left certain 'shadow zones' where the characteristics, emanating from the shelf, did not reach and which, as Cushman-Roisin *et al.* (1989) argue, are determined by diffusion. The question of whether a problem is ill-posed or well-posed is here resolved by identifying, so called, primary (fundamental) intervals on which the solution can be independently specified. This applies both to focusing and to resonant cases. In particular it implies that when the size of the fundamental intervals shrinks to the size of a point, no stationary solution pattern exists at all, because there is apparently just one ray on which the solution can be specified and that ray is 'plane-filling'. Physically the function that leads to the specification within the primary intervals has been shown to be related to the pressure. What happens when the solution is over-specified is not clear yet and awaits a consideration of truly propagating solutions to the problem at hand.

7.3. *Relevance to field observations*

The implications of focusing of an internal wave field for nature are not clear yet. Oceans, or smaller-scale basins, are not two-dimensional, their boundaries are not smooth, the fluid is not uniformly stratified and the forcing field is not monochromatic. However, refraction of incoming waves orients them preferentially in a cross-isobath direction (Wunsch 1969). Also, the large-scale internal waves are presumably not very sensitive to the details of the topography. Non-uniform stratification (neglecting reflection on this non-uniformity) only leads to curved ray paths (Cushman-Roisin & Svendsen 1983). Finally, the theory has been developed with just one single frequency in mind, whereas in reality internal waves are forced over the whole internal-wave band. However, in view of the linearity of the problem, these solutions can all be superposed. Probably, internal waves of tidal origin will be most important owing to the ubiquitous nature of the forcing.

When internal waves do not lose their energy very rapidly by reflection at boundaries and are able to cross the basin back and forth geometric focusing should, in principle, occur. The eventual implication is that it offers a mechanism by which 'mixing at a distance' (along the attractor) can occur directly within the interior of the stratified ocean basin (and subsequently diffuse through the entire ocean along isopycnal surfaces). In this way it may perhaps contribute to mixing, leading to mid-

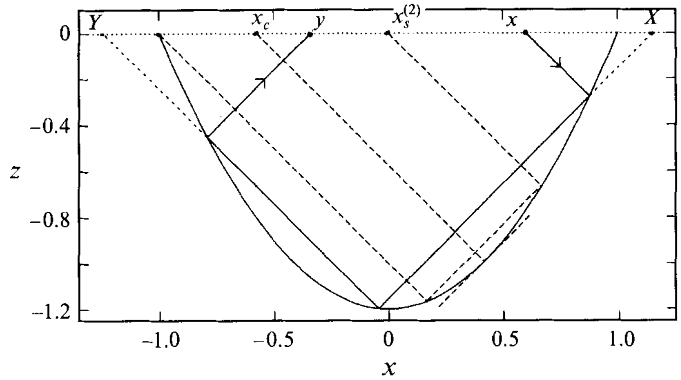

FIGURE 21. Definition sketch of parabolic basin with $1 \leqslant \tau \leqslant 3/2$.

depth ocean diapycnal diffusivities of $10^{-4}\text{m}^2\text{s}^{-1}$ required by global budget studies (Munk 1966) and offer an alternative to boundary mixing (Garrett 1991). In order to quantify the proposed mechanism we need to estimate the amount of energy lost after reflection off a non-critically sloping bottom – the most common situation – and 'surface', and thus assess the number of bounces an internal wave may go through before it is being focused to the extent that the Richardson number becomes subcritical and mixing ensues. Alternatively, the smallest observable scale to which focusing proceeds may be set by the irregularities of smaller-scale topography by which the internal-wave field becomes diffusively scattered (Longuet-Higgins, 1969). (The phenomenon of split-reflection (Baines 1971*b*), which this small-scale process entails, has here in fact been disregarded altogether and needs to be addressed in future studies).

Admittedly, focusing, and therefore mixing, will first take place near the bottom boundary, but, rather than being an isotropic process along the boundary, the above mechanism suggests that there are specific locations where mixing occurs first (and preferentially): i.e. near places where the attractor intersects the bottom. Near-bottom, intermittent, intermediate turbid layers have recently been observed in Emerald Basin on the Scotian Shelf by Azetsu-Scott, Johnson & Petrie (1995) who attribute this to anisotropic mixing due to (near) critical reflection (and amplification) of internal tides (presumably originating at the opposing break in topography). It is intriguing to speculate that the observed layers and inferred anisotropic mixing may, alternatively, be due to geometric focusing of internal tides.

We are indebted to Peter Beerens for help with the numerical formulation of the algorithm, to Erwin Embsen for computer support, to Taco de Bruin for help with preparation of figure 13 and to Huib de Swart, Henk Dijkstra, Kees Vreugdenhil, Ferdinand Verhulst and Matthias Münnich for enlightning conversations. During the course of the refereeing process it appeared that some of the work presented here has been anticipated by Cushman-Roisin (1993), who had presented it at the 1991 IUGG-meeting in Vienna. Also, Dr. V. Shrira pointed out that related work has been done by Bunimovich (1980).

Appendix.

An explicit map for a basin with parabolic cross-section can be constructed for $1 \leqslant \tau \leqslant 3/2$. For this interval each characteristic reflects either two or three times from the bottom prior to reaching the surface. These two regions are separated by $x_s^{(2)}(\tau) \equiv x_l^{\bullet}(x_r(-1)) = 6/\tau - 5$, a generalization of $x_s(\tau)$. Here $x_{l,r}^{\bullet}(x)$ denote the conjugate mappings of those defined in (3.2*a*) and (3.2*b*), obtained by changing the signs in front of the radicals appearing in those expressions. For $-1 \leqslant x \leqslant x_s^{(2)}$ the rightward map is given by (3.4). For $x_s^{(2)} \leqslant x \leqslant 1$ the image is obtained with the aid of figure 21. Let y denote the true image of x after a rightward mapping. Two virtual auxiliary points, denoted X and Y, are related to x and y in the following way. A leftward map (3.2*b*), and its conjugate, applied to X yield x and Y:

$$x = \frac{1}{\tau} - X + \left(\frac{-4X}{\tau} + 4 + \frac{1}{\tau^2}\right)^{1/2}, \tag{A 1a}$$

$$Y = \frac{1}{\tau} - X - \left(\frac{-4X}{\tau} + 4 + \frac{1}{\tau^2}\right)^{1/2}. \tag{A 1b}$$

A rightward map of Y, equation (3.2*a*), and its conjugate, give y and X:

$$y = -\frac{1}{\tau} - Y - \left(\frac{4Y}{\tau} + 4 + \frac{1}{\tau^2}\right)^{1/2}, \tag{A 2a}$$

$$X = -\frac{1}{\tau} - Y + \left(\frac{4Y}{\tau} + 4 + \frac{1}{\tau^2}\right)^{1/2}. \tag{A 2b}$$

Combining this information leads to an explicit dependence $y(x)$. Adding (A 1*a*) and (A 1*b*) yields

$$Y = \frac{2}{\tau} - x - 2X(x), \tag{A 3a}$$

whereas adding the other two gives

$$y = -\frac{2}{\tau} - X(x) - 2Y(x, X(x)). \tag{A 3b}$$

Inserting (A 3*a*) into this finally leads to

$$y = -\frac{6}{\tau} + 2x + 3X(x), \tag{A 4}$$

where $X(x)$, the inverse of (A 1*a*) is given by $x_r(x)$ in (3.3*a*). Sign changes of s should, of course, again be accounted for. The new sign is given by $s^{(n-1)}$, where n signifies the number of times the characteristic has hit the bottom prior to reaching the surface.

REFERENCES

ABDULLAEV, S. S. 1993 *Chaos and Dynamics of Rays in Waveguide Media.* Gordon and Breach.

AZETSU-SCOTT, K., JOHNSON, B. D. & PETRIE, B. 1995 An intermittent, intermediate nepheloid layer in Emerald Basin, Scotian Shelf. *Cont. Shelf Res.* **15**, 281–293.

BAINES, P. G. 1971*a* The reflexion of internal/inertial waves from bumpy surfaces. *J. Fluid Mech.* **46**, 273–291.

BAINES, P. G. 1971*b* The reflexion of internal/inertial waves from bumpy surfaces. Part 2. Split reflexion and diffraction. *J. Fluid Mech.* **49**, 113–131.

BAINES, P. G. 1973 The generation of internal tides by flat-bump topography. *Deep-Sea Res.* **20**, 179–205.

BAINES, P. G. & FANG, X.-H. 1985 Internal tide generation at a continental shelf/slope junction: a comparison between theory and a laboratory experiment. *Dyn. Atmos. Oceans* **9**, 297–314.

Berry, M. V. 1981 Regularity and chaos in classical mechanics, illustrated by three deformations of a circular 'billiard'. *Eur. J. Phys.* **2**, 91–102.

Bunimovich, L. A. 1980 Concerning certain properties of internal ocean waves with horizontally varying Väisälä-Brunt frequency. *Izv. Atmos. Ocean Phys.* **16**, 354–359.

Cacchione, D. & Wunsch, C. 1974 Experimental study of internal waves over a slope. *J. Fluid Mech.* **66**, 223–239.

Cushman-Roisin, B. 1993 Natural resonance of internal tides. *IAPSO Proc., PS-10* **18**, 321.

Cushman-Roisin, B. & Svendsen, H. 1983 Internal gravity waves in sill fjords: vertical modes, ray theory and comparison with observations. In *Coastal Oceanography*, (ed. H. G. Gade, A. Edwards & H. Svendsen), pp. 373–396, Plenum.

Cushman-Roisin, B., Tverberg, V. & Pavia, E. G. 1989 Resonance of internal waves in fjords: a finite-difference model. *J. Mar. Res.* **47**, 547–567.

Defant, A. 1941 *Physical Oceanography*, Vol. II. Pergamon Press.

deWitt, L. M., Levine, M. D., Paulson, C. A. & Burt, W. V. 1986 Semidiurnal internal tide in JASIN: observations and simulation. *J. Geophys. Res.* **91**, 2581–2592.

Garrett, C. 1991 Marginal mixing theories. *Atmos.-Oceans* **29**, 313–339.

Görtler, H. 1943 Über eine Schwingungserscheinung in Flüssigkeiten mit stabiler Dichteschichtung. *Z. Angew. Math. Mech.* **23**, 65–71.

Groen, P. 1948 Two fundamental theorems on gravity waves in inhomogeneous incompressible fluids. *Physica* **14**, 294–300.

Ivey, G. N. & Nokes, R. I. 1989 Vertical mixing due to the breaking of critical internal waves on sloping boundaries. *J. Fluid Mech.* **204**, 479–500.

Krauss, W. 1973 *Interne Wellen.* Gebrüder Bornträger.

Larsen, L. H. 1969 Internal waves incident upon a knife edge barrier. *Deep-Sea Res.* **16**, 411–419.

Lighthill, J. 1978 *Waves in Fluids.* Cambridge University Press.

Longuet-Higgins, M. S. 1969 On the reflexion of wave characteristics from rough surfaces. *J. Fluid Mech.* **37**, 231–250.

Magaard, L. 1962 Zur Berechnung interner Wellen in Meeresräumen mit nicht-ebenen Böden bei einer speziellen Dichteverteilung. *Kieler Meeresforsch.* **18**, 161–183.

Magaard, L. 1968 Ein Beitrag zur Theorie der internen Wellen als Störungen geostrophischer Strömungen. *Deutsche Hydrogr. Z.* **21**, 241–278.

Manton, M. J. & Mysak, L. A. 1971 Construction of internal wave solutions via a certain functional equation. *J. Math. Anal. Appl.* **35**, 237–248.

Mowbray, D. E. & Rarity, B. S. H. 1967 A theoretical and experimental investigation of the phase configuration of internal waves of small amplitude in a density stratified liquid. *J. Fluid Mech.* **28**, 1–16.

Munk, W. H. 1966 Abyssal recipes. *Deep Sea-Res.* **13**, 207–230.

Münnich, M. 1993 On the influence of bottom topography on the vertical structure of internal seiches. PhD thesis, ETH Zürich.

Münnich, M. 1994 The influence of bottom topography on internal seiches in continuously stratified media. In *Preprints of the Fourth Intl Symp. on Stratified Flows*, Vol. 2 (ed. E. Hopfinger, B. Voisin and G. Chavand). LEGI, Institut de Mécanique de Grenoble.

Pingree, R. D. & New, N. L. 1991 Abyssal penetration and bottom reflection of internal tidal energy in the Bay of Biscay. *J. Phys. Oceanogr.* **21**, 28–39.

Robinson, R. M. 1969 The effects of a vertical barrier on internal waves. *Deep-Sea Res.* **16**, 421–429.

Sandstrom, H. 1969 Effect of topography on propagation of waves in stratified fluids. *Deep-Sea Res.* **16**, 405–410.

Sandstrom, H. 1976 On topographic generation and coupling of internal waves. *Geophys. Astrophys. Fluid Dyn.* **7**, 231–270.

Schuster, H. G. 1984 *Deterministic Chaos.* Physik Verlag.

Thorpe, S. A. 1968 On the shape of progressive internal waves. *Phil. Trans. R. Soc. Lond.* A **263**, 563–614.

Turner, J. S. 1973 *Buoyancy Effects in Fluids.* Cambridge University Press.

Wunsch, C. 1968 On the propagation of internal waves up a slope. *Deep-Sea Res.* **15**, 251–258.

Wunsch, C. 1969 Progressive internal waves on slopes. *J. Fluid Mech.* **35**, 131–141.