Physica D 57 (1992) 476–506
North-Holland

PHYSICA D

Flux-minimizing curves for reversible area-preserving maps

R.L. Dewar [a] and J.D. Meiss [b]

[a] *Department of Theoretical Physics and Plasma Research Laboratory, Research School of Physical Sciences, The Australian National University, G.P.O. Box 4, Canberra, A.C.T. 2601, Australia*
[b] *Program in Applied Mathematics, University of Colorado, Box 526, Boulder, CO 80309, USA*

Received 25 November 1991
Revised 7 April 1992
Accepted 7 April 1992
Communicated by R.S. MacKay

Approximately invariant circles for area preserving maps are defined through a mean square flux, (φ_2-) variational principle. Each trial circle is associated with its image under the area preserving map, and this defines a natural one-dimensional dynamics on the x coordinate. This dynamics is assumed to be semiconjugate to a circle map on an angle variable θ. The consequences of parity ($\mathbb{P}$-) and time ($\mathbb{T}$-) reversal symmetries are explored. Stationary points of φ_2 are associated with orbits. Numerical calculations of φ_2 indicate a fractal dependence on rotation number, with minima at irrational numbers and maxima at the rationals. For rational frequency there are at least three stationary points of φ_2: a pair of minimax points, related by $\mathbb{PT}$-reversal, and a stationary, but not minimal, symmetric solution. A perturbation theory without small denominators can be constructed for the symmetric solution. A nonreversible solution for the (0, 1) resonance shows that this solution is highly singular and its circle map is not invertible.

1. Introduction

One of the most appealing ideas of classical Hamiltonian mechanics is the action–angle representation, which reduces the problem of understanding the dynamics of a system to the essentially *geometric* one of transforming to a phase-space coordinate system in which the new level surfaces of momentum (action) are invariant tori of the dynamical system, on which there is a trivial, lower dimensional dynamics involving only the angles.

As is now well known, this idea is fatally flawed, owing to the fact that the invariant sets of a generic Hamiltonian system do *not* form a set of smooth invariant nested tori. However, the Kolmogorov–Arnold–Moser (KAM) theorem [1–3] tells us that, for systems slightly perturbed from one of the exceptional (integrable) ones for which action–angle coordinates exist, there remains a finite measure of invariant tori topologically equivalent to the original ones; but we also know that between these KAM tori there are chaotic regions (invariant sets of the same dimension as the phase space, containing unstable orbits), islands (stable orbits surrounded by invariant tori with different topology from the original ones), and invariant cantor sets (cantori) representing a residue of KAM surfaces which have ceased to exist [4].

In many cases, nevertheless, the action–angle description does provide an approximation, in some sense, to a nonintegrable system. A case in point is the Hamiltonian flow describing the magnetic field lines in nonaxisymmetric plasma containment devices (stellarators). The vacuum fields are carefully designed, by visual inspection of Poincaré plots, or by more sophisticated dynamical methods [5], so as to appear close to integrable. Introduction of a plasma will al-

ter the field somewhat, but for moderate plasma pressures the magnetic field lines will tend to remain close to nested tori. Because the magnetic field causes plasma properties to be highly anisotropic, it is desirable to use a curvilinear coordinate system, with level surfaces of one of the coordinates corresponding to the approximately invariant surfaces of the magnetic field system [6].

Another important application of approximate action–angle transformations is to the estimation of the long-time stability of particle orbits in accelerators [7].

We thus have practical contexts in which it is important to have an action–angle-like coordinate system approximating a nonintegrable system. We also need a measure for the departure of the true system from the approximating one, and in particular a measure for the transport of flux and particles across the "approximately invariant surfaces" since the aim in both the plasma and accelerator applications is to confine particles for a very long time.

Another plasma context where it has been suggested that an action–angle-like transformation would give an improved description of the dynamics is in the field of wave-particle interaction theory, where it is known as the oscillation-centre representation [8]. This has been a very useful technique in regions of phase space where the motion is essentially integrable [9], but remains problematical in chaotic regions where perturbative methods break down [10]. One intriguing question raised by the attempts at a perturbation theory for oscillation-centre variables is that of causality: whether to use acausal [8] or causal [11] propagators for constructing the generating function. We shall find that the issue of time-reversal invariance arises even in the simple mapping problem which is the subject of this paper.

Rather than tackle the full Hamiltonian problem we choose in this paper to restrict attention to reversible, area-preserving (symplectic) twist maps of the plane. These form a prototype for Hamiltonian systems more susceptible to numerical experimentation than continuous time systems, and are consequently already well explored, especially the standard map (see e.g. MacKay et al. [4]). Since these maps are much simpler than full Hamiltonian systems (especially those corresponding to magnetic field lines in plasma confinement problems) their study forms a good test bed for developing numerical methods as well as clarifying conceptual problems.

Our basic idea for constructing approximately invariant curves in these mappings is one of flux minimization. This was suggested by Wigner [12] as a way of partitioning a system into weakly interacting subsystems. More recently [13], a global flux minimization principle was enunciated for defining an optimum action–angle-like representation of a given nonintegrable Hamiltonian system. Another global minimization principle for approximately invariant curves of maps has recently been introduced by Bazzani et al. [14].

A drawback of such global principles is that they involve minimization over the full phase space rather than just the vicinity of an approximately invariant curve. MacKay et al. [4] have studied the flux through "partial barriers" containing quasi-periodic orbits and have observed that the flux appears to be minimal for "noble" winding numbers. This is a local method, but it is based on extremizing action [15,16], not flux. Furthermore, the flux as defined by MacKay et al. [4] is *independent* of choice of partial barrier (provided it contains the orbit) and so minimizing it cannot give rise to any unique approximately invariant curve.

A new variational principle based on extremizing the *mean square flux* through trial partial barriers was recently proposed by Meiss and Dewar [17]. This is similar to the principle advanced by Dewar [13], but is local rather than global, and is formulated for mappings rather than Hamiltonian systems. It was shown [17] that the intersections of unconstrained rota-

tional curves which extremize the mean square flux are orbits, thus showing that the principle is fundamentally related to the invariant sets of the dynamical system, a property which would seem essential for a natural generalization of action–angle coordinates. It is this new principle which we develop further in the present paper, introducing a parametric representation which allows questions of conjugacy and time-reversal invariance to be investigated, and presenting numerical results from its implementation for the standard map.

The numerical results reveal the existence of several fundamentally different classes of solutions. This paper does not give a complete catalogue of possible solutions, but rather explores a few special cases with a view to achieving an initial classification of classes of solutions on the basis of their symmetry under time and parity reversal. We also seek to establish whether some of the one-dimensional maps in the theory can be assumed to be invertible, and find in general that they cannot.

In section 2 we develop the properties of area-preserving twist maps whose generating functions have time reversal and parity symmetries, generalizing work of MacKay [18], and also present a new discussion of discrete translation invariance and discrete Galilean transformation invariance in periodic systems. The standard map is an example of a dynamical system possessing these symmetries.

As in action–angle theory there is a lower dimensional dynamics associated with the approximately invariant curves. In the case of area preserving maps the associated dynamics is one dimensional, and in the case of periodic area preserving maps the one-dimensional dynamics is defined on a circle. In section 3 we review the properties of one-dimensional dynamics on a circle, which we will use to generate the approximately invariant curves of the area-preserving map. In particular, a representation in terms of a semiconjugacy to a circle map is introduced which will allow generalization of the formulation of Meiss and Dewar [17] to the case where rotational circles are not graphs over x. Also, a special parametric representation is defined which allows reversibility to be made manifest.

In section 4 we show how a circle map defines a rotational circle C and its image $C^* \equiv T(C)$ in the two-dimensional cylindrical phase space of the area-preserving map T, and in section 5 we define the quadratic flux φ_2 and the area $\mathcal{A}$. We show that φ_2 is invariant under time and parity reversal provided that the circle map is also reversed, thus showing that extremizing circle maps are either reversible or occur in pairs related by reversal. We derive the Euler–Lagrange equation for the variational principle in section 6 and show that it implies that some sets of intersections of C and C^* associated with unconstrained extrema of φ_2 must be orbits under T.

We present an analytic calculation of φ_2 and $\mathcal{A}$ for a one-mode trial function in section 7 and in section 8 we describe a numerical implementation of a multimode minimization. Results of a scan in rotation number for the standard map show in a striking fashion the hierarchy of nonlinear resonances described by the Farey tree. Local minima of φ_2 appear to be associated with quasiperiodic orbits, corresponding to KAM curves or cantori, where the flux is low, whereas periodic orbits are associated with minimax points of φ_2, where the flux goes through a maximum in one direction, but is minimal in all other orthogonal directions. We present numerical and analytic periodic solutions of the Euler–Lagrange equation for the lowest order rational rotation numbers in section 9, showing that solutions which minimize the quadratic flux (except for the one descending direction due to the flux peaking at rational surfaces) have broken $\mathbb{PT}$-symmetry. Conclusions are given in section 10.

A useful identity for deriving properties of circle maps is presented in appendix A, while a time-symmetric representation for circle maps is given in appendix B and applied to the problem of finding the Euler–Lagrange equation for con-

strained variations in appendix C.

2. Area-preserving maps

2.1. Hamiltonian formulation

An area-preserving map T: $\mathbb{R}^2 \to \mathbb{R}^2$ such that a point (x, y) maps to a new point (x^*, y^*) can be defined implicitly by

$$y^* = y_+(x, x^*) \equiv F_2(x, x^*),$$
$$y = y_-(x, x^*) \equiv -F_1(x, x^*), \tag{2.1}$$

where the functions $y_\pm$ are defined in terms of a generating function of Type 1 [19], $F(x, x^*)$, with $F_1(x, x^*) \equiv \partial F(x, x^*)/\partial x$ and $F_2(x, x^*) \equiv \partial F(x, x^*)/\partial x^*$.

We require the twist condition,

$$\partial y_-(x, x^*)/\partial x^* = -F_{12}(x, x^*) > C > 0, \tag{2.2}$$

for some positive constant C, so that eqs. (2.1) can globally be solved for x^* and y^* in terms of x and y. As T may be thought of as being the return map for a periodically perturbed Hamiltonian system, we shall term x the "position" and y the "momentum".

It can be seen by inspection of eq. (2.1) that the generating function of the inverse map T^{-1} is

$$F^{(-1)}(x, x^*) \equiv -F(x^*, x). \tag{2.3}$$

2.2. Lagrangian formulation

An orbit segment of T is defined as a sequence $(x_0, x_1, \ldots, x_n)$ such that the action

$$W[x_0, x_1, \ldots, x_n] \equiv \sum_{t=0}^{n-1} F(x_t, x_{t+1}) \tag{2.4}$$

is stationary for all variations of $(x_0, x_1, \ldots, x_n)$ with x_0 and x_n held fixed. This yields the Euler–Lagrange equations

$$F_2(x_{t-1}, x_t) + F_1(x_t, x_{t+1}) = 0,$$
$$t = 1, 2, \ldots, n-1, \tag{2.5}$$

which by eqs. (2.1) is the statement that

$$y_-(x_t, x_{t+1}) = y_+(x_{t-1}, x_t). \tag{2.6}$$

By eqs. (2.1) this shows that the x_t do indeed correspond to the x-components of iterates of T applied to (x_0, y_0), with $y_0 \equiv y_-(x_0, x_1)$. Thus the second-order difference equation eq. (2.5) is an alternative, Lagrangian statement of the discrete Hamiltonian dynamics described by T. However, the Lagrangian dynamics defined by eq. (2.5) does not uniquely specify T.

This is a consequence of the fact that the Lagrangian dynamics does not follow from one unique generating function, but, rather, from an equivalence class. We define a *dynamically equivalent* generating function F_e as one which generates the same orbit segments as F for any initial and final positions x_0 and x_n (n arbitrary). Thus all generating functions $F_e(x, x^*)$, defined by

$$\sigma F_e(x, x^*) = F(x, x^*) - U(x) + U(x^*) - K, \tag{2.7}$$

where K and σ are arbitrary constants and U an arbitrary function, are dynamically equivalent since the associated actions differ only by a constant factor and by endpoint contributions (cf. the dynamical equivalence of Lagrangians differing by total time derivatives in the continuous time case). The equivalence can also be demonstrated by direct substitution into eq. (2.5). In the following we shall assume that $\sigma = \pm 1$.

2.3. Lagrangian symmetries

The fundamental dynamical symmetries which we shall use are most easily defined

in the Lagrangian description before translating them into the Hamiltonian representation. We define the *time-reversal* ($\mathbb{T}$) and *parity-reversal* ($\mathbb{P}$) symmetries of a dynamical system as the properties that, given any orbit segment $(x_0, x_1, \ldots, x_n)$, the sequences $(x_n, x_{n-1}, \ldots, x_0)$ and $(-x_0, -x_1, \ldots, -x_n)$, respectively, are also orbit segments. In terms of the action, $\mathbb{T}$-symmetry is the property that

$$W(x_0, x_1, \ldots, x_n) \doteq W(x_n, x_{n-1}, \ldots, x_0), \tag{2.8}$$

and $\mathbb{P}$-symmetry the property that

$$W(x_0, x_1, \ldots, x_n) \doteq W(-x_0, -x_1, \ldots, -x_n), \tag{2.9}$$

where $\doteq$ denotes equality up to a constant plus terms depending only on the endpoints x_0 and x_n. This is similar to the idea of dynamical equivalence introduced in the previous section, except that here we have an active interpretation of the transformation—the generating function remains the same but the orbits are transformed. We have not found it necessary to allow a factor analogous to σ in eq. (2.7) in the definition of equivalence for the maps we use.

These hold for any n if F has the symmetries $\mathbb{T}$:

$$F(x, x^*) = F(x^*, x) + Q(x) - Q(x^*), \tag{2.10}$$

for some function Q and any x and x^*, and $\mathbb{P}$:

$$F(x, x^*) = F(-x, -x^*) + P(x) - P(x^*), \tag{2.11}$$

where P is necessarily odd (apart from a trivial constant). The differences $Q(x) - Q(x^*)$ and $P(x) - P(x^*)$, must be allowed in general because these contribute only endpoint terms to the action, no matter what value of n is taken. Setting $x = x^* = 0$ we see that there can be no constant term in eq. (2.10) or eq. (2.11) analogous to K in eq. (2.7).

We also define the combined $\mathbb{PT}$-symmetry as the condition that

$$F(x, x^*) = F(-x^*, -x) + R(x) - R(x^*), \tag{2.12}$$

for some function R, which must be even for consistency. This is automatic if both $\mathbb{T}$ and $\mathbb{P}$ symmetries apply, in which case $Q(x) = R(x) + P(x)$. This symmetry is particularly useful because, unlike the $\mathbb{T}$ and $\mathbb{P}$ symmetries, it does not reverse the sign of the velocity $x_{t+1} - x_t$ and can thus define a symmetry of a given orbit, rather than relating two generally different orbits.

The $\mathbb{T}$ and $\mathbb{P}$ symmetries are common symmetries of physical Hamiltonians and it is interesting that there are other analogues of physical symmetries which physically motivated maps possess. The physical configuration variable, x is often an angle (with a period we take to be 1) or the position coordinate for motion in a periodic potential. Accordingly we take the configuration space of T to be a covering space for a circle, and the phase space to be a covering space for a cylinder; so that, for example, the line $y = \text{const.}$ represents a circle drawn around the cylinder. Thus we assume our dynamical system to be periodic, a discrete version of the translation invariance of physical systems. By a *periodic system* we mean that, given any orbit segment $(x_0, x_1, \ldots, x_n)$, the sequence $(x_0 + m, x_1 + m, \ldots, x_n + m)$ is also an orbit segment for any integer m. In terms of the generating function, the periodicity assumption implies

$$F(x, x^*) = \tilde{F}(x, x^*) + A(x) - A(x^*) + ax, \tag{2.13}$$

for some function A, and function $\tilde{F}(x, x^*)$ with the periodicity

$$\tilde{F}(x + m, x^* + m) = \tilde{F}(x, x^*). \tag{2.14}$$

The possibility of having the linear term ax in eq. (2.13) must be admitted because actions differing by a constant (in this case, anm) give the same variational equations. The terms $A(\cdot)$ in eq. (2.13) can be removed without affecting the Lagrangian dynamics by a transformation of the type eq. (2.7) with $\sigma = 1$, $K = 0$, but the term ax cannot be. Thus a parametrizes a one-parameter family of fundamentally distinct systems. In the special case $a = 0$ we can reduce F to the periodic form $\tilde{F}$ without affecting the Lagrangian dynamics. In the Hamiltonian framework such systems, with periodic generating functions, are called *exact symplectic* (see section 4). As a consequence of the following lemma we shall need to consider only this class:

Lemma 1. Periodic systems which also have $\mathbb{T}$ and/or $\mathbb{P}$ symmetry must have $a = 0$.

This is a consequence of the fact that the term ax is asymmetric and odd.

The final symmetry which we wish to define is *periodic Galilean invariance*: Given an orbit segment $(x_0, x_1, \ldots, x_n)$, the sequence $(x_0, x_1 - k, \ldots, x_n - nk)$ is also an orbit segment for any integer k. Assuming the system to be exact symplectic, Galilean invariance requires the generating function to have the property

$$F(x, x^*) = F(x, x^* - k) + B_k(x) - B_k(x^*) + C_k, \tag{2.15}$$

for some function B_k and constant C_k.

2.4. Conjugacies

Hoewver, taking $\nu = 0$, we show in this section that the concept of Lagrangian dynamical equivalence introduced in section 2.2 corresponds in the Hamiltonian, phase space description to a special case of equivalence or *conjugacy* of maps related by transformations of the form $T_e \equiv S \circ T \circ S^{-1}$. Here $\circ$ denotes composition of functions, e.g. $f \circ g(\cdot) \equiv f(g(\cdot))$. In this case the position-preserving transformations S are determined by the transformation of the generating function. Replacing F by F_e in eqs. (2.1) we see that eq. (2.7) implies a phase space transformation S:

$$x \mapsto \bar{x} = x,$$
$$y \mapsto \bar{y} = \sigma[y + U'(x)]. \tag{2.16}$$

We consider only two possibilities for σ, the first case being $\sigma = +1$. In this case S is canonical because it conserves the fundamental Poisson bracket $\{x, y\}$ [19] (and hence conserves area). (In geometrical language such transformations are said to be symplectic because they preserve the symplectic two form $\omega = \mathrm{d}x \wedge \mathrm{d}y$ [20].)

The second possibility is $\sigma = -1$. Since it changes the sign of an element of area, but preserves its absolute value, S is in this case *antisymplectic* [20] (or anticanonical [21]). It is also an involution. (An *involution* is a map S such that $S \circ S = \mathrm{Id}$, the identity transformation.) Thus the maps generated by eq. (2.7) with $\sigma = -1$ form the dynamical equivalence class of maps conjugate to T under antisymplectic involutions which preserve x.

We also wish to consider phase space transformations which do not preserve x. These can also be related to the generating function by considering an extended dynamical equivalence under change of variable $\bar{x}_t = f_t(x_t)$, with dynamically equivalent generating function $\bar{F}$ being given by

$$\sigma \bar{F}(\bar{x}_t, \bar{x}_{t+1}) = F(x_t, x_{t+1}) + U(x_{t+1}) - U(x_t) - K, \tag{2.17}$$

which implies the transformation S:

$$\bar{x}_t = f_t(x_t),$$
$$\bar{y}_t = \sigma[y + U'(x)]/f_t'(x_t), \tag{2.18}$$

of which eq. (2.16) is the special case where f is the identity.

For example, in order to handle $\mathbb{PT}$-symmetry we shall need to find the generating function for

the dynamical equivalence class of maps conjugate to T under antisymplectic transformations which change the sign of x, i.e. for which $\overline{x} = -x$. Then the above equations apply with $f_t(x) = -x$ and $f_t'(x) = \sigma = -1$. In this case S is an involution if U is even.

An example where the transformation depends on t is the Galilean transformation $\overline{x}_t = x - ct$ (where c and t are here taken to be integers). In this case $f_t'(x) = 1$.

2.5. Hamiltonian symmetries

Corresponding to the Lagrangian symmetries of section 2.3 there are symmetries in the phase space description. The first type of symmetry is a conjugacy which leaves T invariant: $T = S \circ T \circ S^{-1}$. The transformation S is found from eq. (2.18) by choosing $\sigma = 1$ in eq. (2.17) and U so as to cancel any corresponding function in the Lagrangian symmetry relation, ending up with $\overline{F}(\overline{x}_t, \overline{x}_{t+1}) = F(\overline{x}_t, \overline{x}_{t+1})$ (hence $\overline{T} = T$). As an example $\mathbb{P}$-invariance, eq. (2.11), is of this type, where we take $U(\cdot) = P(\cdot)$. So also is periodic Galilean invariance, eq. (2.15), where we take $U(\cdot) = B_k(\cdot)$. An interesting consequence of this will be discussed in section 2.6.

Of more interest for the moment are cases where T is invariant under a combination of conjugacy and time reversal:

$$T = S \circ T^{-1} \circ S^{-1}, \tag{2.19}$$

which is known as weak reversibility [22]. When S is an antisymplectic involution, so that $S^{-1} = S$, T is said to be *reversible* [23] with respect to the symmetry S.

To find S from eq. (2.18), choose $\sigma = -1$ in eq. (2.17) and U so as to cancel any corresponding function in the Lagrangian symmetry relation, ending up with $\overline{F}(\overline{x}_t, \overline{x}_{t+1}) = -F(\overline{x}_t, \overline{x}_{t+1})$, the generating function for the inverse transformation by eq. (2.3).

Using eq. (2.10) we see that $\mathbb{T}$-symmetry implies weak reversibility under transformations of the type eq. (2.16), with $U(\cdot) = Q(\cdot)$. From eq. (2.12) and eq. (2.18) we see, taking $U(\cdot) = R(\cdot)$ (even), that $\mathbb{PT}$-symmetry implies reversibility under antisymplectic involutions S_1:

$$\begin{aligned} \overline{x} &= -x, \\ \overline{y} &= y + R'(x). \end{aligned} \tag{2.20}$$

The assumption of reversibility under S_1 is often made [4,24,25] in the study of invariant sets of area-preserving maps for convenience in locating periodic points. We have here shown how it is related to the fundamental parity and time-reversal symmetries important in theoretical physics [21].

A *symmetry line* is defined as a curve of fixed points of S. For example the symmetry line of S_1 is the curve $x = 0$.

The utility of reversibility of T under antisymplectic involutions S in locating periodic orbits derives from the following lemma:

Lemma 2. Every orbit which has a point on a symmetry line is symmetric. Conversely, if an orbit is symmetric under S, then it has a point on the symmetry line of either S or TS.

To see this, let (x_0, y_0) be a point on the symmetry line of S. Then, by definition, $(x_0, y_0) = S(x_0, y_0)$. Repeatedly applying eq. (2.19) to $(x_j, y_j) \equiv T^j(x_0, y_0)$ implies that $(x_j, y_j) = S \circ T^{-j}(x_0, y_0) \equiv S(x_{-j}, y_{-j})$. Thus the point (x_{-j}, y_{-j}) is the reflection of (x_j, y_j). Conversely suppose that the orbit is symmetric, then $(x_j, y_j) = S(x_0, y_0)$ for some j. If j is even eq. (2.19) implies that $T^{-j/2}(x_j, y_j) = S \circ T^{j/2}(x_0, y_0)$, while if j is odd we have $T^{-(j-1)/2}(x_j, y_j) = T \circ S \circ T^{(j+1)/2}(x_0, y_0)$, both of which imply there is a point on a symmetry line.

2.6. Examples

As an example, consider the generalized standard map

$$F(x,x^*) = \tfrac{1}{2}(x-x^*)^2 - V(x). \tag{2.21}$$

This satisfies eq. (2.10) with $Q(x) = -V(x)$, so that the map is $\mathbb{T}$-reversible for any V. Furthermore, if V is an even function then $F(x,x^*)$ obeys eq. (2.12), so that the map is then $\mathbb{PT}$-reversible under the antisymplectic involution S_1 of eq. (2.20), with $R \equiv -V$. We also assume that V is periodic, so that F is exact periodic as defined in eq. (2.14). It also satisfies the condition, eq. (2.15), for periodic Galilean invariance with $B_k(x) = -kx$ and $C_k = -\frac{1}{2}k^2$. Thus, from eq. (2.18), T is invariant under the transformation

$$\begin{aligned} \overline{x}_t &= x_t - ct, \\ \overline{y}_t &= y_t - c. \end{aligned} \tag{2.22}$$

Recalling that c is an arbitrary integer and using the exact periodicity of F, we see that the generalized standard map is periodic in the y-direction as well as the x-direction.

The generating function above is not symmetric between x and x^*, but, by choosing $U(x) = -\frac{1}{2}V(x)$ in eq. (2.7) we can transform to the dynamically equivalent, symmetric generating function

$$\begin{aligned} F_{\mathrm{sym}}(x,x^*) &= \tfrac{1}{2}(x-x^*)^2 \\ &\quad -\tfrac{1}{2}[V(x) + V(x^*)]. \end{aligned} \tag{2.23}$$

The corresponding $\overline{y}$ corresponds to the symmetry coordinate y of [4]. In general, generating functions which obey the $\mathbb{T}$-reversal symmetry condition, eq. (2.10), can be symmetrized in this fashion.

The commonly used standard map has a generating function of the form eq. (2.21), with

$$V(x) = -\frac{k}{4\pi^2}\cos 2\pi x, \tag{2.24}$$

where k is the nonlinearity parameter. This is an even function so the map

$$\begin{aligned} x^* &= x + y - \frac{k}{2\pi}\sin 2\pi x, \\ y^* &= y - \frac{k}{2\pi}\sin 2\pi x, \end{aligned} \tag{2.25}$$

generated by F is $\mathbb{PT}$-reversible. For the standard map it is known [24] to high precision that the last rotational invariant circle (a rotational circle has the same topology as the circle $y = \text{const.}$) to break up as k is increased has golden mean rotation number $\gamma \equiv (1+\sqrt{5})/2$. In fact by periodic Galilean invariance and $\mathbb{P}$ symmetry the circles with rotation numbers $\nu = n \pm \gamma$, where n is any integer are equivalent, and are therefore destroyed simultaneously. The golden circle is destroyed at $k = k_c(\gamma) \equiv 0.971635406\cdots$. Above this value, there is still an invariant set with this rotation number, but it is a Cantor set, and hence is called a *cantorus* [26].

3. Circle maps

Just as in conventional action–angle theory the dimensionality of the dynamics is reduced to half that of the phase space (being in terms of the angles alone, on a given invariant torus), so we associate a one-dimensional dynamics with the approximately invariant rotational circles we seek in our two-dimensional, cylindrical phase space. (Note that, unlike the integrable case, most orbits under this one-dimensional dynamics do not correspond to orbits under T in the full phase space.) In this section we discuss the representation of circle map dynamics, and in the next we discuss the phase space representation for rotational circles.

3.1. Representations

A one dimensional map

$$x \mapsto x^* = \alpha(x) \tag{3.1}$$

is called a circle map if $\alpha(x) - x$ is periodic (cf. eq. (2.14)). Thus one representation of the map is

$$\alpha(x) = x + \Omega + \sum_{m=1}^{\infty} a_m \cos 2\pi m x + \sum_{m=1}^{\infty} b_m \sin 2\pi m x. \tag{3.2}$$

As in action–angle theory we represent the one-dimensional dynamics in terms of a new variable θ which evolves as

$$\theta \mapsto \theta^* \equiv \rho(\theta), \tag{3.3}$$

where ρ is a differentiable circle map which in some sense is simpler than α. For example, we will sometimes use the rigid rotation $\rho = R_\nu$ where

$$R_\nu(\theta) \equiv \theta + \nu. \tag{3.4}$$

In any case, the θ-dynamics will be assumed related to that of x by the change of variables,

$$x = X(\theta), \quad x^* = X(\theta^*), \tag{3.5}$$

where X is a differentiable circle map. We regard the θ-dynamics as, in a sense, more fundamental than the x-dynamics, despite the fact that there is considerable freedom in choosing X. In the case where an $X(\theta)$ can be found which is continuous and has a continuous inverse, the change of variables is a homeomorphism (topological equivalence) and the map α is conjugate to the map ρ:

$$\alpha(x) = X \circ \rho \circ X^{-1}(x). \tag{3.6}$$

However when X is not invertible, we can go uniquely from θ-space to x-space, but not vice versa. The relation between the x-space and θ-space dynamics is then called a *semiconjugacy*. Since α is not a single valued function in this case, we shall reinterpret the symbol α to denote the set $\{X(\theta), X \circ \rho(\theta) \mid \theta \in \mathbb{R}\}$, that is, the curve in the x–x^* plane described by eq. (3.5). This causes no problem, since, unlike our previous formulation [17], we now do not actually use α, but rather work purely in terms of X and ρ.

In fact we shall concentrate on two cases in this paper. In the first case we fix $\rho = R_\nu$ and allow X to vary. In the second we fix $X(\theta) = \theta$, and allow ρ to vary. (It would appear however, from an argument in section 9.3, that there are cases where both ρ and X must be allowed to vary in a nontrivial way.)

3.2. Conjugacy to rigid rotation

A diffeomorphism of a circle, α, possesses a rotation number ν (preserved under conjugacy) defined by

$$\nu \equiv \lim_{n\to\infty} \alpha^n(x)/n, \tag{3.7}$$

where $\alpha^n \equiv \alpha \circ \alpha^{n-1}$. By a theorem of Denjoy (see e.g. [27]), if ν is irrational and $\log \alpha'(x)$ is of bounded variation then there exists an invertible function $X(\theta)$ such that $\rho(\theta) = R_\nu(\theta)$; that is, such that α is conjugate to the rigid rotation R_ν. By a theorem of Herman (see e.g. [27]), if ν satisfies a Diophantine condition, $|\exp 2i\pi m\nu - 1| > C_\epsilon |m|^{-1-\epsilon}$ for any nonzero integer m and for any $\epsilon > 0$ and for some C_ϵ (such ν are of full measure), and if α is k-times differentiable, where $k \geq 3$, then α is conjugate to a rigid rotation by an X which is $(k - 1 - \eta)$-times differentiable for any $\eta > 0$. If α is analytic then so is X. (It is shown in [27] that the restriction to $k \geq 3$ is unnecessarily restrictive, but was needed in the original proof for technical reasons). Note that the rotation number is necessarily unique only if α is invertible.

These theorems rely on ν being irrational. However if Ω in eq. (3.2) is varied (keeping the Fourier coefficients fixed), one finds that, for almost all Ω, ν is "locked" to a rational value because of the occurrence of attracting periodic

orbits of the circle map. That is, Ω as a function of ν has a "Devil's staircase" behaviour. Thus it is not advisable to restrict ourselves to circle maps conjugate to rigid rotations and we must assume in general that ρ has nonvanishing Fourier coefficients in the representation analogous to eq. (3.2). One might hope to reduce the nonuniqueness of this representation in a way useful for numerical purposes by choosing X so as to minimize the Fourier spectral width of ρ using, for instance, the "spectral condensation" algorithm of Hirshman and Meier [28]. One may well be able to reduce the amplitude of the low-order Fourier modes of ρ by a good choice of X, but Herman's theorem shows that it would be misguided to try to economize on the overall number of modes used by removing the asymptotic tail of the spectrum of ρ using conjugacy, since the Fourier spectrum of X in general decays *slower* than that of α. Thus more modes would be used to represent X to a given accuracy than would be saved from ρ.

3.3. Reversibility

It will turn out that our circle map can inherit the $\mathbb{PT}$-symmetry of T. We define a $\mathbb{PT}$-reversible circle map α as one for which

$$\alpha(x) = -\alpha^{-1}(-x), \tag{3.8}$$

where α^{-1} is the inverse map (assuming it exists). Since this reversibility condition is preserved under conjugation *provided X is an odd function*, we can extend the definition to the case where X is not monotonic by redefining $\mathbb{PT}$-reversibility of the one-dimensional dynamics to be the condition

$$\rho(\theta) = -\rho^{-1}(-\theta) \tag{3.9}$$

for any X such that ρ^{-1} exists and for which $X(\theta) = -X(-\theta)$. (If no such X exists then the dynamics is *a forteriori* not $\mathbb{PT}$-reversible.) In particular, note that the rigid rotation R_ν is always $\mathbb{PT}$-reversible, so that any α conjugate, with odd X, to a rigid rotation is reversible. We shall always consider X to be odd so that the θ- and x-dynamics have the same symmetry properties.

Although eq. (3.8) is strictly meaningless when X is not a homeomorphism, we can interpret $\mathbb{PT}$-reversibility geometrically as a reflection symmetry of the curve eq. (3.5) about the diagonal lines $x^* + x = k$, with k any integer (see figs. 2 and 4).

The representation eq. (3.2) has the drawback that there is no simple relationship between the Fourier series of the map and that of its inverse, except that their constant parts Ω are equal in magnitude and opposite in sign (see eq. (A.2)). Thus there is no simple way to characterize $\mathbb{PT}$-reversibility, eq. (3.8), in this representation. A representation in which $\mathbb{PT}$-reversibility (or otherwise) of ρ is manifest is given in appendix B.

4. Rotational circles

A *rotational circle* is a loop in the cylindrical phase space of T which cannot be shrunk to a point. That is, it is a curve with the same topology as the invariant curves $y = \text{const.}$ of the unperturbed map. Let $\eta \in \mathbb{R}$ parametrize the position of any point on the rotational circle C. That is C is the map $\eta \mapsto (x,y) \equiv (x_-(\eta), y_-(\eta))$, where the functions $x_-(\eta)$ and $y_-(\eta)$ have the periodicities $(x_-(\eta + n), y_-(\eta + n)) = (x_-(\eta) + n, y_-(\eta))$ for any integer n. C need not be the graph of a function of x: it may have S-shaped sections or even cross itself. The iterate of C, $C^* \equiv T(C)$, similarly defines a rotational circle $C^* : \eta \mapsto (x^*, y^*) \equiv (x_+(\eta), y_+(\eta))$.

From eqs. (2.1) we have

$$y_\pm(\eta) = y_\pm(x_-(\eta), x_+(\eta)). \tag{4.1}$$

Thus, specification of the pair of functions $x_-(\eta)$ and $x_+(\eta)$, which is a parametric description of the *single* circle map $\alpha : x \mapsto x^*$, defines in a natural way the *two* rotational curves C and C^*.

The parametrization of C and C^* is so far arbitrary. We use a representation based on eq. (3.3), setting $\theta = \eta$ for x_- and $\theta = \rho(\eta)$ for x_+ so that

$$x_-(\eta) = X(\eta),$$
$$x_+(\eta) = X\circ\rho(\eta). \tag{4.2}$$

This parametrization treats $x_-(\eta)$ differently from $x_+(\eta)$, so that it is not manifestly symmetric with respect to time reversal. An alternative, time-symmetric parametrization is developed in appendices B and C.

The parametric specification of C and C^* in terms of η is inconvenient for purposes of comparing the two curves because a given value of η corresponds to different values of x on the two curves. To rectify this we introduce an alternative parametric description with a common $x = X(\theta)$ and with $y = Y_-(\theta)$ on C and $y = Y_+(\theta)$ on C^*, with

$$Y_-(\theta) \equiv y_-(\theta),$$
$$Y_+(\theta) \equiv y_+\circ\rho^{-1}(\theta), \tag{4.3}$$

(see fig. 1). Note that $Y_+(\theta)$ is single valued only if ρ is monotone, which we shall assume in this section to be the case. If ρ is monotone, the only allowed turning points of x in this representation come from possible turning points of $X(\theta)$. That is, if one of the curves has a point where its tangent is vertical, then the other also has a corresponding point sharing the same vertical tangent.

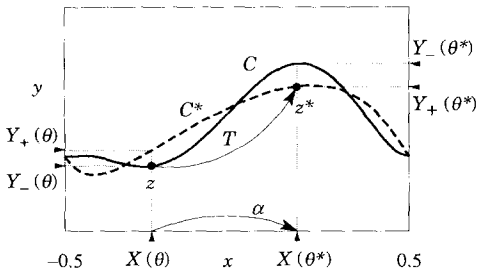

Fig. 1. Phase plane of area preserving map T, showing rotational curve C and its image $C^* = T(C)$ with the θ-parametrization of a point $z = (x,y)$ and its image $z^* = (x^*,y^*)$.

Using eqs. (2.1), (4.1) and (4.2) in eq. (4.3) gives $Y_\pm$ explicitly in terms of the generating function F:

$$Y_+(\theta) = F_2(X\circ\rho^{-1}(\theta), X(\theta)), \tag{4.4}$$

and

$$Y_-(\theta) = -F_1(X(\theta), X\circ\rho(\theta)). \tag{4.5}$$

The area $\mathcal{A}$ under C is defined by

$$\mathcal{A} = \int_0^1 Y_-(\theta)\,X'(\theta)\,\mathrm{d}\theta$$
$$= \int_0^1 y_-(\eta)\,x_-'(\eta)\,\mathrm{d}\eta, \tag{4.6}$$

and similarly for the area $\mathcal{A}^*$ under C^*, with $-$ replaced by $+$.

The upward *flux* φ_1 through C is defined to be the area above C which is below C^*

$$\varphi_1 \equiv \int_0^1 \Delta Y(\theta)\,H(\Delta Y)\,X'(\theta)\,\mathrm{d}\theta, \tag{4.7}$$

where H is the Heaviside step function and

$$\Delta Y(\theta) \equiv Y_+(\theta) - Y_-(\theta) \tag{4.8}$$

is the (signed) vertical distance between C^* and C at $x = X(\theta)$. Substitution of eqs. (4.4) and (4.5) into eq. (4.8) and comparison with eq. (2.5) demonstrates

Lemma 3. Define $x_j \equiv X\circ\rho^j(\theta_0)$. Then $(x_0, x_1, \ldots, x_n)$ is an orbit segment if and only if $\Delta Y\circ\rho^i(\theta_0) = 0$ for $i = 1,2,\ldots,n-1$.

The *net flux* $\mathcal{F}$ through C is defined to be the difference between the upward and downward

fluxes; i.e. $\mathcal{F} = \mathcal{A}^* - \mathcal{A}$. Using the generating function forms in eq. (2.1) we find

$$\mathcal{F} = \int_0^1 [F_1(x_-, x_+)\, x'_-(\eta) + F_2(x_-, x_+)\, x'_+(\eta)]\, d\eta. \tag{4.9}$$

We recognize the integrand as the perfect differential dF, so $\mathcal{F} = F(x_-(0) + 1, x_+(0) + 1) - F(x_-(0), x_+(0))$. Maps for which this vanishes, i.e. which preserve area globally as well as locally, are said to be *exact symplectic*. Since, by lemma 1, either $\mathbb{T}$ or $\mathbb{P}$ symmetry (or both) assures this condition can be imposed without affecting the Lagrangian dynamics, we assume that $\mathcal{F} \equiv 0$. Equivalently, $\mathcal{A} = \mathcal{A}^*$. This allows us to give an alternative definition of the upward flux φ_1:

$$\varphi_1 = \tfrac{1}{2} \int_0^1 |\Delta Y(\theta)|\, X'(\theta)\, d\theta. \tag{4.10}$$

5. Mean square flux

5.1. Definition

As we showed in the previous section, the upward flux is completely balanced by the downward flux, independent of the choice of C. Although φ_1 depends on the choice of C (indeed, it can be made zero if C can be found which is invariant under T), MacKay et al. [4] found that, for a given invariant set, the upward flux through their partial barriers was independent of choice of partial barrier. We therefore conclude that minimizing φ_1 cannot give rise to any unique C.

In seeking a variational principle we are naturally led to consider quadratic functionals. Accordingly we now define a "second moment" of flux, the *mean square flux*, as

$$\varphi_2 \equiv \tfrac{1}{2} \int_0^1 [\Delta Y(\theta)]^2\, X'(\theta)\, d\theta. \tag{5.1}$$

From eq. (4.4) we see that $Y_+(\theta)$, and hence $\Delta Y(\theta)$, are defined only if ρ is invertible. However, we shall find in section 9.3 that unconstrained minimization of φ_2 tends to drive the slope of ρ negative so that, to continue further, we need to extend the definition of φ_2 to include the case where ρ is not invertible. Of course, by $\mathbb{T}$-symmetry, we can equally well regard ρ^{-1} as the fundamental circle map, denoted by σ, say, and consider cases where σ is not invertible. This corresponds to choosing an irreversible one-dimensional dynamics going backward in time. By choosing to work with ρ we choose the forward direction of time. We shall call such solutions "causal" or "retarded" solutions, whereas solutions obtained by going backwards in time by using σ are called "anticausal" or "advanced" solutions.

Although ΔY is not defined as a single-valued function when ρ is not invertible, the function $\Delta Y \circ \rho$ is. Thus to extend the definition of φ_2, all we have to do is to make the change of variable $\theta \mapsto \rho(\theta)$ to put eq. (5.1) in the form

$$\begin{aligned}\varphi_2 &\equiv \tfrac{1}{2} \int_0^1 [\Delta Y \circ \rho(\theta)]^2 \rho'(\theta)\, X' \circ \rho(\theta)\, d\theta, \\ &\equiv \tfrac{1}{2} \int_0^1 [F_2(X(\theta), X \circ \rho(\theta)) \\ &\quad + F_1(X \circ \rho(\theta), X \circ \rho^2(\theta))]^2 \\ &\quad \times \rho'(\theta)\, X' \circ \rho(\theta)\, d\theta,\end{aligned} \tag{5.2}$$

which does not require inversion of ρ.

5.2. Symmetries

We now show that the assumed $\mathbb{P}$ and $\mathbb{T}$ symmetries of $F(\cdot,\cdot)$ (see section 2) imply that φ_2 is symmetric under certain transformations of the circle map $\rho(\theta)$. As always (see section 3)

we assume that X is an odd function. We shall also assume, for the purposes of discussing $\mathbb{T}$-reversibility, that ρ is invertible.

If F obeys the $\mathbb{T}$-reversibility condition eq. (2.10), then the transformation

$$\rho(\theta) \mapsto \rho^{-1}(\theta) \tag{5.3}$$

introduces terms involving Q into ΔY, which however cancel by eqs. (4.4) and (4.5). Thus the transformation leaves φ_2 invariant. Similarly, if F is $\mathbb{P}$-symmetric, eq. (2.11) implies that φ_2 is invariant under the transformation

$$\rho(\theta) \mapsto -\rho(-\theta). \tag{5.4}$$

Given a solution which makes φ_2 stationary, these symmetry operations allow us to generate new solutions, in general distinct from the original one because the operations eq. (5.3) and eq. (5.4) change the sign of the rotation number. However, in the case $\nu = 0$, a transformation reversing the sign of ν does not necessarily generate a new solution and thus when $\nu = 0$ we can have $\mathbb{P}$-symmetric solutions which are not necessarily $\mathbb{T}$-symmetric and vice versa. In the $\mathbb{P}$-symmetric case, ρ is odd, which is useful for numerical purposes since it can be represented with only sine terms. For numerical convenience most studies of $\nu = 0$ solutions were made assuming $\mathbb{P}$-symmetry, but in a case where this was not assumed it was found that, although there was a solution with broken $\mathbb{P}$-symmetry which made φ_2 extremal, the φ_2-*minimizing* solution was in fact $\mathbb{P}$-symmetric (see section 9.3).

Less obvious is the fact that these arguments apply also for the $\nu = \frac{1}{2}$ solutions. This is because parity reversal changes ν to $-\frac{1}{2}$, but the periodicity in the y-direction following from eq. (2.22) allows us to add 1 to ν, bringing it back to $\frac{1}{2}$.

The combined $\mathbb{PT}$-symmetry operation

$$\rho(\theta) \mapsto -\rho^{-1}(-\theta), \tag{5.5}$$

however, always leaves the rotation number invariant. We thus expect to find stationary solutions ρ obeying the one-dimensional $\mathbb{PT}$-reversibility condition eq. (3.9). It is also possible that this symmetry is spontaneously broken and stationary solutions with the same rotation number as the reversible solution occur in pairs related by eq. (5.5). We shall show later both numerically and by explicit construction that this is indeed typically what happens, and that the $\mathbb{PT}$-symmetry breaking solutions have lower φ_2.

6. Euler–Lagrange equations

6.1. First variation of φ_2

To vary X and ρ in eq. (5.2) observe that $\delta[\rho'(\theta) X' \circ \rho(\theta)] = \mathrm{d}\delta[X \circ \rho(\theta)]/\mathrm{d}\theta$. Integrating this term by parts (noting that the variations at the endpoints 0 and 1 are equal, by periodicity) we find that the F_{11} and F_{22} terms cancel, leaving

$$\begin{aligned}\delta\varphi_2 = &\int_0^1 F_{12}(X, X\circ\rho)\,\Delta Y \\ &\circ\rho\,[\rho' X'\circ\rho\,\delta X - X'\delta(X\circ\rho)]\,\mathrm{d}\theta \\ &+\int_0^1 F_{12}(X\circ\rho, X\circ\rho^2)\,\Delta Y\circ\rho \\ &\times[\rho' X'\circ\rho\,\delta(X\circ\rho^2) - (X\circ\rho^2)'\delta(X\circ\rho)]\,\mathrm{d}\theta\end{aligned} \tag{6.1}$$

(the arguments (θ) being implicit). Making the change of variable $\theta \mapsto \rho(\theta)$ in the first integral, we find the first variation in φ_2 to be

$$\begin{aligned}\delta\varphi_2 = -\int_0^1 E(\theta)\,&[X'\circ\rho(\theta)\,\delta X\circ\rho^2(\theta) \\ &-\rho'\circ\rho(\theta)\,X'\circ\rho^2(\theta)\,\delta X\circ\rho(\theta) \\ &+X'\circ\rho(\theta)\,X'\circ\rho^2(\theta)\,\delta\rho\circ\rho(\theta)]\,\rho'(\theta)\,\mathrm{d}\theta,\end{aligned} \tag{6.2}$$

where

$$E(\theta) \equiv F_{12}(X\circ\rho(\theta), X\circ\rho^2(\theta)) \times [\Delta Y\circ\rho^2(\theta) - \Delta Y\circ\rho(\theta)]. \quad (6.3)$$

(For a time-symmetric form of the first variation, including a Lagrange multiplier for constrained variations, see appendix C.2.)

6.2. Stationarity when ρ is invertible

When ρ is invertible we can make the change of variable $\theta \mapsto \rho^{-1}(\theta)$ in eq. (6.2) to make the argument of $\delta\rho$ simply θ. Then the condition that φ_2 be stationary with respect to arbitrary variations in ρ is the Euler–Lagrange equation

$$\Delta Y\circ\rho(\theta) = \Delta Y(\theta), \quad (6.4)$$

except possibly at the turning points of $X(\theta)$ and $X\circ\rho(\theta)$. However, since we assume X and ρ to be differentiable and hence perforce continuous, eq. (6.4) must apply at these turning points also. Observe that if eq. (6.4) is satisfied then φ_2 is automatically stationary with respect to arbitrary variations in X, which is not surprising since almost everywhere a variation $\delta\rho$ can be found which produces the same effect on the relation α of eq. (3.1) as a variation δX. Using eqs. (4.4) and (4.5) we can also express the Euler–Lagrange equation eq. (6.4) in a form similar to that of Meiss and Dewar [17]

$$F_2(X(\theta), X\circ\rho(\theta)) + F_1(X\circ\rho(\theta), X\circ\rho^2(\theta)) = F_2(X\circ\rho^{-1}(\theta), X(\theta)) + F_1(X(\theta), X\circ\rho(\theta)). \quad (6.5)$$

Equations (6.4) and (6.5) have the simple geometrical interpretation in the phase space of the area-preserving map that the distance between C^* and C along the vertical line $x = X(\theta)$ equals the distance between the two curves along the vertical line $x = X(\theta^*)$, where θ^* is the iterate $\theta^* = \rho(\theta)$. Note that, in cases where $X(\theta)$ is not a monotonic function (e.g. fig. 2), there is an apparent ambiguity as to which branches of the curves to measure ΔY between because C and C^* turn over at points sharing a common vertical tangent, so that there are two or more branches of C and C^* lying above a given value of x. The correct branches of C and C^* between which to measure ΔY are obviously those corresponding to the same branches of the multivalued function $\theta(x)$.

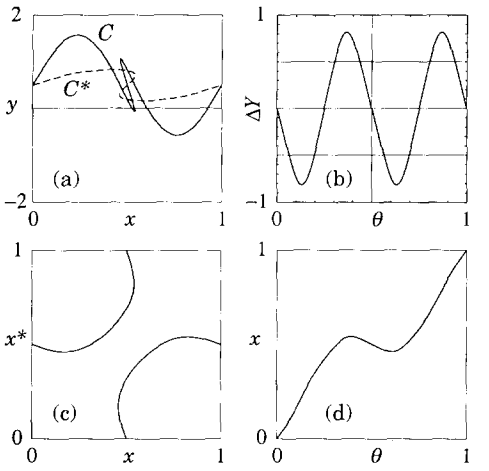

Fig. 2. Reversible solution in the case $k = 8.5$, $\nu = 0.5$ (see sections 8 and 9.2), for which α is not a diffeomorphism because X is not invertible. (a) Curves C and C^* showing both curves folding over at turning points of $x = X(\theta)$. (b) Graph of vertical distance between C and C^* vs. θ. (c) Relation $x^* = \alpha(x)$. Note the symmetries about the diagonals $x + x^* =$ integer, following from reversibility of α. Position coordinates are evaluated mod 1. (d) Semiconjugacy map $x = X(\theta)$.

Consider points on the two curves where $\Delta Y = 0$. These points are necessarily intersections of C and C^*, although if X is not monotone not *all* intersections are points where $\Delta Y = 0$—only intersections of branches C and C^* corresponding to the same branch of $\theta(x)$. We shall term such intersections *true* intersections. For instance, in fig. 2, the intersection at $x = \frac{1}{2}$ is a true intersection because $\Delta Y(\frac{1}{2}) = 0$, but the other six intersections in the range $x \in (x_1, x_2)$, where $x_1 \simeq 0.46$ and $x_2 \simeq 0.54$ are the turning points, are spurious.

We can now enunciate the following important result

Theorem 1. True intersections of φ_2-extremizing rotational curves C and C^* generated by an invertible circle map ρ belong to families which are orbits under the area-preserving map T.

To see this, let there be a true intersection at $\theta = \theta_0$. That is, let $\Delta Y(\theta_0) = 0$. Then the Euler–Lagrange equation eq. (6.4) implies the existence of a family of true intersections at $\theta = \theta_j$ (i.e. $\Delta Y(\theta_j) = 0$) where the θ_j are images of θ_0 under repeated application of ρ,

$$\theta_j \equiv \rho^j(\theta_0), \tag{6.6}$$

for any integer j. Hence, by lemma 3 in section 4 the family

$$x_j \equiv X(\theta_j) \tag{6.7}$$

forms an orbit in the Lagrangian form of the original dynamics, eq. (2.5) (see fig. 3).

Thus, although the one-dimensional dynamics associated with our generalized action–angle representation does not in general have significance for the original two-dimensional dynamics (unlike the integrable case), theorem 1 shows that there is a *subset* of orbits in the one-dimensional dynamics which does correspond one-to-one with orbits under the two-dimensional dynamics.

The curves C and C^* must cross at least twice since the condition of zero net flux implies that the area below C and that below C^* are equal. There must therefore be at least two orbit families, corresponding to intersections where C^* crosses C from below as θ is increased ($\Delta Y'(\theta_j) > 0$), and vice versa [$\Delta Y'(\theta_j) < 0$].

6.3. Periodic orbits and action

We now relate the two classes of orbit discussed above to the action minimizing and minimax orbits used in [4]. Let $(x_0, x_1, \ldots, x_n)$, be an (m, n) periodic orbit, with x_j defined by eq. (6.7) and $x_n \equiv x_0 + m$. Then the first variation of the action

$$W_{m,n} \equiv \sum_{j=0}^{n-1} F(x_j, x_{j+1}) \tag{6.8}$$

is zero because $\Delta Y(\theta_j)$ is zero. Calculating the second variation with the special trial function $\delta x_j = \delta\xi_j$, where

$$\delta\xi_j \equiv \delta\theta_0 \frac{\partial X(\theta_j)}{\partial\theta_0} = \delta\theta_0 X'(\theta_j) \frac{\partial\theta_j}{\partial\theta_0}, \tag{6.9}$$

with

$$\partial\theta_j/\partial\theta_0 \equiv \frac{\partial\rho^j(\theta_0)}{\partial\theta_0} = \prod_{i=0}^{j-1} \rho'(\theta_i), \tag{6.10}$$

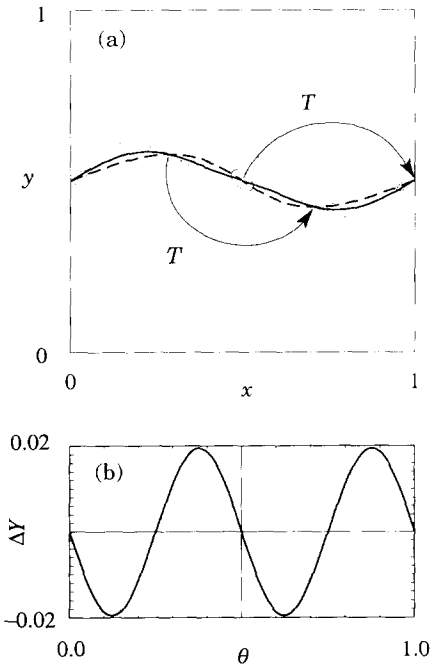

Fig. 3. (a) C (—) and C^* (– –) curves solving the unconstrained Euler–Lagrange equation (see sections 8 and 9.2) which intersect at period-two orbits of the standard map with $k = 1$. Neighbouring orbits are also shown. (b) Graph of $\Delta Y(\theta)$: $\Delta Y'(\theta_j) > 0$ corresponds to an action minimizing orbit (always hyperbolic), while $\Delta Y(\theta_j) = 0$, $\Delta Y'(\theta_j) < 0$ corresponds to an action minimax orbit (elliptic in this case).

we find the second variation

$$\delta^2 W_{m,n} = \sum_{j=0}^{n-1} (\delta\xi_j)^2 \frac{\Delta Y'(\theta_j)}{X'(\theta_j)}. \tag{6.11}$$

Thus if $\Delta Y'(\theta_j)/X'(\theta_j) < 0$ for all j then this is sufficient to ensure that $\delta^2 W_{m,n} < 0$ for the variation eq. (6.9). That is, there is a descending direction for the action. It is shown in [29] that such minimax orbits are either elliptic (residue [24] in [0, 1]) or hyperbolic with reflection (residue > 1), while action minimizing orbits ($\delta^2 W_{m,n} > 0$ for all $\delta\xi_j$) are always hyperbolic (residue < 0). Clearly eq. (6.11) implies, in the case of action minimizing orbits, that $\Delta Y'(\theta_j)/X'(\theta_j) > 0$ for at least *some* j. Since differentiating $\Delta Y \circ \rho(\theta) = \Delta Y(\theta)$ together with our assumption $\rho'(\theta) > 0$ imply that the sign of $\Delta Y'(\theta_j)$ is preserved within an orbit family, the sign of $\Delta Y'(\theta_j)/X'(\theta_j)$ will be the same for *all* j provided $X'(\theta_j) > 0$ for all j. Figure 2 shows however that the sign of $X'(\theta_j)$ is not always conserved.

6.4. *Invariant circles and cantori*

An invariant circle is a closed loop whose image under the application of T is itself. Clearly a rotational invariant circle, if it exists, corresponds to a minimum of φ_2 since $\Delta Y(\theta) \equiv 0$, and hence φ_2 vanishes, when C and C^* coincide. Thus one application of our new variational principle is to the calculation of rotational invariant circles. Just as our principle is an alternative to the principle of stationary action for periodic orbits, so in this application it can be used instead of the variational principle of Percival [30,26] based on averaging the generating function over θ, with ρ prescribed to be a rigid rotation with specified irrational rotation number ν. Unlike the stationary action and Percival's principles, in our case the *one* variational principle gives both periodic and quasiperiodic orbits.

When no invariant circle with the given irrational ν exists, Percival's Euler–Lagrange equation $\Delta Y(\theta) = 0$ can still be solved for $X(\theta)$, with ρ a rigid rotation, but $X(\theta)$ then has [4,26] an infinite number of discontinuities corresponding to the gaps in the cantorus. Such a solution would violate our assumption that X is differentiable, so that it would seem undesirable to use semiconjugacy to rigid rotation for constructing cantori (except possibly through a limiting procedure via a sequence of rationals). We saw in section 3.2 that Herman's theorem implies that the function X making α conjugate to a rigid rotation is essentially one order less smooth than α itself, so that by avoiding the imposition of conjugacy to rigid rotation in our variational principle, it is possible to work with a smoother class of trial functions. If, as we conjecture, cantori are locally φ_2-minimizing then the quadratic flux minimization principle would not require conjugacy to rigid rotation to control the rotation number—the variational principle itself would seek out a cantorus. Which cantorus is found would depend on the initial guess.

Although we saw in section 3.3 that ρ is $\mathbb{PT}$-reversible if it is a rigid rotation, this symmetry carries over to the circle map α only if $X(\theta)$ is an odd function. Thus Percival's description of invariant circles does not necessarily preclude spontaneous symmetry breaking of the type discussed in section 5.2. However, we now show that in fact X *must* be an odd function if $\rho(\theta) = \theta + \nu$ and C is a rotational invariant circle.

Our argument is based on lemma 2 in section 2.5. Every invariant circle must intersect $x = 0$, a symmetry line under S_1, the involution defined by eq. (2.20), so every invariant circle contains a symmetric orbit $x_j = X(j\nu)$. That is, $X(j\nu) = -X(-j\nu)$. This orbit is dense since ν is irrational, so X must be odd and the reversibility of α then follows.

6.5. *Stationarity, ρ not invertible*

Consider for definiteness a function $\rho(\theta)$ which has negative slope in a region $\theta \in [\theta_<, \theta_>]$, say. We assume $\theta_<$, $\theta_>$, $\theta_<^* \equiv \rho(\theta_>)$

and $\theta^*_> \equiv \rho(\theta_<)$ all lie in the interval $[-\frac{1}{2}, \frac{1}{2}]$. Our method of generalizing φ_1 and φ_2 to non-monotonic ρ is to make the change of variable $\theta^* = \rho(\theta)$ so that eq. (4.4) now defines a unique function $Y_+ \circ \rho(\theta) \equiv Y_+^<(\theta^*)$ for $\theta \in [-\frac{1}{2}, \theta_<]$, $Y_+ \circ \rho(\theta) \equiv Y_+^{<>}(\theta^*)$ for $\theta \in [\theta_<, \theta_>]$ and $Y_+ \circ \rho(\theta) \equiv Y_+^>(\theta^*)$ for $\theta \in [\theta_>, -\frac{1}{2}]$. Note that eq. (4.5) shows that $Y_-(\theta^*)$ is a single valued function.

To evaluate the flux we will need to consider integrals which have the general form

$$I = \int_{-1/2}^{1/2} f(\Delta Y \circ \rho(\theta), \rho(\theta))\, \rho'(\theta)\, \mathrm{d}\theta, \quad (6.12)$$

where the dependence on the second argument of f is such that there is no ambiguity if we transform back to the θ^* variable. On the other hand, we know that $\Delta Y(\theta^*) \equiv Y_+(\theta^*) - Y_-(\theta^*)$ has three branches on the interval $\theta^* \in [\theta^*_<, \theta^*_>]$. However, we can transform back to θ^* uniquely by subdividing the region of integration into the subintervals $\theta \in [-\frac{1}{2}, \theta_<]$, $\theta \in [\theta_<, \theta_>]$, and $\theta \in [\theta_>, \frac{1}{2}]$. The inverse function of ρ is then defined on these restricted domains, so we can transform back to θ^*, obtaining

$$\begin{aligned} I = &\int_{-1/2}^{\theta^*_<} f(\Delta Y^<(\theta^*), \theta^*)\, \mathrm{d}\theta^* \\ &+ \int_{\theta^*_>}^{1/2} f(\Delta Y^>(\theta^*), \theta^*)\, \mathrm{d}\theta^* \\ &+ \int_{\theta^*_<}^{\theta^*_>} [f(\Delta Y^<(\theta^*), \theta^*) + f(\Delta Y^>(\theta^*), \theta^*) \\ &- f(\Delta Y^{<>}(\theta^*), \theta^*)]\, \mathrm{d}\theta^*, \end{aligned} \quad (6.13)$$

that is, the overlapping branches contribute additively, but the retrograde branch contributes with negative sign.

This result may be used to show that the first flux moment φ_1 remains the expected area between C and C^*, but we use it here to generalize the Euler–Lagrange equation to the case of non-invertible ρ (the anticausal case can be treated similarly). We use eq. (6.2), assuming that the conjugacy function X is monotonic in the region of interest, and is held fixed so that only $\delta\rho$ contributes. Changing variables so that the argument of $\delta\rho$ is θ^*, using eq. (6.13), we find the modified Euler–Lagrange equation for unconstrained variation of $\delta\rho$

$$\begin{aligned} &\Delta Y \circ \rho(\theta^*) \\ &\quad = \Delta Y^<(\theta^*) + \Delta Y^>(\theta^*) - \Delta Y^{<>}(\theta^*), \end{aligned} \quad (6.14)$$

which applies on the interval $\theta^* \in [\theta^*_<, \theta^*_>]$. Provided $[\theta^*_<, \theta^*_>] \subseteq [\theta_<, \theta_>]$ we can identify ΔY on the left hand side as $\Delta Y^{<>}$ everywhere in the interval.

7. Single mode minimization

Specializing now to the case of the standard map eq. (2.24) we show in this section that it is possible to evaluate φ_2 and Ω analytically if we take as trial function a circle map which is a rigid rotation in θ,

$$\rho(\theta) = \theta + \nu, \quad (7.1)$$

with the conjugacy function X containing a single Fourier mode

$$X(\theta) = \theta + \frac{A}{2\pi} \sin 2\pi\theta. \quad (7.2)$$

It is easily verified that $X'(\theta)$ is strictly positive, that is, X is a diffeomorphism, only if $|A| < 1$.

It is also possible to do the integrals in the case where one mode is used in the representation eq. (3.2), but the rigid rotation representation has the advantage that it is manifestly $\mathbb{PT}$-reversible, eq. (3.9), and that it allows us to specify ν easily.

Substituting the representation eq. (7.2) into eq. (C.4) we have

$$\Omega = \nu + \frac{A^2}{4\pi} \sin 2\pi\nu, \quad (7.3)$$

and from eq. (5.2) we have

$$\varphi_2 = \frac{1}{8\pi^2}\int_0^1 \Big[k \sin 2\pi\Big(\theta + \frac{A}{2\pi}\sin 2\pi\theta\Big) - 4A\sin^2\pi\nu\, \sin 2\pi\theta\Big]^2 \times (1 + A\cos 2\pi\theta)\, \mathrm{d}\theta. \tag{7.4}$$

Expanding the square and using the identity

$$\exp\frac{z}{2}\Big(t - \frac{1}{t}\Big) = \sum_{l=-\infty}^{\infty} J_l(z)\, t^l, \tag{7.5}$$

with $t = \exp \mathrm{i}\theta$, we find

$$\varphi_2 = \frac{k^2}{16\pi^2} + \frac{A^2}{\pi^2}\sin^4\pi\nu - \frac{k}{\pi^2}J_1(A)\,\sin^2\pi\nu. \tag{7.6}$$

Assuming ν is not an integer, this is minimal with respect to A when

$$2A\sin^2\pi\nu = kJ_1'(A). \tag{7.7}$$

For small k, this implies $A \approx k/4\sin^2\pi\nu$, which is what one would find at lowest order from a perturbation theory construction of a KAM curve with rotation number ν [31]. The divergence as $\nu \to 0$ is a consequence of the small denominator problem of perturbation theory. It is encouraging to note that the full expression eq. (7.7) does not suffer from this problem since the vanishing of the right hand side as $\nu \to 0$ is accomplished by A approaching $j'_{1,1}$, the first positive (assuming $k > 0$) zero of J_1'. Thus

$$|A| < j'_{1,1} \approx 1.84118. \tag{7.8}$$

However, this supremum is greater than unity, so that for small ν or large k the circle map α is not a diffeomorphism, giving rise to folding over of the curves C and C^* as discussed in section 4 and similar to that in fig. 2.

If A is large enough that

$$|A\cos\pi\nu| > 1, \tag{7.9}$$

then $X(\theta + \nu) + X(\theta)$ becomes nonmonotonic, and the curves develop loops. Even in this pathological case eq. (5.1) still provides a perfectly workable definition of φ_2. Note that, according to eq. (7.8), when $|\nu - \frac{1}{2}| < \pi^{-1}\cos^{-1}(1/j'_{1,1}) \approx 0.3172$, $|A\cos\pi\nu|$ will not exceed unity, and hence looping cannot occur, independent of the value of k.

Extremizing with respect to both A and ν we find a degenerate saddle point of φ_2 at $\nu = 0$, $A = j'_{1,1}$, corresponding to the large island which contains the fixed point at the origin, and a minimum of φ_2 at $\nu = \frac{1}{2}$, A given by eq. (7.7). Although there is an island containing a period two orbit, the single mode approximation misses this resonance, and the minimum at $\nu = \frac{1}{2}$ can be viewed as a very bad approximation to the minima corresponding to KAM curves or cantori on either side of this resonance [see fig. 6]. (Although, paradoxically, the solution obtained by setting $\nu = \frac{1}{2}$ is also quite a good approximation to one of the isolated reversible solutions to be described in section 9.2.) The different types of solution when $\nu = \frac{1}{2}$ and $\nu \simeq 0$ are illustrated in fig. 4.

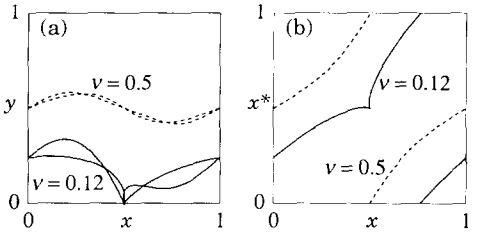

Fig. 4. Single mode minimization in the case $k = k_c = 0.9716$, showing in (a) the curves C and C^* for both the unconstrained minimizing solution at $\nu = 0.5$ (dashed) and the minimizing solution when ν is constrained to be 0.12 (solid). In the latter case $A = 1.0754$, which is just on the threshold for loop formation. The corresponding circle maps are shown in (b). Also shown is a chaotic orbit started near (0.5, 0).

8. Multimode minimization

In order to treat more general trial functions for the circle map than that used in section 7 it is necessary to resort to numerical minimization. For maximum flexibility our algorithm includes an arbitrary number of sine and cosine modes in the conjugacy function X as well as an arbitrary number in the rotation map ρ, which is represented in a manner analogous to eq. (3.2). That is, we expand X and ρ in finite basis sets

$$X(\theta) = \theta + \sum_{n=1}^{N} x_n u_n(\theta) \tag{8.1}$$

and

$$\rho(\theta) = \theta + \sum_{m=0}^{M} \rho_m v_m(\theta), \tag{8.2}$$

where the $\{u_i(\theta)\}$ and $\{v_i(\theta)\}$ are Fourier bases chosen from the sets $\{\sin j\theta\}$ and $\{\cos j\theta\}$. The trial variation is found, by substituting eqs. (8.1) and (8.2) into eq. (6.2), to be, in terms of the variations in the Fourier coefficients x_n and ρ_m, of the form

$$\delta\varphi_2 = \sum_{n=1}^{N} \varphi^x_{,n}\delta x_n + \sum_{m=1}^{M} \varphi^\rho_{,m}\delta\rho_m, \tag{8.3}$$

where the components of the gradient in the trial function space are

$$\begin{aligned}\varphi^x_{,n} \equiv &\int_0^1 \mathrm{d}\theta\, E(\theta)\,\rho'(\theta)\\ &\times[\rho'\circ\rho(\theta)X'\circ\rho^2(\theta)\,u_n\circ\rho(\theta)\\ &-X'\circ\rho(\theta)\,u_n\circ\rho^2(\theta)]\end{aligned} \tag{8.4}$$

and

$$\begin{aligned}\varphi^\rho_{,m} \equiv &-\int_0^1 \mathrm{d}\theta\, E(\theta)\\ &\times\rho'(\theta)\,X'\circ\rho(\theta)\,X'\circ\rho^2(\theta)\,v_m\circ\rho(\theta).\end{aligned} \tag{8.5}$$

Both φ_2 and its gradient were computed for the standard map by numerical integration using the trapezoidal rule on a θ-mesh with the number of grid points taken to be 8 times the maximum number of sine or cosine modes in the representations of X and ρ. Stationary points of φ_2 were sought using the Fletcher–Reeves–Polak–Ribiere conjugate gradient method [32], modified slightly to search for null points of the gradient rather than minima of φ_2 (except, optionally, in the initial bracketing step of each line search). Convergence to high precision was slow in some cases, indicating that a preconditioned algorithm might have been better. The constant term in eq. (8.2) was held fixed to implement the constant Ω constraint discussed in section C.1 (in the case where X is the identity map). The constraint $\sum_m \rho^c_m = \text{const.}$, where ρ^c_m are the coefficients of the cosine terms, was also used to implement the constraint $\rho(0) = \text{const.}$, but was not found to give qualitatively different results. The Fourier modes of X and/or ρ active during the numerical extremization were selected by choosing the order in which they were stored in the state vector, and by selecting the dimension of the state vector. The calculations were performed on an Apple Macintosh IIcx in extended precision (ca. 18 significant digits).

The deviation from $\mathbb{PT}$-reversibility of the circle map was measured by calculating the "irreversibility parameter"

$$\mathcal{I} \equiv \int_0^1 [\theta + \rho(-\rho(\theta))]^2\rho'(\theta)\,X'\circ\rho(\theta)\,\mathrm{d}\theta. \tag{8.6}$$

Clearly, $\mathcal{I} \geq 0$, with equality if and only if the condition eq. (3.9) is satisfied almost everywhere. The rotation number ν of ρ was estimated by performing N (say) preiterates of ρ, then M more iterates, calculating the mean increase in θ per iteration. For low resolution scans, $M = N = 100$ was sufficient.

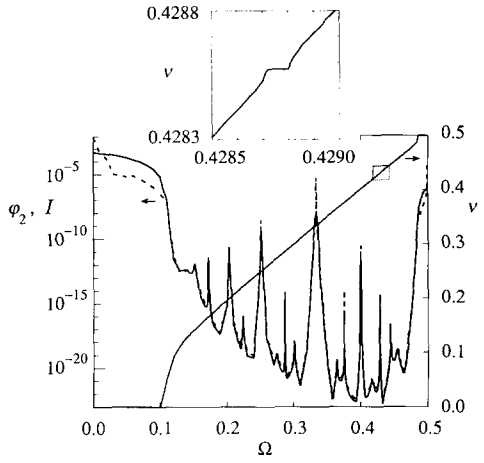

Fig. 5. Plot of numerically calculated quadratic flux φ_2 (solid line), irreversibility $\mathcal{I}$ (dashed line) and rotation number ν vs. the control parameter Ω (constant part of the circle map) for the standard map with $k = 0.3$. The minimization was performed with 13 sine and 13 cosine modes in ρ and with X the identity. Inset: Blowup of the 3/7 resonance.

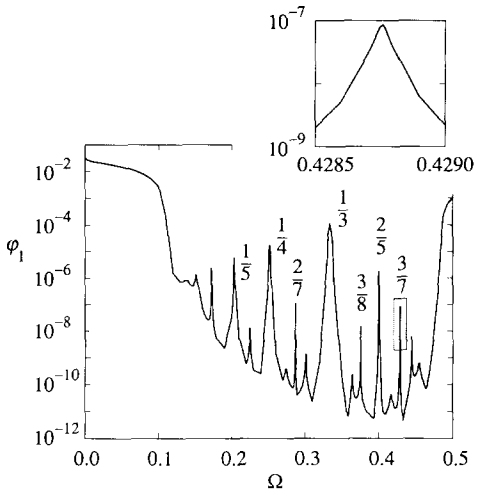

Fig. 6. Plot of numerically calculated flux φ_1 vs. Ω for the standard map with $k = 0.3$. The minimization was performed with 13 sine and 13 cosine modes and no conjugacy. Also shown is a blowup of the (3,7) resonance.

To get an overall picture of the nature of the dependence of φ_2 on Ω, a scan was performed at $k = 0.3$ over the interval $\Omega \in [0, \frac{1}{2}]$. The conjugacy function X was set to be the identity so that α and ρ were the same function, and minimization was performed over a 26-dimensional space made up of the first 13 sine modes and first 13 cosine modes in the Fourier representation eq. (3.2). A plot of ν vs. Ω is shown in fig. 5, showing the mode locking around $\Omega = 0$ and $\frac{1}{2}$. There is actually mode locking at all resonances, so that $\nu(\Omega)$ is a Devil's staircase. The 3/7 resonance is blown up to illustrate this. In this case it was necessary to take the iteration parameters to be $N = M = 10^4$. Only the interval $[0, \frac{1}{2}]$ is shown because of the symmetry about $\nu = \frac{1}{2}$ which was observed numerically, and is expected from the combination of parity or time reversal symmetry and the periodicity in the y-direction following from eq. (2.22).

A plot of φ_2 and $\mathcal{I}$ vs. Ω is also shown in fig. 5, while φ_1 vs. Ω is shown in fig. 6 These figures suggest the fractal nature of the quadratic flux as a function of Ω—a local minimum at each irrational value of Ω and a local maximum at each rational value. With a finite number of modes a numerical minimization can resolve only a finite number of maxima and minima, but as more modes are added one observes more maxima and minima. It is also seen that the irreversibility $\mathcal{I}$ is roughly proportional to φ_2, while the flux φ_1 is roughly proportional to $\varphi_2^{1/2}$. Note finally that the heights of the peaks form a hierarchy corresponding to the Farey tree [33]. The highest is $\nu = 0/1$; level 2: $\nu = 1/2$; level 3: $\nu = 1/3$; level 4: $\nu = 1/4, 2/5$; level 5: 1/5, 2/7, 3/8, 3/7; level 6: 1/6, 2/9, 3/11, 3/10, 4/11, 5/13 ,5/12, 4/9; and so on, with the amplitudes and widths of the resonances decreasing at each level. All resonances up to Level 6 were clearly seen, except 5/13, whose height was presumably below the apparent truncation threshold of 10^{-23}. The variation of φ_2 over twenty orders of magnitude shows that this numerical method is a sensitive tool for investigating flux transport through resonances in such systems.

When instead ρ was constrained to be a rigid rotation R_ν, and the minimization was performed over the amplitudes of the sine modes in the Fourier representation of X a similar pattern was observed in general scans of ν (but

with $\mathcal{I} = 0$ of course). As in the one mode analytic model in section 7, the solutions tended to be such that the C and C^* curves avoided the islands.

However, if we took ν to be an exact rational fraction (to within machine rounding error), solutions such as those shown in figs. 2 and 3 were found. The surprising thing about these solutions is that, although they were found by using a special trial function (ρ a rigid rotation), by serendipity they satisfy to high precision the *unconstrained* Euler–Lagrange equation, eq. (6.4). Thus we have strong numerical evidence that there exist reversible solutions to eq. (6.4) which (trivially) have invertible ρ and thus can be used to find periodic orbits by theorem 1. In section 9.2 we sketch a perturbative construction for these solutions, which it is seen avoids the zero denominator problem usually encountered in perturbation theory at resonances. We have not carried out this expansion to all orders in k, let alone proved convergence, but the numerical evidence that the solutions remain smooth and well-behaved for k-values well beyond the critical value for destabilization of the elliptic action-minimax orbits suggest that the perturbation theory may have a finite, possibly even infinite, radius of convergence. We show in section 9.1 that there is at least one resonance for which an expansion in k trivially converges.

9. Periodic solutions

9.1. Exact reversible fixed point solution

As is apparent in fig. 4, the flux-minimizing solutions as ν tends to zero correspond to curves C and C^* which avoid the interior of the large island surrounding the elliptic fixed points on the x-axis, the $(0, 1)$ resonance. Instead the curves lie in the separatrix region of the island where the flux is much smaller than that through the "turnstile" [4] pivoting about the action-minimizing and minimax fixed points on the x-axis. We show in this section that there is an exact, $\mathbb{PT}$-reversible solution of the unconstrained Euler–Lagrange equation eq. (6.5) which gives curves passing through the fixed points of the island, thus defining turnstiles.

A solution conjugate to rigid rotation with zero frequency has $\rho(\theta) = \theta$. In this case the solution to the Euler–Lagrange eq. (6.4) for the generalized standard map, eq. (2.21), is $\Delta Y(\theta) = -V'(X)$ where $X(\theta)$ is arbitrary. If, for simplicity, we let X be the identity map, so that $\theta \equiv x$, then this solution becomes

$$\Delta Y(x) = -\frac{k}{2\pi} \sin 2\pi x, \tag{9.1}$$

for the standard map. Note that ΔY has a zero at $x = 0$ (action minimax orbit) and at $x = \frac{1}{2}$ (action-minimizing orbit). From eq. (9.1) we can easily evaluate the flux,

$$\varphi_1 = \frac{k}{2\pi^2}, \tag{9.2}$$

through the turnstile. The quadratic flux φ_2 is given by

$$\varphi_2 = \frac{k^2}{16\pi^2}. \tag{9.3}$$

Interestingly, this is precisely the quasilinear diffusion coefficient D_{QL} [34,35]. A connection between a flux minimization principle and quasilinear diffusion was also used in [10].

Although this solution makes φ_2 *stationary* with respect to variations in ρ, it does not follow that it is necessarily *minimal*. In order to investigate the "stability" of φ_2 against small variations, we use the sum-difference representation eq. (B.2) to expand the circle map and its inverse about the identity map

$$\alpha^{\pm 1}(x) = x \pm \omega(x) + \frac{1}{4}\frac{\mathrm{d}}{\mathrm{d}x}[\omega(x)]^2 + \mathcal{O}(\omega^3), \tag{9.4}$$

assuming X to be the identity so that $\rho \equiv \alpha$. For the standard map this gives

$$\varphi_2 = \frac{k^2}{16\pi^2} - \frac{k}{2}\int_0^1 \omega(x)^2 \cos 2\pi x \, dx + \mathcal{O}(\omega^3), \tag{9.5}$$

from which we see that ω's with support on $[0, \frac{1}{4}] \cup [\frac{3}{4}, 1]$ decrease φ_2, while ω's with support on $[\frac{1}{4}, \frac{3}{4}]$ increase φ_2. A numerical search for the $\nu = 0$, φ_2-minimizing solutions would start by incrementing ω in the φ_2-descending direction, thus perturbing ρ away from the identity only outside the interval $x \in [\frac{1}{4}, \frac{3}{4}]$. This suggests that the minimizing solutions will tend to be the same as the reversible solution in a finite neigbourhood of $x = \frac{1}{2}$. We shall see in section 9.3 that this is indeed the case.

It is seen from eq. (9.5) that there is an infinity of descending directions away from this solution, even if we limit attention to $\mathbb{PT}$-reversible (even ω) perturbations. If we constrain the trial function space to be circle maps conjugate to rigid rotation, the variation vanishes at this order but is still not positive definite at higher order if we allow ν to vary. Thus it appears that there is no simple constraint which is transversal to all the descending directions, and we conclude that there is no natural one-parameter family of constrained minimizing solutions in which this solution is an extremum—it is an *isolated* solution.

9.2. Reversible periodic solutions

When we take ρ to be a rigid rotation R_ν, with ν a rational fraction (not an integer), and seek minimax solutions numerically we find solutions with very small harmonic content in X and ΔY, such as in figs. 2 and 3. This suggests that we seek an analogue of the exact trivial fixed point solution of section 9.1 for higher order resonances using perturbation theory, and in this section we sketch how such a theory works.

For definiteness assume T to be a map with a generating function of the form eq. (2.21) with V replaced by ϵV, ϵ being our formal expansion parameter. Then ΔY, eq. (4.8), is of the form

$$\Delta Y(\theta) = -L_\nu X(\theta) - \epsilon V' \circ X(\theta), \tag{9.6}$$

with L_ν a linear difference operator defined by

$$L_\nu X(\theta) \equiv X(\theta+\nu) - 2X(\theta) + X(\theta-\nu), \tag{9.7}$$

where we have assumed that the class of solutions we seek is those which are conjugate to the rigid rotation $\rho = R_\nu$. Taking $\nu = m/n$, m and n being mutually prime integers, observe that eq. (6.4) implies that $\Delta Y(\theta)$ is periodic with period $1/n$. That is,

$$\Delta Y(\theta) = -U(\theta, \epsilon)/2\pi, \tag{9.8}$$

where $U(\theta, \epsilon)$ is a $1/n$-periodic function in θ,

$$U(\theta, \epsilon) = \sum_{l=-\infty}^{\infty} U_{ln}(\epsilon) \exp 2\pi i l n \theta. \tag{9.9}$$

Writing

$$X(\theta) = \theta + \frac{1}{2\pi}\xi(\theta, \epsilon) \tag{9.10}$$

in eq. (9.7), with ξ 1-periodic in θ, we see that eq. (9.8) can be written

$$L_\nu \xi = U - 2\pi\epsilon V' \circ X. \tag{9.11}$$

The right hand side of eq. (9.11) must have no component within the nullspace of L_ν, which is spanned by the functions $\exp 2\pi i l n\theta$, where l is any integer. That is, in terms of the Fourier coefficients U_{ln} defined in eq. (9.9), we require

$$U_{ln} = 2\pi\epsilon \int_0^1 V' \circ X(\theta) \exp(-2\pi i l n \theta) \, d\theta. \tag{9.12}$$

This removes the resonant forcing terms from eq. (9.11) which are normally the source of the small denominator problem in perturbation

theory. This is similar in spirit to the approach adopted in [8,10,11] but without any *ad hoc* filtering function or resonance broadening. It is also a Lagrangian, rather than Hamiltonian, formulation in the spirit of [26,30].

To illustrate this method we shall consider the standard map, for which we have $2\pi\epsilon V'(x) = k \sin 2\pi x$ so that eq. (9.11) becomes

$$L_\nu \xi(\theta,k) = U(\theta,k) - k\sin(2\pi\theta + \xi). \quad (9.13)$$

Expanding in powers of k

$$f(\theta,k) = \sum_{j=1}^{\infty} k^j f^{(j)}(\theta), \quad (9.14)$$

where f is any function, we find to lowest order

$$\xi^{(1)}(\theta) = \frac{\sin 2\pi\theta}{4\sin^2 \pi\nu}, \quad (9.15)$$

which agrees with the small-A limit of the single mode minimization, eq. (7.7). Since ν is assumed not to be an integer (this case is treated in section 9.1), the denominator cannot vanish.

Note that, at each order, an arbitrary $1/n$-periodic function can be included in $\xi^{(n)}$, but that we can make ξ unique by requiring that it have no $1/n$-periodic component. This constraint was not used in generating figs. 2 and 3.

9.3. *Symmetry breaking solutions*

Numerical minimization for $\Omega = 0$, assuming X the identity and using only sine modes in ρ to enforce $\mathbb{P}$-symmetry (recall section 5.2), yields a pair of nonreversible solutions of the unconstrained Euler–Lagrange equation, eq. (6.4). That is, $\mathbb{PT}$-symmetry is spontaneously broken through the breaking of $\mathbb{T}$-symmetry. By scanning in from finite ν using cosine modes as well, a solution with broken $\mathbb{P}$-symmetry was also found, but this had greater φ_2 (at least, for $k = 1$ and 13 sine and cosine modes) and was not studied further. As for the reversible

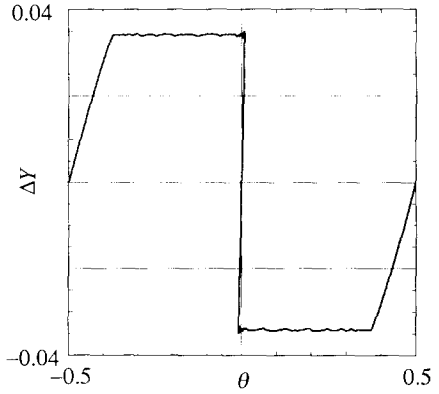

Fig. 7. Plot of numerically calculated vertical distance between C^* and C, $\Delta Y(\theta)$, for the standard map with $k = 0.3$, $\Omega = 0$. The minimization was performed with 64 sine modes and X the identity, so that $\theta \equiv x$.

solutions, C and C^* for the $\mathbb{P}$-symmetric solutions intersect at the corresponding orbits of T, but, as shown in fig. 7, $\Delta Y(x)$ appears to be close to discontinuous at the intersections corresponding to the action-minimax orbits ($x =$ integer, say 0). The function $\Delta Y(x)$ is then roughly constant between 0 and $\frac{1}{2} - a$, where $a < \frac{1}{2}$, consistently with the prediction of the Euler–Lagrange equation when the circle map ρ has attracting fixed points. On the other hand, the part of the ΔY curve passing through the action-minimizing hyperbolic points (i.e. with support on the interval $x \in [\frac{1}{2} - a, \frac{1}{2} + a]$) has finite slope. Similar behaviour was observed for $\Omega = \frac{1}{2}$. Closer examination shows however that the apparently flat top of the ΔY curve is not exactly constant, and that the circle map of the solution with the smaller φ_2 (specification of ρ rather than ρ^{-1} being an asymmetric procedure) is not monotonic.

Owing to the apparently nonanalytic nature of the numerical solutions, we cannot hope to understand these solutions by a conventional perturbation theory. However we can use the numerical solutions to suggest a nonperturbative geometric construction for these orbits. To aid in visualizing the problem we show in figs. 8, 9 and 10 a solution for larger k, where the detailed structure is more apparent. The nonmonotonic-

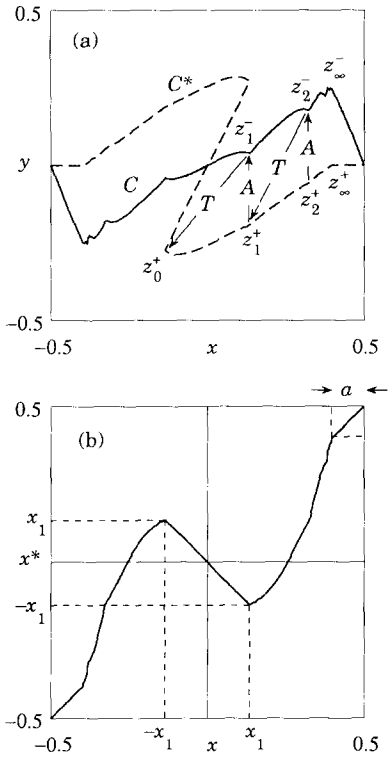

Fig. 8. (a) Curves C and C^* for the causal minimizing solution at the (0,1) resonance. Here A denotes the map A_a of the text. (b) Nonmonotonic circle map α. The minimization was done for the standard map with $k = 2.6$, $\Omega = 0$, 64 sine modes and X the identity, so that $\theta \equiv x$.

ity of ρ and the corresponding multivalued nature of Y_+ in the region $x \in [-x_1, x_1]$ is now clear. We label the left branch $Y_+^<$, the middle branch which (x^*, y^*) traverses in the retrograde direction as θ is increased $Y_+^{<>}$, and the right branch $Y_+^>$.

We now focus on the numerical solution in fig. 8. Since we take X to be the identity we have $\theta \equiv x$ and $\alpha \equiv \rho$. Symmetry allows us to assume $x_< = -x_1$, $x_> = x_1$. The numerical solution motivates us to assume also $x_<^* = -x_1$, $x_>^* = x_1$, that is, that α maps the interval $[-x_1, x_1]$ onto itself. We also assume that α is an odd, analytic function close to the reflection $-x$,

$$\alpha(x) = -x + a_3 x(x^2 - x_1^2) + \mathcal{O}(x^5), \quad (9.16)$$

on the interval $x \in [-x_1, x_1]$. Then eq. (4.5)

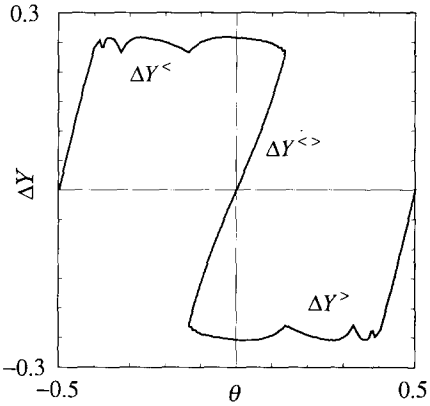

Fig. 9. Plot of numerically calculated vertical distance between C^* and C, $\Delta Y(\theta)$, for the same case as in fig. 8. There are three branches: $y = \Delta Y^<$, $y = \Delta Y^{<>}$ and $y = \Delta Y^>$.

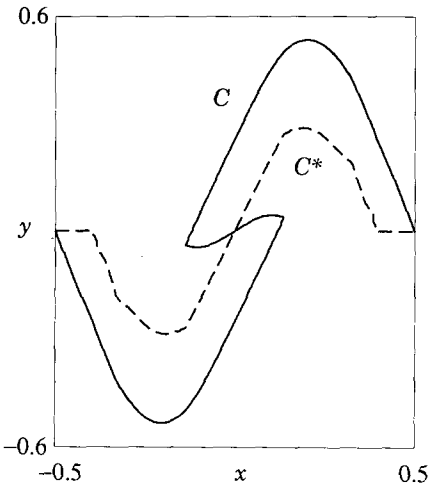

Fig. 10. Curves C and C^* for the anticausal minimizing solution at the (0,1) resonance. The nonmonotonic circle map α^{-1} is the same as the circle map α of the causal solution. The minimization was done for the same case as in fig. 8.

and eq. (4.4) give

$$Y_-(x) = \frac{k}{2\pi}\sin 2\pi x - 2x + a_3 x(x^2 - x_1^2) + \mathcal{O}(x^5),$$
$$Y_+^{<>}(x) = 2x - a_3 x(x^2 - x_1^2) + \mathcal{O}(x^5). \quad (9.17)$$

We shall show later that $\Delta Y^<(x) = -\Delta Y^>(-x)$ by $\mathbb{P}$-symmetry. Thus the even part of $\Delta Y^>(x)$ does not contribute to eq. (6.14). Substituting eq. (9.17) in eq. (6.14) allows us to relate the

odd part of $\Delta Y^{>}(x)$ to the cubic correction in eq. (9.16)

$$\Delta Y^{>}(x) = -c_0 + c_2 x^2 - (2 - \tfrac{1}{2}k)a_3 x(x^2 - x_1^2) + \mathcal{O}(x^4). \tag{9.18}$$

Note that the linear and cubic terms vanish for $k = 4$ (the value at which the fixed point at the origin becomes hyperbolic) provided a_3 remains finite there. Also it turns out that a_3 is very small for moderate values of k.

Figure 8 also suggests that the part of the solution with support in intervals $x \in [-\frac{1}{2}, -\frac{1}{2} + a]$, $x \in [\frac{1}{2} - a, \frac{1}{2}]$, $0 < a < \frac{1}{2}$, is the same as the reversible solution, $\alpha = \mathrm{Id}$, for which C is the curve $y = k \sin(2\pi x)/2\pi$ and C^* is the x-axis, $y = 0$. This assumption implies that α maps the interval $[\frac{1}{2} - a, \frac{1}{2} + a]$ onto itself. This is obviously acceptable provided we can join these sections of C^* to the retrograde branch $y = Y_+^{<>}$ in such a manner as to satisfy the Euler–Lagrange equation eq. (6.4) everywhere.

Since the Euler–Lagrange equation implies that the separation between C and C^* is conserved at values of x which are orbits under the circle map α: $x^* = T(x, Y_-(x))$, the points on C^* connected to the point $(\frac{1}{2} - a, 0)$ by the Euler–Lagrange equation preserve the separation $k \sin(2\pi a)/2\pi$. That is, they must be orbits under the composite map $T \circ A_a$, where A_a is a map giving a constant shift $k \sin(2\pi a)/2\pi$ in the y-direction (which may be thought of as a discrete analogue of a constant acceleration)

$$A_a(x, y) \equiv \left(x, y + \frac{k}{2\pi} \sin 2\pi a\right). \tag{9.19}$$

The corresponding points on C are images under A_a of those on C^*. These points form the endpoints of smooth segments out of which we build the parts of the curves C and C^* whose projection on the x-axis is the interval $[-x_1, \frac{1}{2} - a]$.

It is readily verified that $(\frac{1}{2} - a, 0)$ is a fixed point of $T \circ A_a$, so that the endpoints of the C^* segments which we seek lie on the unstable manifold of this fixed point. (The anticausal solution can be constructed from the stable manifold.) To determine a we have simply to require that the last endpoint, $z_0^+ = (x_0, y_0^+)$, lie on the retrograde solution $y = 2x$, which implies the relation between $x_1 = -x_0$ and a

$$\frac{k}{2\pi} \sin 2\pi a = 4x_1 - \frac{k}{2\pi} \sin 2\pi x_1. \tag{9.20}$$

We can determine a and x_1 iteratively by guessing a, then using eq. (9.20) to calculate x_1 and testing to see whether $z_0^+ = (-x_1, -2x_1)$ is on the unstable manifold by iterating the inverse composite map to form the sequence $\{z_n^+\} \equiv \{(A_a^{-1} \circ T^{-1})^n (z_0^+)\}$. When a is chosen correctly, $z_n^+ \to (\frac{1}{2} - a, 0)$ as $n \to \infty$. The endpoints of the segments from which we make up C are then given by $z_n^- = A_a(z_n^+) = T^{-1}(z_{n-1}^+)$. We then have that α maps the interval $[x_1, \frac{1}{2} - a]$ onto $[-x_1, \frac{1}{2} - a]$ and the interval $[-\frac{1}{2} + a, -x_1]$ onto $[-\frac{1}{2} + a, x_1]$.

The interior points of the segments are determined by requiring that they lie on the unstable manifolds of a family of maps $T \circ A_b$, $b > a$, with fixed points $(\frac{1}{2} - b, 0)$. For instance, the maximum value of b can be determined iteratively by guessing b and testing to see whether the point $z_0^+(b) = (x, Y_-^{<>}(x) - (k/2\pi) \sin 2\pi b)$ lies on the stable manifold by iterating the inverse composite map to form the sequence $\{z_n^+(b)\} \equiv \{(A_b^{-1} \circ T^{-1})^n (z_0^+(b))\}$. When the appropriate value, $b = b(x)$, is chosen, $z_n(b) \to (\frac{1}{2} - b, 0)$ as $n \to \infty$. There is a complicating factor arising from this prescription, however. It is that the limit points $z_\infty(b) \equiv (\frac{1}{2} - b, 0)$ are different for different intermediate points, so that for large n the segments must become folded over. This is not a fatal problem since the folding is the same for both the C and C^* segments so that a (rather singular) conjugacy could be found such that $\rho(\theta)$ is monotonic in the region of the limit points. This region is too small for us so far to have been able to resolve it numerically.

We can now calculate $\Delta Y^{>}(x) = -(k/2\pi) \times \sin 2\pi b(x)$, find its cubic component, and hence deduce a_3 from eq. (9.18). We can use $\mathbb{P}$-symmetry to show that $\{P(z_n(b))\} \equiv \{(A_b \circ T^{-1})^n \circ P(z_0(b))\}$, where $P(x,y) \equiv (-x,-y)$ to prove the assumed relation $\Delta Y^{<}(x) = -\Delta Y^{>}(-x)$, but there is no corresponding proof that $\Delta Y^{>}$ is even, and numerical evidence confirms that it is not, although it is very close to even for moderate k. Thus a_3 is small but nonzero.

We conclude this section by noting that in the limit $k \to 0$, the set $\{x_n\}$ becomes dense and ΔY approaches a step function constant in the intervals $x \in [-\frac{1}{2} + a, 0]$ and $x \in [0, \frac{1}{2} - a]$, while elsewhere being equal to the reversible solution $-(k/2\pi) \sin 2\pi x$ (cf. fig. 7). We can use the fact that the flux φ_1 is the same as that for the reversible solution, $k/2\pi^2$, to give a condition for a in the limit $k \to 0$

$$\cot \pi a = (1 - 2a)\pi, \tag{9.21}$$

which yields $a = 0.128990$. The corresponding limiting behaviour for φ_2 is

$$\varphi_2 \sim 5.5615 \times 10^{-3} k^2, \tag{9.22}$$

which is of course less than the corresponding quadratic flux for the reversible solution given by eq. (9.3) as $6.33257 \times 10^{-3} k^2$.

10. Conclusions

We have established that the quadratic flux minimization principle provides a basis for a sensitive nonperturbative numerical method for investigating flux transport through resonances in area-preserving maps. We have also shown that the method produces a pair of approximately invariant curves which have dynamical significance, in that at least some of their intersections are associated with dynamical orbits (the correspondence between orbits and intersections is not complete when the associated circle map is only semiconjugate to rigid rotation, and also in the general case of noninvertible circle maps where the simple Euler–Lagrange equation eq. (6.4) does not apply). Though we have mainly studied periodic orbits, the Euler–Lagrange equation presumably holds for quasiperiodic orbits as well. The solutions have been classified according to whether or not they break the combined parity and time-reversal symmetry allowed by the class of maps investigated. In some cases we have sketched analytic methods for constructing these solutions. Except on invariant curves there are in general at least three solutions: two quadratic-flux-minimizing solutions with broken symmetry and a minimax solution which has the full symmetry allowed by the map. This has been shown both analytically and numerically for the case $\nu = 0$, and the numerical evidence is that it is generically true.

The symmetry breaking solutions are interesting as an example of a problem where an "arrow of time" arises spontaneously (with no statistical averaging, although there is averaging over angle in the definition of the quadratic flux). Indeed these solutions exhibit *irreversibility* in a very real sense since the iterates of the circle map describing the one-dimensional dynamics form only a semigroup. However these solutions are probably too singular to be useful for constructing a generalized action–angle representation, and in any case it would not appear appropriate for a choice of the direction of time evolution to be necessary in the essentially geometric exercise of constructing a coordinate system.

The reversible solutions corresponding to resonances (rational rotation numbers) are also interesting because, although not flux-minimizing, the approximately invariant curves associated with them are smooth and their calculation is susceptible to an apparently rapidly convergent perturbation theory which does not suffer from the usual small denominator problem. Also, since their circle maps are conjugate to rigid

rotation we have full control over their rotation numbers. It is these solutions which would appear to provide the basis for defining a generalized action–angle representation. One could use a truncated Farey tree construction to define the principal resonances in the domain of interest (and the most noble quasiperiodic surface between each resonance) and use the curves C, C^*, or the time-symmetric curve specified parametrically by $x = X(\theta)$, $y = \frac{1}{2}[Y_+(\theta) + Y_-(\theta)]$ to define a basic ladder of new momentum coordinate surfaces. These curves are assumed not to cross, which will certainly be true for sufficiently small k. The transformation to the new phase-space coordinates would then be completed by interpolation (rather than by using the quadratic-flux-minimizing surfaces corresponding to all irrational rotation numbers since these are not in general smooth and are not continuously connected to the resonance surfaces).

We have studied only the lowest order resonances in detail. It would be interesting to study an infinite sequence of resonances approaching a rotational invariant curve or a cantorus. In the former case φ_2 is obviously a local (and global) minimum on the invariant curve, since it vanishes there. We conjecture, but have not proved, that φ_2 is also a local minimum on a cantorus. The detailed behaviour of the constrained solutions as the control parameter is varied would also be interesting to investigate, as well as the implications of this method for the theory of transport in area preserving maps.

Acknowledgements

One of us (R.L.D.) would like to acknowledge L. Chim for assistance with the analytic single mode calculation, and W.A. Coppel and B.G. Kenny for suggesting useful references. It is a pleasure also to acknowledge useful comments from J.M. Greene and R.S. MacKay on the relation between symmetries and generating functions. One of us (J.D.M.) acknowledges the support of the US National Science Foundation, under grant NSF-DMS9001103.

Appendix A. Circle map identity

We give here a useful identity which may be used to prove relationships (sum rules) between the Fourier coefficients of a circle map, its sum-difference representation and its inverse. We shall work in x-space, though similar relations could equally well be derived in the θ-space representation. The identity is

$$\int_0^1 F(x^* - x)\,[x'_+(\eta) - x'_-(\eta)]\,\mathrm{d}\eta \equiv 0, \quad \text{(A.1)}$$

for *any* integrable function $F(x) = f'(x)$. Here $x^* - x$ is a shorthand for $x_+(\eta) - x_-(\eta)$. Equation (A.1) follows by recognizing that the integrand is the perfect differential $\mathrm{d}f(x^* - x)$ and that the endpoint values of $x^* - x$ are equal. From this it follows that

$$\begin{aligned}
&\int_0^1 F(x^* - x)\,x'_-(\eta)\,\mathrm{d}\eta \\
&\quad \equiv \int_0^1 F(x^* - x)\,x'_+(\eta)\,\mathrm{d}\eta \\
&\quad \equiv \tfrac{1}{2}\int_0^1 F(x^* - x)\,[x'_+(\eta) + x'_-(\eta)]. \quad \text{(A.2)}
\end{aligned}$$

In particular, choosing $F(\cdot) \equiv \cdot$ and $\eta = x_\pm^{-1}(x)$ (assuming the inverse functions exist), we see that the constant in the Fourier representation of α^{-1} corresponding to eq. (3.2) is simply $-\Omega$.

Appendix B. Time-symmetric representation

A representation in which $\mathbb{PT}$-reversibility (or otherwise) of the map $\rho : \theta \mapsto \theta^*$ is manifest is

the *sum–difference representation*

$$\theta^* - \theta = \omega\left(\frac{\theta + \theta^*}{2}\right), \tag{B.1}$$

where ω is a periodic function. In the rigid rotation case $\rho = R_\nu$, $\omega(\cdot) \equiv \nu$ and we have θ^* explicitly in terms of θ. However, in the general case eq. (B.1) defines the map only implicitly.

Defining $\eta \equiv (\theta + \theta^*)/2$, we may rewrite eq. (B.1) as the parametric representation

$$\begin{aligned} \theta^* &= \theta_+(\eta) \equiv \eta + \tfrac{1}{2}\omega(\eta), \\ \theta &= \theta_-(\eta) \equiv \eta - \tfrac{1}{2}\omega(\eta). \end{aligned} \tag{B.2}$$

Solving eqs. (B.2) for θ^* and θ we find

$$\begin{aligned} \rho(\cdot) &= \theta_+ \circ \theta_-^{-1}(\cdot), \\ \rho^{-1}(\cdot) &= \theta_- \circ \theta_+^{-1}(\cdot), \end{aligned} \tag{B.3}$$

respectively. Existence of ρ and ρ^{-1} depends on the existence of the inverse functions $\theta_\pm^{-1}$, for which we need

$$|\omega'(\eta)| < 2, \ \eta \in \mathbb{R};. \tag{B.4}$$

From this point of view the best choice of transformation X is that which minimizes $\sup|\omega'(\eta)|$. However, since this selects ρ which is closest to a rigid rotation, Herman's theorem suggests that aggressive minimization of $\sup|\omega'(\eta)|$ may be undesirable in practice because it leads to poor smoothness properties of X.

If ω is an even function, then the symmetries $\theta_\pm(-\eta) = -\theta_\mp(\eta)$ and $\theta_\pm^{-1}(-\theta) = -\theta_\mp^{-1}(\theta)$ follow readily from eqs. (B.2). It follows from eqs. (3.9) and (B.3) that ρ is $\mathbb{PT}$-reversible.

By substituting $\theta^* = \rho(\theta)$ and $\theta = \rho^{-1}(\theta^*)$ into eq. (B.1), setting $\theta^* = -\theta$, subtracting the resulting equations and using eq. (3.9) we see that $\mathbb{PT}$-reversibility implies that ω is an even function. That is, we have the result

Theorem 2. A circle map $\rho : \theta \mapsto \theta^*$, representable in the sum–difference form eq. (B.1), is $\mathbb{PT}$-reversible if and only if ω *is even.*

In this case only cosine terms appear in the Fourier series for ω.

The sum–difference representation can be used as an alternative, time-symmetric parametrization to that used in section 4, as illustrated in appendix C. It is also sometimes convenient in analytical work, as in section 9.1. However, its implicit nature and the need to assume invertibility of both θ_+ and θ_- make the explicit representation eq. (3.3) preferable for most of this paper.

Appendix C. Constrained variation

C.1. Constraints

Meiss and Dewar [17] derived the Euler–Lagrange equation for extremizing φ_2 defined without using the conjugacy function X (equivalent to taking X to be the identity). Thus the circle map $\alpha(x) = X \circ \rho \circ X^{-1}(x)$ was restricted to be a diffeomorphism. In our present formulation this hard constraint is removed since X need not be monotonic.

However, we shall find that unconstrained minimization of φ_2 selects only irrational values of the rotation number ν of α and ρ (the rational values being obtained as minimax solutions) whereas our desire to construct a generalization of action–angle coordinates leads us to seek a way to scan *continuously* over the phase space as in figs. 5 and 6. Constraining ν directly is not practicable because its value is not simply related to any of the representations of the general circle map, and in any case the circle map mode-locking phenomenon means that ν is not a good control parameter. A practical method of constraint is to hold fixed Ω, the constant in the Fourier representation of α, eq. (3.2) as was done in figs. 5 and 6. We could also constrain the corresponding constant in the Fourier representation of ρ, or, better, in the Fourier representation of $\omega(\eta)$, but these constants have no fundamental significance since they depend on

choice of representation. Another fundamental constraint would be to hold the area $\mathcal{A}$ ($= \mathcal{A}^*$) fixed while minimizing φ_2.

In this appendix we find how such constraints affect the Euler–Lagrange equation by using the method of Lagrange multipliers. We also use this as an opportunity to show how the variational principle can be formulated in a time-symmetric fashion by using the sum–difference representation of appendix B as an alternative parametrization to that used in section 4. Specifically, in place of eq. (4.2) we take

$$x_\pm(\eta) = X \circ \theta_\pm(\eta), \tag{C.1}$$

and in place of eq. (4.3),

$$Y_\pm(\theta) \equiv y_\pm \circ \theta_\pm^{-1}(\theta). \tag{C.2}$$

Introducing the Lagrange multiplier λ to effect the constraint, we obtain the new objective functional

$$\Phi = \varphi_2 - \lambda\psi, \tag{C.3}$$

with ψ taken to be either $\mathcal{A}$ or

$$\begin{aligned}\Omega &\equiv \int_0^1 [x_+(\eta) - x_-(\eta)]\, x'_-(\eta)\, d\eta, \\ &= \tfrac{1}{2}\int_0^1 [x_+(\eta) - x_-(\eta)] \\ &\quad \times [x'_+(\eta) + x'_-(\eta)]\, d\eta,\end{aligned} \tag{C.4}$$

where the second form follows from eq. (A.2). The Lagrange multiplier λ is to be determined by the condition

$$\psi = \psi_0 = \text{const.} \tag{C.5}$$

We can perform an unconstrained minimization by setting $\lambda = 0$.

C.2. Constrained extremization

In order to carry out the variation we use the identity

$$\delta f^{-1}(\cdot) = -\frac{\delta f \circ f^{-1}(\cdot)}{f' \circ f^{-1}(\cdot)}, \tag{C.6}$$

for any function $f(\cdot)$, which can be obtained by varying the identity $f \circ f^{-1}(\cdot) \equiv \cdot$. Using this identity we have

$$\delta\rho^{\pm 1}(\theta) = \frac{\theta'_\mp(\eta_\mp)\delta\theta_\pm(\eta_\mp) - \theta'_\pm(\eta_\mp)\delta\theta_\mp(\eta_\mp)}{\theta'_\mp(\eta_\mp)}, \tag{C.7}$$

where $\eta_\mp \equiv \theta_\mp^{-1}(\theta)$.

Varying $X(\cdot)$ and/or $\theta_\pm(\cdot)$ in eq. (C.3) yields, after changes of variable $\theta = \theta_\pm(\eta)$ as appropriate (cf. section 6.1)

$$\begin{aligned}\delta\Phi = -\int_0^1 & F_{12}(x_-(\eta), x_+(\eta)) \\ & \times [\Delta Y \circ \theta_+(\eta) - \Delta Y \circ \theta_-(\eta)]\delta W(\eta)\, d\eta \\ & -\lambda\delta\psi,\end{aligned} \tag{C.8}$$

where

$$\delta W(\eta) \equiv x'_-(\eta)\delta x_+(\eta) - x'_+(\eta)\delta x_-(\eta). \tag{C.9}$$

We can hold the function X fixed without loss of generality because the same variation in δW can be effected by a suitable variation in ρ as can be effected by varying X except at the turning points of X, which we can ignore by continuity. Using the time-symmetric parametric representation eq. (B.2) the variational form δW depends only on the variation in the single function ω:

$$\delta W = X' \circ \theta_+(\eta)\, X' \circ \theta_-(\eta)\, \delta\omega(\eta). \tag{C.10}$$

This form is pleasing for its symmetry between past and present. However we shall find that

there are situations where the circle map or its inverse is not invertible. For a discussion of this case see sections 8 and 9.3, but for the present we assume ρ to be a diffeomorphism.

The variations $\delta\psi$ in our two constraints are given by

$$\delta\Omega = \int_0^1 \delta W(\eta)\,\mathrm{d}\eta, \tag{C.11}$$

and

$$\delta\mathcal{A} = -\int_0^1 F_{12}(x_-(\eta), x_+(\eta))\,\delta W(\eta)\,\mathrm{d}\eta\,. \tag{C.12}$$

Note that, for maps with generating functions of the form eq. (2.21), we have $F_{12} \equiv -1$ so $\delta\Omega = \delta\mathcal{A}$ (indeed, $\Omega = \mathcal{A}$) so that the two constraints are equivalent in this case.

Requiring $\delta\Phi = 0$ for all $\delta\omega$ in eq. (C.10) yields, using eq. (C.8), the Euler–Lagrange equation

$$\begin{aligned} &X'\circ\theta_+(\eta)X'\circ\theta_-(\eta) \\ &\quad\times[\Delta Y\circ\theta_+(\eta) - \Delta Y\circ\theta_-(\eta) - \kappa(\eta)\lambda] \\ &\quad= 0, \end{aligned} \tag{C.13}$$

where $\kappa = 1$ when we take $\psi = \mathcal{A}$ and $\kappa(\eta) = -1/F_{12}(x_-, x_+)$ when we choose the constraint $\psi = \Omega$. In either case, κ is positive definite.

Equation (C.13) is to be satisfied for all $\eta \in [0,1]$ (except ranges of θ for which ρ or ρ^{-1} has negative slope—see section 9.3), and hence, by periodic extension, all $\eta \in \mathbb{R}$. Setting $[\,\cdot\,] = 0$, we can regard (C.13) as a simple difference equation which connects $\Delta Y(\theta)$ on orbits of the circle map $\rho(\theta)$. Since this implies that the gap between C^* and C increases by λ at every iteration of the circle map, it would appear that, for $\lambda \neq 0$, the distance between the curves would have to diverge at attracting periodic points. The cure for this seeming pathology would appear to lie in the modification to the Euler–Lagrange equation which applies when the assumption of invertibility of the circle map breaks down (see section 6.5).

Certainly, we have found numerically that convergence to a minimum with Ω constrained to be constant is possible using a representation with a finite number of Fourier modes and that the function $\varphi_2(\Omega)$ appears continuous, although there does appear to be some complicated bifurcation behaviour near the $(0,1)$ resonance. We have not done a systematic study of convergence with respect to the number of Fourier modes.

We have also investigated the use of a localized constraint, such as specifying the value of $\alpha(0)$, since the $\lambda = 0$ Euler–Lagrange equation is then satisfied almost everywhere, but numerical evidence indicates that this type of constraint gives similar results to the global constraint of constant area.

References

[1] A.N. Kolmogorov, Akad. Nauk SSSR Dokl. 98 (1954) 527–530 [English translations in Lecture Notes in Physics Vol. 93 (1979) p. 51; B.-l. Hao, ed., Chaos II (World Scientific, Singapore) pp. 107–112.].

[2] V.I. Arnol'd, Usp. Mat. Nauk 18 (1963) 13–39 [English translation: Russian Math. Surv. 18 (1963) 9–36].

[3] J. Moser, Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II (1962) 1–20.

[4] R.S. MacKay, J.D. Meiss and I.C. Percival, Physica D 13 (1985) 55–81.

[5] J.D. Hanson and J.R. Cary, Phys. Fluids 27 (1984) 767–769.

[6] A.H. Reiman and N. Pomphrey, J. Comput. Phys. 94 (1991) 225–249.

[7] R.L. Warnock and R.D. Ruth, Long-term bounds on nonlinear Hamiltonian motion, Physica D 56 (1992) 188–215.

[8] R.L. Dewar, Phys. Fluids 16 (1973) 1102–1107.

[9] J.R. Cary and A.N. Kaufman, Phys. Fluids 24 (1981) 1238–1250.

[10] R.L. Dewar and G.W. Kentwell, Phys. Letters A 111 (1985) 391–395.

[11] R.L. Dewar, J. Phys. A: Math. Gen. 9 (1976) 2043–2057.

[12] E. Wigner, J. Chem. Phys. 5 (1937) 720–725.

[13] R.L. Dewar, Physica D 17 (1985) 37–53.

[14] A. Bazzani, E. Remiddi and G. Turchetti, J. Phys. A: Math. Gen. 24 (1991) L53–L59.
[15] S. Aubry, Physica D 7 (1983) 240–258.
[16] J.N. Mather, Topology 21 (1982) 457–467.
[17] J.D. Meiss and R.L. Dewar, Minimizing flux, in: Proc. Miniconference on Chaos and Order (Centre for Mathematical Analysis, Australian National University, Canberra, Australia, 1991), eds. N. Joshi and R.L. Dewar (World Scientific, Singapore), pp. 97–103.
[18] R.S. MacKay, Renormalisation in area preserving maps, PhD thesis, Princeton University Plasma Physics Laboratory (October 1982).
[19] H. Goldstein, Classical Mechanics, 2nd ed. (Addison-Wesley, Reading, MA, 1950), ch. 9.
[20] R. Abraham and J.E. Marsden, Foundations of Mechanics, 2nd ed. (Addison-Wesley, Redwood City, CA, 1985) ch. 4, pp. 308–309.
[21] E.C.G. Sudarshan and N. Mukunda, Classical Dynamics: A Modern Perspective (Wiley-Interscience, New York, 1974) ch. 16, pp. 275–281.
[22] V.I. Arnol'd and M.B. Sevryuk, Oscillations and bifurcations in reversible systems, in: Nonlinear phenomena in plasma physics and hydrodynamics, ed. R.Z. Sagdeev (Mir, Moscow, 1986) ch. 2, p. 37.
[23] J.A. G. Roberts and G.R. W. Quispel, Phys. Rep. 216 (1992) 63–177.
[24] J.M. Greene, J. Math. Phys. 20 (1979) 1183–1201.
[25] J.M. Greene, R.S. MacKay, F. Vivaldi and M.J. Feigenbaum, Physica D 3 (1981) 468–486.
[26] I.C. Percival, Variational principles for invariant tori and cantori, in: Nonlinear Dynamics and the Beam–Beam Interaction eds. M. Month and J. Herrera (Am. Inst. Phys., New York, 1979) pp. 302–310.
[27] Y. Katznelson and D. Ornstein, Ergod. Theor. Dynam. Syst. 9 (1989) 643–680.
[28] S.P. Hirshman and H.K. Meier, Phys. Fluids 28 (1985) 1387–1391.
[29] R.S. MacKay and J.D. Meiss, Phys. Lett. A 98 (1983) 92–94.
[30] I.C. Percival, J. Phys. A: Math. Gen. 12 (1979) L57–L60.
[31] E. Rosengaus and R.L. Dewar, J. Math. Phys. 23 (1982) 2328–2338.
[32] W.H. Press, B.P. Flannery, S.A. Teukolsky and W.T. Vetterling, Numerical Recipes, (Cambridge Univ. Press, Cambridge, 1986).
[33] G.H. Hardy and E.M. Wright, An Introduction to the Theory of Numbers (Oxford Univ. Press, Oxford, 1979).
[34] A.B. Rechester, M.N. Rosenbluth and R.B. White, Phys. Rev. A 23 (1981) 2664–2672.
[35] J.R. Cary, J.D. Meiss and A. Bhattacharjee, Phys. Rev. A 23 (1981) 2744–2746.