

The Periodic Orbits of an Area Preserving Twist Map

S. B. Angenent*

Department of Mathematics, University of Leiden, Niels Bohrweg 1, Leiden, The Netherlands

Abstract. We study the oscillation properties of periodic orbits of an area preserving twist map. The results are inspired by the similarity between the gradient flow of the associated action-function, and a scalar parabolic PDE in one space dimension. The Conley-Zehnder Morse theory is used to construct orbits with prescribed oscillatory behavior.

1. Introduction

We shall consider a C^1 area preserving diffeomorphism \tilde{F} of the cylinder $S^1 \times R$ onto itself. Such a diffeomorphism can be described by a mapping $F: R^2 \rightarrow R^2$ (its lift) given by $F(x, y) = (f(x, y), g(x, y))$, where x is the angle coordinate. The components of F satisfy the periodicity conditions

$$f(x+1, y) = f(x, y) + 1, \quad g(x+1, y) = g(x, y).$$

The map F is said to be a *twist diffeomorphism* if $f(x, y)$ is an increasing function of y , and in fact

$$\partial_2 f(x, y) > 0 \tag{1.1}$$

holds for all (x, y) in R^2 . Here ∂_k denotes differentiation with respect to the k -th argument.

We shall consider twist diffeomorphisms which satisfy the *infinite twist condition*, i.e.

$$\lim_{y \rightarrow \pm \infty} f(x, y) = \pm \infty$$

for any $x \in \mathbb{R}$. We shall study the set of periodic orbits of \tilde{F} .

The main feature which distinguishes twist maps from other area preserving maps is that they have a single valued generating function, i.e. there is a C^2 function

* Current address: Department of Mathematics, University of Wisconsin, Van Vleck Hall, 480 Lincoln Drive, Madison, WI 53706, USA

$h(x, x')$ on R^2 such that

$$(x', y') = F(x, y) \quad \text{iff} \quad y = \partial_1 h(x, x') \quad \text{and} \quad y' = -\partial_2 h(x, x').$$

This function also satisfies

$$\partial_1 \partial_2 h(x, x') > 0 \quad \text{for all } x \text{ and } x'. \tag{1.2}$$

The generating function h is uniquely determined (up to a constant) by the map F . Its construction is given by Mather [10]. See also Aubry and le Daeron [2].

The existence of h implies that any orbit of the map F is completely determined by its sequence of x coordinates $\{x_n : n \in Z\}$. One easily verifies that a sequence $\{x_n : n \in Z\}$ can only be the sequence of x coordinates of an orbit of F if

$$\partial_i W(\dots, x_{-1}, x_0, x_1, \dots) = 0,$$

where

$$W(x) = \sum_{i=-\infty}^{+\infty} h(x_i, x_{i+1})$$

(this sum is not well defined, however its partial derivatives are).

Orbits of F are therefore critical points of the “function” W . This idea becomes useful if we restrict ourselves to periodic orbits of \tilde{F} . Let $X_{p,q}$ be defined by

$$X_{p,q} = \{(x_i)_{i \in Z} : x_{i+q} = x_i + p \text{ for all } i \in Z\},$$

where p and q are integers and $q > 0$. For any x in $X_{p,q}$ we define

$$W(x) = \sum_{i=1}^q h(x_i, x_{i+1}).$$

Now W is a well defined C^2 function on $X_{p,q}$ and its critical points are periodic orbits of \tilde{F} with period q , which “go around the cylinder” p times after q iterations.

The usual way of constructing such orbits is to prove that W takes its maximum at some point in $X_{p,q}$. This maximum must of course be a critical point of W . Most of the critical points of W are not maxima however. It is our intention to study these other critical points.

Our approach is based on the similarity between the present problem and that of finding solutions of

$$(I) \quad \begin{cases} x''(t) + f(t, x(t)) = 0, & -\infty < t < +\infty, \\ x(t+p) = x(t) + q, \end{cases}$$

where $f(t, x)$ is a C^1 function satisfying

$$f(t+1, x) = f(t, x+1) = f(t, x).$$

Indeed, this problem is variational, with potential function

$$\tilde{W}(x) = - \int_0^q [\frac{1}{2}x'(t)^2 - g(t, x(t))] dt,$$

where

$$g(t, x) = \int_0^x f(t, \xi) d\xi.$$

In this case the gradient flow of \tilde{W} (with respect to the $L_2(0, q)$ innerproduct) is a semilinear parabolic partial differential equation:

$$(II) \quad \begin{cases} x_s = x_{tt} + f(t, x), & s > 0, & -\infty < t < +\infty, \\ x(s, t + q) = x(s, t) + p. \end{cases}$$

The flow induced by (II) on some function space has a number of geometric properties, the most prominent of which is the maximum principle. The gradient flow of W on $X_{p,q}$ turns out to have similar properties, and in fact our main results will concern systems which contain (I) as a special case.

To illustrate the analogy we have in mind we shall compute $W, \partial_k W,$ and F in the following specific example. Let $g(x)$ be a C^2 function such that $g(x) = g(x + 1)$ holds for all real x , and define $f(x) = g'(x)$, and

$$h(x, x') = -\frac{1}{2}(x - x')^2 + g(x').$$

Then $h(x, x')$ is the generating function of a twist map $F: (x, y) \rightarrow (x', y')$ given by

$$\begin{aligned} y &= \partial_1 h(x, x') = x' - x, & \text{i.e. } x' &= x + y, \\ y' &= -\partial_2 h(x, x') = x' - x - f(x'), & \text{i.e. } y' &= y - f(x + y), \end{aligned}$$

so $F(x, y) = (x + y, y - f(x + y))$. If one takes $g(x) = (k/2\pi) \cos(2\pi x)$ ($k > 0$), one obtains a map known as the “standard map.”

We see that the orbit $\{(x_k, y_k)\}_{k \in \mathbb{Z}}$ is indeed completely determined by its sequence of x coordinates, since we have $y_k = x_{k+1} - x_k$. Furthermore a sequence $\{x_k\}_{k \in \mathbb{Z}}$ is the sequence of x coordinates of an orbit of the map F if and only if it satisfies the second order difference equation

$$\partial_k W = x_{k+1} - 2x_k + x_{k-1} + f(x_k) = 0, \quad k \in \mathbb{Z},$$

where

$$W = \sum_k \left\{ -\frac{1}{2}(x_k - x_{k+1})^2 + g(x_{k+1}) \right\}.$$

Note that this is the numerical analyst’s (simplest) version of the differential equation occurring in I! The gradient flow of W is given by

$$\dot{x}_k = x_{k+1} - 2x_k + x_{k-1} + f(x_k),$$

which is the discretized version of II.

Seen in this way, it should not come as a surprise that the gradient flow of W has much in common with the parabolic PDE II.

2. The Gradient Flow of W

Let p and q be given integers, with $q > 0$, and let F_1, \dots, F_q be C^1 area preserving twist maps of R^2 with generating functions h_1, \dots, h_q respectively. We define h_j for $j < 1$ and $j > q$ by requiring $h_j = h_{j+q}$ for all j . We shall assume that the maps F_j satisfy the following condition:

The no flux condition. Let C be a curve which winds around the cylinder once. Then the total area between C and $F(C)$ vanishes, i.e.

$$\int_C y dx = \int_{F(C)} y dx.$$

This condition is equivalent to the following condition on the generating functions $h_j(x, x')$:

$$(H) \quad h(x + 1, x' + 1) = h(x, x') \quad \text{for all } x, x'.$$

We recall that the twist property (1.1) of the maps F_j implies that

$$(Tw) \quad \partial_1 \partial_2 h(x, x') > 0 \quad \text{for all } x \text{ and } x'.$$

We refer to Mather [10, 11] and Aubry and le Daeron [2] for more details.

Given the generating functions h_1, \dots, h_q we define W in $C^2(X)$ as follows:

$$W(x) = \sum_{j=1}^q h_j(x_{j-1}, x_j),$$

where

$$X = X_{p,q} = \{(x_j)_{j \in \mathbb{Z}} : x_{j+q} = x_j + p \text{ for all } j\}$$

is the space we defined in the introduction. This function will enable us to study the fixed points of the composite map $F_q \circ F_{q-1} \circ \dots \circ F_1$. If $F_j = F$ for all j then such fixed points are of course q periodic points of the single map F . Note that the composite map $F_q \circ F_{q-1} \circ \dots \circ F_1$ need not be a twist map.

Lemma 2.1 *Any point (a, b) in R^2 satisfies*

$$F_q \circ \dots \circ F_1(a, b) = (a + p, b)$$

if and only if there is a critical point x in X of W which satisfies

$$\begin{aligned} x_1 &= a, \\ -\partial_2 h_1(x_0, x_1) &= b. \end{aligned}$$

Proof. The gradient of W is given by

$$\partial_j W = \partial_2 h_j(x_{j-1}, x_j) + \partial_1 h_{j+1}(x_j, x_{j+1}).$$

Hence if one defines $y_j = \partial_1 h_{j+1}(x_j, x_{j+1})$, then

$$(x_j, y_j) = F_j(x_{j-1}, y_{j-1})$$

if and only if $\partial_j W = 0$, which proves the lemma.

In view of this result we shall identify any orbit of F_1, \dots, F_q with its sequence of x coordinates.

There are several structures on X which are relevant for our problem. First of all X is a manifold diffeomorphic to R^q . Furthermore X has a partial ordering defined by

$$x \leq y \quad \text{iff} \quad x_i \leq y_i \quad \text{for all } i.$$

We also define

$$\begin{aligned} x < y &\quad \text{iff} \quad x \leq y \quad \text{and} \quad x \neq y, \\ x \ll y &\quad \text{iff} \quad x_i < y_i \quad \text{for all } i. \end{aligned}$$

With this ordering X becomes a lattice.

The gradient flow of W will be denoted by $\{\phi_t\}_{t \geq 0}$. It is defined by the differential equation

$$x'(t) = \text{grad } W(x(t))$$

on X .

Lemma 2.2. *For all x, y in X and $t > 0$ we have:*

- (a) *if $x \leq y$ then $\phi_t(x) \leq \phi_t(y)$,*
- (b) *if $x < y$ then $\phi_t(x) \ll \phi_t(y)$.*

Proof. Let V_j denote $\partial_j W$, then

$$\partial_i V_j = \begin{cases} 0 & \text{if } j < i - 1 \text{ or } j > i + 1 \\ \partial_1 \partial_2 h_i(x_i, x_i) & \text{if } j = i + 1 \\ \partial_1 \partial_2 h_j(x_i, x_j) & \text{if } j = i - 1. \end{cases}$$

Hence all the off diagonal elements of the Jacobian of $\text{grad } W$ are nonnegative, which implies (a). The other assertion follows from the fact that the upper and lower diagonal only contain strictly positive elements (see Hirsch [8]).

The analogue of this lemma for the parabolic equation (II) mentioned in the introduction is of course the comparison principle. Lemma 2.2 may be restated as follows:

If the graphs of x and y (think of them as piecewise linear functions on R) do not intersect, then neither do the graphs of $\phi_t(x)$ and $\phi_t(y)$.

For parabolic equations this has been generalized by Matano [9]. His result states roughly that, if $x(s, t)$ and $y(s, t)$ are solutions of the parabolic equation (II), then the number of intersections of the graphs of $t \rightarrow x(s, t)$ and $t \rightarrow y(s, t)$ is a nonincreasing function of s (note that for the PDE we have called the time variable s)¹. The discrete version of this principle is given in Smillie [13] in the context of competitive and cooperative differential equations.

To define the “number of intersections” we introduce some more notation.

For x, y in X we say that x intersects y transversally, in symbols $x \bar{\cap} y$, if, for any integer i , $x_i = y_i$ implies that $x_{i-1} - y_{i-1}$ and $x_{i+1} - y_{i+1}$ have opposite signs.

If x and y intersect transversally then we define $I(x, y)$, the number of intersections of x with y , to be the largest integer k for which there are

$$i_0 < i_1 < \dots < i_k = i_0 + q,$$

such that

$$(x_{i_j} - y_{i_j})(x_{i_{j+1}} - y_{i_{j+1}}) < 0$$

holds for $j = 0, 1, 2, \dots, k - 1$.

Clearly k must be even, and $k \leq q$.

The following result was proved by Smillie [13]:

Proposition 2.3. *Given x and y in X , the set of t in R for which $\phi_t(x)$ and $\phi_t(y)$ do not intersect transversally is discrete and $I(\phi_t(x), \phi_t(y))$ is a nonincreasing function of t*

¹ Actually Matano doesn't prove this. Our statement is an easy consequence of his results however

which has a jump discontinuity exactly at those t for which $\phi_t(x)$ and $\phi_t(y)$ do not intersect transversally.

3. The Twist Number of an Orbit

Let x in X be a critical point of W , and put $y_j = \partial_1 h_{j+1}(x_j, x_{j+1})$ and $P_j = (x_j, y_j)$. Then $P_j = F_j(P_{j-1})$ for any integer j (where we have defined $F_{j+q} = F_j$).

An intuitive description of the twist number can be given as follows. Take a vector u_0 in $T_{P_0}R^2$, and define u_j in $T_{P_0}R^2$ by

$$u_j = dF_j(P_{j-1})u_{j-1} \quad \text{for all } j.$$

Identify the tangent spaces $T_{P_j}\mathbb{R}^2$ with \mathbb{R}^2 in the obvious way, and let the vector u_j have components (ξ_j, η_j) . Assume that $u_0 \neq 0$.

For each integer j we define θ_j to be the angle between u_{j-1} and u_j , oriented in the clockwise sense. This angle is only defined up to a multiple of 2π , so we have to specify which multiple we mean. For this we use the following rule:

$$\begin{aligned} \text{if } \xi_{j-1} \cdot \xi_j \geq 0, & \quad \text{then } -\pi < \theta_j < +\pi, \\ \text{if } \xi_{j-1} \cdot \xi_j < 0, & \quad \text{then } 0 < \theta_j < +2\pi. \end{aligned}$$

Then we define the twist number of the orbit x , $\tau(x)$, to be

$$\tau(x) = \lim_{n \rightarrow \infty} (2n)^{-1} \sum_{j=-nq}^{+nq} \theta_j / 2\pi.$$

Roughly speaking, $2\pi\tau(x)$ is the average angle about which $dF(P_0)$ rotates the vector u_0 . Or, alternatively, $\tau(x)$ is the average number of times the sequence ξ_n changes sign, in an interval of length q .

Instead of proving that this limit exists and is independent of u_0 we shall give another definition. Our definition of $\tau(x)$ will involve the Morse index of x as a critical point of W . It must be seen as a discrete version of Floquet theory for a Hill’s operator, which reflects the analogy of our problem with the problem (I) of the introduction.

In a recent paper Mather [11] defines “the amount of rotation” of an orbit of a C^1 map of the cylinder. He also shows how this amount of rotation is connected with the Morse index of a critical point of W . It follows from his results that the twist number $\tau(x)$ coincides with Mather’s amount of rotation in the case we study.

First we take a look at the connection between the derivatives $dF_j(P_{j-1})$ and the second derivative of W at x .

Lemma 3.1. *Let $u_j = (\xi_j, \eta_j)$ in $T_{P_j}R^2$ be given. Then $u_j = dF(P_{j-1}) \cdot u_{j-1}$ holds for all j if and only if*

$$\begin{aligned} \alpha_i \xi_{i-1} + \beta_i \xi_i + \alpha_{i+1} \xi_{i+1} &= 0, \\ \eta_i &= \gamma_i \xi_i + \alpha_{i+1} \xi_{i+1} \end{aligned}$$

hold for all integers i . Here α_i , β_i and γ_i are defined by

$$\begin{aligned} \alpha_i &= \partial_1 \partial_2 h_i(x_{i-1}, x_i), \\ \beta_i &= \partial_2^2 h_i(x_{i-1}, x_i) + \partial_1^2 h_{i+1}(x_i, x_{i+1}), \\ \gamma_i &= \partial_1^2 h_{i+1}(x_i, x_{i+1}). \end{aligned}$$

Proof. If one differentiates the relations

$$y_j = -\partial_2 h_j(x_{j-1}, x_j) = \partial_1 h_{j+1}(x_j, x_{j+1}),$$

one gets

$$\begin{aligned} \eta_j &= -\partial_2^2 h_j(x_{j-1}, x_j) \xi_j - \partial_1 \partial_2 h_j(x_{j-1}, x_j) \xi_{j-1} \\ &= \partial_1^2 h_{j+1}(x_j, x_{j+1}) \xi_j + \partial_1 \partial_2 h_{j+1}(x_j, x_{j+1}) \xi_{j+1}, \end{aligned}$$

which proves the lemma.

After we have identified $T_x X$ with R^q (coordinates: ξ_j with $\xi_{j+q} = \xi_j$) we see that the linear part of the vectorfield $\text{grad } W$ at its critical point x is given by L ,

$$(L\xi)_i = \alpha_i \xi_{i-1} + \beta_i \xi_i + \alpha_{i+1} \xi_{i+1},$$

where α_i and β_i are defined as in Lemma 3.1. Thus L is a Jacobi matrix, and the following is known (see van Moerbeke [12]):

Proposition 3.2. *The spectrum of L is given by*

$$\sigma(L) = \{ \lambda_0 > \lambda_1 \geq \lambda_2 > \lambda_3 \geq \dots \geq \lambda_{q-1} \}.$$

So, for all i we have $\lambda_{2i} > \lambda_{2i+1}$.

For each real λ we consider the difference equation

$$\alpha_i \xi_{i-1} + \beta_i \xi_i + \alpha_{i+1} \xi_{i+1} = \lambda \xi_i.$$

Given ξ_0 and ξ_1 then all other ξ_i 's can be computed and we have

$$\begin{pmatrix} \xi_q \\ \xi_{q+1} \end{pmatrix} = M(\lambda) \cdot \begin{pmatrix} \xi_0 \\ \xi_1 \end{pmatrix},$$

where the two by two matrix $M(\lambda)$ is given by

$$M(\lambda) = \begin{pmatrix} 0 & 1 \\ -\alpha_q & \lambda - \beta_q \\ \alpha_{q+1} & \alpha_{q+1} \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ -\alpha_1 & \lambda - \beta_1 \\ \alpha_2 & \alpha_2 \end{pmatrix}.$$

Note that $\det M(\lambda) = 1$. Clearly λ is an eigenvalue of L if and only if 1 is an eigenvalue of $M(\lambda)$.

Since $M(\lambda)$ is a linear mapping on R^2 it maps half lines to half lines and therefore induces a mapping on the set of half lines with the origin as end point. This set (with the appropriate topology) is naturally identified as a circle ("polar coordinates"), and therefore we may regard $M(\lambda)$ as an orientation preserving diffeomorphism of the circle. For such a diffeomorphism the rotation number is defined up to an integer (see Coddington and Levinson [3]), and it depends continuously on the mapping, i.e. on λ .

Now it is known that $M(\lambda)$ depends on λ in the following manner:

- (1) $\lambda > \lambda_0$ or $\lambda_{2i-1} > \lambda > \lambda_{2i}$. In this case $M(\lambda)$ is conjugate to²

$$\begin{pmatrix} \mu & 0 \\ 0 & \mu^{-1} \end{pmatrix}$$

² Within $SL(2, R)$, i.e. there is a positively oriented base of R^2 with respect to which M has the specified matrix. The orientation explains the occurrence of the ± 1 in the Jordan normal forms

for some $\mu > 1$. As a map of the circle it has four hyperbolic fixed points, whence its rotation number is an integer.

(2) $\lambda_{2i} > \lambda > \lambda_{2i+1}$. There are several possibilities: $M(\lambda)$ is conjugate to one of the following standard forms:

$$\begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \quad 0 < \phi < 2\pi,$$

$$\begin{pmatrix} -\mu & 0 \\ 0 & -\mu^{-1} \end{pmatrix} \quad \mu > 1,$$

or

$$\begin{pmatrix} -1 & \pm 1 \\ 0 & -1 \end{pmatrix}.$$

In the first case the rotation number is $k + (\phi/2\pi)$ for some integer k , and in the other cases it is $\frac{1}{2} + k$ for some k .

(3) $\lambda = \lambda_j$. Again there are several possibilities. If λ_j is a double eigenvalue of L , then $M(\lambda)$ is the identity matrix. Otherwise $M(\lambda)$ is conjugate to

$$\begin{pmatrix} 1 & \varepsilon \\ 0 & 1 \end{pmatrix},$$

where $\varepsilon = (-1)^j$. Here we see that the conjugacy class of $M(\lambda)$ within $SL(2, R)$ allows us to distinguish between λ_{2i} and λ_{2i+1} .

Furthermore, the rotation number is a nondecreasing function of λ . Indeed, all the factors

$$\begin{pmatrix} 0 & 1 \\ -\alpha_j & \lambda - \beta_j \\ \alpha_{j+1} & \alpha_{j+1} \end{pmatrix}$$

of the product that defines $M(\lambda)$ represent diffeomorphisms of the circle which are strictly increasing in the parameter λ . Using this fact one can see what happens to $M(\lambda)$ (considered as a map of the circle) as λ decreases from $+\infty$ to $-\infty$ (see Fig. 1). In this way one verifies the classification of the $M(\lambda)$'s given above.

Hence, if we normalize the rotation number of $M(\lambda)$ by requiring it to be zero for $\lambda > \lambda_0$, there is a unique function $\varrho(\lambda)$ representing the rotation number of $M(\lambda)$.

Lemma 3.3. *The rotation number $\varrho(\lambda)$ of $M(\lambda)$, with $\varrho(\lambda) = 0$ for large λ is nondecreasing in λ . The eigenvalues of L are those $\lambda_j \in R$ for which $\varrho(\lambda_j)$ is an integer, and for which $\varrho(\lambda)$ is not constant on a neighbourhood of λ_j .*

For $\lambda \leq \lambda_{q-1}$ we have $\varrho(\lambda) = -\frac{1}{2}q$, and for $\lambda_{2j-1} \leq \lambda \leq \lambda_{2j}$ we have $\varrho(\lambda) = -j$.

We can now give a precise description of the twist number $\tau(x)$.

Definition. $\tau(s) = -\varrho(0)$.

Remark. The matrix $M(0)$ is conjugate to the Jacobian of F at P_0 ,

$$dF(P_0) = dF_q(P_{q-1}) \circ \dots \circ dF_1(P_0),$$

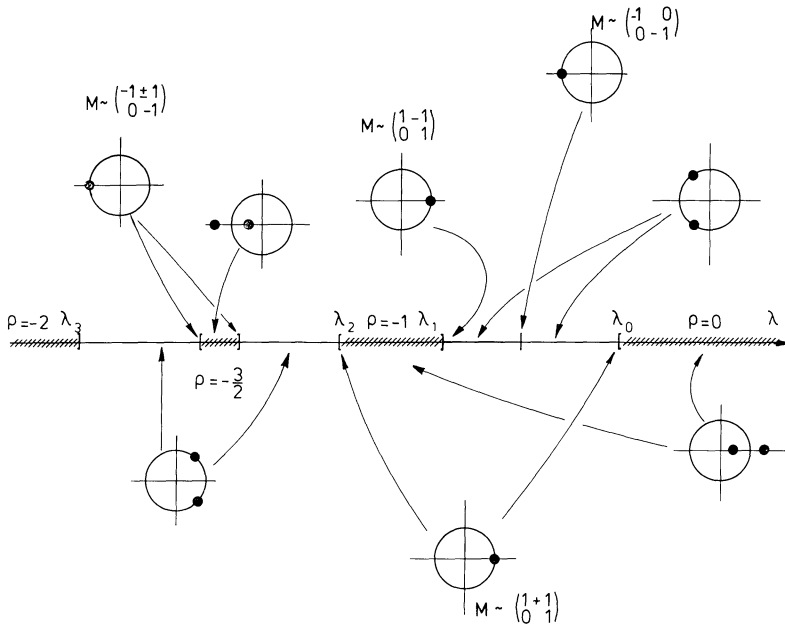


Fig. 1. $\rho(\lambda)$ and the spectrum of $M(\lambda)$ for various values of $\lambda \in \mathbb{R}$

so that the rotation number of $M(0)$ and $dF(P_0)$ coincide. However, as maps of the circle the integer part of their rotation numbers are not defined, and it is this integer part which we shall need. We have defined it in such a way that it fits in with the intuitive description of $\tau(x)$ given at the beginning of this section.

We conclude this section with a discussion of the number of sign-changes of the eigenvectors of L .

Let ξ^i be an eigenvector of L corresponding to the eigenvalue λ_i . Define the number of signchanges of ξ^i to be $I(\xi^i, 0)$ (note that eigenvectors of L always intersect 0 transversally). Although the following lemma can be proven in a more direct way, we shall sketch a proof which shows that the oscillation properties of eigenvectors of L may be regarded as consequences of Smillie's result.

Lemma 3.4. ξ^i has $2[(i + 1)/2]$ signchanges³.

Proof. The unit sphere S^{q-1} in R^q is naturally identified with the set of half lines in R^q with the origin as endpoint. Since the flow on R^q given by e^{tL} is linear, this identification gives us a flow on S^{q-1} . A straightforward calculation shows that this flow is gradientlike with respect to

$$V(\xi) = (\xi, L\xi).$$

It follows from Proposition 2.3 that the following sets are isolating neighbourhoods for the flow e^{tL} on S^{q-1} :

$$N_k = \text{closure of } \{ \xi \text{ in } S^{q-1} : \xi \neq 0 \text{ and } \xi \text{ has } 2k \text{ sign changes} \}.$$

³ For any real number x $[x]$ denotes the integer part of x , i.e. the largest integer which is not larger than x

Let the maximal invariant set in N_k be I_k . Since the isolating neighbourhood N_k is independent of the Jacobi matrix L , the homotopy index of I_k is also independent of L .

For the particular Jacobi matrix one gets by putting $\alpha_i \equiv 1$ and $\beta_i \equiv 0$, one can explicitly determine the I_k and their homotopy indices. It turns out that, if $1 < 2k < q$, I_k is a hyperbolic circle of equilibria for the flow, given by

$$I_k = \left\{ (\zeta_l) = \left(\cos \left(\frac{2\pi kl}{q} + \theta \right) \right) \mid \theta \in \mathbb{R}/2\pi\mathbb{Z} \right\}.$$

The corresponding homotopy index is

$$h(I_k) = [S^1 \times S^{2k-1}/S^1 \times \{p\}],$$

where we use [...] to denote the homotopy type of a pointed topological space, and p is any point in S^{2k-1} .

The other invariant set I_0 (and in case q is even also $I_{q/2}$) consists of a hyperbolic fixed point whose homotopy index is $[S^0]$ (and $[S^{q-1}]$ respectively). The reader who wants to verify this should bear in mind that, since the flow is gradient like, all isolated invariant sets consist of fixed points and orbits connecting them.

Any critical point of V is an eigenvector of L , so we see that all I_k contain at least one eigenvector. Since the flow is equivariant with respect to the transformation $\xi \rightarrow -\xi$, it follows that each I_k contains at least one pair of eigenvectors $\{\xi, -\xi\}$.

We claim that the I_k with $1 < 2k < q$ contain at least two such pairs. Indeed, suppose that one of the I_k contains only one such pair, say $\{\xi^j, -\xi^j\}$. Then I_k contains no orbits of e^{tL} other than the constant orbits ξ^j and $-\xi^j$. The fact that I_k is an isolated invariant set implies that λ_j is a simple eigenvalue of L (otherwise there would be eigenvectors of L , i.e. critical points of V , arbitrarily close to ξ^j). By choosing local coordinates on S^{q-1} near ξ^j one can compute the Hessian of V at ξ^j and conclude that ξ^j and $-\xi^j$ are nondegenerate critical points of V , and that the homotopy index of I_k is $[S^j \vee S^j]$. This contradicts the fact that $h(I_k) = [S^1 \times S^{2k-1}/S^1 \times \{p\}]$.

Smillie's result (Proposition 2.3) implies that $(I_0, I_1, \dots, I_{[q/2]})$ is an admissible ordering of a Morse decomposition of $(S^{q-1}, \exp(tL))$. After some thought one concludes from this that:

$$I_0 = \{\xi^0, -\xi^0\},$$

$$I_k = \{\alpha \xi^{2k-1} + \beta \xi^{2k} : \alpha^2 + \beta^2 = 1\},$$

and if q is even,

$$I_{q/2} = \{\xi^q, -\xi^q\},$$

which proves the lemma.

Lemma 3.5. *Let $1 \leq i \leq j \leq q$ be given. Then any nonzero linear combination of $\xi^i, \xi^{i+1}, \dots, \xi^j$ has at least $2[(i+1)/2]$ and at most $2[(j+1)/2]$ sign changes.*

Proof. Without loss of generality we may assume that there are numbers c_k with c_i and c_j nonzero, such that

$$\eta = \sum_i^j c_k \xi^k$$

is the given linear combination. Define

$$\eta(t) = e^{tL}\eta = \sum_k^j e^{\lambda_k t} c_k \xi^k.$$

Then the number of sign changes of $\eta(t)$ is nonincreasing (by Proposition 2.3). However we have

$$\eta(t)/\|\eta(t)\| \rightarrow \xi^i \quad \text{as } t \rightarrow \infty,$$

and

$$\eta(t)/\|\eta(t)\| \rightarrow \xi^j \quad \text{as } t \rightarrow -\infty,$$

so the number of sign changes of η lies between that of ξ^i and ξ^j .

4. Main Results

Having defined the twist number we can now state our main results. Let $F = F_q \circ F_{q-1} \circ \dots \circ F_1$, W and X be defined as before. Then we have the following theorems.

Theorem 1. *Let x in X be an orbit with $\tau(x) > 0$. Then a and b , defined by*

$$\begin{aligned} a &= \sup \{c < x : \text{grad } W(c) = 0\}, \\ b &= \inf \{d > x : \text{grad } W(d) = 0\}, \end{aligned}$$

are critical points of W with $\tau(a) = \tau(b) = 0$.

It follows from the discussion in Sect. two that two different orbits x, y always intersect transversally. Therefore it makes sense to talk about $I(x, y)$ for any couple of orbits x and y (i.e. critical points of W).

Theorem 2. *Let $a < x < b$ be as in Theorem 1. For any integer k with $0 < k < \tau(x)$ there are at least two orbits y^1, y^2 in X such that*

$$a < y^i < b \quad i = 1, 2,$$

and

$$I(y^i, x) = 2k \quad i = 1, 2.$$

From this theorem we will be able to derive:

Theorem 3. *There is an orbit y (not periodic, i.e. y is not in X) with $a < y < b$, such that y is homoclinic to a , or homoclinic to b , or y is a heteroclinic orbit from a to b . In the last case there is another orbit $a < y' < b$, which is heteroclinic from b to a .*

In general, a and b will be saddle points (for the map F), and the homoclinic orbits will be transversal. It is well known that this implies the existence of a ‘‘Smale horseshoe,’’ and its consequences.

We do not make any genericity assumptions however. The Theorems 1, 2 and 3 are consequences of the variational structure on the set of orbits of F .

Also it should be noted that we do not require the orbits a, x, b and y to be monotone (i.e. $x_i < x_{i+1}$ for all i). Concerning monotone orbits we have the following result:

Theorem 4. *If all the maps F_j are the same and x is a monotone orbit, then*

- (i) $a_i > \max(x_i - 1, x_{i-1}),$
- (ii) $b_i < \min(x_i + 1, x_{i+1}),$
- (iii) $a_i \geq b_{i-1},$
- (iv) *any orbit y with $a < y < b$ is monotone.*

Proof of Theorem 4. The first two inequalities follow from the definition of a and b given in Theorem 1, and the fact that, if $(x_i)_{i \in \mathbb{Z}}$ is an orbit, then so are $(x_i + 1)_{i \in \mathbb{Z}}$ and $(x_{i+1})_{i \in \mathbb{Z}}$.

By Theorem 1 $\tilde{a} = (a_{i+1})_{i \in \mathbb{Z}}$ is an orbit, and by (i) $\tilde{a} > x$. Hence we have $\tilde{a} \geq b$, which implies the third inequality.

Finally, if $a < y < b$, then $y_i < b_i \leq a_{i+1} < y_{i+1}$, which completes the proof.

5. Proof of Theorem 1

Let L be the linear part of the vectorfield $\text{grad } W$ at x . Since $\tau(x) > 0$, it follows from the discussion in Sect. 3 that the first eigenvalue of L, λ_0 , is strictly positive. Indeed, we have $\varrho(0) = -\tau(x)$, so $\varrho(0) < 0$. Therefore the fact that $\varrho(\lambda)$ is a nondecreasing function of λ and that $\varrho(\lambda_0) = 0$ implies that $\lambda_0 > 0$.

This eigenvalue is simple, so there are two unique orbits $y_{\pm}(t)$ of the gradient flow $\{\phi_t : t \geq 0\}$ of W such that

$$\begin{aligned} y_-(t) < x < y_+(t) & \text{ for all } t, \\ y_{\pm}(t) \rightarrow x & \text{ as } t \rightarrow -\infty, \end{aligned} \tag{5.1}$$

and

$$y'_-(t) < 0 < y'_+(t).$$

These orbits lie on the so called “fast unstable manifold” of x , corresponding to the eigenvalue λ_0 of L .

Let $x \pm 1$ denote the points in X defined by $(x \pm 1)_i = x_i \pm 1$. Our assumptions on the twistmaps F_1, \dots, F_q imply that $x + 1$ and $x - 1$ are critical points of W too. Furthermore, for large negative t (5.1) implies that

$$x - 1 < y_-(t) < x < y_+(t) < x + 1. \tag{5.2}$$

Since the gradient flow ϕ_t is order preserving (Lemma 2.2), and $x, x + 1$ and $x - 1$ are fixed points of this flow, (5.2) holds for all t in \mathbb{R} (use the fact that $y_{\pm}(t + s) = \phi_s(y_{\pm}(t))$ holds for all $s \geq 0$).

Define

$$\begin{aligned} c &= \inf\{y_-(t) : -\infty < t < \infty\}, \\ d &= \sup\{y_+(t) : -\infty < t < \infty\}. \end{aligned}$$

Then $y_-(t) \downarrow c$ and $y_+(t) \uparrow d$ as $t \rightarrow \infty$.

Claim. $a = c$ and $b = d$.

Indeed c and d are critical points of W (they are ω -limit sets for the gradient flow of W !), so $c \leq a$ and $b \leq d$ hold.

Now let $e < x$ be any critical point of W . It follows from Lemma 2.2 that $e \ll x$, since $e = \phi_t(t) \ll \phi_t(x) = x$ holds for all $t > 0$. Hence for large negative t we have $e < y_-(t) < x$, and, arguing as before, this must hold for all t .

So we see that $e \leq c$ whenever e is a critical point below x . Combining this with the definition of a (in Theorem 1) we conclude that $a \leq c$, and therefore $a = c$. The same argument will prove that $b = d$.

Now suppose that $\tau(a) > 0$. Then, as we saw before for x , there would be an orbit $z(t)$ of the gradient flow such that $z(t) \downarrow a$ as $t \rightarrow -\infty$, and $z'(t) > 0$. Consider $z^* = \sup_t z(t)$. Again we have that z^* is a critical point of W . It clearly satisfies $a < z^* \leq x$, which implies that $z^* = x$. So we see that $z(t) \uparrow x$ as $t \rightarrow -\infty$. On the other hand, for large negative t , $z(t)$ is close to a and $y_-(t)$ is close to x , so $z(t) < y_-(t)$ must hold for all t . Hence, letting t tend to infinity, we see that $z^* \leq \inf_t y_-(t) = a$, which is a contradiction. It follows that $\tau(a) = 0$. The same argument also shows that $\tau(b) = 0$, which concludes the proof of Theorem 1.

6. Outline of Proof of Theorem 2

Since $a < b$ are fixed points of ϕ_t , and ϕ_t is orderpreserving, it follows that

$$Q = \{y \in X : a < y < b\}$$

is positively invariant under the gradient flow ϕ_t :

$$\phi_t(Q) \subset Q \quad \text{for all } t \geq 0.$$

Let k be a given integer such that $0 < k < \tau(x)$. Then, we must show that there are at least two critical points y^1, y^2 in Q such that $I(y^i, x) = 2k$ holds for $i = 1, 2$. In order to do this we define

$$O_k = \{y \in Q : y \bar{\cap} x \text{ and } I(y, x) = 2k\}$$

and

$$M_k = \{y \in O_k : \overline{\gamma(y)} \subset O_k\},$$

where $\gamma(y)$ denotes the orbit of y under the gradient flow, i.e.

$$\gamma(y) = \{\phi_t(y) : -\infty < t < \infty\}.$$

Clearly, the critical points y^1 and y^2 we are looking for should be contained in M_k . The study of sets like M_k belongs to Conley's generalised Morse theory. We shall assume the reader to be familiar with the Morse-Conley index as it is described in Conley [4], or in Conley and Zehnder [5, 6].

Theorem 2 will be a direct consequence of the next result.

Theorem 5. *If $0 < k < \tau(x)$, then M_k is an isolated invariant set for the gradient flow of W . Its homotopy index is*

$$h(M_k) = [(S^1 \times S^{2k-1}) / (S^1 \times \{p\})],$$

where p is a point in S^{2k-1} .

Proof that Theorem 5 implies Theorem 2. Since $h(M_k)$ is nontrivial, M_k cannot be empty. It therefore contains a complete orbit of ϕ_t and its closure. Since ϕ_t is a gradient flow M_k must contain at least one critical point of W .

Suppose that M_k contains only one critical point, say y . Then M_k must be $\{y\}$, because any nontrivial orbit of ϕ_t in M_k would have to be homoclinic to y which is impossible for gradient flows.

Let L be the linear part of grad W at y . Then $\lambda = 0$ is at most a double eigenvalue of L (recall that $L\xi = 0$ is a second order difference equation). In other words the rank of the Hessian of W at y is at least $q - 2$. A result of Dancer [7] then tells us that, if $b_j (j > 0)$ are the Betti numbers of $h(M_k) = h(\{y\})$, at most one of the b_j is nonzero.

It is an easy exercise in algebraic topology however to compute the Betti numbers of $h(M_k)$. One finds that $b_j = 1$ if $j = 2k - 1$ or $j = 2k$, whereas the other b_j are zero. So we see that, if there is only one critical point in O_k , $h(M_k)$ can never be the homotopy type of $S^1 \times S^{2k-1} / S^1 \times \{p\}$. Hence O_k contains at least two critical points.

It remains to prove Theorem 5. This we do in three steps. First we prove that M_k is an isolated invariant set, and we construct some isolating neighbourhoods for M_k (note that O_k is not compact, and therefore can never be an isolating neighbourhood of anything!).

Then we show that for all different choices of the twist maps F_1, \dots, F_q , the integer p , and the triple $a < x < b$, the resulting M_k 's are always related by continuation and therefore have the same Conley index (always assuming that $0 < k < \tau(x)$).

Finally we compute $h(M_k)$ for a specific example. In this example we let F_1, \dots, F_q be the same integrable twist map. It is at this point that we see that our theorems also apply to Problem (I) of the introduction.

7. Proof that M_k is an Isolated Invariant Set

Lemma 7.1. *The following statements are equivalent:*

- (a) M_k is compact.
- (b) M_k is an isolated invariant set.
- (c) For all y on ∂O_k there is a neighbourhood U_y of y such that U_y and M_k are disjoint.

Proof. “(a) implies (c)” is obvious from the fact that M_k and ∂O_k are disjoint.

“(c) implies (b)”. Since the boundary of O_k is compact a finite number of neighbourhoods U_y cover ∂O_k . Let $K \subset O_k$ be the complement of these U_y 's. We may assume that M_k is contained in the interior of K . By the choice of the

neighbourhoods U_y , one easily verifies that M_k is the maximal invariant set in K . Hence K is an isolating neighbourhood of the isolated invariant set M_k .

“(b) implies (a)”. By definition isolated invariant sets are compact, so M_k is compact.

Lemma 7.2. *If $0 < k < \tau(x)$, then M_k is an isolated invariant set.*

Proof. We verify condition (c) of the foregoing lemma.

The boundary of O_k contains three kinds of points. First there are those y in ∂O_k for which $y \nabla x$, $a < y < b$, but not $a \ll y$ or not $b \gg y$ holds. It follows from Lemma 2.2 that the orbit of y cannot lie in \bar{O}_k . By continuity of the flow there is a neighbourhood of y , U say, such that U and M_k are disjoint.

The second kind of points are those y on the boundary of O_k for which $a \ll y \ll b$, $y \neq x$ but $y \nabla x$ does not hold. Proposition 2.3 tells us that, either for all $t > 0$ or for all $t < 0$, $\phi_t(y)$ does not lie in \bar{O}_k . Again, by continuity, some neighbourhood of y is then disjoint from M_k .

We are now left with one other boundary point of O_k , namely x itself. This point deserves its own lemma.

Lemma 7.2. *If $0 < k < \tau(x)$, then there is a neighbourhood U of x such that U and M_k are disjoint.*

Proof. Identify X and R^q via

$$y \in X \rightarrow (y_1 - x_1, y_2 - x_2, \dots, y_q - x_q) \in R^q,$$

so that x becomes the origin.

The linear part of $\text{grad } W$ at x , L , can be written as $L = L_+ + L_-$, where L_+ and L_- are self adjoint, $L_+ L_- = L_- L_+ = 0$ and

$$(u, L_+ u) > 0, \quad (u, L_- u) \leq 0$$

holds for all nonzero u in R^q .

Let $\{v_0, \dots, v_{q-1}\}$ and $\{\lambda_0 > \lambda_1 \geq \lambda_1, \dots, \lambda_{q-1}\}$ be the eigenvectors and values of L . The null space of L_+ is spanned by $\{v_m, v_{m+1}, \dots, v_{q-1}\}$, where $m > 2k$. Indeed, $\tau(x) > k$ implies $g(0) < -k$ (notation as in Sect. two), so L must have at least $2k + 1$ positive eigenvalues.

Hence $u \neq 0$ and $L_+ u = 0$ imply that u has at least $2k + 2$ sign changes (by Lemma 3.5) and therefore does not lie in \bar{O}_k . Thus there is a constant $K > 0$ such that

$$(u, L_+ u) \geq K \|u\|^2$$

holds for all u in \bar{O}_k ($\|u\|$ is the Euclidean norm of u). This also implies that $\|L_+ u\| \geq K \|u\|$ for all u in \bar{O}_k .

Consider the function $G(u) = \frac{1}{2}(u, L_+ u)$. Close to $u = 0$ the gradient flow is given by

$$u'(t) = Lu(t) + o(\|u(t)\|),$$

and so we have

$$\frac{d}{dt} G(u) = (L_+ u, L_+ u) + o(\|u\|^2) \geq (K^2 + o(1)) \|u\|^2 > 0,$$

whenever u lies in O_k .

Now let $\varepsilon > 0$ be so small that $G_t > 0$ whenever $G(u) \leq \varepsilon$ and $u \in O_k$, and define $U = \{u : G(u) < \varepsilon\}$. If $u(t)$ is an orbit of the gradient flow which lies in O_k , and for some t_0 in R lies in U , then for all $t \leq t_0$ we have $u(t) \in U$ and $\frac{d}{dt} G(u(t)) > 0$. It follows that $u(t) \rightarrow 0$ as $t \rightarrow -\infty$. Hence the closure of this orbit does not lie in O_k which proves that U and M_k are disjoint.

8. Continuation of $h(M_k)$

Identify X with R^q as in the proof of Lemma 7.3. In these coordinates we have

$$W(u) = \sum_1^q h_j(x_{j-1} + u_{j-1}, x_j + u_j).$$

This sum represents the Lagrangian W we would get if we had taken $p = 0$, and had studied the maps G_j defined by

$$G_j(u, v) = (f_j(x_{j-1} + u) - x_j, g_j(u, v)),$$

where $F_j = (f_j, g_j)$.

Without loss of generality we may therefore assume that $p = 0$ and $x = 0$.

Our main tool in computing $h(M_k)$ will be its invariance under continuation.

Let $F_1(\lambda), \dots, F_q(\lambda)$ be a continuous family of twist maps such that $a(\lambda) < 0 < b(\lambda)$ are critical points of the corresponding Lagrangians $W(\lambda)$. Also assume that $a(\lambda)$ and $b(\lambda)$ depend continuously on λ .

Lemma 8.1. *If for all λ in (λ_1, λ_2) we have $0 < k < \tau(0)$, then the isolated invariant sets $M_k(\lambda)$ are related by continuation, and $h(M_k(\lambda))$ is independent of λ .*

Proof. Let λ_0 in (λ_1, λ_2) be given. In the proof of Lemma 7.1 we constructed an isolating neighbourhood K for $M_k(\lambda)$. Using the postulated continuity of $a(\lambda)$ and $b(\lambda)$ one easily proves that K is an isolating neighbourhood of $M_k(\lambda)$ for all λ close to λ_0 . Since the interval (λ_1, λ_2) is connected, this proves the lemma.

We shall prove that all the isolated invariant sets M_k corresponding to different sets of maps F_1, \dots, F_q are related by continuation. Our proof is divided into two steps. First we show that if we keep $a < 0 < b$ fixed any M_k can be related to one particular $M_k(a, b)$. Then we show that these special $M_k(a, b)$'s are all related to each other.

Construction of $M_k(a, b)$. Let $a < 0 < b$ be given and consider

$$U(y) = \sum_1^q \left\{ -\frac{1}{2}(y_i - y_{i-1})^2 + V_i(y_i) \right\}, \tag{8.1}$$

where the functions V_i in $C^2(R)$ have the properties

$$\begin{aligned} &V_i'(0) = 0, \\ (1) \quad &V_i'(a_i) + a_{i-1} - 2a_i + a_{i+1} = 0, \\ &V_i'(b_i) + b_{i-1} - 2b_i + b_{i+1} = 0, \end{aligned}$$

for $i = 1, \dots, q$, and

(2) $V_i''(0)$, $-V_i'''(a_i)$ and $-V_i'''(b_i)$ are chosen so large that 0, a and b are nondegenerate minima respectively maxima of U . Furthermore the matrix $U''(0) - W''(0)$ should be strictly positive definite.

With some effort one can even construct polynomials which have these properties.

Define $W_\lambda = (1 - \lambda)W + \lambda U = W + \lambda(U - W)$, and $a(\lambda) = a$, $b(\lambda) = b$ for $0 \leq \lambda \leq 1$.

Lemma 8.2. *The family $\{(a(\lambda), b(\lambda), W_\lambda) : 0 \leq \lambda \leq 1\}$ satisfies the hypothesis of Lemma 8.1.*

Proof. We only have to check that for $0 \leq \lambda \leq 1$ the twist number $\tau(0)$ is larger than k .

By assumption, this certainly holds if $\lambda = 0$, for then $W_\lambda = W$. The construction of the twist number given in Sect. three shows that $\tau(0)$ increases if the Hessian of W at $x = 0$ increases. Since we have chosen $U'' - W''$ positive definite, the twist number is an increasing function of λ . Hence $\tau(0) > k$ for all λ in $[0, 1]$.

Lemma 8.2 shows that, if $\tau(0) > k$, $h(M_k)$ only depends on a and b , and in fact that $h(M_k) = h(M_k(a, b))$, where we define $M_k(a, b)$ to be the M_k corresponding to the Lagrangian U introduced above.

Continuation of the $M_k(a, b)$. Let U_0 and U_1 be two functions given by (8.1) with different a and b : $a(0) < 0 < b(0)$, and $a(1) < 0 < b(1)$, and different V_i : $V_i(0, y)$ and $V_i(1, y)$ (which satisfy the conditions (1) and (2) we introduced in the construction of the $M_k(a, b)$).

Then again we have two isolated invariant sets, $M_k(0)$ and $M_k(1)$ (assuming $\tau(0) > k$ for both cases of course).

Lemma 8.3. *$M_k(0)$ and $M_k(1)$ are related by continuation.*

Proof. Without loss of generality one may assume that $U_1''(0) - U_0''(0)$ is positive definite. Define $a(\lambda)$, $b(\lambda)$ and $V_i(\lambda, y)$ by linear interpolation:

$$\begin{aligned} a(\lambda) &= (1 - \lambda)a(0) + \lambda a(1), \\ b(\lambda) &= (1 - \lambda)b(0) + \lambda b(1), \\ V_i(\lambda, y) &= (1 - \lambda)V_i(0, y) + \lambda V_i(1, y). \end{aligned}$$

Using smooth “bump-functions” (a partition of unity) one can now redefine the $V_i(\lambda, y)$ in a neighbourhood of $a_i(\lambda)$ and $b_i(\lambda)$ in such a way that $a(\lambda)$ and $b(\lambda)$ become nondegenerate local maxima of $U_\lambda = (1 - \lambda)U_0 + \lambda U_1$, and without changing U_λ near $x = 0$ ($0 \leq \lambda \leq 1$), or near $a(\lambda)$ and $b(\lambda)$ ($\lambda = 0, 1$).

The family $\{(a(\lambda), b(\lambda), U_\lambda) : 0 \leq \lambda \leq 1\}$ satisfies the hypothesis of Lemma 8.1, so $M_k(0)$ and $M_k(1)$ are related by continuation.

Corollary. *There is a universal homotopy type, h_k , such that whenever $\tau(x) > k$ one has $h(M_k) = h_k$, where M_k is defined as in Sects. 6 and 7.*

All that remains to be done to compute h_k is to find a specific example for which $h(M_k)$ can be computed explicitly.

9. An Explicit Example

Let F_1 be the time one map of the following system of ordinary differential equations:

$$\begin{aligned} \dot{x} &= y, \\ \dot{y} &= -(\omega^2/\pi)\sin \pi x. \end{aligned} \tag{9.1}$$

This system is Hamiltonian with

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

so the map F_1 is integrable, and area preserving.

We shall put $F_2 = F_3 = \dots = F_q = F_1$.
Define the function $h(x_0, x_1)$ on R^2 by

$$h(x_0, x_1) = \sup \left\{ \int_0^1 \left\{ \frac{1}{2}u'(s)^2 + (\omega/\pi)^2 \cos \pi u(s) \right\} ds \right\},$$

where the supremum is taken over all u in the Sobolev space $H^1(0, 1)$ which satisfy the boundary conditions $u(0) = x_0$ and $u(1) = x_1$.

Lemma 9.1. *If $0 < \omega < \pi$, then h is an analytic function, and it is the generating function of the map F_1 . The map F_1 is an analytic area-preserving twist diffeomorphism.*

Proof. Let $F_1(x, y) = (f(x, y), g(x, y))$. Then

$$\frac{\partial f}{\partial y}(x_0, y_0) = v(1),$$

where v is the solution in $C^2([0, 1])$ of

$$\begin{aligned} v''(t) + \omega^2 \cos(\pi x(t))v(t) &= 0, & 0 < t < 1, \\ v(0) = 0, \quad v'(0) &= 1, \end{aligned} \tag{9.2}$$

and x is the C^2 solution of

$$\begin{aligned} x''(t) + (\omega^2/\pi)\sin \pi x(t) &= 0, & 0 < t < 1, \\ x(0) = x_0, \quad x'(0) &= y_0. \end{aligned} \tag{9.3}$$

The inequality $-\pi^2 < \omega^2 \cos \pi x(t) < \pi^2$ on $[0, 1]$ implies that $v(t) > 0$ for all t in $(0, 1]$. Hence $\frac{\partial f}{\partial y} > 0$, and F_1 is a twist map. The same inequality implies that

$$\begin{aligned} x''(t) + (\omega^2/\pi)\sin \pi x(t) &= 0, & 0 < t < 1, \\ x(0) = x_0, \quad x(1) &= x_1, \end{aligned} \tag{9.4}$$

has a unique solution which depends analytically on x_0 and x_1 . Indeed, the inequality states that the first t value conjugate to $t = 0$ lies outside of the interval $[0, 1]$, independent of the initial data $x(0)$ and $x'(0)$.

The solution $x(x_0, x_1; t)$ of (9.4) maximizes the ‘‘action integral’’

$$-\int_0^1 \left\{ \frac{1}{2}u'(s)^2 + (\omega/\pi)^2 \cos \pi u(s) \right\} ds,$$

so $h(x_0, x_1)$ is just this integral with $u(t) = x(x_0, x_1; t)$ inserted. Hence, h is also analytic. It is well known from classical mechanics that $h(x_0, x_1)$ is the generating function for the time one map F_1 (see Arnold [1]).

Having specified the maps F_1, \dots, F_q , and $F = (F_1)^q$, we define $a < 0 < b$ by

$$a_i = -1, \quad b_i = +1, \quad i = 1, \dots, q.$$

Let $k > 0$ be an integer. Note that the twist number of an orbit in any $X_{p,q}$ is never larger than $\frac{1}{2}q$ (compare with Lemma 3.3). Therefore we only have to compute h_k when $0 < k < \frac{1}{2}q$.

To compute h_k we must have $0 < k < \tau$ in our example.

Lemma 9.2. *For $0 < \omega < \pi$ the twist number of the zero solution of our example is given by $\tau = q\omega/2\pi$.*

Proof. In Sect. three a straightforward method of calculating τ is given. We leave it to the reader to perform these calculations. It should be noted however that $dF_1(0, 0)$ is conjugate to a rotation over an angle ω , so $(dF_1)^q = dF(0)$ is conjugate to a rotation over $q\omega$. Hence its twist number should be $q\omega/2\pi$, according to the intuitive description given in Sect. three.

So we see that, if we choose ω close enough to π , then we will certainly have $0 < k < \tau$.

To find M_k we have to locate the critical points of

$$W = \sum_1^q h(x_{j-1}, x_j)$$

in O_k , and the orbits of the gradient flow that connect these critical points. A phase-plane analysis shows that there is a unique solution $X(t)$ of

$$\begin{aligned} X''(t) + \omega^2/\pi \sin \pi X(t) &= 0, & -\infty < t < \infty, \\ X(t+q) &= X(t), & X(-t) &= X(t), \end{aligned} \tag{9.5}$$

which has $2k$ sign changes in the interval $0 < t < q$ (the assumption $0 < k < \tau$ is essential here!). This provides us with a one parameter family, $x(s)$, of critical points of W , namely

$$x_j(s) = X(j+s), \quad j = 1, 2, 3, \dots, q, \quad s \text{ in } \mathbb{R}.$$

Lemma 9.3. $M_k = \{x(s) : s \in \mathbb{R}\}$ is diffeomorphic to a circle.

Proof. The system (9.1) has a one parameter family of periodic solutions whose period is a monotone increasing function of the amplitude.⁴ Using this fact one proves that the $x(s)$ are the only critical points of W in O_k . Since W is constant on $\{x(s) : -\infty < s < \infty\}$, there are no orbits between the $x(s)$, so that $M_k = \{x(s) : -\infty < s < \infty\}$.

The fact that $x(s)$ is periodic in s ($x(s+q) = x(s)$, so the period divides q) shows that M_k is the image of a circle. It follows from our assumption $0 < \omega < \pi$ that, if $X'(s) = 0$ for some s , then $X'(s+1) \neq 0$. Hence M_k is an immersed circle. We leave it

⁴ The period can be explicitly computed using elliptic integrals. That the period is a monotone function of the amplitude is a consequence of the fact that $(\sin x)/x$ is a decreasing function of x ($x > 0$)

to the reader to verify that M_k is an embedded circle (one has to check that $x(s_0 + i_0) = x(s_0)$ for some s_0 and some integer i_0 implies that $x(s + i_0) = x(s)$ holds for all s).

The final step in computing $h_k = h(M_k)$ consists of studying the gradient flow near M_k .

Lemma 9.4. *M_k is a hyperbolic invariant set of the gradient flow. Its unstable manifold has dimension $2k$.*

Proof. From $\text{grad } W(x(s)) = 0$ for all s one derives

$$W''(x(s)) \cdot x'(s) = 0 \quad \text{for all } s.$$

To prove hyperbolicity we have to prove that $x'(s)$ spans the null space of $W''(x(s))$.

Apply the construction (and notation) of Sect. three to this situation. Using the fact that the periods of the solutions of

$$X''(t) + \omega^2/\pi \sin \pi X(t) = 0$$

increase with amplitude one sees that the matrix $M(0)$ is conjugate to

$$\begin{pmatrix} 1 & -\alpha \\ 0 & 1 \end{pmatrix}$$

with $\alpha > 0$. Indeed, this is just the matrix of $d(F_1^{-q})$ at $(X(s), X'(s))$ with respect to the (positively oriented) basis

$$\{(X'(s), X''(s)), (-X''(s), X'(s))\}.$$

The null space of $W''(x(s))$ must therefore be one dimensional which proves the hyperbolicity of M_k .

From $x'_j(s) = X(j+s)$ one sees that $x'(s)$ has $2k$ sign changes. Thus $W''(x(s))$ has $2k$ or $2k - 1$ positive eigenvalues. It follows from the discussion in Sect. three (after Lemma 3.2) that $\alpha > 0$ implies that $W''(x(s))$ has $2k - 1$ positive eigenvalues, and the unstable manifold of M_k is $2k$ -dimensional.

Corollary. *$h(M_k) = h_k$ is the homotopy type of $S^1 \times S^{2k-1}/S^1 \times \{p\}$.*

Proof. By the unstable manifold theorem the flow near M_k is the product of that near a hyperbolic fixed point with index $2k - 1$, and the trivial flow on a circle. Now use the product rule to obtain $h(M_k)$ (see Conley [4]).

This corollary finishes the proof of Theorem 5.

10. Proof of Theorem 3

Let $a < x < b$ be defined as in Sect. 4, and put

$$Y = \{(x_j)_{j \in \mathbb{Z}} : \sup |x_j - jp/q| < \infty\}.$$

This set is a Banach manifold diffeomorphic to l_∞ , the space of bounded sequences with the supremum norm. Furthermore it is ordered in the same way as $X_{p,q}$.

The function W cannot be defined on Y , however its gradient flow can. It is simply given by the system of ordinary differential equations

$$\dot{x}_j = \partial_2 h_j(x_{j-1}, x_j) + \partial_1 h_{j+1}(x_j, x_{j+1}). \tag{10.1}$$

Thus we still have a C^1 flow, ϕ_t ($-\infty < t < \infty$), which is orderpreserving for $t \geq 0$. Indeed, the variational equation corresponding to (10.1) is of the form $\delta\dot{x}(t) = A(t)\delta x(t)$. Here $A(t)$ is an infinite matrix whose off diagonal elements are nonnegative, and whose sub and super diagonal elements are strictly positive. Therefore the derivative $D\phi(t)$ is a strictly positive operator.

The flow $\phi(t)$ extends the gradient flow of W on $X_{p,q}$ which we have used up to now.

On Y we also define the left translation $T:(Ty)_i = y_{i+q} - p$.

Lemma 10.1 *There is no orbit y in Y of the maps F_i lying strictly between x and b .*

Proof. Let y be such an orbit. Then we have $x \ll y \ll b$.

Using the fact that $\tau(x) > 0$, an analysis of the map F around the periodic point corresponding to x will reveal that, if for some integer j , y_j is close enough to x_j , the orbit y will oscillate about x (near j). This contradicts the fact that $y > x$ and therefore we have $\inf(y_i - x_i) > 0$.

Now let $y_+(t)$ be the orbit of ϕ_t we had constructed in the proof of Theorem 1 (Sect. 5). Then again we have $y_+(t) < y$ for large negative t . Since the flow ϕ_t is order preserving we get $\sup y_+(t) \leq y$, i.e. $b \leq y$. This contradicts $y < b$ so the orbit y cannot exist.

Lemma 10.2. *If y in Y is an orbit of the maps F_i satisfying $x_i < y_i \leq b_i$ for all $i \geq 0$, then $\lim_{i \rightarrow \infty} y_i - b_i = 0$.*

Proof. Consider the translates $y^n = T^n(y)$. Then we have $x_i < y_i^n \leq b_i$ for all $i > -nq - 1$.

Each y^n is an orbit of the maps F_i which satisfies $x_i < y_i^n \leq b_i$ for all $i \geq -nq$. Extract a subsequence y^{n_k} of the y^n such that the $(y^{n_k})_i$ converge as $k \rightarrow \infty$ (not necessarily uniformly in i). This can be done since for each i , $(y^n)_i$ is a bounded sequence. Denote the limit by z . Then z is an orbit of the maps lying between x and b . It follows from the previous lemma that $z = x$ or $z = b$. If $z = x$, then the sequence y_i would contain elements which come arbitrarily close to x_i . As we remarked in the proof of the previous lemma this is impossible, since y_i would have to oscillate around x_i in some interval $i_0 < i < i_1$, so $z = b$.

We have shown that any convergent subsequence of the translates y^n converges to b . This implies that the whole sequence y^n converges to b , i.e. that $y_i - b_i \rightarrow 0$ as $i \rightarrow \infty$.

Lemma 10.3. *There is an orbit y in Y such that $a < y < b$ and $y - x$ has one or two sign changes*

Proof. By Theorem 2 we have orbits y^n in $X_{np, nq}$ such that $a < y^n < b$ and $(y^n - x)_i$ has two sign changes for $0 < i < nq + 1$. We may assume that one of these sign changes occurs in the interval $0 < i < q + 1$.

Now take a convergent subsequence of the y^n . Its limit, which we call z , is an orbit of the maps. There are two possibilities:

(1) $z = x$.

(2) $a < z < x$ is an orbit with at least one sign change (in the interval $0 < i < q + 1$), and at most two sign changes.

We can exclude the first alternative for the following reasons. If $z = x$ were to hold, then for large n the orbit $(y^n)_i$ would come arbitrarily close to x_i on any

bounded interval of the form $-N < i < N$. The twist number of the orbit x , $\tau(x)$, is positive however, so that for large n the orbit y^n intersects x approximately $2\tau(x)N/q$ times on the interval $-N < i < N$. This contradicts the fact that y^n intersects x exactly twice on any interval of length nq .

Thus we are left with the second alternative, which proves the lemma.

These lemmas enable us to prove Theorem 3. Let y be the orbit constructed in the last lemma. If y and x intersect twice, then it follows from Lemma 10.2 that y is homoclinic to a or to b .

Suppose that y and x only intersect once, and that y is a heteroclinic orbit from a to b . In the proof of Lemma 10.3 we had a sequence of orbits, y^n , intersecting x twice in the interval $0 < i < nq + 1$. By fixing one of these intersections in the fixed interval $0 < i < q + 1$ and taking a limit, we constructed y . Since y is heteroclinic from a to b we have obviously fixed the left intersection of the y^n with x . By fixing the other intersection and repeating the argument in Lemma 10.3, we then get a heteroclinic orbit from b to a .

Acknowledgement. I would like to thank L.A. Peletier for critically reading the original manuscript and pointing out to me some of the vague and inaccurate "arguments" (all of them I hope) it contained.

References

1. Arnold, V.I.: Mathematical methods of classical mechanics. Berlin, Heidelberg, New York: Springer 1978
2. Aubry, S., le Daeron, P.Y.: The discrete Frenkel-Kontorova model and its extensions. *Physica* **8 D**, 381–422 (1983)
3. Coddington, E.A., Levinson, N.: Theory of ordinary differential equations. New York: McGraw-Hill 1955
4. Conley, C.: Isolated invariant sets and the Morse index. C.B.M.S. Reg. Conf. Ser. Math. **38**, Published by the AMS (1978)
5. Conley, C., Zehnder, E.: The Birkhoff-Lewis fixed point theorem and a conjecture of V.I. Arnold. *Invent. Math.* **73**, 33–49 (1983)
6. Conley, C., Zehnder, E.: Morse type index theory for flows and periodic solutions for Hamiltonian equations. *Commun. Pure Appl. Math.* **37**, 207–253 (1984)
7. Dancer, E.N.: Degenerate critical points, homotopy indices and Morse inequalities. *J. Reine Angew. Math.* **350**, 1–22 (1984)
8. Hirsch, M.: Systems of differential equations which are competitive or cooperative, I: Limit sets. *SIAM J. Math. Anal.* **13**, 167–179 (1982)
9. Matano, H.: Nonincrease of the lap-number of a solution for a one dimensional semilinear parabolic equation. *J. Fac. Sci. Univ. Tokyo* **29**, 401–441 (1982)
10. Mather, J.N.: Existence of quasiperiodic orbits for twist homeomorphisms of the annulus. *Topology* **21**, 457–467 (1982)
11. Mather, J.N.: Amount of rotation about a point and the Morse index. *Commun. Math. Phys.* **94**, 141–153 (1984)
12. Moerbeke, P. van: The spectrum of Jacobi matrices. *Invent. Math.* **37**, 45–81 (1976)
13. Smillie, J.: Competitive and cooperative tridiagonal systems of differential equations. *SIAM J. Math. Anal.* **15**, 531–534 (1984)

Communicated by J. N. Mather

Received February 11, 1985; in revised form November 3, 1987