Invent. math. 88, 225–241 (1987)

Inventiones mathematicae

Birkhoff periodic orbits for small perturbations of completely integrable Hamiltonian systems with convex Hamiltonians

David Bernstein[1] and Anatole Katok[2]*

[1] Department of Mathematics, Northeastern University, Boston, MA 02115, USA
[2] Department of Mathematics, California Institute of Technology, Pasadena, CA 91125, USA

Dedicated to the memory of C.C. Conley

1. Introduction

By far the most remarkable property of small perturbations of completely integrable Hamiltonian systems is the preservation of invariant tori corresponding to irrational frequency vectors which are not too well approximable by rationals. This fact and its various ramifications was discovered by A.N. Kolmogorov, V.I. Arnold and J. Moser and became commonly known as KAM theory. This theory leaves open the basic question, namely, what happens to the rest of the invariant tori of the unperturbed completely integrable system? It is relatively easy to show that generically most of those tori, both rational and irrational, disappear. Thus, a more precise formulation of the above question should be like this. Are there special sufficiently simple motions in the perturbed system which are similar to the periodic and quasi-periodic motions on the destroyed tori, and therefore can be viewed as "traces" or "ghosts" of those tori?

The study of small perturbations of non-degenerate completely integrable systems with two degrees of freedom can be reduced to the consideration of area-preserving twist maps of the annulus or the cylinder [15]. For such maps S. Aubry [4, 5] and J. Mather [17] (cf. also [13]) established the existence of special invariant sets which are projected injectively to the circle and carry motions with any given admissible rotation number. Furthermore, the map preserves the cyclic order of points on any of those invariant sets. For any irrational rotation number such set is either an invariant circle, or, if the invariant circle for the given rotation number does not exist, it is a Cantor set with the motion described by A. Denjoy [11]. Moreover, the Cantor sets are always accompanied by order-preserving orbits doubly asymptotic to them [6, 14, 18]. As might be expected, for any rational rotation number one has a collection of at least two order-preserving (Birkhoff) periodic orbits together with

* Partially supported by NSF Grant DMS 8414400

homo- or heteroclinic orbits [14, 18]. All those objects possess certain continuity properties with respect to the rotation number [18].

Translating the above-mentioned results to the case of small perturbations of non-degenerate completely integrable Hamiltonian systems one obtains a more than satisfactory solution to the problem of vanishing tori for systems with two degrees of freedom. Namely, there are always traces of those tori present in the form of either Denjoy type minimal sets together with double asymptotic orbits (for any admissible irrational rotation number), or of at least two Birkhoff periodic orbits accompanied by homo- or heteroclinic orbits (for any admissible rational rotation number). Furthermore, the Denjoy minimal sets appear as limits of Birkhoff periodic orbits.

The Aubry-Mather approach is based on two key ingredients: the variational principle for finding desired motions and the regularity of the projection of any order-preserving orbit to the circle. Let us point out that the variational principle for finding order-preserving periodic orbits for twist maps can be substituted by certain topological arguments [7, 12]. The regularity of the projection allows us to take limits with respect to rotation number, thus, producing invariant circles or Denjoy type Cantor sets without any use of invariant measures or variational methods.

On the other hand, solutions representing global minima in various variational problems, associated to a twist map and posed without assuming preservation of order, turn out to be order preserving [6].

The earliest result concerning the preservation of some periodic orbits for Hamiltonian systems with more than two degrees of freedom is the Birkhoff-Lewis theorem [8, 9], (cf. also [2]) whose accurate proof was given by Moser [19]. Applying the method from [9] and [19] to our situation, one can find periodic orbits for a perturbation of a completely integrable system in a neighborhood of an invariant torus of the unperturbed system filled by periodic orbits of period, say, T, if the size of the perturbation is so small, that the orbits of the perturbed system stay sufficiently close to the original orbits during time T. We do not see how that method can be made to work uniformly in T for perturbations of fixed size.

Conley and Zehnder [10] discovered a remarkable global method for finding periodic orbits for sympletic maps and Hamiltonian systems. In this paper we use a version of their main trick. The result of the [10] most relevant for our discussion is Theorem 3 which represents a global generalization of the Birkhoff-Lewis theorem. This theorem depends only on a sort of a boundary condition rather than on closeness of perturbed and unperturbed systems. However, it can not be directly applied to the finding of very long orbits for a perturbed system which stay near the original torus.

Any attempt to carry out the Aubry-Mather approach to the case of more than two degrees of freedom faces the obvious problem that the arguments based on the preservation of order are no longer available. On the other hand, the variational arguments can be used, at least under some extra assumptions on the unperturbed system. In the present paper we make the first modest, but we believe non-trivial, step in that direction. Our main result for systems with n degrees of freedom is the existence of at least n distinct periodic orbits with any admissible rational frequency vector near the corresponding torus of the unperturbed system under the extra assumption that the Hamiltonian of the unperturbed system has convex energy surfaces in the action-angle variables.

The reduction to the discrete-time case described in Sect. 7 allows obtaining this result from the corresponding result for sympletic maps (Theorem A, Sect. 1). This latter theorem follows immediately from Proposition 2 (Sect. 4) which provides certain estimates for specific critical points of the Lagrangian introduced in Sect. 2, and from Proposition 4 (Sect. 5) which establishes the existence of required critical points.

In Sect. 6 we show that at least one of the orbits described in Theorem A satisfies certain regularity conditions (Theorem B) which allows one to take limits in frequency vectors (Theorem C). Unfortunately, the structure of the limit objects corresponding to irrational frequencies is not completely clear at that stage.

One of the crucial ingredients of our method is a global topological trick which is very similar to the one used by C.C. Conley and E. Zehnder in [10]. This trick allows us to show the existence of solutions of the variational problem with sufficiently low values of our Lagrangian. Then the estimates of the Lagrangian show that these solutions correspond to periodic orbits with desired properties.

We work with discrete time symplectic maps and in the last section show how the results about continuous time Hamiltonian systems are derived from those for symplectic maps.

2. Preliminaries and formulation of main result

Let us consider the space $M=\mathbb{T}^n\times\mathbb{R}^n=\{(\varphi_1\dots\varphi_n, r_1\dots r_n),\ \varphi_i\in\mathbb{R}/\mathbb{Z},\ r_i\in\mathbb{R}\}$ with the natural sympletic 2-form

$$\Omega=\sum_{i=1}^{n} d\varphi_i\wedge dr_i$$

and let $f_0\colon\mathbb{T}^n\times U\to\mathbb{T}^n\times U$ be an integrable symplectic diffeomorphism, i.e. an Ω-preserving diffeomorphism of the form

$$f_0(\varphi, r)=(\varphi+\alpha(r), r), \qquad \varphi\in\mathbb{T}^n,\ r\in U.$$

Here $U\subset\mathbb{R}^n$ is a set diffeomorphic to an open n-disc.

Let, furthermore $F_0\colon\mathbb{R}^n\times U\to\mathbb{R}^n\times U$ be a lift of f_0 to the universal cover, so that for $x\in\mathbb{R}^n$, $r\in U$

$$F_0(x, r)=(x+a(r), r). \tag{1}$$

Throughout this paper we will assume the following non-degeneracy condition

(i) *a: $U\to\mathbb{R}^n$ is a regular injective map.*

Then the map F_0 can be represented via a generating function $H_0(x, x')$ so that if $F_0(x, r)=(x', r')$ then

$$r=\frac{\partial H_0}{\partial x}, \qquad r'=-\frac{\partial H_0}{\partial x'} \tag{2}$$

It follows immediately from (1) and (2) that the function H_0 actually depends only on the difference $x'-x$, $H_0(x,x')=h(x'-x)$. Let $b: a(U)\to U$ be the map inverse to a. Then from (2)

$$dh(\delta)=b(\delta)\,d\delta.$$

Condition (i) is a discrete-time equivalent of the standard non-degeneracy condition in KAM theory, cf. e.g. [1] or [3], Apprendix 8. Next we will introduce an additional non-trivial restriction which will also be assumed in all subsequent considerations, usually without any separate mentioning.

(ii) *h is a strictly convex function on $a(U)$, i.e. the Hessian of h at every point $\delta\in a(U)$ is a positive definite quadratic form.*

In Sect. 7 we will interpret conditions (i) and (ii) in terms of continuous time Hamiltonian systems.

Suppose that a lift F of f can be represented by a generating function $H(x,x')$ so that $F(x,r)=(x',r')$ if and only if

$$\frac{\partial H}{\partial x}(x,x')=r \qquad \frac{\partial H}{\partial x'}(x,x')=-r'.$$

Suppose in addition that the perturbation f preserves the r-component of the center of masses on each torus $\mathbb{T}^n\times\{r_0\}$ for $r_0\in U$, or, equivalently, that for any $m\in\mathbb{Z}^n$

$$H(x+m,x'+m)=H(x,x'). \tag{3}$$

So we can write $H(x,x')=h(x'-x)+P(x,x')$ where P satisfies (3). In order for $H(x,x')$ to exist, it suffices to assume that f is C^1-close to f_0. Then H is C^2-close to H_0. Conversely, if $H(x,x')$ is a small C^2-perturbation of H_0 then it defines a small C^1-perturbation of the map f_0. However, our results depend only on the smallness of the C^1 size of the perturbation $P(x,x')$ of the generating function. So it suffices to assume that f is C^0 close to f_0 provided that f can be represented by a generating function.

Let $r_0\in U$ and $s_0=a(r_0)$. We will study orbits of the map f which stay sufficiently close to the torus $\mathbb{T}^n_{r_0}=\mathbb{T}^n\times\{r_0\}$.

Our first main result establishes the existence of such orbits if the vector s_0 has rational coordinates.

Let $(\varphi,r)\in\mathbb{T}^n\times\mathbb{R}^n$ be a periodic orbit of the map f with the prime period q and let $(x,r)\in\mathbb{R}^n\times\mathbb{R}^n$ be a lift of the point (φ,r). Then there exists a vector $w\in\mathbb{Z}^n$ such that $F^q(x,r)=(x+w,r)$. We will call the vector $\frac{w}{q}$ the *rotation vector* of the point (φ,r). The rotation vector depends on the choice of the lift F but it is uniquely defined modulo $\mathbb{Z}^n$.

Theorem A. *Let f be a perturbation of an integrable symplectic map f_0 satisfying* (i) *and* (ii). *Let $w=(w_1\dots w_n)\in\mathbb{Z}^n$, q be a positive integer such that $w_1\dots w_n, q$ are relatively prime and the vector $\frac{w}{q}$ belongs to $a(U)$. Let furthermore $r_{w,q}=a^{-1}\left(\frac{w}{q}\right)$. There exists a constant Δ depending on f_0 but not on w and q such that for any $\delta<\Delta$ if the map f is defined by the generating function $H=h+P$*

where the C^1 norm of the perturbation part P of H is equal to δ, then the map f has at least $n+1$ different periodic orbits with rotation vector $\frac{w}{q}$ which lie completely inside the $C\delta^{\frac{1}{3}}$ neighborhood of the torus $\mathbb{T}^n \times \{r_{w,q}\}$ and at least one of those orbits lies inside the $C\delta^{\frac{1}{2}}$ neighborhood of that torus. Here C depends only on the unperturbed map f_0.

This theorem follows immediately from Proposition 2 which is proved in Sect. 4 and Proposition 4 proved in Sect. 5.

We conclude this section with the description of a reduction of our problem to another one which is defined globally in $\mathbb{T}^n \times \mathbb{R}^n$ and which coincides with our problem in a neighborhood of $\mathbb{T}^n_{r_0}$. We use the word "problem" instead of "map" because we are going to modify the generating function H and we do not care whether the perturbed generating function defines a map.

Let V be a neighborhood of δ_0 such that the function h in V is sufficiently close to its second Taylor polynomial T_2. By the convexity assumption (ii), T_2 is a convex second degree polynomial which can be defined in $\mathbb{R}^n$. Thus, it is easy to see that one can construct a C^2 small function $\bar{q}$ on $\mathbb{R}^n$ which coincides with $h-T_2$ on V, and vanishes outside of a compact set. Therefore, $\bar{h} \stackrel{\text{def}}{=} T_2+\bar{q}$ is a strictly convex function which coincides with h in V and with T_2 outside of a compact set. Similarly we modify the generating function $H(x,x')=h(x'-x)+P(x,x')$ of the diffeomorphism f into a function $\bar{H}(x,x')=\bar{h}(x'-x)+\bar{P}(x,x')$ where the function $\bar{P}$ is uniformly C^1 small, coincides with P for $x'-x\in V$ and vanishes when $x'-x$ lies outside of a certain compact set. Naturally, we can make $\bar{H}$ satisfy periodicity condition (3).

Obviously, if $x'-x\in V$ then setting

$$r=\frac{\partial \bar{H}}{\partial x}, \qquad r'=-\frac{\partial \bar{H}}{\partial x'}$$

we have $F(x,r)=(x',r')$ because locally $\bar{H}=H$. Furthermore if $x_i\in\mathbb{R}^n$ is a finite or infinite sequence of vectors such that

$$\frac{\partial \bar{H}}{\partial x}(x_i, x_{i+1}) = -\frac{\partial \bar{H}}{\partial x'}(x_{i-1}, x_i) \stackrel{\text{def}}{=} r_i$$

and

$$x_{i+1}-x_i\in V \quad \text{for all } i \tag{4}$$

then the sequence (x_i, r_i) is an orbit or an orbit segment for F.

3. Periodic states with given rotation vector

Let us fix $w=(w_1 \dots w_n)\in\mathbb{Z}^n$ and $q\in\mathbb{Z}_+$ such that $w_1,\dots,w_n,q$ are relatively prime. Let us consider the space $\Psi_{w,q}$ of all double-infinite sequences $x=(\dots,x_{-1},x_0,x_1,\dots)$ of vectors from $\mathbb{R}^n$ satisfying the following periodicity condition

$$x_{i+q}=x_i+w \quad \text{for all } i\in\mathbb{Z}. \tag{5}$$

Every sequence $\{x_i\}$ is uniquely determined by a q-tuple of vectors $(x_1, \ldots, x_q)$, so $\Psi_{w,q}$ can be naturally identified with $(\mathbb{R}^n)^q$. We introduce two kinds of identifications in $\Psi_{w,q}$. First, for any $m \in \mathbb{Z}^n$ we identify $\{x_i\}$ with $\{x_i + m\}$. Secondly we identify every sequence $\{x_i\}$ with its shift $\{y_i\}$, where $y_i = x_{i+1}$. In other words, our identifications are generated by the translations $T_m: \{x_i\} \mapsto \{x_i + m\}$ and the shift $S: \{x_i\} \mapsto \{x_{i-1}\}$. The quotient space of $\Psi_{w,q}$ corresponding to the first identification will be denoted by $\Phi^*_{w,q}$, and the result of both identifications will be denoted by $\Phi_{w,q}$.

A convenient coordinate system in $\Psi_{w,q}$ is given by the parameters $(v, t) = (v, t_1, \ldots, t_{q-1})$ where

$$v = \frac{x_0 + \ldots + x_{q-1}}{q}, \qquad t_i = x_i - x_{i-1} - \frac{w}{q}, \qquad i = 1, \ldots, q-1.$$

In terms of these coordinates we have

$$T_m(v, t) = (v + m, t)$$

$$S(v, t_1, \ldots, t_{q-1}) = \left(v + \frac{w}{q}, t_2, \ldots, t_{q-1}, -t_1 - t_2 - \ldots - t_{q-1}\right). \tag{6}$$

Thus, the space $\Phi^*_{w,q}$ is diffeomorphic to $T^m \times \mathbb{R}^{n(q-1)}$ and it represents a q-fold covering of $\Phi_{w,q}$. By (6) the latter space is an $\mathbb{R}^{n(q-1)}$ bundle over the torus $\mathbb{T}^n$ and thus it is homotopically equivalent to $\mathbb{T}^n$.

We define the function $L_{w,q}$ on $\Psi_{w,q}$ by

$$L_{w,q}(x) = \sum_{i=1}^{q} \bar{H}(x_i, x_{i+1}). \tag{7}$$

Obviously, the function $L_{w,q}$ is both T_m and S invariant, so it defines a function on $\Phi_{w,q}$ which we denote by the same symbol $L_{w,q}$ and will sometimes call the Lagrangian.

A point $x \in \Phi_{w,q}$ is called *an equilibrium state* if x is a critical point of $L_{w,q}$. Such a state must satisfy the following conditions:

$$0 = \frac{\partial L_{w,q}}{\partial x_i} = \frac{\partial \bar{H}(x_i, x_{i+1})}{\partial x} + \frac{\partial \bar{H}(x_{i-1}, x_i)}{\partial x'}.$$

If an equilibrium $\{x_i\}$ satisfies additional conditions: $x_{i+1} - x_i \in V$ for all i then, as we noted at the end of Sect. 2, it corresponds to an orbit of F which by the periodicity condition (5) can be projected onto a periodic orbit of f with rotation vector $\frac{w}{q}$.

4. Critical points of the Lagrangian $L_{w,q}$

Proposition 1. *$L_{w,q}$ is a proper function on $\Phi_{w,q}$ and is bounded from below.*

Proof. Since $\bar{H}(x, x')$ is bounded from below and goes to $+\infty$ as $|x' - x| \to \infty$, the statement follows from (7). $\square$

Thus we have the following

Corollary 1. *The function $L_{w,q}$ reaches its absolute minimum l_0 at some equilibrium state x^0.*

In our next statement we will show that if the rotation vector $\frac{w}{q}$ is sufficiently close to $a(r_0)$, $\left(\text{in particular, if } a(r_0)=\frac{w}{q}\right)$, then this equilibrium state satisfies the condition (4) and determines a periodic orbit of f with rotation vector $\frac{w}{q}$ which lies near the torus $\mathbb{T}^n_{r_0}$. Moreover, under a somewhat stronger assumption on the size of the perturbation we can prove (4) for every critical point x of $L_{w,q}$ with the critical value less than l_0+C, where C does not depend on w or q.

We begin with some general estimates. We will use the letter C with various indices to denote various unspecified constants which may depend on the map f_0 but not on w or q. Let $x=\{x_i\}$ be an arbitrary state, set $a_i=x_i-x_{i-1}$ and let $\delta_k(k=0,1)$ be the C^k-norm of the perturbation $\bar{P}$ of the generating function. We will assume throughout the rest of the paper that $\delta_k<\mathcal{D}$, $k=0,1$ where the constant $\mathcal{D}$ is chosen once and for all for a given map f_0.

Lemma 1. *If x is a critical point of $L_{w,q}$ then $|a_{i+1}-a_i|<C\delta_1$ for every i. Besides, if x is a point of absolute minimum for $L_{w,q}$ then $|a_{i+1}-a_i|<C\delta_0^{\frac{1}{2}}$.*

Proof. Let $a_i=a+\tau$, $a_{i+1}=a-\tau$, and let $y=\frac{1}{2}(x_{i+1}+x_{i-1})$ so that $x_i=y+\tau$. Consider a family of states $x(\sigma)$, $\sigma\in\mathbb{R}^n$ defined by

$$x_j(\sigma)=\begin{cases} x_j & \text{for } j\not\equiv i(\operatorname{mod} q) \\ \frac{1}{2}(x_{j+1}+x_{j-1})+\sigma & \text{for } j\equiv i(\operatorname{mod} q). \end{cases}$$

We have

$$\begin{aligned} &L_{w,q}(x(\sigma))-L_{w,q}(x) \\ &\quad=\bar{H}(x_{i-1},y+\sigma)+\bar{H}(y+\sigma,x_{i+1})-\bar{H}(x_{i-1},y+\tau)-\bar{H}(y+\tau,x_{i+1}) \\ &\quad=\bar{P}(x_{i-1},y+\sigma)+\bar{P}(y+\sigma,x_{i+1})-\bar{P}(x_{i-1},y+\tau)-\bar{P}(y+\tau,x_{i+1}) \\ &\qquad+\bar{h}(a+\sigma)+\bar{h}(a-\sigma)-\bar{h}(a+\tau)-\bar{h}(a-\tau). \end{aligned} \tag{8}$$

The sum of the first four terms is bounded by $4\delta_0$. If x is a point of absolute minimum then for $\sigma=0$

$$L_{w,q}(x(0))-L_{w,q}(x)\geqq 0$$

so we have

$$\bar{h}(a+\tau)+\bar{h}(a-\tau)-2\bar{h}(a)\leqq 4\delta_0.$$

On the other hand, by the convexity of $\bar{h}$

$$\bar{h}(a+\tau)+\bar{h}(a-\tau)-2\bar{h}(a)\geqq C_1|\tau|^2.$$

Therefore, $C_1|\tau|^2\leqq 4\delta_0$ and $|a_{i+1}-a_i|<C_2\delta_0^{\frac{1}{2}}$. This proves the second statement of the lemma.

If x is a critical point then $\left(\frac{\partial}{\partial\sigma} L_{w,q}(x(\sigma))|_{\sigma=\tau}\right)=0$. The differential of the sum of the first four terms in (8) at $\sigma=\tau$ is bounded by $2\delta_1$. The last two terms do not depend on σ and the differntial of $\bar{h}(a+\sigma)+\bar{h}(a-\sigma)$ at $\sigma=\tau$ equals $d\bar{h}(a+\tau)-d\bar{h}(a-\tau)$. Let $Q(s)$ be the matrix of second derivatives of $\bar{h}$ at s. By our construction of the function $\bar{h}$ (cf. Sect. 2) $\langle Q(s)\tau,\tau\rangle \geqq C_3|\tau|^2$. On the other hand by the Mean Value theorem $\langle d\bar{h}(a+\tau)-d\bar{h}(a-\tau),\tau\rangle=\langle Q(a+\lambda\tau)2\tau,\tau\rangle$ for some $\lambda\in(-1,1)$. Therefore, $|d\bar{h}(a+\tau)-d\bar{h}(a-\tau)|\geqq 2C_3|\tau|$.

Combining all terms in (8) we have $C_3|\tau|<\delta_1$ or $|a_{i+1}-a_i|<C_4\delta_1$. □

Lemma 2. *Suppose x is a point of absolute minimum for $L_{w,q}$. Then for any i, j*

$$|a_i-a_j|<C\min(\delta_1^{\frac{1}{2}},\delta_0^{\frac{1}{4}}).$$

Proof. Let $\beta_0=\max_l|a_l-a_{l+1}|$, β be any number greater than β_0, $\tau=|a_i-a_j|$ and suppose that k is an integer such that $k\beta<\tau/5$. We can assume that $i<j\leqq i+\frac{q}{2}$. Then the sets $\{i,\dots,i+k\}$ and $\{j,\dots,j+k\}$ are disjoint mod q so $i_1=i+k<j$ and $j_1=j+k<i+q$. Next we perform a sort of surgery and construct for every $m\in\mathbb{Z}^n$ a state $y(m)$ such that

$$y_l(m)=\begin{cases} x_i+\dfrac{l-i}{k}[x_{i_1}-m-x_i] & i\leqq l\leqq i_1\\ x_l-m & i_1\leqq l\leqq j\\ x_j-m+\dfrac{l-j}{k}[x_{j_1}-(x_j-m)] & j\leqq l\leqq j_i\\ x_l & j_1\leqq l\leqq i+q\end{cases}$$

and extend it for all l by periodicity.

We will try to choose m to make the vectors $(y_{i_1}-y_i)$ and $(y_{j_1}-y_j)$ as close to each other as possible. For that purpose choose as m any of the integer lattice vectors nearest to $\frac{1}{2}(x_{i_1}-x_i-x_{j_1}+x_j)$. Let us estimate the difference $L_{w,q}(x)-L_{w,q}(y(m))$ from below. This difference is equal to

$$\begin{aligned}\sum_{l=1}^{k}&[\bar{H}(x_{i+l},x_{i+l-1})+\bar{H}(x_{j+l},x_{j+l-1})\\ &-\bar{H}(y_{i+l},y_{i+l-1})-\bar{H}(y_{j+l},y_{j+l-1})]\\ =\sum_{l=1}^{k}&[\bar{P}(x_{i+l},x_{i+l-1})+\bar{P}(x_{j+l},x_{j+l-1})\\ &-\bar{P}(y_{i+l},y_{i+l-1})-\bar{P}(y_{j+l},y_{j+l-1})]\\ +\sum_{l=1}^{k}&\left[\bar{h}(a_{i+l})+\bar{h}(a_{j+l})-\bar{h}\left(\frac{y_{i_1}-y_i}{k}\right)-\bar{h}\left(\frac{y_{j_1}-y_j}{k}\right)\right].\end{aligned}$$

The absolute value for the first sum is obviously bounded by $4k\delta_0$. Let us rewrite each term of the second sum. Let $a_{i+l}=b+\rho_l$ where b is the average of $\{a_{i+1},\dots,a_{i+k}\}$ and $\sum\rho_l=0$. Similarly let $a_{j+l}=c+\sigma_l$, $\sum\sigma_l=0$. Obviously

$|\rho_l|<k\beta$, $|\sigma_l|<k\beta$. In particular, $\tau=|a_i-a_j|=|b-c+\rho_0-\sigma_0|$ and $|\tau-|b-c||\leqq 2k\beta\leqq\frac{2}{5}\tau$ so $|b-c|>\dfrac{\tau}{2}$. Also

$$\frac{y_{i_1}-y_i}{k}=\frac{b+c}{2}+\zeta \quad \text{and} \quad \frac{y_{j_1}-y_j}{k}=\frac{b+c}{2}-\zeta \quad \text{where } |\zeta|<\frac{d}{k}$$

(d depends only on the dimension n of $\mathbb{T}^n$). Furthermore

$$\sum_{l=1}^{k}\bar{h}(a_{i+l})=k\bar{h}(b)+\langle d\bar{h}(b),\textstyle\sum\rho_l\rangle+O(k(k\beta)^2)=k\bar{h}(b)+O(k^3\beta^2).$$

Similarly

$$\sum_{l=1}^{k}\bar{h}(a_{j+l})=k\bar{h}(c)+O(k^3\beta^2),$$

$$\bar{h}\left(\frac{b+c}{2}+\zeta\right)+\bar{h}\left(\frac{b+c}{2}-\zeta\right)=2\bar{h}\left(\frac{b+c}{2}\right)+O(k^{-2}).$$

By the convexity of $\bar{h}$

$$\bar{h}(b)+\bar{h}(c)-2\bar{h}\left(\frac{b+c}{2}\right)>C_1\left|\frac{b-c}{2}\right|^2>C_5\tau^2.$$

Combining all these inequalities we have

$$L_{w,q}(x)-L_{w,q}(y)>C_5k\tau^2-4k\delta_0-C_6k^3\beta^2-C_7k^{-1}. \tag{9}$$

Since x is a point of absolute minimum

$$\tau^2<C_8\delta_0+C_9k^2\beta^2+C_{10}k^{-2}. \tag{10}$$

Now we are ready to complete the proof. By Lemma 1 we can choose $\beta=C_4\delta_1$ or $C_2\delta_0^{\frac{1}{2}}$. Take $k\approx\beta^{-\frac{1}{2}}$. If $k\beta>\tau/5$ then $\tau<C_{11}\beta^{\frac{1}{2}}$ and for $\beta=C_4\delta_1$ (resp. $C_2\delta_0^{\frac{1}{2}}$) $\tau<C_{12}\delta_1^{\frac{1}{2}}$ (resp. $\tau<C_{13}\delta_0^{\frac{1}{4}}$). If however $k\beta<\tau/5$ then we can apply the surgery described above and from (10) obtain the estimate $\tau^2<C_{14}\beta$ and $\tau<C_{15}\beta^{\frac{1}{2}}$ and we repeat the preceding argument. □

Lemma 3. *For every $E>0$ there is $C>0$ such that if x is a critical point of $L_{w,q}$ with $L_{w,q}(x)<E+\inf(L_{w,q})$ then for any i, j $|a_i-a_j|<C\delta_1^{\frac{1}{3}}$.*

Proof. By Lemma 1 we can choose $\beta=C_4\delta_1$. As before $\tau=|a_i-a_j|$. Take $k\approx\beta^{-\frac{2}{3}}$. If $k\beta>\tau/5$ then $\tau<C_{16}\delta_1^{\frac{1}{3}}$ and we are done. Otherwise we can apply the surgery from the proof of Lemma 2 and from (9) obtain the estimate:

$$\tau^2<(C_{17}E+C_{18})\beta^{\frac{2}{3}}.$$

Since $\beta=C_4\delta_1$ we have $\tau<C_{19}\delta_1^{\frac{1}{3}}$. □

Proposition 2. *Suppose $\dfrac{w}{q}\in V$ and $r_{w,q}=a^{-1}\left(\dfrac{w}{q}\right)$. Let δ be the C^1-norm of the perturbation $\bar{P}$ of the generating function. Then*

a) *any point in* $\Phi_{w,q}$ *of absolute minimum for* $L_{w,q}$ *determines a periodic orbit with rotation vector* $\frac{w}{q}$ *which lies completely inside the* ε*-neighborhood of the torus* $\mathbb{T}^n \times \{r_{w,q}\}$, *where* $\varepsilon = C\delta^{\frac{1}{2}}$.

b) *for any* $E>0$ *there is* $\Delta>0$, *independent on* w *and* q, *such that if* $\delta<\Delta$ *then any critical point* $x \in \Phi_{w,q}$ *of* $L_{w,q}$ *with* $L_{w,q}(x)<E+\inf(L_{w,q})$ *determines a periodic orbit with rotation vector* $\frac{w}{q}$ *which lies completely inside the* ε*-neighborhood of the torus* $T^n \times \{r_{w,q}\}$ *where* $\varepsilon = C\delta^{\frac{1}{3}}$.

Proof. Let x be a critical point of $L_{w,q}$ satisfying conditions of a) or b). Let

$$r_i = \frac{\partial \bar{H}}{\partial x}(x_i, x_{i+1}) = d\bar{h}(a_{i+1}) + \frac{\partial \bar{P}}{\partial x}(x_i, x_{i+1}).$$

By Lemmas 2 and 3 for any i, j $|a_i - a_j| < \varepsilon$. Since $w = \sum_1^q a_i$ we have $\left|a_i - \frac{w}{q}\right| < \varepsilon$. $\frac{\partial \bar{P}}{\partial x} = O(\delta)$ by definition of δ. But $\delta = O(\varepsilon)$ so $\frac{\partial \bar{P}}{\partial x} = O(\varepsilon)$. On the other hand $d\bar{h}\left(\frac{w}{q}\right) = r_{w,q}$ and $d\bar{h}(a_{i+1}) - d\bar{h}\left(\frac{w}{q}\right) = O\left(a_{i+1} - \frac{w}{q}\right) = O(\varepsilon)$. Therefore $|r_i - r_{w,q}| = O(\varepsilon)$. □

5. Existence of $(n+1)$ critical points

Let us recall that we denoted the quotient space of $\Psi_{w,q}$ modulo the group of integral translations by $\Phi^*_{w,q}$. The function $L_{w,q}$ is naturally defined on $\Phi^*_{w,q}$. Let as before $l_0 = \inf(L_{w,q})$ and let

$$M_t^* = \{x \in \Phi^*_{w,q} \mid L_{w,q}(x) \leqq t\},$$
$$M_t = \{x \in \Phi_{w,q} \mid L_{w,q}(x) \leqq t\}.$$

Proposition 3. *For a given* $\bar{h}$ *and* δ_0 *there exists a constant* E *such that for any perturbation* $\bar{P}$ *uniformly bounded by* δ_0 *there exists a continuous map* $\Gamma: \mathbb{T}^n \to M^*_{l_0+E}$, *such that for any* w *and* q *the composition of* Γ *with the inclusion* $i_1: M^*_{l_0+E} \to \Phi^*_{w,q}$ *is a homotopy equivalence between* $\mathbb{T}^n$ *and* $\Phi^*_{w,q}$.

Proof. We begin with defining some preliminary tools. Let

$$\varphi(t) = \begin{cases} 1, & \text{if } \ t - kq \in [0,1] \text{ for some } k \in \mathbb{Z} \\ 0, & \text{otherwise} \end{cases},$$

$$a_i(t) = \int_{i-t}^{i} \varphi(t)\,dt, \qquad i \in \mathbb{Z}.$$

It can be easily seen that

(a) $a_i(t+q) = a_i(t) + 1$

(b) $a_{i+q}(t) = a_i(t)$

(c) $a_i(t)=a_{i+1}(t)\in\mathbb{Z}$ for all but at most three values of $i\in[1,q]$ and for these exceptional i's $|a_i(t)-a_{i+1}(t)|\leqq 1$.

Let furthermore $A(t)=\{a_i(t),\ i\in\mathbb{Z}\}$ be the sequence of numbers $a_i(t)$. For any vector $z=(t_1,\dots,t_n)\in\mathbb{R}^n$ let $V(z)=\{v_i(z),\ i\in\mathbb{Z}\}\overset{\text{def}}{=}\sum_{j=1}^{n}A(t_j)\,e_j$ where $e_1,\dots,e_n$ is the standard basis of $\mathbb{R}^n$. So $V(z)$ is a family of double-infinite sequences of vectors depending on z as a parameter. Conditions (a), (b), (c) immediately imply that

(a′) $v_i(z+qm)=v_i(z)+m,\ m\in\mathbb{Z}^n$

(b′) $v_{i+q}(z)=v_i(z)$

(c′) $v_i(z)=v_{i+1}(z)\in Z^n$ for all but at most $3n$ values of $i\in\{1,\dots,q\}$ and for those exceptional i each coordinate of $v_{i+1}(z)-v_i(z)$ does not exceed one in absolute value.

Now we are ready to define the map Γ. Let x be a point of absolute minimum of $L_{w,q}$ in $\Psi_{w,q}$. Let $x(z)$ be the following n-parameter family of states: $x(z)=x+V(z)$. By property (b′) $x(z)\in\Psi_{w,q}$. Denote the map $z\mapsto x(z)$ by $X:\mathbb{R}^n\to\Psi_{w,q}$. Furthermore by property (a′) the map $X:\mathbb{R}^n\to\Psi_{w,q}$ can be projected to a map $\Gamma:\mathbb{T}^n=\mathbb{R}^n/q\mathbb{Z}^n\to\Phi^*_{w,q}$. First we show that Γ is homotopic to the standard embedding $\Gamma_1: z\mapsto\{w_i(z)\}$ where $w_i(z)=\frac{z}{q}+i\frac{w}{q}$ and therefore effects a homotopy equivalence between $\mathbb{T}^n$ and $\Phi^*_{w,q}$. This homotopy is given by the explicit formula Γ_ε: $z\mapsto\{w_i(\varepsilon,z)\}$ where $w_i(\varepsilon,z)=(1-\varepsilon)\,x_i+\varepsilon\left(i\frac{w}{q}\right)+(1-\varepsilon)\,v_i(z)+\varepsilon\frac{z}{q}$. Next we show that $\Gamma(\mathbb{T}^n)\subset M^*_{l_0+E}$. To that end we estimate the difference $L_{w,q}(x(z))-L_{w,q}(x)=L_{w,q}(x(z))-l_0$. By property (c′) $\bar{H}(x_i(z),x_{i+1}(z))$ coincides with $\bar{H}(x_i,x_{i+1})$ for all but $3n$ values of $i\in\{1,\dots,q\}$. Let i be one of the exceptional values. Then all coordinates of $v_i(z)-v_{i+1}(z)$ are less than or equal to one in absolute value. Since by Lemma 2 for a given δ_0 the differences $|x_{i+1}-x_i|$ are uniformly bounded, $|x_{i+1}(z)-x_i(z)|=|(x_{i+1}-x_i)+(v_{i+1}(z)-v_i(z))|$ are also uniformly bounded and

$$|\bar{H}(x_i(z),x_{i+1}(z))-\bar{H}(x_i,x_{i+1})|\leqq|\bar{H}(x_i(z),x_{i+1}(z))|+|\bar{H}(x_i,x_{i+1})|\leqq C.$$

Therefore $|L_{w,q}(x(z))-L_{w,q}(x)|\leqq 3nC$. □

Proposition 4. *For given $\bar{h}$ and δ_0 there exists a constant E independent of w,q such that for any perturbation $\bar{P}$ uniformly bounded by δ_0 there are at least $n+1$ different critical points of $L_{w,q}$ in $\Phi_{p,w}$ with critical values less than l_0+E.*

Proof. We will show that the Lusternik-Shnirelman category (for definition see, e.g. [16], Sect. 3 or [20], Chap. 5), of the set M_{l_0+E} is at least $n+1$. This implies the existence of $n+1$ critical points ([16], Sect. 3). The category of M is greater than k if there are $\omega_1\dots\omega_k\in H^1(M,\mathbb{Z})$ such that $\omega_1\cup\omega_2\cup\dots\cup\omega_k\neq 0$ (cf. [20], Theorem 5.14). We will derive the existence of such cohomology classes from the following commutative diagram.

$$\begin{tikzcd} \mathbb{T}^n \arrow[r, "\Gamma"] \arrow[dr, "\gamma"'] & M^*_{l_0+E} \arrow[r, "\pi"] \arrow[d, "i_1"] & M_{l_0+E} \arrow[d, "i_2"] \\ & \Phi^*_{w,q} \arrow[r, "\pi"] & \Phi_{w,q} \end{tikzcd}$$

Here i_i, i_2 are inclusions, π is the projection, and Γ is the map constructed in Proposition 3. Let $\omega_1 \dots \omega_n$ be generators of $H^1(\Phi_{w,q}; \mathbb{Z})$. Then $\omega = \omega_1 \cup \dots \cup \omega_n$ is a generator of $H^n(\Phi_{w,q}; \mathbb{Z})$ and $\pi^* \omega = q\omega_0$ where ω_0 is a generator of $H^n(\Phi^*_{w,q}; \mathbb{Z})$. Here $\cup$ means cup-product in cohomology groups. Since γ is a homotopy equivalence we have $0 \neq \gamma^*(q\omega_0) = \gamma^* \pi^* \omega = \Gamma^* \pi^* i_2^*(\cup \omega_i) = \Gamma^* \pi^*(\cup(i_2^* \omega_i))$. So if we denote $i_2^* \omega_i$ by $\psi_i \in H^1(M_{l_0+E}, \mathbb{Z})$ and their cup product by ψ we have $\Gamma^* \pi^* \psi \neq 0$, thus $\psi \neq 0$. And the category of M_{l_0+E} is greater than n. □

6. Weak regularity of minimal orbits

Let ω: $[0, 1) \to \mathbb{R}_+$, $\omega(0) = 0$, be a modulus of continuity i.e. a non-decreasing continuous non-negative function which is strictly positive outside 0 and let $(\varphi_i, r_i) = f^i(\varphi_0, r_0) \in \mathbb{T}^n \times \mathbb{R}^n$ be an orbit of f. We say that this orbit is *ω-regular* if for any $i, j \in \mathbb{Z}$

$$|r_i - r_j| \leq \omega(\operatorname{dist}(\varphi_i, \varphi_j)).$$

In particular, ω-regularity implies that two different points on the orbit do not have the same φ coordinates. If $\omega(t) = Lt$ for some constant L we call ω-regular orbits *Lipschitz regular.* For $n=1$ all orbits corresponding to the absolute minimum of the Lagrangian are Lipschitz regular [13] with a fixed constant L which depends only on the map. As we mentioned in the introduction, regularity plays the key role in the whole theory of twist maps. We are not able to prove any ω-regularity for the minimal orbits for $n > 1$. However, a slightly weaker property does hold in that case.

Proposition 5. *Under the assumptions of Theorem A let* $(\varphi_i, r_i) = f^i(\varphi_0, r_0)$, $(\varphi_q, r_q) = (\varphi_0, r_0)$ *be any periodic orbit of f corresponding to a state which minimizes the functional* $L_{w,q}$. *Assume in addition that the second derivatives of H are bounded. Then there exist ε_0 and C depending only on f such that if i, j, k are different* $\operatorname{mod} q$ *and* $\max(\operatorname{dist}(\varphi_i, \varphi_j), \operatorname{dist}(\varphi_j, \varphi_k)) < \varepsilon_0$ *then*

$$|r_i - r_k| < C(\operatorname{dist}(\varphi_i, \varphi_j) + \operatorname{dist}(\varphi_i, \varphi_k))^{\frac{1}{2}}.$$

Proof. As usual we pass to the universal cover and consider the state $x = \{x_n\}_{n \in \mathbb{Z}} \in \Psi_{w,q}$ generating the given periodic orbit. We will assume that $i=0$ and $0 < j < k < q$. The case $0 < k < j < q$ is considered similarly. From the conditions of the proposition we have for some $m_1, m_2 \in \mathbb{Z}^n$

$$|x_j + m_1 - x_k| < \varepsilon_0, \qquad |x_k + m_2 - x_q| < \varepsilon_0.$$

The idea of the proof is to rearrange the pieces $x_0, \dots, x_{j-1}$; $x_j, \dots, x_{k-1}$ and $x_k, \dots, x_{q-1}$ of the state x into a new state y and then estimate both the value of the Lagrangian and its derivatives at y. Thus, we define

$$y_l=\begin{cases} x_l, & 0\leqq l<j \\ x_{l+k-j}-m_1, & j\leqq l<j+q-k \\ x_{l+k-q}+m_2, & j+q-k\leqq l<q \end{cases}$$

and then extend y_l to other l according to (5) in order to obtain a state $y\in\Psi_{w,q}$.

Let us estimate $L_{w,q}(y)$ from above. We have

$$\begin{aligned} L_{w,q}(y)-l_0 &= L_{w,q}(y)-L_{w,q}(x) \\ &= H(x_{j-1},x_k-m_1)-H(x_{j-1},x_j) \\ &\quad +H(x_{k-1},x_q-m_2)-H(x_{k-1},x_k) \\ &\quad +H(x_{q-1},x_j+m_1+m_2)-H(x_{q-1},x_q) \\ &< C_1(|x_j+m_1-x_k|+|x_k+m_2-x_q|+|x_q-m_1-m_2-x_j|) \\ &= C_1(\operatorname{dist}(\varphi_j,\varphi_k)+\operatorname{dist}(\varphi_k,\varphi_q)+\operatorname{dist}(\varphi_q,\varphi_j)) \\ &< C_2(\operatorname{dist}(\varphi_0,\varphi_j)+\operatorname{dist}(\varphi_0,\varphi_k)). \end{aligned} \tag{11}$$

On the other hand

$$\frac{\partial L_{w,q}}{\partial x_0}(y)=\frac{\partial \bar H}{\partial x'}(x_{k-1}-m_2-m_1,x_0)+\frac{\partial \bar H}{\partial x}(x_0,x_1)$$

and

$$\frac{\partial \bar H}{\partial x'}(x_{k-1},x_k)=-r_k,\qquad \frac{\partial \bar H}{\partial x}(x_0,x_1)=r_0.$$

Since the second derivatives of H are bounded, we can assume that the second derivatives of $\bar H$ are bounded too. Thus, we have

$$\left|\frac{\partial L_{w,q}}{\partial x_0}(y)-(r_0-r_k)\right|<C_3|x_k-x_0-m_1-m_2|=C_3\operatorname{dist}(\varphi_0,\varphi_k). \tag{12}$$

Using the fact that x is a point of absolute minimum for $L_{w,q}$ and boundness of the second derivatives of $\bar H$ we obtain from (12) that for all positive t and for some $C_4>0$

$$L_{w,q}(y)-(|r_k-r_0|-C_3\operatorname{dist}(\varphi_0,\varphi_k))\,t+C_4t^2\geqq L_{w,q}(x)$$

and from (11)

$$C_4t^2-(|r_k-r_0|-C_3\operatorname{dist}(\varphi_0,\varphi_k))\,t+C_2(\operatorname{dist}(\varphi_0,\varphi_j)+\operatorname{dist}(\varphi_0,\varphi_k))\geqq 0.$$

Consequently

$$(|r_k-r_0|-C_3\operatorname{dist}(\varphi_0,\varphi_k))^2<C_5(\operatorname{dist}(\varphi_0,\varphi_j)+\operatorname{dist}(\varphi_0,\varphi_k))$$

which implies the statement of the proposition. □

Weak regularity does not imply regularity for two reasons. First, it is possible that the orbit in consideration contains two points whose φ-coor.

dinates are very close or even equal, while all other points have φ-coordinates far away from the two. In this case the assertion of Proposition 5 is vacuous. Second, in the situation of Proposition 5 one may have $\operatorname{dist}(\varphi_i, \varphi_k) \ll \operatorname{dist}(\varphi_i, \varphi_j)$. In this case the estimate for $|r_i - r_k|$ becomes very weak, although the corresponding estimate for $|r_i - r_j|$ which can be obtained by re-naming indices is good. The last observation is refined in the following statement.

Theorem B. *Under the assumptions of Proposition 5 for every point (φ_i, r_i) of the orbit the inequality*

$$|r_i - r_j| < c(\operatorname{dist}(\varphi_i, \varphi_j))^{\frac{1}{2}}$$

holds for all except probably one $j \in \{0, \ldots, q-1\}$. The constant c depends only on f, but not on w and q.

Proof. Let k, $i < k < i+q$ be such that

$$\operatorname{dist}(\varphi_i, \varphi_k) = \min_{i<j<i+q} \operatorname{dist}(\varphi_i, \varphi_j).$$

Then for every $j \neq k$, $i < j < i+q$ one has from Proposition 5

$$|r_i - r_j| < c(\operatorname{dist}(\varphi_i, \varphi_j) + \operatorname{dist}(\varphi_i, \varphi_k))^{\frac{1}{2}} < c'(\operatorname{dist}(\varphi_i, \varphi_j))^{\frac{1}{2}}. \quad \square$$

Theorem B allows us to make the first modest step toward the extension of results by Aubry and Mather to higher dimensions. Namely we will show that limits of minimal periodic orbits are regular.

Theorem C. *Under the assumptions of Proposition 5 let $\frac{w^{(n)}}{q^{(n)}}$, $n = 1, 2 \ldots$ be a sequence of rotation vectors and $(\varphi^{(n)}, r^{(n)})$ be a sequence of points whose f-orbits correspond to absolute minima of the functionals $L_{w^{(n)}, q^{(n)}}$. Suppose that the sequence $(\varphi^{(n)}, r^{(n)})$ converges to a point (φ, r) which is not an isolated point of its orbit. Then the orbit of (φ, r) is ω-regular where $\omega(t) = ct^{\frac{1}{2}}$.*

Proof. Let us note first that (φ, r) is not a periodic point and all points of its orbit are not isolated. Consider any two points on the orbit of (φ, r), say $(\varphi_i, r_i) = f^i(\varphi, r)$ and $(\varphi_j, r_j) = f^j(\varphi, r)$. Let us show first that $\varphi_i \neq \varphi_j$. For otherwise one can find k such that $\operatorname{dist}(\varphi_i, \varphi_k)$ is very small compared to $|r_i - r_j|^2$, then approximate the piece of orbit containing all three points by a piece of orbit of $(\varphi^{(n)}, r^{(n)})$ with very high precision and use Proposition 5 for the approximation.

Similarly if $\varphi_i \neq \varphi_j$, let us find k such that $\operatorname{dist}(\varphi_i, \varphi_k) \ll \operatorname{dist}(\varphi_i, \varphi_j)$. Assume $i < j < k$. Other cases can be treated similarly. Approximate the piece of orbit from i to k by a piece of orbit of $(\varphi^{(n)}, r^{(n)})$ and again apply Proposition 5 for the approximation. $\square$

At the current stage we are not able to make a comprehensive description of orbits which appear as limits defined in the last theorem. We hope that if $\frac{w^{(n)}}{q^{(n)}} \to \alpha$ then the limit orbits have rotation vector α and are fairly similar to quasi-periodic orbits with that rotation vector.

Let us note that in the situation described in Theorem C even if (φ, r) happens to be a periodic orbit, it satisfies the same regularity condition if the approximating orbits pass near that orbit several times.

7. Continuous time Hamiltonian systems

Let us consider a completely integrable Hamiltonian system with n degrees of freedom. Let $(I, \varphi)=(I_1 \dots I_n, \varphi_1 \dots \varphi_n)$, $I \subset U$ be an open set in $\mathbb{R}^n$, $\varphi \in \mathbb{T}^n$ be the action-angle coordinates for that system ([3], Sect. 50) so that the Hamiltonian which we will denote by H_0 depends only on I and the time evolution leaves every torus $\{I^{(0)}\} \times \mathbb{T}^n$ invariant. Let us fix $I^{(0)} \in U$ and assume that

$$\frac{\partial H_0}{\partial I_1}(I^{(0)}) \neq 0. \tag{13}$$

Then in the neighborhood of the invariant torus $\{I^{(0)}\} \times \mathbb{T}^n$ one can define the Poincaré map T on the hypersurface $\varphi_1=0$. This map has the following form

$$T(I, \varphi_2 \dots \varphi_n)=\left(I, \varphi_2+\frac{\partial H_0}{\partial I_2} \bigg| \frac{\partial H_0}{\partial I_1}, \dots, \varphi_n+\frac{\partial H_0}{\partial I_n} \bigg| \frac{\partial H_0}{\partial I_1}\right). \tag{14}$$

The restriction of the map T to the $2n-2$-dimensional invariant manifold $N=\{H_0(I)=H_0(I^{(0)}),\ \varphi_1=0\}$ is a symplectic map with respect to the induced symplectic form $\sum_{i=2}^{n} dI_i \wedge d\varphi_i$.

By the implicit function theorem, on the hypersurface $H_0(I)=H_0(I^{(0)})$ one can locally express

$$I_1=S(I_2 \dots I_n). \tag{15}$$

We will make the following assumption:

(iii) *the hypersurface in I-space $H_0(I)=H_0(I^{(0)})$ is strictly differentiably convex at the point $I^{(0)}$.*

This condition is satisfied for example, if the function H_0 is strictly convex in I at $I^{(0)}$, i.e. if the Hessian $\left.\frac{\partial^2 H_0}{\partial I^2}\right|_{I=I^{(0)}}$ is a positive definite quadratic form. Condition (iii) implies that S is a strictly convex function of $(I_2 \dots I_n)$ near $(I_2^{(0)}, \dots, I_n^{(0)})$.

Proposition 6. *The generating function for the lift of the map $T: N \to N$ to the universal cover has the form $\mathscr{H}(x, x')=h(x'-x)$ where h is the Legendre transformation of the function S.*

Proof. One has from (15) for $i=2, \dots, n$

$$\frac{\partial S}{\partial I_i}=-\frac{\partial H_0}{\partial I_i} \cdot \left(\frac{\partial H_0}{\partial I_1}\right)^{-1}$$

so that from (2) if $\delta_i=\frac{\partial S}{\partial I_i}$ then $\frac{\partial h}{\partial \delta_i}=I_i$. $\square$

By Proposition 6 if assumption (iii) is satisfied, h is a strictly convex function because it is a Legendre transform of a strictly convex function. Thus condition (iii) for H_0 implies that conditions (i) and (ii) from Sect. 2 are locally satisfied for the Poincaré map T.

Consider now a C^2 small Hamiltonian perturbation $H(I, \varphi)$ of the Hamiltonian H_0. The first-return map on the manifold $N_H = \{\varphi_1 = 0,\ H(I, \varphi) = H_0(I^{(0)})\}$ is still symplectic. Since manifolds N_H and N are close and the projection along I_1 direction is a symplectic map, the new map can be viewed as a C^1 small symplectic perturbation of the map T. Thus, this map is determined by a generating function which is C^2 (and hence C^1) close to h so that all results from the previous sections can be applied to this case. The properties of the Poincaré map can be in the obvious way extended to the continuous time systems. We leave exact formulations of the results for Hamiltonian systems corresponding to Theorems A, B, C to the reader.

Remark. The C^2 closeness of the perturbed Hamiltonian to the one for the integrable system is only sufficient but not necessary for application of our results because we only need C^1 closeness for the generating function (compare with discussion in Sect. 2). However since the relationship between the Hamiltonian and the generating function of a Poincaré map is rather complicated for the non-integrable case, we do not try to interpret the weaker condition in terms of the Hamiltonian.

Acknowledgements. This paper was written during the visit of the first author to California Institute of Technology in December 1984. We would like to thank Caltech's Mathematics Department for the support for that visit. We also feel a pleasant obligation to thank the referee for his thoughtful remarks.

References

1. Arnold, V.I.: Small denominators II. Proof of a theorem of A.N. Kolmogorov on the preservation of conditionally-periodic motions under a small perturbation of the Hamiltonian. Russ. Math. Surv. **18**, 9–36 (1963)
2. Arnold, V.I.: Sur une propriété topologique des applications globalement canoniques de la mécanique classique. C.R. Acad. Sci. Paris **261**, 3719–3722 (1965)
3. Arnold, V.I.: Mathematical methods of classical mechanics. Berlin-Heidelberg-New York: Springer 1978
4. Aubry, S.: Theory of the devil's staircase. Seminar on the Riemann Problem and Complete Integrability. Lect. Notes Math. **925**. Berlin-Heidelberg-New York: Springer 1978–79
5. Aubry, S., La Daeron, P.Y.: The discrete Frenkel-Kontorova model and its extensions. Physica 8D, 381–422 (1983)
6. Aubry, S., La Daeron, P.Y., André, G.: Classical ground-states of a one-dimensional model for incommensurate structures. (Preprint, 1982)
7. Bernstein, D.: Birkhoff periodic orbits twist maps with graph intersection property. Ergodic Theory Dyn. Syst. **5**, 531–532 (1985)
8. Birkhoff, G.D.: Une généralisation à n dimension du dernier théorème géométrique de Poincaré. C.R. Acad. Sci. Paris **192**, 196–198 (1931)
9. Birkhoff, G.D., Lewis, D.C.: On the periodic motions near a given periodic motion of a dynamical system. Ann. Mat. **12**, 117–133 (1933)
10. Conley, C.C., Zehnder, E.: The Birkhoff-Lewis fixed point theorem and a conjecture of V.I. Arnold. Invent. Math. **73**, 33–49 (1983)

11. Denjoy, A.: Sur les courbes définies par les équations différentielles à la surface du tore. J. Math. Pure Appl. (Ser. 9), **11**, 333–375 (1932)
12. Hall, G.R.: A topological version of a theorem of Aubry-Mather on twist maps. Ergodic Theory Dyn. Syst. **4**, 585–603 (1984)
13. Katok, A.: Some remarks on Birkhoff and Mather twist map theorems. Ergodic Theory Dyn. Syst. **2**, 185–194 (1982)
14. Katok, A.: More about Birkhoff periodic orbits and Mather sets of twist maps. (Preprint with continuation, 1982)
15. Katok, A.: Periodic and quasi-periodic orbits for twist maps. Dynamical Systems and Chaos, Proceedings, Sitges, Lect. Notes Phys. **179** (Springer, 1982)
16. Ljusternik, L.: The topology of calculus of variations in the large. AMS, Translation of Math Monograph v. 16, 1966
17. Mather, J.N.: Existence of quasi-periodic orbits for twist homeomorphisms of the annulus. Topology **21**, 457–467 (1982)
18. Mather, J.N.: A criterion for the non-existence of invariant circles. Publ. Math. IHES, **63**, 153–204 (1982)
19. Moser, J.: Proof of a generalized form of a fixed point theorem due to G.D. Birkhoff. Geometry and Topology, Rio de Janeiro. Lect. Notes Math. **597**, 464–494 (1976)
20. Schwartz, J.T.: Nonlinear functional analysis, 1963–64. Courant Inst. of Math. Science 1965

Oblatum 14-IV-1985