6

INVARIANT MANIFOLDS

34. The Theory of Kolmogorov–Arnold–Moser

KAM theory, which proves the existence of many invariant tori for systems close to integrable, is one of the greatest achievements in Hamiltonian dynamics. It has historical roots going back to Weierstrass who, in 1878, wrote to S. Kovalevski that he had constructed formal power series for quasi-periodic solutions to the planetary problem. The denominators of the coefficients of these series involved integer combinations of the frequencies of rotation of the planets around the sun. These denominators could be close to zero and hence impede the convergence of the series. Weierstrass advised Mittag-Leffler to make this problem of convergence the subject for a prize sponsored by the king of Sweden. In the 271 pages work (Poincaré (1890)) for which he won the prize, Poincaré does not solve the problem completely, and his tentative answer to the convergence is negative. In Poincaré (1899) , he speculates on the possibility of such a convergence, given appropriate number theoretic conditions, but still deems it improbable. It was therefore a significant event when Arnold (1963) (in the analytic, Hamiltonian context) and Moser (1962) (in the differentiable twist map context) gave, following the ideas of Kolmogorov (1954), a proof of existence of quasi-periodic orbits on invariant tori filling up a set of positive measure in the phase space. We can only give here a very limited account of this complex theory, and refer to Moser (1973) and de la Llave (1993) for introductions as well as Bost (1986) for an excellent survey and bibliography. There are many KAM theorems, the most applicable ones being often the hardest ones to even state. We present here a relatively simple statement, cited in Bost (1986).

Theorem 34.1 (KAM for symplectic twist maps) *Let f_0 be an integrable symplectic twist map of $\mathbb{T}^n \times \mathbb{D}^n$ of the form:*

$$f_0(\boldsymbol{q}, \boldsymbol{p}) = (\boldsymbol{q} + \omega(\boldsymbol{p}), \boldsymbol{p})$$

where $\mathbb{D}^n$ is a disk in $\mathbb{R}^n$ and $\omega : \mathbb{D}^n \mapsto \mathbb{R}^n$ is C^∞ (since f_0 is twist, $D\omega$ is invertible). Let $\boldsymbol{p}_0$ be an interior point of $\mathbb{D}^n$. Suppose that the following condition is satisfied:

Diophantine condition: *there are positive constants τ and c such that:*

$$\forall \boldsymbol{k} \in \mathbb{Z}^{n+1} \backslash \{0\}, \qquad \left| \sum_{j=1}^{n} k_j \omega_j(\boldsymbol{p}_0) + k_{n+1} \right| \geq c \left(\sum_{j=1}^{n+1} |k_j| \right)^{-\tau} \tag{34.1}$$

Then there is a neighborhood W of f_0 of C^∞ exact symplectic maps such that, for each $f \in W$, there exists an embedded invariant torus $\mathbb{T}_f \simeq \mathbb{T}^n$ in the interior of $\mathbb{T}^n \times \mathbb{D}^n$ such that:

(i) $\mathbb{T}_f$ is a C^∞ Lagrangian graph over the zero section,

(ii) $f\big|_{\mathbb{T}_f}$ is C^∞ conjugated to the rigid translation by $\omega(\boldsymbol{p}_0)$,

(iii) $\mathbb{T}_f$ and the conjugacy depend C^∞ on f.

Moreover the measure of the complement of the union of the tori $\mathbb{T}_f(\boldsymbol{p}_0)$ goes to 0 as $\|f - f_0\|$ goes to 0.

Remark 34.2

1) The diophantine condition (34.1) is shared by a large set of vectors in $\mathbb{R}^n$. As an example, when $n = 1$, the set of real numbers $\mu \in [0, 1]$ such that $|\mu - p/q| > K/q^\tau, \tau > 2$ for some K is dense in $[0, 1]$ and has measure going to 1 as K goes to 0.

2) The most common versions of KAM theorems concern Hamiltonian systems with a Legendre condition. In Chapter 7 we show the intimate relationship of such Hamiltonian systems with symplectic twist maps. It therefore comes as no surprise that KAM theorems have equivalents in both categories of systems. Note that there are *isoenergetic* versions of the KAM theorem for Hamiltonian systems, where the existence of many invariant tori is proven in a prescribed energy level (see Broer (1996), Delshams & Gutiérrez (1996a)).

3) One important contribution in Moser (1962) was his treatment of the finitely differentiable case: he was able to show a version for $n = 1$ (twist maps) where f_0 and its perturbation

are $C^l, l \geq 333$ instead of analytic. This was later improved to $l > 3$ by Herman (1983) and in higher dimension n, to $l > 2n + 1$ (at least if the original f_0 is analytic).

4) There is a version of the KAM for *non symplectic* perturbations of completely integrable maps of the annulus, called the *Theorem of translated curves*, due to Rüssmann (1970). It states that, around an invariant circle for f_0 whose rotation number ω satisfies the diophantine condition (34.1) (only one j in this case), there exists a circle invariant by $t_a \circ f$ for a perturbation f of f_0 and $t_a(x, y) = (x, y + a)$, which has same rotation number as the original.

5) One may wonder if, among all invariant tori of a symplectic twist map close to integrable, the KAM tori are typical. KAM theory says that in measure, they are. However Herman (1992a) (see also Yoccoz (1992)) shows that, for a generic symplectic twist map close to integrable, there is a residual set of invariant tori on which the (unique) invariant measure has a support of Hausdorff dimension 0 (and hence cannot be a KAM torus). Things get even worse when the differential $D\omega$ in Theorem 34.1 is not positive definite: there may be many invariant tori that project onto, but are not graphs over the 0–section, and this for maps arbitrarily close to integrable (see Herman (1992 b)).

6) KAM theory implies the stability of orbits on the KAM tori, hence stability with high probability. But in “real situations” it is impossible to tell, for lack of infinite precision on the knowledge of initial conditions, whether motion actually takes place on a KAM torus. Nekhoroshev (1977) provides an estimate of how far a trajectory can drift in the momentum direction over long periods of time: If $H(\boldsymbol{q}, \boldsymbol{p}) = h(\boldsymbol{p}) + f_\varepsilon(\boldsymbol{q}, \boldsymbol{p})$ is a real analytic Hamiltonian function on $T^*\mathbb{T}^n$ with $f_\varepsilon < \varepsilon$ (a small parameter) and $h(\boldsymbol{p})$ satisfies a certain condition (steepness) implied by convexity, then there exist constants ε_0, R_0, T_0 and a such that, if $\varepsilon < \varepsilon_0$, one has:

$$|t| \leq T_0 \exp[(\varepsilon_0/\varepsilon)^a] \Rightarrow |\boldsymbol{p}(t) - \boldsymbol{p}(0)| \leq R_0(\varepsilon/\varepsilon_0)^a.$$

With a (quasi) convexity condition instead of the steepness condition, Lochak (1992) and Pöshel (1993) showed that the optimal a is $\frac{1}{2n}$. Delshams & Gutiérrez (1996a) present unified proofs of the KAM theorem and Nekhoroshev estimates for analytic Hamiltonians.

Whereas we cannot give a proof of the KAM theorem in this book, the following theorem (Arnold (1983)) offers a simple model in a related situation in which the KAM method can be applied in a less technical way.This will allow us to sketch very roughly the central ideas of the method.

Theorem 34.3 *There exists $\varepsilon > 0$ depending only on K, ρ and σ such that, if a is a 2π –periodic analytic function on a strip of width ρ, real on the real axis with $a(z) < \varepsilon$ on the strip and such that the circle map defined by*

$$f : x \mapsto x + 2\pi\mu + a(x)$$

is a diffeomorphism with rotation number μ satisfying the diophantine condition:

$$|\mu - p/q| > \frac{K}{q^{2+\sigma}}, \quad \forall\ p/q \in \mathbb{Q}$$

then f is analytically conjugate to a rotation R of angle $2\pi\mu$

Sketch of proof: We seek a change of coordinates $H : \mathbb{S}^1 \to \mathbb{S}^1$ such that:

$$H \circ R = f \circ H \tag{34.2}$$

write $H(z) = z + h(z)$, with $h(z + 2\pi) = h(z)$. Then (34.2) is equivalent to

$$h(z + 2\pi\mu) - h(z) = a(z + h(z)). \tag{34.3}$$

Since $a(z) < \varepsilon$, h must be of order ε as well and thus, in first approximation, (34.3) is equivalent to:

$$h(z + 2\pi\mu) - h(z) = a(z) \tag{34.4}$$

Decomposing $a(z) = \sum a_k e^{i2\pi kz}, h(z) = \sum b_k e^{ikz}$ in their Fourier series and equating coefficients on both sides of (34.4) we obtain:

$$b_k = \frac{a_k}{e^{i2\pi k\mu} - 1}$$

where we see the problem of small divisors arise: the coefficients b_k of h may become very big if μ is not sufficiently rational.

It turns out that, assuming the diophantine condition and using an infinite sequence of approximate conjugacies given by solutions of (34.4), one obtains sequences h_n, a_n and corresponding $H_n, f_n = H_n^{-1} \circ f \circ H_n$ which converge to H, R for some H. The domain of h_n and f_n is a strip that shrinks with n but in a controllable way. This iterative process of "linear" approximations to the conjugacy can be interpreted as a type of Newton's method for the implicit equation $\mathcal{F}(f, H) = H^{-1} \circ f \circ H = R$ (given f, find H) and inherits the quadratic convergence of the classical Newton's method: $R - \mathcal{F}(f_n, H_n) = O(\varepsilon^{2n})$ (see Hasselblat & Katok (1995) Section 2.7.b). $\square$

35. Properties of Invariant Tori

The previous section showed the existence of many invariant tori for symplectic twist maps close to integrable. These tori are Lagrangian graphs with dynamics conjugated to quasi–periodic translations. In dimension 2, the Aubry-Mather theorem gives an answer to the question of what happens to these tori when they break down, *eg.* in large perturbations of integrable maps. In higher dimension, Mather's theory of minimal measure also provides an answer to that question (see Chapter 9). In this section, we look for properties that invariant tori may have whether they arise from KAM or not. We will see that certain attributes of KAM tori (*eg.* graphs with recurrent dynamics) imply their other attributes (*eg.* Lagrangian), as well as other properties not usually stated by the KAM theorems (minimality of orbits).

A. Recurrent Invariant Toric Graphs Are Lagrangian

Theorem 35.1 (Herman (1990)) *Let T be an invariant torus for a symplectic twist map f of $T^*\mathbb{T}^n$ and suppose $f\big|_T$ is conjugated by a diffeomorphism h to a an irrational translation R on $\mathbb{T}^n$. Then T is Lagrangian.*

Proof. Since the restriction of the symplectic 2-form $\omega\big|_T$ is invariant under $f\big|_T$ and since $R = h^{-1} \circ f\big|_T \circ h$, the 2-form $h^*\omega\big|_T$ is invariant under R. Since R is recurrent, $h^*\omega\big|_T = \sum_{i,j} a_{kj} dx_k \wedge dx_j$ must have constant coefficients a_{kj}. Integrating $h^*\omega\big|_T$ over the x_k, x_j subtorus yields on one hand a_{kj}, on the other hand 0 by Stokes' theorem since $h^*\omega\big|_T = dh^*\lambda\big|_T$ is exact. Hence $h^*\omega\big|_T = 0 = \omega\big|_T$ and the torus T is Lagrangian. $\square$

B. Orbits on Lagrangian Invariant Tori Are Minimizers

The following theorem is attributed to Herman by MacKay & al. (1989), whose proof we reproduce here.

Theorem 35.2 *Let T be a Lagrangian torus, C^1 graph over the zero section of $T^*\mathbb{T}^n$ which is invariant for a symplectic twist map f whose generating function S satisfies the following superlinearity condition:*

$$\lim_{\|\boldsymbol{Q}-\boldsymbol{q}\|\to\infty} \frac{S(\boldsymbol{q},\boldsymbol{Q})}{\|\boldsymbol{Q}-\boldsymbol{q}\|} \to +\infty \tag{35.1}$$

Then any orbit on T is minimizing.

Note that Condition (35.1) is implied by the convexity condition $\langle \partial_{12}S(\boldsymbol{q},\boldsymbol{Q}).\boldsymbol{v},\boldsymbol{v}\rangle \leq -a\|\boldsymbol{v}\|^2$ as can easily be seen by the proof of Lemma 27.2.

Proof. Since T is Lagrangian, it is the graph of the differential of some function plus a constant 1-form: $T = dg(\mathbb{T}^n) + \beta$ (see 57.4). Change coordinates so that T becomes the zero section: $(\boldsymbol{q},\boldsymbol{p}') = (\boldsymbol{q},\boldsymbol{p} - dg(\boldsymbol{q}) - \beta)$. If $F_0(\boldsymbol{q},\boldsymbol{p}) = (\boldsymbol{Q},\boldsymbol{P})$, then, in the coordinates $(\boldsymbol{q},\boldsymbol{p}')$, we have $\boldsymbol{Q}' = \boldsymbol{Q}, \boldsymbol{P}' = \boldsymbol{P} - dg(\boldsymbol{Q}) - \beta$. Thus a possible generating function is:

$$R(\boldsymbol{q},\boldsymbol{Q}) = S(\boldsymbol{q},\boldsymbol{Q}) + g(\boldsymbol{q}) - g(\boldsymbol{Q}) + \beta(\boldsymbol{q}-\boldsymbol{Q}).$$

Indeed

$$\begin{aligned} -\boldsymbol{p}' &= \partial_1 R(\boldsymbol{q},\boldsymbol{Q}) = \partial_1 S(\boldsymbol{q},\boldsymbol{Q}) + dg(\boldsymbol{q}) + \beta \\ \boldsymbol{P}' &= \partial_2 R(\boldsymbol{q},\boldsymbol{Q}) = \partial_2 S(\boldsymbol{q},\boldsymbol{Q}) - dg(\boldsymbol{Q}) - \beta \end{aligned}$$

We now show that R is constant on T, where it attains its minimum. We first note that:

$$\partial_1 R(\boldsymbol{q},\boldsymbol{Q}) = 0 \Leftrightarrow \boldsymbol{p} = dg(\boldsymbol{q}) + \beta \Leftrightarrow \boldsymbol{Q} = l(\boldsymbol{q})$$
$$\partial_2 R(\boldsymbol{q},\boldsymbol{Q}) = 0 \Leftrightarrow \boldsymbol{P} = dg(\boldsymbol{Q}) - \beta \Leftrightarrow \boldsymbol{Q} = l(\boldsymbol{q})$$

where $l(\boldsymbol{q}) = \pi \circ f(\boldsymbol{q}, dg(\boldsymbol{q}) + \beta)$ and π is the canonical projection. Hence $R(\boldsymbol{q}, l(\boldsymbol{q})) = R_0$ is constant , and $DR(\boldsymbol{q},\boldsymbol{Q})$ is non zero if $\boldsymbol{Q} \neq l(\boldsymbol{q})$. Since g is periodic, the superlinearity of S implies that R is coercive, *i.e.* $\lim_{\|\boldsymbol{Q}-\boldsymbol{q}\|\to\infty} R(\boldsymbol{q},\boldsymbol{Q}) \to \infty$. Since R has all its critical points on T, it must attain its minimum R_{min} there. It is now easy to see that the

$\boldsymbol{q}$ coordinates $\boldsymbol{q}_n, \ldots, \boldsymbol{q}_k$ of any orbit segment on T must minimize the action. Indeed, let $\boldsymbol{r}_n, \ldots, \boldsymbol{r}_k$ be another sequence of points of $\mathbb{T}^n$ with $\boldsymbol{q}_n = \boldsymbol{r}_n$, $\boldsymbol{q}_k = \boldsymbol{r}_k$. Then:

$$W(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_k) = \sum_{j=n}^{k-1} R(\boldsymbol{r}_j, \boldsymbol{r}_{j+1}) + g(\boldsymbol{q}_k) - g(\boldsymbol{q}_n) + \beta(\boldsymbol{q}_k - \boldsymbol{q}_n)$$
$$\geq (k-n)R_{min} + g(\boldsymbol{q}_k) - g(\boldsymbol{q}_n) + \beta(\boldsymbol{q}_k - \boldsymbol{q}_n) = W(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_k)$$

□

Remark 35.3 Arnaud (1989) (see also Herman (1990)) has interesting examples which show that the condition that the graph be Lagrangian is essential in Theorem 35.2. Consider the Hamiltonians on $T^*\mathbb{T}^2$ is given by:

$$H_\varepsilon(q_1, q_2, p_1, p_2) = \frac{1}{2}(p_1 - \varepsilon\cos(2\pi q_2))^2 + \frac{1}{2}p_2^2.$$

The torus $\{(q_1, q_2, \varepsilon\cos(2\pi q_2), 0)\}$ is made of fixed points for the corresponding Hamiltonian system, but it is not Lagrangian (exercise). A further perturbation $G_{\varepsilon,\delta}(\boldsymbol{q},\boldsymbol{p}) = H_\varepsilon(\boldsymbol{q},\boldsymbol{p}) + \delta\sin(2\pi q_2), 0 < \delta \leq \varepsilon$ of these Hamiltonians also provide counterexamples to the strict generalization of the Aubry-Mather theorem to higher dimensions: such systems have no minimizers of rotation vector 0. All the fixed points for the time 1 map have non trivial elliptic part.

C. Birkhoff's Graph Theorem

We now present a theorem of Birkhoff for two dimensional twist map which shows that invariant circles must be graphs. The topological proof we give, due to Katznelson & Ornstein (1997) is interesting in that it also offers a method of proof for the Aubry-Mather theorem, which we present in the next subsection. In subsection E, we sketch the generalization to higher dimensions of the graph theorem by Bialy and Polterovitch.

Theorem 35.4 (Birkhoff) *Let f be a twist map of the cylinder $\mathcal{A}$. Then:*

(1) **(Graph Theorem)** *Any invariant circle which is homotopic to the circle $C_0 = \{y = 0\}$ is a (Lipschitz) graph over C_0.*

(2) If two invariant circles C_- and C_+ homotopic to C_0 bound a region without other invariant circles, for any ϵ, there are (uncountably many) orbits going from ϵ-close to $C_\pm$ to ϵ-close to $C_\mp$.

This theorem was proved as two independent theorems by Birkhoff (1920).

Proof (after Katznelson & Ornstein (1997)). For both (1) and (2), we can assume the existence of an invariant circle, call it C_+. Take any circle C which is a graph over C_0 and which lies under C_+ . The image $f(C)$ of this circle by f may not be a graph anymore, but one can make a *pseudograph* $Uf(C)$ by *trimming* it, a process that we denote by U. We now describe pseudo graphs and the trimming map. Take all the points of $f(C)$ that can be "seen" vertically from above. This set forms the graph of a function which is continuous except for at most countably many jump discontinuities. Because of the positive twist condition, these jumps must always be downward as x increases: if C is given the rightward orientation, a vector tangent to C must avoid the cone Θ_v^+, by the ratchet phenomenon (see Lemma 12.1 in Chapter 2). Make a circle out of this graph by adjoining vertical segments at the jumps. This is $Uf(C)$. We call such a curve a *right pseudograph*: a curve made of the graph of a function $y = h(x)$ which is continuous except for downward jump discontinuities (the limit to the right $h(x^+)$ and the left $h(x^-)$ exist at each point and $h(x^-) \geq h(x^+)$), and by adjoining to this graph vertical segments to close the jumps.

We can apply f to a pseudograph C and trim it as we did for a graph. Because of the positive twist condition, the horizontal part of $Uf(C)$ is made of images under f of horizontal parts of C. Given a (right pseudo) graph C, we obtain a sequence of curves $C_n = (Uf)^n C$.

Lemma 35.5 $C_\infty = \lim_{n\to+\infty} \sup C_n$ *is an f-invariant graph, where* $\lim\sup$ *is taken in the sense of functions $y = h(x)$ with the obvious allowance for vertical segments.*

Proof. After one iteration of Uf on a (right pseudo) graph C, we get a pseudograph with a *downward modulus of continuity*: the ratchet phenomenon and the vertical cuts implies that, for any pair of points z and z' in the lift of $Uf(C)$, $z' - z$ is in a cone V of vectors

(x, y) with $y \geq \delta x$ if $x \leq 0$ and $y \leq \delta x$ if $x > 0$ (see Figure 35.1). This implies that C_∞ also has this modulus of continuity, and hence is a pseudograph.

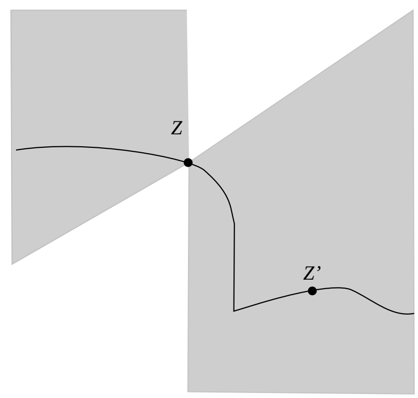

Fig. 35.1. The cone defining the modulus of continuity at a point z of $Uf(C)$.

There is a partial order on circles homotopic to $C_0 = \{y = 0\}$: we say that $C \preceq C'$ if C' is in the closure of the upper component of $\mathcal{A} \backslash C$, which we denote by $\mathcal{A}_+(C)$. Clearly f and U preserve this order, and $C \preceq U(C)$ for any circle C homotopic to $\{y = 0\}$. This implies that $f^n(C) \preceq Uf^n(C) \preceq C_\infty$ for all n, and hence $f(C_\infty) \preceq Uf(C_\infty) \preceq C_\infty$. By area preservation $f(C_\infty) = Uf(C_\infty) = C_\infty$.

If C_∞ were not a graph, its vertical segments would be mapped by f inside $\mathcal{A}_-(C_\infty) = \mathcal{A}_-(Uf(C_\infty))$, and $\mathcal{A}_-(C_\infty)$ would contain $\mathcal{A}_-(f(C_\infty))$ as a proper subset. This contradicts the fact that f has zero flux. Hence C_∞ is an f-invariant graph. $\square$

We now finish the proof of Birkhoff's theorems. Suppose that f admits an invariant circle C_* homotopic to the zero section C_0. We show that it is a (Lipschitz) graph. The region below C_* is invariant. Let C_{max} be the supremum of the invariant *graphs* in this region (under the partial order $\prec$). By continuity, C_{max} is an invariant circle which is a graph. But Proposition 12.3 implies that all invariant circles that are graphs are in fact Lipschitz graphs (again, the ratchet phenomenon). If $C_{max} \neq C_*$, then there exist a (not invariant) graph C with $C_{max} \prec C \prec C_*$. Applying the trimming iteration process to C, we get an invariant (Lipschitz) graph C_∞ with $C_{max} \prec C_\infty \prec C_*$. This contradicts the maximality of C_{max}. Hence $C_* = C_{max}$ is a Lipschitz graph.

If f does not admit any invariant circle homotopic to C_* other than the boundaries, the iteration process performed on any (right) pseudograph must converge to the upper boundary: we have $C \prec Uf(C)$. Since $C_\infty \subset closure(\cup f^n(C_*))$, on any graph ϵ close to the lower boundary, there is a point whose ω-limit set is in the upper boundary. We could have defined a trimming L of curves homotopic to the boundaries by taking their lower envelope (the points seen from below) instead of U. Then $L(C)$ is a *left pseudograph* and L preserves the order of circles and $L(C) \prec C$ for any curve C homotopic to the boundaries. Using $\liminf$ instead of $\limsup$ in the argument above, we get an iteration process $L \circ f$ which converges to an invariant graph, which must be the lower boundary this time. And on any graph ϵ close to the upper boundary, there is a point whose ω-limit set is in the lower boundary. $\square$

Remark 35.6 Performing both the Uf and Lf trimming processes on the same curve C yields points that come arbitrarily close to both boundaries in $forward\ time$. This fact was proven by Mather (1993) variationally and Hall (1989) topologically. See also LeCalvez (1990). The results of Mather and Hall are actually sharper. Mather also finds orbits whose α-limit set is in one boundary, the ω- limit set in the same or the other boundary . Hall uses the existence of such solutions asymptotic to the boundaries to replace the area preserving condition. Both authors find orbits "shadowing" any prescribed sequence of Aubry-Mather sets in a region of instability (technically, Hall shadows periodic orbits). It would be interesting to find a new proof of these results based on the trimming technique used above.

D*. Aubry-Mather Theorem Via Trimming

The above proof of Birkhoff's theorems appears as an aside in Katznelson & Ornstein (1997). They also recover the Aubry-Mather theorem with their trimming method. For this they define, abstractly, a different type of trimming operator, which they call *proper trimming*. Under proper trimming, the area below a curve is preserved. The main difficulty is to show the existence of such an operator. Once the existence is established, one takes limits of iterations under the map and the trimming operator. The limit is a pseudograph whose horizontal parts are $forward$ invariant under f. The Aubry-Mather sets are the intersection of all the forward images (by the map) of these horizontal parts. Finally, they show the existence of Aubry-Mather sets of all rotation numbers by applying this trimming procedure simultaneously to all the horizontal circles in the annulus. Fathi (1997) offers some analog to this in higher dimension, where he considers a certain semiflow on graphs of differentials on cotangent bundles (which are necessarily Lagrangian). In the limit, he recovers the generalized Aubry-Mather sets of Mather (which are described in Section 49).

E*. Generalizations of Birkhoff's Graph Theorem to Higher Dimensions

This section surveys the work of Bialy, Polterovitch and, indirectly, Herman on invariant Lagrangian tori. It will require from the reader knowledge of material dispersed throughout the book, and more. Bialy & Polterovitch (1992a) prove the following generalization to Birkhoff's Graph Theorem. We explain the terminology in the sequel.

Theorem 35.7 *Let F be the time one map of an optical Hamiltonian system of $T^*\mathbb{T}^n$, and let L be a smooth invariant Lagrangian torus for F which satisfies the following conditions:*
1) L is homologous to the zero section of $T^*\mathbb{T}^n$.
2) $F\big|_L$ is either chain recurrent or preserves a measure which is positive on open sets.
Then L is a smooth graph (i.e. a section) over the 0-section.

Optical (see Chapter 7) means that the Hamiltonian H is time periodic and convex in the fiber: H_{pp} is positive definite. Homologous to the zero section means that, together, the 0-section and the invariant torus bound a chain of degree $n+1$, presumably some smooth manifold of dimension $n+1$ in our case. As for Condition 2), it suffices here to say that either chain recurrence or existence of an invariant Borel measures are satisfied when the invariant torus is of the type exhibited by the KAM theorem, where the map $F\big|_L$ is conjugated to an irrational translation. In their paper, the authors use a condition that implies 2), as we show at the end of this section:
2') the suspension of $F\big|_L$ admits no transversal codimension 1 cocycle homologous to zero.
This theorem is a culmination of efforts by these authors, as well as by Herman (1990) who gives a perturbative version of this result as well as some important Lipschitz estimates for invariant Lagrangian tori. We now give a very rough idea of the proof of Theorem 35.7. First reduce the theorem to the case of an autonomous Hamiltonian on $T\mathbb{T}^{n+1}$ by viewing time as an extra $\mathbb{S}^1$ dimension, with the energy as its conjugate momentum (extended phase space).

Assume by contradiction that the invariant torus L is *not* a graph. Consider the set $S(L)$ of critical points of the projection $\pi\big|_L$. Generically, $S(L)$ consists of an $n-1$ dimension submanifold of L whose boundary is of dimension no more than $n-3$. Assume we are in the generic case. Then $S(L)$ can be cooriented by the flow: the Hamiltonian vector field is transverse to it on the invariant torus. This makes $S(L)$ a cocycle, i.e. a representative of a cohomology class of the torus. It turns out that this cohomology class is dual to the Maslov class of the torus L. [The *Maslov class* of L is the pull-back of the generator of $H^1(\Lambda(n))$ by the Gauss map, where $\Lambda(n)$ is the (Grassmanian) space of all Lagrangian planes in $\mathbb{R}^{2n}$. Prosaically, this means the following: *the oriented intersection of $S(L)$ with any closed curve on L counts how many "turns" the Lagrangian tangent space of L makes along the curve*. We explain that a little. The number of turns can be made quite precise because $\Lambda(n)$ has one "hole" around which Lagrangian spaces can turn ($H_1(\Lambda(n)) = \mathbb{Z}$)]. $S(L)$ is the set of points on L where the Lagrangian tangent space becomes vertical in some direction. The tangent space, seen as a graph over the vertical fiber, is given by a bilinear form which is degenerate at points of $S(L)$ and, thanks to the optical condition, decreases index (*i.e.* the dimension of the positive definite subspace increases) when following the flow at those points. Bialy and Polterovitch refer to Viterbo (1989) who proves that tori homologous to the zero section have Maslov class zero. Condition 2') now concludes: since it is homologous to zero, the cocycle $S(L)$ must be empty, *i.e.* there is no singularity in the projection $\pi\big|_L$ and the torus is a graph. The non generic case follows by making a limit argument using uniform Lipschitz estimates for invariant tori proven by Herman (1990).

Finally, let us show how the fact that $F\big|_L$ is measure preserving implies Condition 2'). Assume F is the time 1 map of an autonomous Hamiltonian system on $T^*\mathbb{T}^n$, L is an invariant torus and Ω is the volume form on L preserved by the Hamiltonian vector field X_H. The Homotopy Formula (see 59.7) $L_{X_H}\Omega = di_{X_H}\Omega + i_{X_H}d\Omega$ implies that $di_{X_H}\Omega = 0$. Assume X_H is transversal to S, a codimension 1 cocycle homologous to zero and let C be an n-dimensional chain that S bounds. Transversality implies $\int_S i_{X_H}\Omega \neq 0$. On the other hand, Stokes' Theorem yields $\int_S i_{X_H}\Omega = \int_C di_{X_H}\Omega = 0$. This contradiction implies that $S = \emptyset$. $\square$

Remark 35.8 As noted by Bialy and Polterovitch, it is not clear that Theorem 35.7 is optimal: Condition 2) may be unnecessary, as is the case in dimension 2. One could imagine a new proof of this theorem using higher dimensional trimming on Lagrangian pseudographs, which would not need this hypothesis...

36. (Un)Stable Manifolds and Heteroclinic orbits

A. (Un)Stable Manifolds

Consider two hyperbolic fixed point $\boldsymbol{z}^* = (\boldsymbol{q}^*, \boldsymbol{p}^*), \boldsymbol{z}^{**} = (\boldsymbol{q}^{**}, \boldsymbol{p}^{**})$ for a symplectic twist map F of $T^*\mathbb{T}^n$. We remind the reader that the *stable and unstable manifolds* at any fixed point $\boldsymbol{z}^*$ are defined as:

$$\mathcal{W}^s(\boldsymbol{z}^*) = \{\boldsymbol{z} \in T^*\mathbb{T}^n \mid \lim_{n\to+\infty} F^n(\boldsymbol{z}) = \boldsymbol{z}^*\},$$
$$\mathcal{W}^u(\boldsymbol{z}^*) = \{\boldsymbol{z} \in T^*\mathbb{T}^n \mid \lim_{n\to+\infty} F^{-n}(\boldsymbol{z}) = \boldsymbol{z}^*\}$$

Moreover the tangent space to $\mathcal{W}^s$ at $\boldsymbol{z}^*$ is given by the vector subspace $E^s(\boldsymbol{z}^*)$ of eigenvectors of eigenvalue of modulus less than 1, with a similar fact for $\mathcal{W}^u$ and E^u. In our case, the differential DF at the points $\boldsymbol{z}^*$ and $\boldsymbol{z}^{**}$ has as many eigenvalues of modulus less than 1 as it has of modulus greater than 1. Hence the stable and unstable manifolds at these points have both dimension n. The following appears in Tabacman (1993):

Proposition 36.1 *The (un)stable manifolds of a hyperbolic fixed point for a symplectic twist map are Lagrangian. Close to the hyperbolic fixed point, they are graphs of the differentials of functions.*

Proof. Consider a point $\boldsymbol{z}$ on the stable manifold of the hyperbolic fixed point $\boldsymbol{z}^*$, and two vectors $\boldsymbol{v}, \boldsymbol{w}$ tangent to that manifold at $\boldsymbol{z}$. Then:

$$\omega_{\boldsymbol{z}}(\boldsymbol{v}, \boldsymbol{w}) = \omega_{F^k(\boldsymbol{z})}(DF^k(\boldsymbol{v}), DF^k(\boldsymbol{w})) \to \omega^*_{\boldsymbol{z}}(0,0) = 0, \text{as } k \to \infty.$$

which, since it has dimension n in $T^*\mathbb{T}^n$, proves that the stable manifold is Lagrangian. The same argument, using F^{-k}, applies to show that the unstable manifold is Lagrangian. We leave the proof of the second statement to the reader (Exercise 36.2). □

In Exercise 36.3, the reader will show a generalization of this fact that makes it applicable to exact symplectic maps (not necessarily twist) of general cotangent bundles. The exercise will show that the (un)stable manifolds are in fact *exact Lagrangian*, *i.e.* the restriction of the canonical 1-form λ to the (un)stable manifolds is exact.

Exercise 36.2
b) Prove that the local (un)stable manifold of a hyperbolic fixed point $\boldsymbol{z}^*$ for a symplectic twist map F is a graph over the zero section (*Hint.* use the formula for the differential of F given in 25.5, and the twist condition det $(\partial_{12}S) \neq 0$ to show that the (un)stable subspace of $DF_{\boldsymbol{z}^*}$ cannot have a vertical vector. To do this, expend $\omega_{\boldsymbol{z}^*}(DF\boldsymbol{w}, \boldsymbol{w})$ assuming $\boldsymbol{w} = (0, w)$ and show that necessarily $w = 0$.)
c) Deduce from this that the (un)stable manifolds are graphs of differentials of functions Φ^u, Φ^s defined on a neighborhood of $\pi(\boldsymbol{z}^*)$ in the zero section.

Exercise 36.3 Let F is an exact symplectic map (not necessarily twist) of the cotangent bundle T^*M of some manifold: $F^*\lambda - \lambda = dS$ for some function $S : M \to \mathbb{R}$ (λ is the canonical 1 form on T^*M). In Appendix 1, it is shown that any Hamiltonian map is exact symplectic, and any composition of exact symplectic map is exact symplectic.
a) Show that the (un)stable manifolds $\mathcal{W}^{s,u}$ of a fixed point $\boldsymbol{z}^*$ are *exact Lagrangian* (immersed) submanifolds, *i.e.* $\lambda\big|_{\mathcal{W}^{s,u}} = dL^{s,u}$ for some functions $L^{s,u} : \mathcal{W}^{s,u} \to \mathbb{R}$. (*Hint.* Show that the integral of λ over any loop on $\mathcal{W}^{s,u}$ is 0).
b) Show that if and $\mathcal{W}$ is an exact Lagrangian manifold invariant under the exact symplectic map F, then:

$$S(\boldsymbol{z}) + constant = L(F(\boldsymbol{z})) - L(\boldsymbol{z}), \quad \forall p \in \mathcal{W}$$

c) Conclude that

$$L^u(\boldsymbol{z}^u) = \sum_{k<0} \left[S(F^k(\boldsymbol{z}^u)) - S(\boldsymbol{z}^*)\right], \quad L^s(\boldsymbol{z}^s) = -\sum_{k\geq 0} \left[S(F^k(\boldsymbol{z}^s)) - S(\boldsymbol{z}^*)\right].$$

For more on this approach, see Delshams & Ramírez-Ros (1997).

B. Variational Approach to Heteroclinic Orbits

As a consequence of Proposition 36.1, we obtain a variational approach to heteroclinic orbits. Let $\boldsymbol{z}^* = (\boldsymbol{q}^*, \boldsymbol{p}^*)$ be a hyperbolic fixed point. Let Φ^u, Φ^s defined on a neighborhood U^* of $\boldsymbol{q}^*$ be the functions whose differentials have for graphs the (un)stable manifolds of $\boldsymbol{z}^*$. We can add appropriate constants to these functions and get $\Phi^s(\boldsymbol{q}^*) = \Phi^u(\boldsymbol{q}^*) = 0$. In the proof of Theorem 35.2, we showed that the function $R(\boldsymbol{q}, \boldsymbol{Q}) = S(\boldsymbol{q}, \boldsymbol{Q}) + g(\boldsymbol{q}) - g(\boldsymbol{Q}) + \beta(\boldsymbol{q} - \boldsymbol{Q})$ was constant on the Lagrangian manifold $Graph(dg + \beta)$. Applying this to $g = \Phi^s$ or Φ^u, $\beta = 0$, we obtain

$$S(\boldsymbol{q}, \boldsymbol{Q}) = \Phi^s(\boldsymbol{Q}) - \Phi^s(\boldsymbol{q}) + \text{constant},$$

where $F(\boldsymbol{q}, \Phi^s(\boldsymbol{q})) = (\boldsymbol{Q}, \Phi^s(\boldsymbol{Q}))$ (this makes sense in a subset of U^*). Applying the equation to $(\boldsymbol{q}^*, \boldsymbol{q}^*)$ shows that the constant is $S(\boldsymbol{q}^*, \boldsymbol{q}^*)$. Hence

$$S(\boldsymbol{q}, \boldsymbol{Q}) - S(\boldsymbol{q}^*, \boldsymbol{q}^*) = \Phi^s(\boldsymbol{Q}) - \Phi^s(\boldsymbol{q})$$

for a point $(\boldsymbol{q}, \boldsymbol{Q})$ on the local stable manifold of $\boldsymbol{z}^*$. We now sum over the orbit $(\boldsymbol{q}_k, \boldsymbol{q}_{k+1})$ of the point $(\boldsymbol{q}, \boldsymbol{Q}) = (\boldsymbol{q}_0, \boldsymbol{q}_1)$ to get:

$$\sum_{k=0}^{N-1} [S(\boldsymbol{q}_k, \boldsymbol{q}_{k+1}) - S(\boldsymbol{q}^*, \boldsymbol{q}^*)] = \sum_{k=0}^{N-1} [\Phi^s(\boldsymbol{q}_{k+1}) - \Phi^s(\boldsymbol{q}_k)] = \Phi^s(\boldsymbol{q}_N) - \Phi^s(\boldsymbol{q}_0)$$

As $N \to \infty, \Phi^s(\boldsymbol{q}_N) \to \Phi(\boldsymbol{q}^*) = 0$ and thus the sum converges to $-\Phi(\boldsymbol{q}_0)$:

$$\sum_{k=0}^{\infty} [S(\boldsymbol{q}_k, \boldsymbol{q}_{k+1}) - S(\boldsymbol{q}^*, \boldsymbol{q}^*)] = -\Phi^s(\boldsymbol{q}_0). \tag{36.1}$$

Applying the same manipulations to the unstable manifold, using the fact that the generating function for F^{-1} is $-S(\boldsymbol{Q}, \boldsymbol{q})$, this leads to:

Proposition 36.4 *Let $\boldsymbol{z}^* = (\boldsymbol{q}^*, \boldsymbol{p}^*), \boldsymbol{z}^{**} = (\boldsymbol{q}^{**}, \boldsymbol{p}^{**})$ be two hyperbolic fixed points for the symplectic twist map F. Let U^* and U^{**} be neighborhoods of $\boldsymbol{q}^*$ and $\boldsymbol{q}^{**}$ on which the differentials of the functions Φ^u and Φ^s respectively give the unstable manifold of $\boldsymbol{z}^*$ and the stable manifold of $\boldsymbol{z}^{**}$. Then critical points of the function*

$$W(\boldsymbol{q}_0, \dots, \boldsymbol{q}_N) = \Phi^u(\boldsymbol{q}_0) + \sum_{k=0}^{N-1} S(\boldsymbol{q}_k, \boldsymbol{q}_{k+1}) - \Phi^s(\boldsymbol{q}_N), \qquad \boldsymbol{q}_0 \in U^*, \boldsymbol{q}_N \in U^{**}$$

are segments of heteroclinic orbits.

Proof. Left to the reader. □

With this set-up, Tabacman (1995) shows that, in the 2 dimensional case, any two local minima (*i.e.* fixed points) ξ and η of $\phi(x) = S(x, x)$ such that $\phi(\xi) = \phi(\eta) < \phi(x)$ for all $x \in (\xi, \eta)$, are joined by some trajectory.

Here is a sketch of a numerical algorithm also proposed (and used) by E. Tabacman to find heteroclinic orbits between two given hyperbolic fixed points $\boldsymbol{z}^*, \boldsymbol{z}^{**}$:

(1) Find a basis for the unstable plane E^u of DF at $\boldsymbol{z}^*$, and display the basis vectors as columns of a $2n \times n$ matrix $\begin{pmatrix} A \\ B \end{pmatrix}$

(2) The matrix $M = BA^{-1}$ is symmetric and E^u is the graph of the differential of the quadratic form $\boldsymbol{q} \mapsto \boldsymbol{q}^t M \boldsymbol{q}$. This function is an approximation to Φ^u (see 55.6.)

(3) Perform similar steps to approximate Φ^s at $\boldsymbol{z}^{**}$.

(4) Pick N (large enough) and use your favorite numerical method to search for critical points of the function W defined above, with points $\boldsymbol{q}_0, \boldsymbol{q}_N$ suitably close to $\boldsymbol{z}^*$ and $\boldsymbol{z}^{**}$ respectively.

(5) For more precision, make $\boldsymbol{q}_0$ and $\boldsymbol{q}_N$ closer to $\boldsymbol{z}^*$ and $\boldsymbol{z}^{**}$ (resp.) and increase N.

Note that this algorithm can be substantially improved by starting, using normal forms, with an approximation of higher degree than linear for the local (un)stable manifolds (see Simó (1990)).

C. Splitting of Separatrices and Poincaré–Melnikov Function

In Hamiltonian systems, the *Poincaré–Melnikov function* (actually an integral), measures how much the intersecting stable and unstable manifolds of two hyperbolic fixed points split. This kind of function has a long and rich history: Poincaré (1899) introduced it as a way to prove non-integrability in Hamiltonian systems. It has then been used to prove the existence of chaos (transverse intersections of stable and unstable manifolds often lead to “horseshoe” subsystems), and to estimate the rate of diffusion of orbits in the momentum direction. The discrete, two dimensional case was considered by Easton (1984), Gambaudo (1985), Glasser & al. (1989), Delshams & Ramírez-Ros (1996). Here, following Lomeli (1997), we give a formula for a Poincaré–Melnikov function for a higher dimensions symplectic twist map in terms of its generating function. A more general treatment, valid in general cotangent bundles, and which does not assume that the separatrix is a graph over the zero section, is given in Delshams & Ramírez-Ros (1997).

Theorem 36.5 *Let F_0 be a symplectic twist map of $T^*\mathbb{T}^n$ with hyperbolic fixed points $\boldsymbol{z}^* = (\boldsymbol{q}^*, \boldsymbol{p}^*), \boldsymbol{z}^{**} = (\boldsymbol{q}^{**}, \boldsymbol{p}^{**})$ such that a subset of $\mathcal{W}^u(\boldsymbol{z}^*) = \mathcal{W}^s(\boldsymbol{z}^{**}) = \mathcal{W}$ containing $\boldsymbol{z}^*, \boldsymbol{z}^{**}$ is the graph $\boldsymbol{p} = \psi(\boldsymbol{q})$ of a function ψ over some open set. Let S_0 be the generating function of F_0. Consider a perturbation F_ε of F_0 with generating function $S_\varepsilon = S_0 + \varepsilon S_1$ such that $S_1(\boldsymbol{q}^*, \boldsymbol{q}^*, 0) = S_1(\boldsymbol{q}^{**}, \boldsymbol{q}^{**}, 0) = 0$ and $\frac{d}{dq}\big|_{\boldsymbol{q}=\boldsymbol{q}^*} S_1(\boldsymbol{q}, \boldsymbol{q}, \varepsilon) = 0 = \frac{d}{dq}\big|_{\boldsymbol{q}=\boldsymbol{q}^{**}} S_1(\boldsymbol{q}, \boldsymbol{q}, \varepsilon)$. Then the function $L : \mathcal{W} \to \mathbb{R}$:*

$$L(\boldsymbol{q}) = \sum_{k\in\mathbb{Z}} S_1(\boldsymbol{q}_k, \boldsymbol{q}_{k+1}, 0) \quad \text{where } \boldsymbol{q}_k = \pi \circ F^k(\boldsymbol{q}, \psi(\boldsymbol{q})) \tag{36.2}$$

is well defined and differentiable. If L is not constant then, for ε small enough, the (un)stable manifolds of the perturbed fixed points of F_ε split. Their intersection is transverse at nondegenerate critical points of L.

Note that, whereas the condition $S_1(\boldsymbol{q}^*, \boldsymbol{q}^*, 0) = S_1(\boldsymbol{q}^{**}, \boldsymbol{q}^{**}, 0)$ is essential, the fact that their value is 0 is just normalization. Also, the condition on the nullity of derivatives can be discarded (see Delshams & Ramírez-Ros (1997)).

Proof. Work in the covering space $\mathbb{R}^{2n}$ of $T^*\mathbb{T}^n$. Let $\Phi : U \subset \mathbb{R}^n \to \mathbb{R}$ and $\psi = d\Phi$ be such that $Graph(\psi) = \mathcal{W}$. As in the proof of Theorem 35.2, change coordinates so that $\mathcal{W}$ lies in the zero section: $(\boldsymbol{q}, \boldsymbol{p}') = (\boldsymbol{q}, \boldsymbol{p} - \psi(\boldsymbol{q}))$. If $F_0(\boldsymbol{q}, \boldsymbol{p}) = (\boldsymbol{Q}, \boldsymbol{P})$, then, in the coordinates $(\boldsymbol{q}, \boldsymbol{p}')$, we have $\boldsymbol{q} = \boldsymbol{q},\ \boldsymbol{p}' = \boldsymbol{p} - \psi(\boldsymbol{q}), \boldsymbol{Q}' = \boldsymbol{Q}, \boldsymbol{P}' = \boldsymbol{P} - \psi(\boldsymbol{Q})$. Thus the generating function becomes:

$$S_{new}(\boldsymbol{q}, \boldsymbol{Q}) = S_{old}(\boldsymbol{q}, \boldsymbol{Q}) + \Phi(\boldsymbol{q}) - \Phi(\boldsymbol{Q}).$$

Note that the first order term S_1 remains the same under this change of coordinates, since we only added terms which are independent of ε. For ε small enough, the (un)stable manifolds $\mathcal{W}^u_\varepsilon, \mathcal{W}^s_\varepsilon$ of the perturbed fixed points $\boldsymbol{z}^*_\varepsilon, \boldsymbol{z}^{**}_\varepsilon$ (respectively) are graphs of the differentials $\psi^{u,s}_\varepsilon = d\Phi^{u,s}_\varepsilon$ for some functions $\Phi^{u,s}_\varepsilon$ of the base variable $\boldsymbol{q}$. Clearly, the manifolds $\mathcal{W}^{u,s}_\varepsilon$ split for ε small enough whenever the following *Poincaré–Melnikov* function:

$$M(\boldsymbol{q}) = \left.\frac{\partial}{\partial\varepsilon}\right|_{\varepsilon=0} (\psi^u_\varepsilon(\boldsymbol{q}) - \psi^s_\varepsilon(\boldsymbol{q}))$$

is not constantly zero, and their intersection is transverse if the differential DM is invertible at the zeros. We will now show that:

$$M(\boldsymbol{q}) = \frac{\partial L}{\partial \boldsymbol{q}}$$

where $L(\boldsymbol{q})$ is the function defined in (36.2) , expressed in our new coordinates. Formula (36.1) gives us expressions for $\Phi^{u,s}_\varepsilon$:

$$\Phi_\varepsilon^u(\boldsymbol{q}) = \sum_{k<0} \left[S_\varepsilon(\boldsymbol{q}_k^u(\varepsilon), \boldsymbol{q}_{k+1}^u(\varepsilon)) - S_\varepsilon(\boldsymbol{q}^{**}, \boldsymbol{q}^{**}) \right],$$
$$\Phi_\varepsilon^s(\boldsymbol{q}) = -\sum_{k\geq 0} \left[S_\varepsilon(\boldsymbol{q}_k^s(\varepsilon), \boldsymbol{q}_{k+1}^s(\varepsilon)) - S_\varepsilon(\boldsymbol{q}^{*}, \boldsymbol{q}^{*}) \right]$$

where $\boldsymbol{q}_k^u(\varepsilon)$ (resp $\boldsymbol{q}_k^s(\varepsilon)$) is the $\boldsymbol{q}$ coordinate of $F_\varepsilon^k(\boldsymbol{q}, \psi_\varepsilon^u(\boldsymbol{q}))$ (resp. of $F_\varepsilon^k(\boldsymbol{q}, \psi_\varepsilon^s(\boldsymbol{q}))$). We change the order of differentiation:

$$M(\boldsymbol{q}) = \left.\frac{\partial}{\partial\varepsilon}\right|_{\varepsilon=0} \frac{\partial}{\partial \boldsymbol{q}}(\Phi_\varepsilon^u(\boldsymbol{q}) - \Phi_\varepsilon^s(\boldsymbol{q})) = \frac{\partial}{\partial \boldsymbol{q}} \left.\frac{\partial}{\partial\varepsilon}\right|_{\varepsilon=0} (\Phi_\varepsilon^u(\boldsymbol{q}) - \Phi_\varepsilon^s(\boldsymbol{q})),$$

and compute one of these terms, using Formula (36.1) and abbreviating $\boldsymbol{q}_k^u = \boldsymbol{q}_k^u(0)$:

$$\begin{aligned}
&\left.\frac{\partial}{\partial\varepsilon}\right|_{\varepsilon=0} \Phi_\varepsilon^u(\boldsymbol{q}) \\
&= \left.\frac{\partial}{\partial\varepsilon}\right|_{\varepsilon=0} \sum_{k<0} \left[S_\varepsilon(\boldsymbol{q}_k^u(\varepsilon), \boldsymbol{q}_{k+1}^u(\varepsilon)) - S_\varepsilon(\boldsymbol{q}^{**}, \boldsymbol{q}^{**}) \right] \\
&= \sum_{k<0} \left[\partial_1 S_0(\boldsymbol{q}_k^u, \boldsymbol{q}_{k+1}^u) \frac{\partial}{\partial\varepsilon} \boldsymbol{q}_k^u(0) + \partial_2 S_0(\boldsymbol{q}_k^u, \boldsymbol{q}_{k+1}^u) \frac{\partial}{\partial\varepsilon} \boldsymbol{q}_{k+1}^u(0) + S_1(\boldsymbol{q}_k^u, \boldsymbol{q}_{k+1}^u) \right] \\
&= \sum_{k<0} S_1(\boldsymbol{q}_k, \boldsymbol{q}_{k+1}),
\end{aligned}$$

where in the last line we took advantage of $\partial_1 S_0(\boldsymbol{q}_k, \boldsymbol{q}_{k+1}) = \partial_2 S_0(\boldsymbol{q}_k, \boldsymbol{q}_{k+1}) = 0$: these are the $\boldsymbol{p}$ coordinates of an orbit on the zero section in our new coordinates. In the line before the last, the terms involving $S_\varepsilon(\boldsymbol{q}^{**}, \boldsymbol{q}^{**})$ disappeared because of our assumption on S_1. The same computation shows that $\left.\frac{\partial}{\partial\varepsilon}\right|_{\varepsilon=0} \Phi_\varepsilon^s(\boldsymbol{q}) = -\sum_{k\geq 0} S_1(\boldsymbol{q}_k, \boldsymbol{q}_{k+1})$. The proof that $\frac{\partial L}{\partial \boldsymbol{q}} = M(\boldsymbol{q})$ follows. □

Remark 36.6 We have only touched the surface of a vast subject here. Once a Melnikov function is found, one has to be able to show that it is non zero on specific examples. This is usually hard, even in dimension 2. Explicit computations often utilizes the fact that, in good situations, the complexified Melnikov function (think of $\boldsymbol{q}$ as complex in the above) is an elliptic function. As a result of such computations, one often finds (*eg.* for standard like maps) that the angle of splitting of the separatrices are *exponentially small in the perturbation parameter* ε, making numerical methods inapplicable. We let the reader consult Delshams & Ramírez-Ros (1996b), Delshams & Ramírez-Ros (1998), Glasser &

al. (1989), Gelfriech & al. (1994). In Gelfreich (1999), a precise formula for the estimate for the exponential splitting of separatrices in the standard map is proven, a culmination of years of work initiated by Lazutkin.

37.* Instability, Transport and Diffusion

A*. Some Questions About Stability

If one thinks of twist maps as local models for symplectic maps around elliptic fixed points, the problem of stability of these fixed points is directly related to the obstruction orbits of twist maps may encounter to drifting in the vertical (momentum) direction. In dimension 2, invariant circles obviously offer such obstructions. Three natural questions arise:

Question 1 *Are the invariant circles the only obstruction for orbits of twist maps to drift vertically? What if there are no invariant circles at all, can orbits drift to infinity on the cylinder?*

Question 2 *Do the invariant tori of higher dimensional (eg. KAM) tori offer any obstruction to the drift of orbits in the momentum direction, at least close to integrable?*

Question 3 *How can we detect when a system does not have invariant tori?*

B*. Answer to Question 1: Shadowing of Aubry-Mather Sets

Are the invariant circles the only obstruction for orbits of twist maps to drift vertically? What if there are no invariant circles at all, can orbits drift to infinity on the cylinder?

The answer is: Yes and Yes. The answer to the first part of the question is already given by Part (2) of Birkhoff's Theorem 35.4 , which says that, in a region bounded by two invariant circles, which contains no other invariant circle, there exist orbits going from one circle to the other (in whichever order). This fact gave rise to the the name *(Birkhoff) region of instability* for such regions.

The answer to the second part of the question (again Yes) follows from Mather (1993) and Hall (1989), who show that given any (infinite) sequence of Aubry-Mather sets, one can find an orbit that shadows it, *i.e.* stays at a prescribed distance from each Aubry-Mather set for a prescribed amount of time (the transition time is not controlled). In particular, for twist maps of the cylinder without any invariant circles, there exist orbits that are unbounded on the cylinder (take a sequence of Aubry-Mather sets going to infinity: such a sequence must exist if the twist is bounded from below). Note that Slijepčević (1999a) has recently given a proof of these results using the gradient flow of the action methods of Chapter 3.

Another approach to instability uses *partial barriers*: invariant sets made of stable and unstable manifolds of hyperbolic periodic orbits or Cantori. The *theory of transport* seeks to study the rate at which points cross these barriers. This theory was initiated by MacKay, Meiss & Percival (1984). The survey Meiss (1992) is beautifully written and encompasses the theory of twist maps of the annulus and transport theory. For other developments, see Rom-Kedar & Wiggins (1990) and Wiggins (1990). MacKay suggested that (the projection in the annulus of) ghost circles could be used as partial barriers.

C*. Partial Answer to Question 2: Unbounded Orbits

Do the invariant tori of higher dimensional (eg. KAM) tori offer any obstruction to the drift of orbits in the momentum direction, at least close to integrable?

The answer is: an ambiguous "No". Topologically, it is clear that, for $n > 1$, an n dimensional torus in $T^*\mathbb{T}^n$ does not separate the space into two disjoint components. However, this does not offer a guarantee that orbits *will* drift in the momentum direction, especially in the presence of a set of high measure of invariant tori. But right after proving his version of the KAM theorem, Arnold (1964) gave a stunning example of a possible limitation to the stability offered by KAM tori. This was an example of a family of Hamiltonian systems, in which, even close to integrable where the KAM theorem implied the existence of many invariant tori, some orbits drifted in the momentum direction. One of the fundamental questions remained: could the mechanism of diffusion detected in the highly nongeneric example of Arnold be found in generic systems? This paper has since generated an immense amount of work from mathematicians and physicists. In particular, much of the recent study of splitting of separatrices can be traced down to this paper, where chains of stable and unstable manifolds of invariant tori of lower dimensions ("whiskered tori") are used to construct drifting orbits. Also, as already mentioned, Nekhoroshev's theory measures the "exponential stickiness" of KAM tori, which implies that orbits must spend a very long time around these tori if they are caught in their neighborhood.

To my knowledge, the first decisive result in the direction of generic diffusion was given by Mather, who announced a striking result for Hamiltonian systems on $T^*\mathbb{T}^2$: For a C^r ($r \geq 2$) generic Riemannian metric g on $\mathbb{T}^2$ and C^r generic potential V periodic in time, the classical Hamiltonian system $H(\boldsymbol{q}, \boldsymbol{p}, t) = \frac{1}{2}\|\boldsymbol{p}\|_g^2 + V(\boldsymbol{q}, t)$ possesses unbounded orbits. Mather's proof departs from the traditional methods used in this problem. It powerfully brings together the constrained variational methods developed in Mather (1993) , the theory of minimal measures of Mather (1991b) (see also Chapter 9) as well as hyperbolic techniques. Delshams, de la Llave & Seara (2000) have given recently an alternate proof to this result, using hyperbolic techniques and methods of geometric perturbation. See also the related results of Bolotin & Treschev (1999), which use some mixture of variational and hyperbolic methods. Finally, de la Llave announced (Fall 1999) a generalization of this theorem to cotangents bundles of arbitrary compact manifolds. His method uses a generalizations of Fenichel's theory of perturbation of normally hyperbolic sets. Interestingly, the orbits found start at high energy levels, where the system is close to integrable. See also Xia (1998) , for some recent developments closer to the spirit of the original problem of diffusion and Bessi (1996), (1997) and (1998), a series of article where the full force of Mather's variational techniques are used to study diffusion, including the classical case of Arnold (1964) . As of this writing, these problems are the subject of some of the most active research in Hamiltonian dynamics.

D*. Partial Answer to Question 3: Converse KAM Theory

How can we detect when a system does not have invariant tori?

Since invariant tori have been associated to long term stability, many people have studied the mechanisms that leads to the breaking of such tori. There is an extensive body of work in dimension 2. A popular tool is Green's criterium, which associates the breaking of an invariant torus of a certain rotation number with the instability (hyperbolicity) of periodic orbits of nearby rotation numbers. Thanks to this criterium, MacKay & Percival (1985) showed that the standard map f_k (see Section 6) does not have any invariant circle for $k > 63/64$). Boyland & Hall (1987) shows that if there is a non cyclically ordered periodic orbit of rotation number ω, then there are no invariant circles of rotation number in an interval given by the Farey neighbors of ω. Mather (1986) relates the non existence of invariant circles to the variation of the action function on Aubry-Mather sets (see also Golé (1992 a) for a related result which uses the total variation of the action on ghost circles). MacKay (1993), studies the question in great length and gives a renormalization method that explains the universality of scaling of the gaps of Aubry-Mather sets near the breaking point of an invariant circle. Finally, MacKay & al. (1989) starts the study of converse KAM theory in symplectic twist maps. Haro (1999) provides some more, interesting results.