FOREWORD

8/31/99

To the reader of this draft: this is a preliminary version of the book, which needs to undergo a good amount of revision. Comments such as this one appear at the top of each chapter to warn my early reader, and remind myself, of some of the work needed. My homespun reference system is not completed yet. Chapters come with names as this point (because the numbers were variable). As an example, Proposition TOPOpropcz belongs to the appendix named TOPO and the number it corresponds to is at the end of that appendix. I have marked points that need my attention in the future by: ???. Where concepts are defined for the first time, I use the definition style. I will build an index listing all terms that appear in that font. This is not the format required by my publisher World Scientific. In the interest of saving paper, I kept the magnification low (the book would have 250 pages in the final format). As a result, some page transition could have been smoother. Finally, this version does not contains all the pictures.

Area preserving maps of the annulus first appeared in the work of Henri Poincaré (ref???) on the three-body problem. As two dimensional discrete dynamical models, they offered a handle for the study of a complicated Hamiltonian system. Since then, these maps and their more specialized offspring called twist maps, have offered many opportunities for rigorous analysis of aspects of Hamiltonian systems, as well as an ideal test ground for important theories in that field (eg. Moser (1962) proved the first differentiable version of the KAM theorem in the context of twist maps).

This book is intended for graduate students and researchers in mathematics and mathematical physics interested in the interplay between the theories of twist maps and Hamiltonian dynamical systems. The original mandate of this book was to be an edited version of the author's thesis on periodic orbits of symplectic twist maps of $\mathbb{T}^n \times \mathbb{R}^n$. While it now comprises substantially more than that, the results presented, especially in the higher dimensional case, are still very much centered around the author's work.

At the turn of the 1980's, the theory of twist maps received a tremendous boost from the work of Aubry and Mather. Aubry, a solid-state physicist, had been led to twist maps in his work on ground states for the Frenkel-Kontorova model. This system, which models deposition on periodic 1-dimensional crystals, while not dynamical, provides a variational approach which is surprisingly relevant to twist maps. Mather, a mathematician who had worked on dynamical systems and singularity theory, gave a proof of existence of orbits of all rotation numbers in twist maps, what is now known as the Aubry-Mather theorem, using a different variational approach proposed by Percival. Aubry, who had conjectured the result, gave a proof using his approach. Both researchers then developed a sophisticated theory using an interplay of their two approaches. This lead to a flurry of work in mathematics and physics.

4 FOREWORD

At about the same time, Conley & Zehnder (1983) gave a proof of the Arnold conjecture on the the torus, which heralded the birth of symplectic topology. This conjecture (now a theorem) states that the number of fixed point for a Hamiltonian map on a closed manifold is regimented in the same way as the number of critical points of real valued functions on that manifold. The proof involved Conley's generalized Morse theory for the study of the gradient flow of the Hamiltonian action functional in loop space. Later, with the influx of Gromov's holomorphic curve theory, this gave rise to Floer cohomology (Floer (???)). Interestingly, Arnold (1978) introduced his conjecture as a generalization of the famous fixed point theorem for annulus maps of Poincaré and Birkhoff, by gluing two annuli into a torus.

This book, while establishing a firm ground in the classical theory of twist maps, reaches out, via generalized symplectic twist maps, to Hamiltonian systems and symplectic topology. One of the approaches used throughout is that of the gradient flow of the action functional stemming from the twist maps' generating functions. We hope to convey that symplectic twist maps offers a relatively simple, often finite dimensional, interface to the variational and dynamical study of Hamiltonian systems on cotangent bundles.

Results for the two dimensional theory presented here include the classical theorems by Poincaré, Birkhoff (Chapter 7 and INVchapter), Aubry and Mather (Chapter AM). A joint work of the author with Sigurd Angenent on the vertical ordering of Aubry-Mather sets appears for the first time here (GCchapter). The approach of this book to the two dimensional theory is deliberately variational (except for Katznelson and Ornstein recent proof of Birkhoff's Graph Theorem in INVchapter) as I sought continuity between the low and high dimensions. Unfortunately, this choice leaves out the rich topological theory of twist maps and, more generally two dimensional topological dynamics. I refer the reader interested in the topological approach to Hall & Meyer (???), LeCalvez (1990) and the bibliography therein.

In higher dimensions, results by the author form the main focus of attention. These results are about existence of periodic orbits and their multiplicity for both symplectic twist maps and Hamiltonian systems on cotangent bundles (Chapter 4 and Chapter 7). The results on Hamiltonian systems use techniques of decompositions of these systems into symplectic twist maps. In Chapter 6, we provide the necessary connections between these maps and Hamiltonian and Lagrangian systems, some for the first time in the literature. In particular, M. Bialy and L. Polterovitch were kind (and patient!) enough to allow me to include their proof of suspension of a symplectic twist map by an optical Hamiltonian flow. Appendix 2 or TOPO establishes the parts of Conley's theory needed in the book, including some refinements that, to my knowledge, never appeared before. For readers uncomfortable with these topics, I try to motivate this appendix (chapter ???) by a hands—on introduction to homology and Morse theory. Chapter 9 presents Chaperon's proof of Arnold's conjecture on the torus, and the commonality between our methods and those of generating phases used in symplectic topology. Appendix 1 or SG, a self contained introduction of symplectic geometry, gathers (and proves most of) the results of symplectic geometry needed in the book.

The results in this book do not make minimizing orbits their central item. In fact, they often deliberately concern systems that cannot have minimizers (non positive definite twist). However, Chapter AMG is devoted to surveying the state of affairs in the generalizations of the Aubry-Mather theory to higher dimensions, where minimizers play a fundamental role. INVchapter, a poor substitute to a treatment that should occupy a volume on its own, surveys the theories of invariant tori (KAM theory and generalizations of Birkhoff's Graph Theorem by Bialy, Polterovitch and Herman), as well as that of splitting of separatrices.

The different topics in this book require different background from the reader. I have striven to make it possible for readers only interested in twist maps of the annulus or of $\mathbb{T}^n \times \mathbb{R}^n$ to read the sections pertaining to these topics with a minimum of reference to symplectic or Riemannian geometry, or to Conley's theory. On the other hand the appendices on symplectic geometry and topology are written, at least in part, with the novice in mind.

In this introduction, we tell three mathematical stories which introduce themes that are interwoven throughout the book. The first one is the evolution of the dynamics of conservative systems (the standard map here) as one pertubs them away from completely integrable. The second story is about the relationship between Lagrangian or Hamiltonian systems and symplectic twist maps, illustrated here by the connection between the billiard map and the geodesic flow on a sphere. The third story relates Poincaré's last geometric theorem to symplectic topology.

1. Fall From Paradise

Consider the map $F_0: \mathbb{R}^2 \mapsto \mathbb{R}^2$ given by:

$$F_0(x,y) = (x+y,y).$$

 F_0 shears any vertical line $\{x=x_0\}$ into the line $\{y\mapsto (x_0+y,y)\}$, of slope 1: as y increases, the image point moves to the right. We say that F_0 satisfies the *twist condition*. F_0 is linear with determinant 1 and hence is area preserving. Since $F_0(x+1,y)=F_0(x,y)+(1,0)$, this map descends to a map f_0 of the cylinder $S^1\times IR$. There, the x variable is seen as an angle. f_0 is called an area preserving twist map of the cylinder, or twist map in short. See Chapter 1 for a more detailed definition of twist maps. The map f_0 has an additional property that makes it special among twist maps: it preserves each circle $\{y=y_c\}$, on which it induces a rotation of angle y_c (measured in fraction of circumference). We say that f_0 is completely integrable. Completely integrable maps are the paradise lost of mathematicians, physicists and astronomers. Not only are the dynamics of such maps entirely understood, but the invariance of each circle $\{y=y_c\}$ assures that no point drifts in the vertical direction. In their original celestial mechanics settings, twist maps appeared as local models of sections of the Hamiltonian flow around an elliptic periodic orbit. In this setting, this lack of drift means stability of the orbit (and by extension, one hoped to establish the stability of the solar system...). Nearby points stay nearby under iteration of the map. Of course "real" systems are rarely completely integrable. But one of the driving paradigms in the theory of Hamiltonian dynamics is the study of how one falls from this completely integrable paradise, and how many of its idyllic features survive the fall.

Falling is easy. Perturb F_0 ever so slightly into an F_{ϵ} :

$$F_{\epsilon}(x,y) = \left(x + y - \frac{\epsilon}{2\pi} \sin(2\pi x), \ y - \frac{\epsilon}{2\pi} \sin(2\pi x)\right),$$

called the standard map. As the reader may check, the vertical lines are still twisted to the right, and the area is still preserved under F_{ϵ} . Looking at the computer pictures of orbits of F_0 and F_{ϵ} in Figure 1.1, we see what appear as invariant circles. We also see new features in the orbits of F_{ϵ} : some structures resembling collars of pearls (elliptic periodic orbits and their "islands"), interspersed with regions filled with clouds of points (chaos and diffusion due to intersecting stable and unstable manifolds of hyperbolic periodic orbits). We also see some "broken" circles made of dashed lines (Cantori or Aubry-Mather sets). These new features become more and more predominant as the value of ϵ increases: the elliptic islands bulge, the chaotic regions spread, and less and less circles appear unbroken. In fact, if $\epsilon > 4/3$, a theorem of Mather (1986) says that no invariant circle survives. However, the deep theory of Kolmogorov-Arnold-Moser (KAM, see INVchapter) implies that uncountably many invariant circles remain for small ϵ , those that have a very irrational rotation angle. In fact these circles occupy a set of large relative measure in the cylinder. A natural question arises: what happens to invariant circles once they break? The answer to this question, given by the Aubry-Mather theorem (see Chapter AM), is that invariant circles are replaced by invariant sets called Aubry-Mather sets whose orbits retain most of the features of those of invariant circles (cyclic order, Lipschitz graph regularity, rotation number and minimization of action). The Aubry-Mather sets with orbits of irrational rotation numbers form Cantor sets, sometimes called Cantori; those with rational rotation numbers usually contain hyperbolic periodic orbits and, depending on the authors' conventions, associated elliptic orbits. Of course the Aubry-Mather sets with their gaps form no topological obstruction to the vertical drift of orbits. In fact Mather (1991a) and Hall (1989) prove that, in a region with no invariant circle, one can find orbits visiting any prescribed sequence of Aubry-Mather sets. Hence these vestiges of stability have now become a stairway to drift and instability! The theory of transport (see Meiss (1992)) points at the regulatory role Aubry-Mather sets have on the rate of vertical diffusion of points.

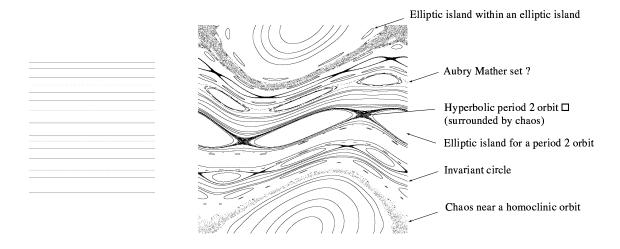


Fig. 1. 1. The different dynamics in the standard map: the left hand side shows a selection of orbits for the completely integrable F_0 , whereas the right hand side displays orbits for F_{ϵ} with $\epsilon = .817$. Of course, one has to take computer generated figures with a grain of salt: computers cannot deal with irrational numbers...

Higher Dimensions

Make $F_0:(x,y)\mapsto (x+y,y)$ defined above into a map of $\mathbb{R}^n\times\mathbb{R}^n$ by having x,y be vector variables. In analogy to the former situation, F_0 descends to a map f_0 from $\mathbb{T}^n\times\mathbb{R}^n$ to itself (x is now a vector of n angles). This space can be interpreted as the cotangent bundle of the torus, an important space in classical mechanics. Not only has the differential DF_0 determinant 1, but it also preserves the symplectic 2-form $\sum_k dx_k \wedge dy_k$ (the two notions were indistinguishable in dimension 2). The vertical fibers $\{x=x_c\}$ are still sheared, in a way made precise in Chapter STM. The map f_0 is called a *symplectic twist map* in this book. Our new f_0 is again called completely integrable as it preserves the tori $\{y=y_c\}$, and induces a translation by the vector y_c on each one. One can perturb f_0 (in the realm of symplectic twist maps) and ask the same kind of questions as in the 2-dimensional case: what of the well understood, stable dynamics of f_0 survives a perturbation of the map, small or large?

It turns out that KAM theory still holds in this case, and guarantees the existence of many invariant tori whose dynamics is conjugated to the translation by (very) irrational vectors. One of the results central to this book is that for arbitrary perturbations, periodic orbits of any rational rotation vector exist for all symplectic twist maps of a large class, and a lower bound on their number is related to the topology of \mathbb{T}^n (see Chapter 4). What about orbits of irrational rotation vector? Strictly speaking, there cannot be a full analog of the Aubry-Mather theorem in higher dimensions. Mather (1991b) developed a powerful theory of minimal invariant measures and their rotation vectors on cotangent bundles of arbitrary compact manifolds. This theory proves the existence and regularity of many minimizing orbits. But in the case where the manifold is \mathbb{T}^n with n > 3, the theory cannot guarantee that more than n directions be represented in the set of all rotation vectors of minimizing orbits. And indeed, some examples exist of maps (or Lagrangian systems) of $\mathbb{T}^3 \times \mathbb{R}^3$ all of whose recurrent minimizing orbits have rotation vector restricted to exactly 3 axes. If one lets go of the requirement that the orbits be action minimizers, then in certain examples, orbits of all rotation vectors can be found. The work of MacKay & Meiss (1992) points to a general theory for maps very far from integrable, but the case of maps moderately close to integrable, where less help from chaos can be expected, is not understood. Interestingly, if one trades the cotangent of a torus for that of a hyperbolic manifold, a large amount of the Aubry-Mather theory can be recovered: minimizing orbits of all rotation "direction", and of at least countably many possible speed in each direction exist (see Boyland & Golé (1996b)). Also, full fledge generalizations of the Aubry-Mather theorem exist in higher dimensional, but non dynamical settings generalizing the Frenkel-Kontorova model, as well as for some PDE's (de la Llave (1999)). We survey all these questions in greater detail in Chapter AMG.

2. Billiards and Broken Geodesics

Symplectic twist maps have rich ties with Hamiltonian and Lagrangian systems. They often appear as cross sections or discrete time snapshots of these systems. In Lagrangian systems, a trajectory γ is an extremal of an action functional $\int_{\gamma} Ldt$. In twist maps, this relates to an action function which is a discrete sum of the form $\sum S_k(x_k, x_{k+1})$ where x_k is a sequences of points of the configuration manifold and S_k are generating functions of twist maps. We explore this relationship in Chapter 6. A beautiful illustration of this occurs in the billiard map. The billiard we consider is planar, convex, and trajectories of a ball inside it are subject to

the law of equality between angle of reflection and angle of incidence. Since we know that it is a straight line between rebounds, a trajectory is prescribed by one of its points of rebound and the angle of incidence at this rebound. In this way, we obtain a map $f:(x,y)\mapsto (X,Y)$, where x is the coordinate of the point of rebound and $y=-cos(\theta)$, where θ is the angle of incidence (see Figure 2. 1). Since x is the point of a (topological) circle, and y is in the interval (-1,1), the map f acts on the annulus $\$^1\times (-1,1)$. The choice of y instead of θ insures that f preserves the usual area in these coordinates (see Section TWISTsecexamples). The twist condition for f is a consequence of the convexity of the billiard: if one increases y (i.e. increases θ) leaving x fixed, X increases.

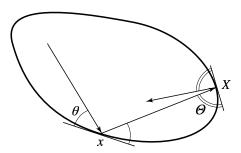


Fig. 2. 1. In a convex billiard, the point x and angle θ at a rebound uniquely and continuously determines the next point X and incidence angle Θ .

The map f can be seen as a limit of section maps for the geodesic flows of a sphere that is being flattened until front and back are indistinguishable. The boundary of the billiard is the (not so round in our illustration) fold of the flattened sphere. [To define the geodesic flow on the unit tangent bundle of the sphere, take a point on the sphere and a unit tangent vector (parameterized by its angle with respect to some tangent frame). Now travel at constant speed along the unique geodesic passing through this point and in the direction prescribed by the vector]. Draw on the sphere the closed curve C which eventually becomes the fold as one flattens the sphere. For a sufficiently flat sphere, all the geodesics on the sphere (except for maybe C, if it is a geodesic) eventually cross C transversally, and one can construct a section map which to one crossing at a certain point and angle makes correspond the next crossing point and angle. Seen in the three dimensional unit tangent bundle, the curve C lifts to a surface parameterized by points in C and all possible crossing angles in $(0, \pi)$, i.e. an annulus, which all trajectories (except maybe for C) of the geodesic flow eventually cross transversally. [Poincaré initiated a similar section map construction in a 3-dimensional energy manifold for the restricted 3-body problem]. The annulus maps that one obtains in this fashion limit, as one flattens the sphere, to the billiard map. To see this, note that the geometry of the flat sphere near a point not on the fold is that of the Euclidean plane, where geodesics are straight lines. At a fold point, the law of reflexion is a simple consequence of what happens to a straight line segment as it is folded along a line transverse to it (see Figure 2.2).

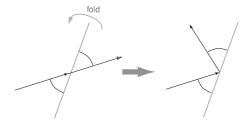


Fig. 2. 2. The law of reflexion as a consequence of folding.

Geodesics are length extremals among all (absolutely continuous) curves on the sphere. It therefore comes as no surprise that orbits of the billiard map are extremals of the length on the space of polygonal lines with vertices on the boundary (see Section TWISTsecexamples). If we inflate our billiard back a little, polygonal lines become broken geodesics. Indeed, the straight line segments can be replaced by segments of geodesic which, since the law of reflexion is not observed at a rebound for a general polygonal line, meet at an angle. In this space of broken geodesics, parameterized by the break points, geodesics are critical for the length function. To see why this is not only a beautiful, but also useful idea, consider the special case of periodic orbits of a certain period for the billiard map and geodesic flow. In the billiard, these correspond to closed polygons (see Figure 2.3), parameterized by their vertices which form a finite dimensional space, whose topology clearly has to do with that of the circle. The same holds for geodesics of our almost flat sphere. In fact, when studying closed geodesics (or geodesic between two given points) on any compact manifold one can restrict the analysis from the infinite dimensional loop space to a finite subspace of broken geodesics. This was a key idea in Morse's analysis of the path space of a manifold (see Milnor (1969)). And, more generally applied to Hamiltonian systems, it is one of the important themes of this book: symplectic twist maps can be used to break down the infinite dimensional variational analysis of Hamiltonian systems to a finite dimensional one. This is discussed in detail in Chapter 6, and again in Chapter 9.

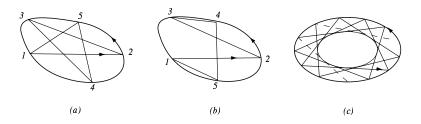


Fig. 2. 3. Different polygonal configurations in billiards: (a) is of period 5, rotation number 3/5 and is cyclically ordered. (b) is also of period 5, but of rotation 1/5 and is not cyclically ordered. Note that neither (a) nor (b) represent orbits since the law of reflexion is not satisfied. (c) is a configuration corresponding to an orbit on an invariant circle for the completely integrable elliptic billiard map. Its rotation number is presumably irrational.

Rotation Number and Ordered Configurations

The billiard map also provides a nice illustration of the notion of rotation number of periodic orbits (see Figure 2. 3 (a) and (b)). A consequence of the Aubry-Mather theorem is that any convex billiard has orbits of all rotation number in (-1,1). Polygonal curves corresponding to orbits on an invariant circle with irrational rotation numbers are all tangent to a circle or caustic inside the billiard (see Figure 2. 3 (c)). Polygonal curves corresponding to Aubry-Mather sets are "tangent" to a Cantor set. Finally, the billiard gives us an illustration of the notion of order for configurations of points. In Example (a) of Figure 2.3, the configuration is eyelically ordered, in that the cyclic order of rebound points is conserved on the boundary after following them to their next rebound. Example (b) is, on the other hand not cyclically ordered. This notion of order is crucial to the Aubry-Mather theory. In Chapter AM, we present a proof of the Aubry-Mather theorem similar to that of Aubry's, in which one finds (cyclically ordered) orbits of irrational rotation numbers by taking limits of cyclically ordered periodic orbits. In GCchapter, we make use of the fact that the gradient flow for the action function (the length in the billiard map) of a twist map preserves the set of cyclically ordered configurations to give another proof of the Aubry-Mather theorem. We use a stronger order property of the flow as well in our proof that Aubry-Mather sets are vertically ordered. Unfortunately, there is no natural order for orbits of higher dimension twist maps. But the same kind of ordering exists in higher dimensional Frenkel-Kontorova models, for which the Aubry-Mather holds, as well as for certain PDEs (see de la Llave (1999)). The gradient flow in the PDE setting corresponds to generalized heat flows. The analogy to the preservation of order is given by theorems of comparison. This analogy, which was already noticed by Angenent (1988), inspired him to introduce the gradient flow of the action in twist maps.

3. An Ancestor of Symplectic Topology

At the end of his life, Poincaré (1912) published a theorem, sometimes called his last geometric theorem, that can be simply stated as: Let f be an area preserving map of a compact annulus, which moves points in opposite directions on the two boundary circles. Then f must have at least two fixed points.

Poincaré gave an incomplete proof of this theorem, writing a moving letter of apology to the editor which mentions his bad health and expresses his desire that his work on this problem not be lost for posterity. Birkhoff gave a substantially different proof, which was also somewhat incomplete as to the existence of at least two fixed points (it did prove the existence of at least one). Since then, a number of proofs have appeared (Brown (???pb), Fathi (1983), Franks (1988), as well as Golé & Hall (1992), where the original proof of Poincaré is completed). We now sketch a proof of the theorem, in the very simple case where the map f also satisfies the twist condition. The ideas involved connect the original proof of Poincaré, the proof of LeCalvez (astérisque) we present in Section PBsecpb and the modern theory of symplectic topology.

Let F be the lift of f to the strip $\mathcal{A} = \{(x,y) \mid x \in \mathbb{R}, y \in [0,1]\}$, which moves boundary points in opposite directions. Such a lift always exists. Denote by (X,Y) the image of a point (x,y) by F. Consider

$$\Gamma = \{(x, y) \in \mathcal{A} \mid X(x, y) = x\},\$$

which is the set of points that only move up or down under the $map^{(1)}$. The twist condition means that the image of each vertical segment $\{x=x_0\}$ by F intersects that segment exactly at one point. This implies that

¹ Poincaré considered the similar set of points that only moved left or right, see Golé & Hall (1992)

 Γ is a graph over the x-axis, and, by periodicity, the lift of a circle γ enclosing the annulus. Clearly, $f(\gamma)$ must also be a circle, graph over the x-circle. Any point in the intersection $\gamma \cap f(\gamma)$ is necessarily fixed by f: such points move neither left, right, nor up, nor down. This intersection is not empty, by area conservation. If $\gamma = f(\gamma)$ (as is the case if f is a completely integrable map), f has infinitely many fixed points. If not, area preservation dictates that there must be points of $f(\gamma)$ strictly above γ and others strictly below. Since both these sets are circles, this implies the existence of at least two points in the intersection, i.e. two fixed points for f.

We now show the connection between fixed points of f and critical points of a real valued function on the circle. As we will see in Chapter 1, the map F comes equipped with a generating function S(x,X) which satisfies S(x+1,X+1)=S(x,X) and YdX-ydx=dS. This derives directly from area preservation and conservation of boundaries. Consider the restriction f of f to f, i.e. f to f, i.e. f to f to f, i.e. f to f the critical points of f to f to f the critical points of f to f the critical points, unless f is shown that a generic symplectic maps has infinitely many periodic orbits around an elliptic fixed point. Arnold (1978) also motivates his famous conjecture on fixed points on symplectic manifolds by an argument similar to this one.

In the coordinates (x, y') = (x, y - y(x)), Γ becomes the 0-section $\{(x, 0)\}$, and $\Gamma(\Gamma) = \{(x, Y(x) - y(x))\}$ y(x) is the graph of the differential of w. The function w is called a generating (phase) function for the manifold $F(\Gamma)$. This is a simple instance of a more general situation: Γ and its image are Lagrangian manifolds, as is any 1-dimensional manifold in a 2-dimensional symplectic manifold (see Appendix 1 or SG). Important theorems in symplectic topology can be expressed, as this one, in terms of intersections of a Lagrangian manifold with the 0-section in some cotangent bundle. To find such intersections, one looks for critical points of generating phase functions for this manifold. As we have seen in the above example, it is easy to do so when the manifold is a graph over the 0-section. The first challenge is to deal with cases where a Lagrangian manifold is not a graph [This will occur for our sets Γ and $F(\Gamma)$ when the map f is not twist, for example]. One then seeks generating phase functions with extra variables [This is in effect what the proof of LeCalvez does: one obtains a generating phase function for the set $F(\Gamma)$ by adding the generating functions of the twist maps that decompose f]. The second challenge is to show that these general generating phase functions have the requisite number of critical points. This is done in this book using Conley's theory on the gradient flow of the generating phase function. One difficulty arises from the non compactness of the space on which this function is defined. One resolves that by seeking compact invariant set of a sufficiently complicated topology for the gradient flow. We called some of these sets "ghost tori" in Golé (1989). These sets have their analogs in the sets of connecting orbits between critical points of the action functional in loop space that Floer based his cohomology on. Although implicit in several parts of the book, we will not use the language of Lagrangian intersection and generating phase function before Chapter 9.

Section INTROsecpb is 3.0