Aubry-Mather Sets and a Problem of Three Bodies

Harish Bhat

April 5, 2000

To Amma and Appa

Contents

Chapter 1

Introduction

The three-body problem of celestial mechanics has, over the course of the past three centuries, given rise to many incredible developments in mathematics. The modern theory of dynamical systems grew out of such developments, initiated mainly by Henri Poincaré and further developed by G.D. Birkhoff and others. Their work was motivated by classical problems in celestial mechanics which classical methods were not powerful enough to solve. These classical problems and methods stemmed from the search for explicit solutions to systems of differential equations, particularly those systems which govern the dynamics of mutually interacting heavenly bodies subject only to Newtonian gravitational forces. These explicit solutions, when they can be found, tell us exactly what will happen to our system as time passes.

What if an explicit "closed-form" solution cannot be found? In some situations, one can even prove rigorously that such solutions do not exist. Either way, we need to reevaluate what we wish to know about our system. The dynamicist would like to be able to say whether a given initial configuration leads to periodic orbits, quasiperiodic orbits, repeated collisions between the bodies, the escape of one or more bodies, or some other situation. We would also like to know how much the long-term evolution of the system is changed by small changes in the system's initial configuration. The "dynamical systems" approach to celestial mechanics consists of analyzing the long-term qualitative and asymptotic behavior of this system of differential equations, in such a way that the explicit solution of the system never needs to be written down.

1.1 From Poincaré maps to twist maps

One way this is done is by looking at the time-evolution of regions of the system's phase space, rather than of individual points. The evolution of a system with given initial conditions is given by a curve in phase space. If one initial condition leads to periodic behavior, then nearby initial conditions will yield trajectories that almost close on themselves. Suppose we fix some point on the periodic orbit and take a codimension one slice of phase space (transverse to the periodic trajectory). Define a new mapping as follows: given a point on the codimension one slice, follow its trajectory until it intersects the slice again. This point of intersection is the image of the original point; the resulting map is called the Poincaré return map for the system. Now instead of a continuously evolving system of differential equations, we have a discrete mapping defined in a small neighborhood of a point on a manifold of one less dimension than our original phase space.

The Poincaré map often gives us deep insights into the system's dynamics. In the case of the planar isosceles three-body problem, two scientists were able to couple a Poincaré map for the system with a novel partition of phase space to give a global analysis of the system's dynamics [13]. They were able to do all of that without solving the rather nasty nonlinear differential equations for this Hamiltonian system.

Continuing with this problem, which will contextualize the discussion of twist map dynamics that is to follow, we find that almost every trajectory that begins with a *syzygy crossing* (which occurs when all three masses are collinear) eventually returns to a syzygy crossing. We will see that the phase space is four-dimensional, implying that the fixed energy surfaces are three-dimensional. Thus a codimension one slice (obtained by fixing the y-coordinate of the third mass to equal the constant y-coordinate of the other two masses) is a two-dimensional surface, which we will see is actually just the right half of the unit disc in the plane. Liouville's theorem on the volume-preservation of Hamiltonian flows implies that the Poincaré map on this slice will preserve area.

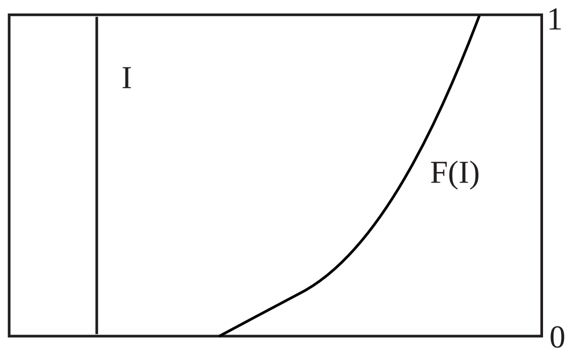
Suppose we can show the existence of an elliptic fixed point p of the Poincaré map f, a fixed point for which Df has a pair of complex conjugate eigenvalues. Then in a neighborhood of p there exists a symplectic change of variables which transforms f into the *Birkhoff normal form.* As long as the coefficients in this expansion are not all zero, f turns out to be a specific kind of area-preserving diffeomorphism of the annulus which fits the following definition:

Definition 1.1.1. Consider a diffeomorphism f of the closed annulus (or cylinder) $A = S^1 \times [0, 1]$. Then f is an *area-preserving twist map* if

1. when we lift $f : A \to A$ to $F : S \to S$, where S is the universal cover $S = \mathbb{R} \times [0, 1]$ of A, $\frac{\partial}{\partial y} F_1(x, y) > 0$[1]
2. f preserves the boundary of the annulus[2]
3. f preserves orientation.
4. f preserves area[3]

We will call such maps "twist maps" for short. The first condition above is known as the "twist condition." The geometric meaning of the twist condition is that any vertical segment on S is rotated to the right by the lift F, which means that on the annulus, any radial segment is rotated counterclockwise.

Twist maps do not arise only from theoretical problems of celestial mechanics. Such maps arise naturally in the classical billiards problem, dynamical analyses of particle beam orbits in high-energy accelerators, and in the plasma physics of fusion reactors. Finally, any Lagrangian system with two degrees of freedom, integrable or non-integrable, gives rise to geodesic flow on a smooth surface. After "squashing" this surface to a plane convex curve C, one can show that the squashed geodesic flow is just Birkhoff's billiards problem with boundary C [5].



[1]For a twist *homeomorphism*, we reformulate the first condition as follows: for each fixed $x \in \mathbb{R}$, $F_1(x, y)$ is a strictly increasing function of y.

[2]This is for convenience again, and not strictly necessary. In case f does not preserve the boundary components, f can be extended to a twist map on the infinite cylinder. We describe this construction at the end of Chapter 2.

[3]David Bernstein [3] showed that this condition is unnecessary for the existence of special orbits (such as Birkhoff periodic orbits and Aubry-Mather-like Cantor sets) to be guaranteed.

1.2 Integrable twist maps

For now, let us explore the most elementary classification of twist maps. The simplest twist maps are the integrable twist maps. A twist map $f(x,y)$ is integrable if its lift can be written as $F(x,y) = (x + h(y), y)$ for some function h. We can say quite a bit about such maps.

For a fixed value of y, $f(x,y)$ rotates each point on the circle $S^1 \times \{y\}$ through an angle of $h(y)$. Consider a lift of f to $S = \mathbb{R} \times [0,1]$, denoted by $F(x,y) = (F_1(x,y), F_2(x,y))$. By the twist condition, for a fixed x, $F_1(x,y)$ must be an increasing function of y. Therefore, $h(y)$ must be a strictly increasing function on $[0,1]$. To proceed any further with our analysis of this integrable twist map, we must look at what exactly f does to each of the circles $S^1 \times \{y\}$. Specifically, we would like to know if for a fixed y, whether there are any periodic points (x,y) such that $f^q(x,y) = (x,y)$. More generally, we would like to determine the *orbit* $\omega(x,y) = \{(a,b) : (a,b) = f^m(x,y) \text{ for some } m \in \mathbb{Z}\}$. This involves the theory of circle maps, the first step of which is to define the *rotation number* for a homeomorphism of the circle.

We will see later how to precisely define this rotation number. Given a homeomorphism $g : S^1 \to S^1$, the rotation number $\rho(g)$ measures the average amount that a point on the circle is rotated by g. For a circle mapping, the rotation number is measured by looking at successive iterates of *one point* on the circle. It turns out that the rotation number number is a function only of the mapping itself, not of which point of S^1 we focus our attention on. For any simple rotation map whose lift to $\mathbb{R}$ is the translation $x \mapsto x + \sigma$, the rotation number of the map is just σ. For a circle map g which has a periodic point z, we would have $g^q(z) = z$ for some $q \in \mathbb{Z}$. Let G be a lift of g to $\mathbb{R}$, and let w denote the lift of z. Then $G^q(w) = w + p$ for some $p \in \mathbb{Z}$, and it turns out that the rotation number of g equals p/q. The analogue of a rotation number for a twist map is the *twist interval*, which tells us the range of rotation numbers that make sense for a twist map on the annulus. Think of the annulus as being made up of a continuum of circles; the restriction of the integrable twist map to each such circle yields a rotation number, and the collection of all such rotation numbers gives the twist interval. For a general twist map τ, note that the restriction of τ to $S^1 \times \{0\}$, denoted by τ_0, is a homeomorphism of the circle—the same is true for the restriction of τ to $S^1 \times \{1\}$, denoted by τ_1. From the definition of the rotation number and the twist condition, we will see that the twist interval for a general twist map is just $[\rho(\tau_0), \rho(\tau_1)]$.

Returning to our integrable twist map f, note that for each y, f restricted to $S^1 \times \{y\}$ is a simple rotation with rotation number $h(y)$. Since h is strictly increasing on $[0,1]$, the twist interval for an integrable twist map is just $[h(0), h(1)]$. For each $r \in [h(0), h(1)]$, there is a corresponding y_r such that $h(y_r) = r$. We are now in a position to answer conclusively the question about orbits posed earlier. If r is rational, then on the corresponding circle $S^1 \times \{y_r\}$, each orbit $\omega(x)$ is a periodic orbit. To see this, write $r = p/q$; then $f^q(x, y_r) = (x + p, y_r) = (x, y_r)$ where we take S^1 to be $\mathbb{R}/\mathbb{Z}$. If r is irrational, then each orbit $\omega(x)$ is a dense subset of the circle.

Putting all of this information together, we can summarize the dynamics of the integrable twist map as follows: each value of $y \in [0,1]$ corresponds to an invariant circle $S^1 \times \{y\}$, and the orbit of a point on a given circle $S^1 \times \{y\}$ is either periodic, if $h(y)$ is rational, or a dense subset of the circle, if $h(y)$ is irrational. We say that the annulus is foliated by invariant circles.

In the classical theory of mechanics, the concept of "integrability" is associated with the existence of conserved quantities, or integrals of motion. The conserved quantity for the integrable twist map is simply h—along any invariant circle, h is conserved. Moreover, to find the invariant circles, we simply to have to search for sets on which h is constant. The existence of this conserved quantity allows us to *reduce* the two-dimensional twist map dynamics down to the one-dimensional dynamics along each invariant circle. We will encounter this theme later.

The results we have for the integrable twist map motivates questions concerning more general twist maps. For a general twist map τ, when are we assured of an invariant circle? Are there other kinds of invariant sets? Given a particular rotation number α in the twist interval, is there an invariant set of a particular kind on which τ has rotation number α? These qualitative questions are typical questions for a dynamicist; for example, the answers to such questions assisted Birkhoff, Arnold, and other dynamicists in gaining qualitative knowledge about the orbits of the three-body problem.

1.3 Non-integrable twist maps

We will now describe some of the results which come to bear on these questions. Recall that the twist interval for a general twist map τ is $[\rho(\tau_1), \rho(\tau_2)]$. In the non-integrable case, we are no longer assured that each circle $S^1 \times \{y\}$ is an invariant circle. This means that if we are to try to show the existence of periodic orbits that correspond to rational numbers in the twist interval, we have to take a different path. This path cannot rely on the invariant circles that came so easily to us in the integrable case.

1.3.1 Birkhoff periodic orbits

In search of periodic orbits of τ, we would like to look for an orbit which behaves under τ as if it was a periodic orbit of a simple rotation map. In other words, we seek a point $w = (x, y) \in A$ such that $\tau^q(w) = w$ for some $q \in \mathbb{N}$. This is just the general definition of a periodic orbit for τ. What would make it more like the periodic orbit of a rotation map? Think of the rotation map on $S^1 = \mathbb{R}/\mathbb{Z}$ whose lift is given by $\theta \mapsto \theta + p/q$. Here p is the integer such that for a lift T of τ, $T^q(z) = z + (p, 0)$, where $\pi(z) = w$. For a periodic orbit of a twist map, if we ignore the radial coordinates of the points, we are left with the projection of the orbit onto the circle. Suppose the resulting orbit on the circle is in the same order on the circle as the set of points $\{x, x + (p/q), x + (2p/q), ...\}$. If, furthermore, the orbit can be parameterized in a particular way by its geometric ordering on the circle, then the periodic orbit of the twist map is a *Birkhoff periodic orbit of type* (p, q).

The concept of rotation number for circle maps extends naturally to give us a rotation number of any periodic orbit on the annulus. For example, given a Birkhoff periodic orbit of type (p, q), we have that $F^q(x_0, y_0) = (x_0 + p, y_0)$, using the same notation as above. Ignoring the y-coordinates, we have a periodic orbit on the circle such that q iterations of x_0 bring it around the circle p times before returning. The earlier discussion of rotation number applies, and we say that the Birkhoff periodic orbit has rotation number p/q.

Once we define Birkhoff periodic orbits, we can prove their existence. It turns out that any rational number p/q in the twist interval has two[4] associated Birkhoff periodic orbits of type (p, q). The proof of this theorem involves two stages: first we set up an action functional on a space of sequences, and then we show that this functional has both a global minimum and a saddle point. The procedure generalizes "Hamilton's principle of least action" from classical mechanics. These two critical points correspond to two Birkhoff periodic orbits. It turns out that the existence of a Birkhoff periodic orbit for each rational number in the twist interval will enable us to show the existence of a particular kind of quasiperiodic orbit associated with each irrational number in the twist interval.

Is this the only kind of periodic orbit that corresponds to the rational rotation number p/q? It turns out that the answer is no. The existence of non-Birkhoff periodic orbits with rotation number p/q, however, is equiv-

[4]For a twist *homeomorphism* which is not of class C^1, only one such orbit of type p/q is guaranteed. However, the Aubry-Mather theorem still holds for such maps.

alent to the non-existence of an invariant circle with rotation number α, where α is an irrational such that a suitable truncation of its continued fraction expansion equals p/q. This result was obtained by Boyland and Hall [6]. (Note that the rotation number of an invariant circle of a twist map is easy to define. By an invariant circle we mean a simple closed curve γ in the annulus such that $\tau(\gamma) = \gamma$. Hence the restriction of τ to γ is a circle homeomorphism with well-defined rotation number; this fixes the rotation number of γ.)

One way to decide the question of whether an invariant circle of a particular rotation number exists is to appeal to KAM theory. The standard KAM theorem shows the existence of invariant tori and quasiperiodic orbits for near-integrable Hamiltonian systems. The version of the KAM theorem that applies in this case is Moser's twist theorem [12] which applies to sufficiently small and sufficiently smooth perturbations of integrable twist maps. For such maps, the theorem guarantees the existence of invariant circles that correspond to "sufficiently irrational" rotation numbers.

1.3.2 Aubry-Mather orbits

Aside from the small class of twist maps that are near-integrable, what can we say about the irrational numbers in the twist interval? The natural pairing between rational numbers and periodic orbits finds its analogue in the pairing between irrational numbers and Aubry-Mather sets. The definition of an Aubry-Mather set is motivated by similar concerns as for the Birkhoff periodic orbit. Specifically, we want the twist map τ to preserve the cyclic order of points in the Aubry-Mather set. This is to say that if (x_i, y_i) and (x_j, y_j) are two distinct points on the orbit, and $\hat{x}_i > \hat{x}_j$ where $\pi(\hat{x}_l) = x_l$, then we want $\pi_1 \circ \tau(x_i, y_i) > \pi_1 \circ \tau(x_j, y_j)$, where π_1 denotes projection onto the first factor.

Definition 1.3.1. Given a twist map τ, the closed invariant set I is an *Aubry-Mather set* if:

1. if for all $(x_1, y_1), (x_2, y_2) \in I$, $y_1 = y_2$ whenever $x_1 = x_2$,
2. τ preserves the S^1-ordering of the x-coordinates of all points in I, and
3. I is minimal, i.e. any other closed invariant set of τ contains I.

Using the fact that any Aubry-Mather set I is a subset of the graph of a Lipschitz function, we can show that the successive x-coordinates of points in I, $\{x_0, x_1, x_2, ...\}$, are successive iterants of a circle homeomorphism. In symbols, there exists a homeomorphism $g : S^1 \to S^1$ such that $g^n(x_0) = x_n$.

As with Birkhoff periodic orbits, we can define the rotation number of an Aubry-Mather set. Suppose we start with an Aubry-Mather set E in the annulus and remove any information about the y-coordinates of the set's points, thus projecting the set onto the circle. Denote this set by $\tilde{E}$. Then a lemma shows that the restriction of the twist map to E projects in the same way to a homeomorphism of $\tilde{E}$. By extending the map linearly on the gaps between points of $\tilde{E}$, we obtain a homeomorphism of S^1. The rotation number of this circle homeomorphism is defined to be the rotation number of the Aubry-Mather set E. Another way to state this is that the restriction of the twist map to the Aubry-Mather set E is topologically conjugate to an appropriate restriction of an (orientation-preserving) homeomorphism of the circle. Again, the rotation number of this circle homeomorphism is the rotation number of E.

We are now in a position to state the main Aubry-Mather theorem:

Theorem 1.3.1. *Given an area-preserving twist map τ, for any irrational number α in the twist interval $[\rho_0, \rho_1]$, there exists either an Aubry-Mather set E of τ with rotation number α.*

With this theorem, we see that the rational/irrational distinction between numbers in the twist interval corresponds to the distinction between (Birkhoff) periodic and Aubry-Mather orbits of the twist map.

To illustrate these results, we focus our attention on a special problem of celestial mechanics, the planar isosceles three-body problem. The specific problem we consider is concrete and easy to visualize, and it simultaneously yields a clear picture of the Aubry-Mather sets we wish to focus on. After briefly tracing the reduction of the problem, the Poincaré map construction, and the behavior of this map in a neighborhood of an elliptic fixed point, we will be in a position to return to circle maps, twist maps, and Aubry-Mather sets. Having understood these ideas, we will be in a position to explain in greater detail the dynamics of the planar-isosceles three body problem near the elliptic fixed point.

Chapter 2

Celestial Mechanics

2.1 The Planar Isosceles Three-Body Problem

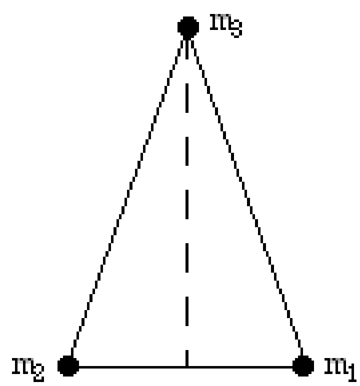
We begin with the equations of motion for the general problem of three bodies lying in a plane. The mass, position and velocity of the i-th body are denoted m_i, (x_i, y_i) and (u_i, v_i), respectively. Then for $i = 1, 2, 3$, we have:

$$\dot{x}_i = u_i \qquad m_i \dot{u}_i = -\frac{\partial U}{\partial x_i} \tag{2.1}$$
$$\dot{y}_i = v_i \qquad m_i \dot{v}_i = -\frac{\partial U}{\partial y_i}$$
$$U = -\sum_{1 \leq i < j \leq 3} \frac{G m_i m_j}{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}}$$

The path from (1) to a twist map which illustrates the hypotheses of the Aubry-Mather theorem is straightforward but long. First we must reduce the system to one with less degrees of freedom, using the isosceles triangle condition that we assume the three bodies satisfy. Having done this, we find that the phase space of our system is four-dimensional. Next we fix the total energy of the system. Now the orbits evolve on a three-dimensional subset of the four-dimensional phase space. We call this subset the fixed-energy surface. Fixing one of the three remaining variables to be zero picks out a two-dimensional slice of the fixed-energy surface. We denote this slice, or Poincaré section, as S. Later we will see that S has the convenient property that almost all orbits which begin on S eventually return to S. Before going any further with this explanation, let us examine some of the details described above.

The first step is to suppose that $m_1 = m_2 = m$, and then impose the conditions which restrict the three bodies to initially lie in an isosceles triangle configuration:

$$\begin{aligned} -x_2(0) &= x_1(0) \quad &\text{and}\quad &-u_2(0) = u_1(0) \\ y_2(0) &= y_1(0) \quad &\text{and}\quad &v_2(0) = v_1(0) \\ x_3(0) &= 0 \quad &\text{and}\quad &u_3(0) = 0 \end{aligned}$$



The above conditions stipulate that the isosceles triangle condition holds at $t = 0$. Additionally, the $t = 0$ conditions guarantee that the masses will be moving, at $t = 0$, in exactly the right way for the isosceles triangle configuration to be preserved for $t > 0$.[1]

The reason we impose these simplfying conditions is to reduce the dimension of the system's phase space. For fixed values of m_i, the general planar three-body problem requires four degrees of freedom (x_i, y_i, u_i, v_i) for each body, and hence a 12-dimensional phase space. The isosceles conditions imply that six of these 12 variables, the ones on the left-hand-sides of the conditions given above, drop out of the picture. Finally we switch to a reference frame in which $y_1 = y_2 = 0$ for all t, and replace y_3 by a new variable y. Now y gives the y-coordinate of m_3 with respect to the center of mass of m_1 and m_2. This allows us to forget about y_1. Similarly, we apply the conservation of momentum in the y-direction for the three bodies to obtain an expression for v_3 in terms of v_2. Then application of a Galilean transformation to translate v_1 to zero means that we need only keep track of $v = v_3 - v_2$.

[1] The details, which involve calculations of various Poisson brackets, are given in Appendix A.

Now our phase space is down to four dimensions: the only variables left are x, u, y, and v. ($x = x_1$ and $u = u_1$.) The new system is:

$$\dot{x} = u \qquad \dot{u} = -\left(\frac{Gm_1}{4x^2} + Gm_3 \frac{x}{(x^2+y^2)^{3/2}}\right) \tag{2.2}$$

$$\dot{y} = v \qquad \dot{v} = -G(2m_1 + m_3)\frac{y}{(x^2+y^2)^{3/2}} \tag{2.3}$$

2.2 Hamiltonian formulation

The next step is to pass to the Hamiltonian formalism, to allow us to construct the convenient Poincaré section discussed earlier. The Hamiltonian for the system, which corresponds to the total energy of the system (2-3) above, is:

$$H(\mathbf{q}, \mathbf{p}) = \frac{1}{2}\left({p_1}^2 + \frac{{p_2}^2}{\alpha}\right) - \frac{Gm_1}{4q_1} - \frac{Gm_3}{\sqrt{{q_1}^2 + {q_2}^2}} \tag{2.4}$$

Writing the Hamiltonian vector field as

$$X_H = \left(\frac{\partial H}{\partial p_1}, -\frac{\partial H}{\partial q_1}, \frac{\partial H}{\partial p_2}, -\frac{\partial H}{\partial q_2}\right), \tag{2.5}$$

we have that the equations of motion are

$$\frac{d}{dt}(\mathbf{q}, \mathbf{p}) = X_H(\mathbf{q}, \mathbf{p}) \tag{2.6}$$

Here $q_1 = x$, $q_2 = y$, $p_1 = u$, and $p_2 = \alpha v$, where $\alpha = m_3/(m_1 + m_2 + m_3)$.

Our problem fits naturally into the four-dimensional symplectic manifold M given by $\{(\mathbf{q}, \mathbf{p}) \in \mathbb{R}^4 : q_1 > 0\}$. The Hamiltonian H is an analytic real-valued function on M. The reason why we do not need to consider negative values of q_1 is simple. Suppose we start with a configuration for which $q_1 > 0$. This means that the position of the first mass is at the lower-right corner of the isosceles triangle. The position of the second mass, denoted here by x_2, is the reflection of q_1 across the y-axis, i.e. $x_2(t) = -q_1(t)$. There is no way that q_1 could ever become negative—if and when q_1 does reach zero, the symmetry of the problem forces it to collide with the second mass. Assuming the collision is elastic, q_1 will subsequently become positive once again. Finally, there is no loss of generality in assuming $q_1 > 0$. If $q_1 < 0$ then $q_2 > 0$ and we may as well keep track of the second mass to begin with.

Note that M has a very simple structure. Our requirement that $q_1 > 0$ removes additive inverses from what is otherwise the vector space $\mathbb{R}^4$ with the plane given by $q_1 = q_2 = 0$ deleted. We have a special reason for taking away the vector space structure of the problem's phase space—it allows us to apply a transformation to regularize the Hamiltonian at $q_1 = 0$. The charts and coordinates on M are trivial. Because of this, the symplectic form, the nondegenerate skew-symmetric two-form Ω, can be represented very simply on M as a canonical symplectic form:

$$\Omega = \sum_{i=1}^{2} dq_i \wedge dp_i.$$

Recall that a two-form on a manifold assigns a skew-symmetric bilinear form to each point on the manifold. Hence given $x \in M$, we see that $\Omega(x) : T_xM \times T_xM \to \mathbb{R}$. In our case, it is clear that $T_xM \cong \mathbb{R}^4$ for all $x \in M$. Hence the dq_i and dp_i, the basis of T_x^*M, can be identified with the basis of $\mathbb{R}^{4\,*}$. For example, dq_1 is just the linear functional on $\mathbb{R}^4$ which, when applied to $z = (z_1, z_2, z_3, z_4) \in \mathbb{R}^4$, simply gives z_1. Given two vectors $v_1 = (a_1, b_1, c_1, d_1)$ and $v_2 = (a_2, b_2, c_2, d_2)$ in T_xM, we write

$$\begin{aligned} \Omega(v_1, v_2) &= (dq_1 \wedge dp_1)(v_1, v_2) + (dq_2 \wedge dp_2)(v_1, v_2) \\ &= (a_1c_2 - a_2c_1) + (b_1d_2 - b_2d_1) \\ &= (a_1\, b_1\, c_1\, d_1) \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{pmatrix} \begin{pmatrix} a_2 \\ b_2 \\ c_2 \\ d_2 \end{pmatrix} \\ &= (v_1)^T J v_2, \text{ where } J = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{pmatrix} \end{aligned}$$

We would like to transform H, given in (4), so that it is defined at $Q_1 = 0, Q_2 \neq 0$. We will obtain a new Hamiltonian Γ defined on $\mathbb{R}^2 \times N$, where $N = M \cup \{(\mathbf{q}, \mathbf{p}) \in \mathbb{R}^4 \,:\, q_1 = 0,\, q_2 \neq 0\}$. N is a manifold with boundary, and the interior of N is just the symplectic manifold M.

2.3 Regularization at a binary collision

H has two singular terms, the first one singular at $q_1 = 0, q_2 \neq 0$ and the second at $(q_1, q_2) = (0, 0)$. The first singularity corresponds to the collision of m_1 and m_2, which we will call a *binary collision.* The second singularity corresponds to the *triple-collision* of the masses. Upon application of a suitable symplectic transformation, we will find that the first singularity disappears—and we will be left with a Hamiltonian defined on all of N. This turns out to be very useful in the analysis of this problem's dynamics. The Poincaré section, expressed in new coordinates after the symplectic transformation is applied, will turn out to be the interior of the right half of the unit disk in the plane. (The regularization of the problem allows us to partition the Poincaré section in a novel way, revealing a great deal of hidden structure in the problem. However, we will not pursue that partition of phase space here.)

The symplectic transformation we need is given by:

$$\varphi : M \to M$$
$$\varphi(q_1, q_2, p_1, p_2) = (\sqrt{q_1}, q_2, 2p_1\sqrt{q_1}, p_2)$$

Its derivative is given by

$$\mathbf{D}\varphi(z) = \begin{pmatrix} (1/2){q_1}^{-1/2} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ p_1 {q_1}^{-1/2} & 0 & 2{q_1}^{1/2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

To show that φ is symplectic, we must show that $\varphi^*\Omega = \Omega$, or in other words, for all $x \in M$, and for all $z^1, z^2 \in T_xM$,

$$\Omega(\mathbf{D}\varphi(z) \cdot z^1, \mathbf{D}\varphi(z) \cdot z^2) = \Omega(z^1, z^2)$$

We now verify that φ is symplectic:

$$\begin{aligned}
&\Omega\left(\left(\frac{z_1^1}{2\sqrt{q_1}}, z_2^1, \frac{p_1 z_1^1}{\sqrt{q_1}} + 2\sqrt{q_1}z_3^1, z_4^1\right), \left(\frac{z_1^2}{2\sqrt{q_1}}, z_2^2, \frac{p_1 z_1^2}{\sqrt{q_1}} + 2\sqrt{q_1}z_3^2, z_4^2\right)\right) \\
&\quad = \left(\frac{z_1^1}{2\sqrt{q_1}}\left(\frac{p_1}{\sqrt{q_1}}z_1^2 + 2\sqrt{q_1}z_3^2\right) - \frac{z_1^2}{2\sqrt{q_1}}\left(\frac{p_1}{\sqrt{q_1}}z_1^1 + 2\sqrt{q_1}z_3^1\right)\right) \\
&\qquad + (z_2^1 z_4^2 - z_2^2 z_4^1) \\
&\quad = (z_1^1 z_3^2 - z_1^2 z_3^1) + (z_2^1 z_4^2 - z_2^2 z_4^1) = \Omega(z^1, z^2)
\end{aligned}$$

We obtain a new Hamiltonian $K : M \to \mathbb{R}$ which satisfies $K \circ \varphi(x) = H(x)$ for all $x \in M$. Writing $\varphi(x) = (Q_1, Q_2, P_1, P_2)$,

$$K(Q_1, Q_2, P_1, P_2) = \frac{1}{2}\left(\frac{P_1^2}{4Q_1^2} + \frac{P_2^2}{\alpha}\right) - \frac{Gm_1}{4Q_1^2} - \frac{Gm_3}{\sqrt{Q_1^4 + Q_2^2}} \tag{2.7}$$

The last step is to scale the Hamiltonian by ${Q_1}^2$ and thus remove the singularity at $Q_1 = 0, Q_2 \neq 0$. We accomplish this by extending phase space by two-dimensions, creating a new Hamiltonian $\Gamma : \mathbb{R}^2 \times M \to \mathbb{R}$ where the two additional coordinates are denoted Q_0 and P_0. $Q_0 = t$ and $P_0 = -h$, where t is time and h is the total energy of the system. Our new Hamiltonian function is written as

$$\begin{aligned}\Gamma &= Q_1^2(P_0 + K) \\ &= \frac{1}{2}\left(\frac{P_1^2}{4} + \frac{P_2^2 Q_1^2}{\alpha}\right) - \frac{Gm_1}{4} - \frac{Gm_3 Q_1^2}{\sqrt{Q_1^4 + Q_2^2}} + P_0 Q_1^2\end{aligned} \tag{2.8}$$

Finally we rescale time to recover the same equations of motion for the system which we had for K and H. Let $dt = Q_1^2 dT$. Using Hamilton's equations for K and the chain rule,

$$\frac{dQ_i}{dT} = Q_1^2 \frac{dQ_i}{dt} = Q_1^2 \frac{\partial K}{\partial P_i} = \frac{\partial \Gamma}{\partial P_i}$$

where the last equality follows from (3). Similarly,

$$\frac{dP_i}{dT} = Q_1^2 \frac{dP_i}{dt} = Q_1^2 - \frac{\partial K}{\partial Q_i} = -\frac{\partial \Gamma}{\partial Q_i}$$

Stipulating that Q_0 and P_0 evolve according to Hamilton's equations implies $dQ_0/dT = \partial\Gamma/\partial P_0 = Q_1^2$ and $dP_0/dT = 0$. This implies $dQ_0/dt = 1$, $Q_0(t) = t$, and $P_0(t) = -h$ for all t. Hamilton's equations are consistent for our new Hamiltonian Γ, which as promised has no singularities on all of $\mathbb{R}^2 \times N$. The regularity of Γ at $Q_1 = 0, Q_2 \neq 0$ shows that the singularity of our original Hamiltonian system at $q_1 = 0, q_2 \neq 0$ was unessential. The dynamics do *not* break down when the binary pair collides.

2.4 The Poincaré map

Now that we know that the physics make sense at a binary collision, we can proceed to the construction of the Poincaré section. The main idea is to use the conserved energy to reduce the order of the problem from four dimensions to three.

Since the Hamiltonian in (7) is equivalent to the energy of the system, we have for any trajectory $(\mathbf{Q}(t), \mathbf{P}(t))$,

$$K(\mathbf{Q}(t), \mathbf{P}(t)) = h, \text{ a constant.} \tag{2.9}$$

From (7) we can solve this equation explicitly for P_2, obtaining

$$P_2 = \pm\frac{\sqrt{\alpha}}{2Q_1}\left(8hQ_1^2 + 2Gm_1 + \frac{8Gm_3Q_1^2}{\sqrt{Q_1^4+Q_2^2}} - P_1^2\right)^{1/2} \tag{2.10}$$

If the main quantity (inside parentheses) is negative, then there is no real solution for P_2. But this just means that such a configuration violates conservation of energy for the fixed h with which we started. Therefore, fixing the value of h means that the motion must lie on the three-dimensional manifold given by

$$\{(Q_1, Q_2, P_1) \in \mathbb{R}^3 : 8hQ_1^2 + 2Gm_1 + \frac{8Gm_3Q_1^2}{\sqrt{Q_1^4+Q_2^2}} - P_1^2 \geq 0,\ Q_1 > 0\}$$

We will refer to this manifold as the *fixed-energy manifold.* We normalize units with $G = 1$, $m_1 = 1/2 - 4m_3$, and $h = -1/8$. Then the intersection of the fixed-energy manifold with the plane $Q_2 = 0$ gives us our Poincaré section:

$$\{(Q_1, P_1) \in \mathbb{R}^2 : Q_1^2 + P_1^2 \leq 1 \text{ and } Q_1 > 0\}$$

Physically, the restriction of Q_2 to zero means that we are taking a snapshot of a trajectory as it passes through a *syzygy crossing.* The syzygy is just the center of mass of the binary pair of masses m_1 and m_2, and Q_2 measures the y-position of the third mass. Hence when $Q_2 = 0$, the three masses are collinear and the third mass is said to be at the syzygy.

Equation (3) shows that $y > 0$ implies $dv/dt < 0$, and that $y < 0$ implies $dv/dt > 0$. Excluding trajectories that lead to triple-collision, this observation implies that all other trajectories have at least one syzygy crossing.

The only trajectories that we need to delete are those at the boundary, i.e. when $Q_1^2 + P_1^2 = 1$. In this case, $P_2 = 0$, which means that the third mass starts at the syzygy and stays there for all t. The dynamics in this case are very simple and we do not need to include them in the Poincaré section. Then we are left with the open subset of the plane given by

$$S = \{(Q_1, P_1) \in \mathbb{R}^2 : Q_1^2 + P_1^2 < 1 \text{ and } Q_1 > 0\}$$

This is our Poincaré section. Each point on S determines a trajectory with the given initial values of Q_1 and P_1 along with $Q_2 = 0$ and P_2 computed using (10).

Let the *Poincaré map* for our problem, $f : S \to S$, be defined as follows. Given a point $\mathbf{x}$ in S, f returns the first subsequent $(t > 0)$

intersection of S with the trajectory that starts at $t = 0$ with position $(Q_1, P_1, Q_2, P_2) = (\mathbf{x}, 0, P_2(\mathbf{x}, 0))$ and evolves according to Hamilton's equations with Hamiltonian K: $(\dot{\mathbf{Q}}, \dot{\mathbf{P}}) = X_K(\mathbf{Q}, \mathbf{P})$. Denoting the *flow* of the Hamiltonian vector field X_K by $\psi^t : M \to M$, we can express the definition more formally:

Definition 2.4.1. Given a point $\mathbf{x} = (Q, P) \in S$ and $\mathbf{z} = (Q, P, 0, P_2(Q, P, 0))$, define $f(\mathbf{x}) = \Pi \circ \psi^{r(\mathbf{z})}(\mathbf{z})$, where $r(\mathbf{z}) = T$ is the least T greater than zero such that $Q_2(T) = 0$, and Π denotes the projection of a four-vector onto its first two factors.

2.5 Further reduction

At the moment we will take a short detour in order to reduce the system still further. This reduction will reveal an important symmetry property that will allow us to deduce an important property of the Poincaré map f. This follows the ideas discussed in Arnold, Section 45-B [1], who presents the reduction from a geometric point of view, different from ours. Consider (9) along with the condition that $\partial K/\partial P_2 \neq 0$. Then the implicit function theorem implies that in a neighborhood of (Q_1, P_1, Q_2, h) there exists a differentiable function L such that

$$P_2 = -L(Q_1, P_1, Q_2, h)$$

Substitution into (9) yields

$$K(Q_1, P_1, Q_2, -L(Q_1, P_1, Q_2, h)) = h \tag{2.11}$$

Differentiating this equation with respect to Q_1 and P_1 and using Hamilton's equations gives

$$\begin{aligned} \frac{\partial K}{\partial Q_1} - \frac{\partial K}{\partial P_2}\frac{\partial L}{\partial Q_1} = 0 &\implies \frac{dP_1}{dt} = -\frac{\partial L}{\partial Q_1}\frac{dQ_2}{dt} \\ \frac{\partial K}{\partial P_1} - \frac{\partial K}{\partial P_2}\frac{\partial L}{\partial P_1} = 0 &\implies \frac{dQ_2}{dt} = \frac{\partial L}{\partial P_1}\frac{dQ_2}{dt} \end{aligned}$$

Then by the chain rule we have

$$\frac{dQ_1}{dQ_2} = \frac{\partial L}{\partial P_1}, \quad \frac{dP_1}{dQ_2} = -\frac{\partial L}{\partial Q_1} \tag{2.12}$$

In other words, once we fix our energy h, we can locally substitute Q_2 as our independent variable and solve the above two-dimensional system for (Q_1, P_1). (At $Q_2 = 0$ we have to choose our initial condition vector $(Q_1, P_1) \in S$.) This will allow us to locally solve for P_2, thus reconstructing the dynamics for the whole problem from the two-dimensional system above. The form of the two-dimensional system shows that L is a Hamiltonian function and, moreover, differentiating (11) with respect to Q_2 and using Hamilton's equations shows that L is in fact conserved along orbits:

$$\frac{\partial K}{\partial Q_2} - \frac{\partial K}{\partial P_2}\frac{\partial L}{\partial Q_2} = 0 \implies -\frac{dP_2}{dt} = \frac{dQ_2}{dt}\frac{\partial L}{\partial Q_2} \implies \frac{dL}{dQ_2} = \frac{\partial L}{\partial Q_2}$$

Observe that the system (12) is invariant under the reflection $(Q_1, P_1, Q_2) \mapsto (Q_1, -P_1, -Q_2)$. This is true in every neighborhood for which we can write the system in the above form; hence it is true everywhere where $\partial K/\partial P_2 \neq 0$. A trivial calculation shows that this conditions is equivalent to $P_2 = 0$. The exclusion of the semicircle boundary from S has already stopped this possibility from occurring; hence for any $p \in S$, we can, in some neighborhood of $(p, 0, P_2)$, write the system in the form of (12). We will now proceed to exploit this symmetry.

2.6 Poincaré map symmetry

Denote the solution of system (12) by $\phi^{Q_2(t)}$. $\phi^s(Q, P)$ is the solution (Q_1, P_1), which passes through (Q, P) when $Q_2 = 0$, evaluated at $Q_2 = s$. To be precise, it is this solution $(Q_1(t), P_1(t))$ evaluated at t such that $Q_2(t) = s$, where we choose $Q_2(0) = 0$ as part of the initial conditions for the problems. Then ϕ satisfies

$$\frac{d}{dQ_2}\phi^{Q_2}(Q, P) = (\frac{\partial L}{\partial P_1}, -\frac{\partial L}{\partial Q_1}) \circ \phi^{Q_2}(Q, P) \qquad (2.13)$$

The symmetry mentioned above implies that if $\phi^{Q_2}(Q, P)$ satisfies (10), then $\phi^{-Q_2}(\rho(Q, P))$ should satisfy (10) as well. Here ρ denotes the reflection across the Q_1 axis in the (Q_1, P_1) plane. Then

$$\phi^{Q_2}(Q, P) = (Q_1, P_1) \iff \phi^{-Q_2}(\rho(Q, P)) = (Q_1, -P_1)$$

by the symmetry property. Hence $\phi^{-Q_2}(\rho(Q, P)) = \rho \circ \phi^{Q_2}(Q, P)$. Given a point $(Q, P) \in S$, $f(Q, P) = \phi^{Q_2(t_+)}(Q, P)$ where t_+ is the least $t > 0$ such that $Q_2(t) = 0$. Similarly, we can define $f^{-1}(Q, P) = \phi^{Q_2(t_-)}(Q, P)$ where t_- is the greatest $t < 0$ such that $Q_2(t) = 0$. The symmetry property of ϕ then gives a similar law for f:

$$f^{-1} \circ \rho = \rho \circ f \qquad (2.14)$$

2.7 Existence of an elliptic fixed point of f

Let A be the Q_1-axis. It is easy to prove that the set $B = f(A) \cap A$ is invariant, i.e. $f(B) = B$. Hence we search for fixed points of f on the Q_1-axis. Using reliable numerical methods, we have been able to verify that there is in fact a fixed point of f on this axis at $Q \approx 0.762967$. We will refer to this fixed point as $\zeta = (\zeta_q, 0)$. In fact ζ is an elliptic fixed point. To show this, we must calculate the eigenvalues of $\mathbf{D}f(\zeta)$.[2] First note that if we write $\rho(v) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} v$, then differentiation of $f^{-1} = \rho \circ f \circ \rho$ gives

$$\mathbf{D}f^{-1}(\zeta) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \mathbf{D}f(\zeta) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

The flow of a Hamiltonian system is a symplectic transformation of phase space. We can use that fact to show that f is symplectic as well. Since f is symplectic, it must preserve the symplectic form $dQ_1 \wedge dP_1$ on S. This is just the area two-form on S; hence f is locally area-preserving and $\det \mathbf{D}f(p) = 1$.

Putting together all of our knowledge about $\mathbf{D}f(\zeta)$, we conclude that

$$\mathbf{D}f(\zeta) = \begin{pmatrix} a & b \\ \frac{a^2-1}{b} & a \end{pmatrix}$$

for some $a, b \in \mathbb{R}$. The characteristic polynomial for $\mathbf{D}f(\zeta)$ is just $p(\lambda) = \lambda^2 - 2a\lambda + 1$, and its discriminant is $4(a^2 - 1)$. Now we use further numerical computations performed by Chesley and Zare [13] which show that $a \approx -0.9882$. Hence the discriminant of the characteristic polynomial is negative and $\mathbf{D}f(\zeta)$ has a pair of complex eigenvalues $\lambda_\pm = a \pm \sqrt{a^2 - 1}$. The eigenvalues are complex conjugates of each other and lie on the unit circle in the complex plane. Therefore, ζ is an elliptic fixed point of f.

2.8 In a neighborhood of ζ, f is a twist map

The eigenvalues of $\mathbf{D}f(\zeta)$ can be written as $\lambda_\pm = \exp(\pm i\alpha_0)$ where $\alpha_0 = 2.9971....$ Suppose $\alpha_0/2\pi$ is irrational[3] Then in a neighborhood of ζ, we can use the *Birkhoff normal form* to write f as a twist map. First we switch to canonical polar coordinates[4] by applying the transformation $Q_1 =$

[2] We show that the Poincaré map is a smooth diffeomorphism of S in Appendix B.

[3] We cannot say this for sure, but it is a safe bet.

[4] The transformation to (τ, θ) is symplectic, unlike the transformation to standard polar coordinates given by $x = r\cos\theta$, $y = r\sin\theta$.

$\sqrt{2\tau}\cos\theta + \zeta_q$, $P_1 = \sqrt{2\tau}\sin\theta$ to obtain $\hat{f}(\theta,\tau) = f(Q_1 + \zeta_q, P_1)$. The last step is to define $g(\theta,\tau) = \hat{f}(\theta,\tau) - \zeta$ so that the fixed point of g is at the origin.

Definition 2.8.1. A Birkhoff normal form of degree s for a symplectic map of the plane g in a neighborhood of an elliptic fixed point is

$$g(\theta,\tau) = (\theta + \alpha_0 + \alpha_1\tau + \cdots + \alpha_M\tau^M, \tau),$$

where M is not more than $[s/2] - 1$.

The idea behind a typical normal form expansion is to use a symplectic change of coordinates to write f as a formal power series from which useful dynamical information can be extracted. The convergence of the series is not necessary for the first few terms to give us information that is sufficiently accurate for practical purposes such as the prediction of the planets' positions. This is one way to justify the use of formal power series—however, it will not be our approach. Instead, we will use a symplectic change of variables to a normal form which is a finite sum as defined above. There will be a small remainder term left over. Combining an estimate on the order of this remainder term with the fact that a small perturbation of a twist map is still a twist map, we will be able to show that f is a twist map in a suitable neighborhood of ζ.

Arrowsmith and Place [2] have the most useful statement and proof of the result we need:

Theorem 2.8.1. *If the eigenvalue of the linear part of an area-preserving map at an elliptic fixed point is not a root of unity of degree s or less, then the map can be locally reduced by a symplectic change of variables to a Birkhoff normal form of degree s plus higher order terms.*

It follows from our assumption that $\alpha_0/2\pi$ is irrational that the eigenvalues are not roots of unity for any s. Therefore, we have

$$g(\theta,\tau) = (\theta + \alpha_0 + \alpha_1\tau + \cdots + \alpha_M\tau^M + R_\theta(\theta,\tau), \tau + R_\tau(\theta,\tau))$$

for any positive integer M, where R_θ and R_τ are $o(\tau^M)$. So long as the polynomial $\alpha_1\tau + \cdots + \alpha_M\tau^M$ does not vanish identically for all τ, there is a least N such that α_N is nonzero. Once we know that $\alpha_N \neq 0$, we see that $\alpha_N > 0$ implies

$$\frac{\partial g_1}{\partial \tau} > 0$$

and the twist condition is satisfied. Otherwise $\alpha_N < 0$ and g^{-1} would satisfy the twist condition.

The only technicality we must deal with is the issue of boundary circles. The only reason we required that the boundary circles remain invariant in our definition of twist maps from Chapter 1 was to ensure that the twist interval was well-defined and simple to determine. Unfortunately, there is no way to guarantee that g has any invariant circles at all. Note, however, that the annulus $S^1 \times [0,1]$ can also be thought of as a cylinder with a height of one unit. Any twist homeomorphism of the unit cylinder can be extended to a homeomorphism of the infinite cylinder $S^1 \times \mathbb{R}$. Hence we can make a twist map of the infinite cylinder using g. Now we do not have to worry about the twist interval—since the mapping keeps twisting and twisting in both directions, the twist interval is just $(-\infty, +\infty)$, the whole real line.

The strategy we choose to overcome the boundary circle problem, then, is to note that without the requirement of boundary circle invariance, g is a twist diffeomorphism on $0 < a < \tau < b$, $\theta \in [0, 2\pi]$ for some a and b. We just extend g to a twist diffeomorphism on the infinite cylinder, and we have our twist map.

Chapter 3

Circle Maps

3.1 From twist maps to circle maps

The theorem from the last section tells us that if we are to understand the dynamics of the Poincaré map f in a neighborhood of the elliptic fixed point ζ, we must make a careful examination of the dynamics of twist maps on the annulus. We saw earlier that in the case of an integrable twist map, the dynamics of the map reduce to the dynamics along each invariant circle $S^1 \times \{y\}$ for each y in $[0, 1]$. In fact, the dynamics of circle maps and circle diffeomorphisms form a model problem for the dynamics of twist maps. We see related behaviors in both settings, but the circle map is easier to analyze because the circle has one less dimension than the annulus. Furthermore, the concepts of lift and rotation number are critical to an understanding of both problems. Therefore we will now take a quick tour of the theory of circle maps. First, a few definitions are necessary:

Definition 3.1.1. A *circle map* f is an orientation-preserving homeomorphism of the circle S^1.

3.2 Lift and rotation number

For convenience we will work with maps on the real line instead of maps on the circle, simply because the machinery of analysis is fully developed for maps from $\mathbb{R}$ to $\mathbb{R}$. To pass from a circle map f to a map on the real line, we need the idea of a lift:

Definition 3.2.1. Given a circle map f, a map $F : \mathbb{R} \to \mathbb{R}$ is a *lift* of f if $\pi \circ F = f \circ \pi$, where $\pi : \mathbb{R} \to S^1$ is the standard projection map.

It turns out that for a given circle map, there are many different lifts. For example, consider the circle map r which rotates each point on the circle 90 degrees counterclockwise. The maps $R(x) = x + 1/4$, $R(x) = x + 5/4$, $R(x) = x + 9/4$, etc. are all lifts of r. In fact, this situation is generic:

Proposition 3.2.1. *Given a circle map f, all lifts of f differ only by an integer translation.*

Proof: Consider two different lifts of f, denoted by F and G. For all $x \in \mathbb{R}$, $\pi \circ F(x) = \pi \circ G(x)$. Hence $F(x) = G(x) + N(x)$ where $N(x) \in \mathbb{Z}$. Since $F - G$ is continuous, $N(x)$ is continuous also, implying that it is a constant. This proves the proposition.

A very useful tool for classifying circle maps is the rotation number. The rotation number is a function from the space of circle maps to the real numbers. For a circle map f, the rotation number of f tells us the average amount that a point on the circle S^1 is rotated by f. To arrive at this concept, we must first consider a preliminary notion:

Definition 3.2.2. Let F be the lift of a circle map f. Then define

$$\rho_0(F, x) = \lim_{n \to \infty} \frac{F^n(x)}{n}$$

for $x \in \mathbb{R}$.

Our first task is to show that this limit actually exists for all lifts of circle maps. First we must show the following:

Lemma 3.2.1. *For a lift F of a circle map f, $(F - \mathrm{id})(x) = F(x) - x$ is periodic with period 1.*

Proof: $\pi \circ F(x+1) = f \circ \pi(x+1) = f \circ \pi(x)$. By the first proposition, $F(x+1) = F(x) + n$ for some integer n. But $F(x+1) - F(x)$ is just the degree of the mapping f. Since f is an orientation-preserving homeomorphism, it is homotopic to the identity, hence $\deg(f) = 1$. Then $F(x+1) - (x+1) = F(x) - x$, proving the lemma.

Remark. Since F^n is the lift of f^n, which itself is a circle map, we also have that $F^n - \mathrm{id}$ is periodic with period 1.

Proposition 3.2.2. *For a lift F of a circle map f, $\rho_0(F, x)$ exists and is independent of x.*

Proof: If f has periodic points, then take a periodic point x. Denote its period by q. Lift x to the point $y \in \mathbb{R}$. Then

$$\pi \circ F^q(y) = f^q(x) = x \text{ so } F^q(y) = y + p$$

This implies that

$$\lim_{n\to\infty} F^{qn}(y)/qn = \lim_{n\to\infty} (y + np)/qn = p/q$$

Every integer k can be written as $k = qn + j$ for some integer $j < q$. Then let $M = \max_{j<q} |F^j(y) - y|$ and note that $|F^k(y) - F^{qn}(y)|/k \leq (M+1)/k$. Finally,

$$\lim_{k\to\infty} \frac{F^k(y)}{k} = \lim_{n\to\infty} \frac{F^{qn}}{qn} = \frac{p}{q}$$

If f has no periodic points, then $F^m(x) - x \notin \mathbb{Z}$ for all $m > 0$, any $x \in \mathbb{R}$. Then for each m, there is an integer k_m such that

$$k_m(x) < F^m(x) - x < k_m(x) + 1$$

In fact k_m is not dependent on x by continuity of $F^m - \text{id}$. Consider the sequence of inequalities:

$$\begin{aligned} k_m < F^m(0) - 0 < k_m + 1 \qquad (3.1)\\ k_m < F^m(F^m(0)) - F^m(0) < k_m + 1 \\ k_m < F^m(F^{2m}(0)) - F^{2m}(0) < k_m + 1 \end{aligned}$$

and so on. Adding the first M inequalities of the sequence gives

$$Mk_m < F^{mM}(0) < M(k_m + 1) \qquad (3.2)$$

Dividing (2) by mM and combining this inequality with (1) divided by m gives

$$\left| \frac{F^{mM}(0)}{mM} - \frac{F^m(0)}{m} \right| < \frac{1}{m}$$

Running the same argument again with m and M interchanged and using the triangle inequality, we obtain

$$\left| \frac{F^M(0)}{M} - \frac{F^m(0)}{m} \right| < \frac{1}{m} + \frac{1}{M}$$

This shows that the sequence $\{F^k(0)/k\}$ is Cauchy and therefore, has a limit.

The two arguments above show that for all lifts F of circle maps, there is some x such that $\rho_0(F, x)$ exists.

The last step is to show that the limit is independent of the point x. Consider $|F^n(x) - F^n(y)| \leq |(F^n(x) - x) - (F^n(y) - y)| + |x - y|$. From the periodicity of $F^n - \text{id}$ we can deduce that the first term in the sum is bounded by 1. (The periodicity implies that the term equals $|(F^n - \text{id})(x) - (F^n - \text{id})(y_0)|$ where $|x - y_0| \leq 1$ and the bound follows immediately.) Then

$$\lim_{n\to\infty} \left| \frac{F^n(x)}{n} - \frac{F^n(y)}{n} \right| = 0$$

and hence $\rho_0(F)$ is in fact independent of x, finishing the proof of the proposition.

Observe that ρ_0 is related to a circle map f only through its lifts. For a certain circle map f, if we have two lifts F_1 and F_2, then for all x, $F_1(x) = F_2(x) + n$ for some integer n. Then $\rho_0(F_1) = \rho_0(F_2) + n$. This motivates the following definition.

Definition 3.2.3. Given a circle map f, the *rotation number* $\rho(f)$ is the fractional part of $\rho_0(F)$ where F is any lift of f.

Remark. Since ρ_0 is defined up to an integer for a circle map f, taking the fractional part ensures that it is a well-defined real-valued function on the space of all circle maps.

3.3 Another way to approach rotation numbers

To see why we would want to define the rotation number in the way described above, consider the following scenario. Suppose a teacher were to give his student three points $\{\theta_0 = 0.0, \theta_2 = 0.3, \theta_3 = 0.7\}$ as depicted below. The teacher then asks, "If we consider the points in the sequence to be successive iterations of a circle map, which rotation numbers is the set of three points consistent with?" There are many possible rotation numbers—how is the student to narrow down the options? The first point that comes to mind is that any rotation number greater than 1/2 is inconsistent with these three points, for in this case, the third iterate would have to lie between the first two iterates. Rotation numbers greater than 1/2 do not preserve the order of the three points in the set. The following definition, inspired by [9], captures the essence of this property:

Definition 3.3.1. The rotation number $\alpha \in \mathbb{R}\backslash\mathbb{Z}$ is *consistent* with the finite orbit segment $\{f^j(z)\}_{\{j=0,1,\ldots,N\}}$ if the mapping $f^j(z) \mapsto j\alpha$ is orientation preserving.

For the three points in our example, any $\alpha \in (0, 0.5)$ works, since in this case $\{0 \cdot \alpha, 1 \cdot \alpha, 2 \cdot \alpha\}$ is in the same order on the circle as the three points in the set.

What if the teacher adds an extra point to the set, say $\theta_3 = 0.2$? Now which rotation numbers are consistent? Some of the rotation numbers which were consistent before are no longer consistent. Consider, for example, $\alpha = 0.3$. We have $3 \cdot \alpha = 0.9$. This does not preserve the ordering since $\theta_2 > \theta_3$ but $2 \cdot \alpha < 3 \cdot \alpha$. Similar calculations show that any $\alpha \in (1/3, 1/2)$ is consistent with the four points in the set. For if $\alpha > 1/3$, then $3 \cdot \alpha \pmod 1 > 0$. Meanwhile, $\alpha < 1/2$ implies $3 \cdot \alpha \pmod 1 < 1 \cdot \alpha$ as well.

We see that as additional points are added to the set, the set of consistent rotation numbers shrinks. If the set is expanded to an infinite set of points on the circle, there may not be a rotation number consistent with that set of points. In our example, we might have a sequence of open intervals $(0, 1, 2), (1/3, 1/2), ...$ with a nonempty set consisting of one point as a "limit," or the sequence might approach the empty set. What we can say is that if there is a rotation number consistent with an infinite set of distinct points on the circle, it must be irrational. For a rational $\alpha \in \mathbb{R}/\mathbb{Z}$, the iterations $j\alpha$ will eventually repeat, making it impossible to preserve the order of the infinite subset.

In fact, the only way there could be an infinite sequence of points on the circle consistent with a rational rotation number would be if the sequence is a periodic orbit.

If there is an α that is consistent with an infinite subset of points, then this equivalent to saying that $f^j(z) \mapsto j\alpha$ preserves orientation for all j. Then as $j \to \infty$, the quantity $F^j(z)/j$ must approach a limit, where F is a lift of f. We are back to the original definition of rotation number.

Now that we understand the rotation number, we can use it to classify circle maps.

3.4 Circle maps with rational rotation number

From the last section, we know that if a circle map has a periodic point, then its rotation number must be rational. More precisely, if x is a q-periodic point of a circle map f, then $F^q(y) = y + p$ for some integer p, where F is a lift of f and $\pi(y) = x$. In this case, $\rho(f) = p/q$.

In fact, the converse is true as well:

Proposition 3.4.1. *If the rotation number of a circle map is rational, say p/q, then the circle map has a periodic point.*

Proof: Let f be a circle map such that $\rho(f) = p/q \in \mathbb{Q}$. Fix $m \in \mathbb{Z}$. Then $\rho(f^m) = \lim_{n\to\infty}(F^{mn}(x) - x)/n$. Let $l = mn$, then as $n \to \infty$, $l \to \infty$ so

$$\rho(f^m) = \lim_{l\to\infty} \frac{F^l(x) - x}{l/m} = m\rho(f) \pmod 1$$

Hence $\rho(f^q) = 0$. We claim that if $\rho(f) = 0$, then f must have a fixed point.

Suppose f has no fixed point and let F be a lift with $F(0) \in [0, 1)$. Suppose $F(x) = x + n$ for some $x \in \mathbb{R}$ and some integer n—then $\pi \circ F(x) = \pi(x) = f \circ \pi(x)$ and $\pi(x)$ is a fixed point of f.

Hence $F(x) - x \notin \mathbb{Z}$ for all $x \in \mathbb{R}$. Then $F(0) > 0$. The periodicity of $F - \mathrm{id}$ and the Intermediate Value Theorem imply that for all x, $F(x) - x \in (0, 1)$. $F - \mathrm{id}$ is continuous on the compact set $[0, 1]$, hence there must exist $\delta > 0$ such that for all x, $\delta \leq F(x) - x \leq 1 - \delta$.

In the same manner as the proof of Proposition 2.2, we obtain a sequence of inequalities by inserting $x = 0, F(0), F^2(0), \ldots$ into the above inequality. Adding the first n such inequalities yields $n\delta \leq F^n(0) \leq (1 - \delta)n$. Dividing by n and taking the limit as $n \to \infty$ shows that $\rho(f) \neq 0$, and the proposition is proved.

This shows that a circle map f with rotation number p/q must have a periodic point x such that $f^q(x) = x$. Furthermore, for a lift F, there exists $y \in \mathbb{R}$ such that $F^q(y) = y + p$. We can strengthen this result by showing the following:

Proposition 3.4.2. *If f is a circle map with rational rotation number, then all periodic orbits have the same period.*

Proof: Suppose $\rho(f) = p/q$, where the fraction is written in lowest terms. Suppose that for a lift f there exists a periodic point $\pi(x)$, then $F^r(x) = x + s$ for some integers r and s. Using x to calculate the rotation number, we find that $p/q = s/r$ so that $s = mp$ and $r = mq$ for some integer m. We claim that $F^q(x) - p = x$. Suppose that $F^q(x) - p > x$. Then since $F^q - \mathrm{id}$ is nondecreasing,

$$F^{2q}(x) - 2p = F^q(F^q(x) - p) - p \geq F^q(x) - p > x$$

Iterating this argument, we have $F^r(x) - s = F^{mq}(x) - mp > x$, a contradiction. The argument against $F^q(x) - p < x$ is similar. Hence the claim and the proposition are proved.

We have a nearly complete description of the dynamics for a circle map with rational rotation number. The only task remaining is to describe the orbits of non-periodic points. This turns out to be rather simple. Either the circle map has one periodic orbit or more than one. If the circle map has more than one periodic orbit, then the orbit of any non-periodic point is forward-asymptotic to one periodic orbit and backward-asymptotic to another. Another way of saying this is that every non-periodic point is heteroclinic to two periodic orbits. When there is only one periodic orbit, we amend this statement to say that every non-periodic orbit is homoclinic to the periodic orbit.

To summarize, we can say that for a circle map with rational rotation number, every orbit is either (i) periodic or (ii) asymptotic to either one or two periodic orbits.

3.5 Circle maps with irrational rotation number

The situtation for these maps is slightly more complicated. Given a circle map f such that $\rho(f) \in \mathbb{R}/\mathbb{Q}$, we can show that the ω-limit set of x is the same for all $x \in S^1$. Furthermore, this set is either the entire circle or it is a Cantor (perfect, nowhere dense) subset of the circle. Recall that y is an ω-limit point of x for f if there is a sequence n_k such that $\lim_{k\to\infty} n_k = \infty$ and $\lim_{k\to\infty} d(f^{n_k}(x), y) = 0$. In other words, starting with x and applying f repeatedly, we select a subset of those iterations. If the distance between y and the iterations goes to zero then y is an ω-limit point. Then the set of all ω-limit points of x for f is the ω-limit set of x.

In the first case, all points have orbits which are dense subsets of the circle, and the circle map f is topologically conjugate to a rotation by $\rho(f)$.

In the second case, the orbit of a given point is either in the Cantor set, or it is asymptotic to the Cantor set orbit. In this case, there is no conjugacy to a rotation.

3.6 Denjoy's theory

Denjoy showed that if instead of a circle homeomorphism we consider a C^1 circle diffeomorphism whose derivative has bounded variation, then we are guaranteed that the map is conjugate to a rotation.

Definition 3.6.1. A map $g : S^1 \to \mathbb{R}$ has *bounded variation* if the

$$\sup \sum_{k=1}^{n} \left| g(x_k) - g(x'_k) \right|$$

over all finite collections $\{I_k\}_{k=1}^n$ of disjoint intervals where $I_k = [x_k, x'_k]$.

The requirement that the map has bounded variation is critical. In fact, there is a beautiful example, called the *Denjoy counterexample*, of a C^1 diffeomorphism of the circle which is not conjugate to a rotation. We will briefly describe this map since it will play a role in our discussion of the Aubry-Mather results.

Let us start with a simple rotation by an irrational number α, that is, a circle map f whose lift is $F(x) = x + \alpha$. The orbit of any point on the circle is a dense subset of the circle. Therefore, pick any point on the circle and consider its orbit. At each point on the orbit, delete the point from the circle and replace it with a small interval. That is, given $x_0 \in S^1$, replace $f^n(x_0)$ with an interval I_n for each $n \in \mathbb{Z}$. So long as the infinite sum of the lengths of all the intervals converges, and so long as we glue in the intervals smoothly, the image of our mutated rotation map is diffeomorphic to a circle.

For a point not on the orbit of x_0, the map remains an irrational rotation by α. Given this mutated rotation map from S^1 to M where M is the mutated circle, we can extend the map to be a diffeomorphism of M by defining it on each of the I_n. We simply choose a collection of orientation-preserving diffeomorphisms h_n mapping each I_n onto I_{n+1}. If we do this carefully enough, the resulting map is in fact a C^1 diffeomorphism of M, which is itself diffeomorphic to a circle. Hence we actually have a C^1 diffeomorphism of the circle. We will call this map a Denjoy counterexample.

Now, a point not on the orbit of x_0 has the same orbit as before, a dense subset of the circle. But note that a point in the interior of a given I_n never returns, under iteration of the Denjoy map, to I_n. Hence the orbits of such points are not dense subsets of the circle. This means that the map is not transitive and therefore, not conjugate to a rotation, even though it is a C^1 diffeomorphism of the circle with irrational rotation number. The existence of such a map will play a role in our discussion of twist maps in the next chapter.

Chapter 4

Twist Maps

4.1 From circle maps to twist maps

The questions we asked about circle maps have simple analogues in the theory of twist maps. Just as a circle map f has an associated rotation number $\rho(f)$, a twist map has a range of rotation numbers which we call the twist interval.

Definition 4.1.1. For a twist map τ, the twist interval is $[\rho(\tau_0), \rho(\tau_1)]$ where τ_j is the circle map obtained by restricting τ to the boundary circle $S^1 \times \{j\}$ for $j = 0, 1$.

For a rational number in the twist interval, is there a periodic orbit of the twist map in the annulus? Are there orbits associated with irrational numbers in the twist interval? To answer these questions, we must first extend our notion of a rotation number. Currently we have only defined the rotation number of a circle map. In a discussion of twist maps, however, it is useful to associate rotation numbers with certain kinds of countable subsets of the annulus. More specifically, we will first extend our definition of rotation number so that we can speak of the rotation number of an orbit of f.

In the results that follow, we will need to work with the lift F of a twist map f. This is defined intuitively: the lift $F : \mathbb{R} \times [0,1]$ satisfies $\pi \circ F_1(x,y) = f_1 \circ (\pi(x), y)$ where $\pi : \mathbb{R} \to S^1$ is the covering map and $F = (F_1, F_2), f = (f_1, f_2)$. Furthermore, it is helpful to have the following definitions in mind:

Definition 4.1.2. Given a twist map f, the *orbit* of a point $z \in A$ is the set $\omega(z) = \{f^j(z)\}_{j \in \mathbb{Z}}$. Take a lift F such that $F_1(0,0) \in [0,1)$. Let x be the point in $\mathbb{R} \times [0,1]$ such that $\pi(x) = z$. Then the *lift of the orbit* $\omega(z)$ is $\Omega(z) = \{F^j(x) + (i,0) : i,j \in \mathbb{Z}\}$.

4.2 Ordered orbits

First we should be clear about which subsets of the annulus we wish to deal with. It will not be possible to define a rotation number for any subset of the annulus—a radial segment $\{\theta\} \times [0,1]$, for example, cannot have a "rotation number" associated with it. However, note that such a segment could not possibly be the orbit of a single point. Our first restriction is to restrict our attention to orbits of a given twist map—when can we assign a rotation number to an orbit? Take the orbit $\{(x_0, y_0), (x_1, y_1), (x_2, y_2), ...\}$ where $(x_{n+1}, y_{n+1}) = f(x_n, y_n)$. For the moment assume that all the x coordinates are distinct. If we toss out all the y coordinates, the resulting set, $\{x_0, x_1, x_2, ...\}$, is a countable subset of S^1, and we can apply our discussion of rotation number from the previous chapter. To actually calculate the rotation number, we could lift the points to $\mathbb{R} \times [0,1]$ and calculate $\lim_{n\to\infty}(\hat{x}_n/n)$ where $\pi(\hat{x}_n) = x_n$. Unfortunately we have no guarantee that this limit exists. As we noted earlier, there are sequences of points on the circle which are inconsistent with all rotation numbers.

To remedy this situation, let us add the condition that the the orbit must be such that f preserves its cyclic order. The following definition makes this idea more precise:

Definition 4.2.1. Let f be a twist map. The orbit of z is an *ordered orbit* if for all $(x_1, y_1), (x_2, y_2) \in \Omega(z)$, $F_1(x_1, y_1) < F_1(x_2, y_2)$ whenever $x_1 < x_2$. Furthermore, if (x_1, y_1) and (x_2, y_2) are in the ordered orbit and $x_1 = x_2$, then y_1 must equal y_2.

The second condition makes sure that when we project the orbit onto S^1, we do not lose any information. Periodic orbits satisfy this condition, as do the quasiperiodic orbits we will encounter shortly.

4.3 Rotation number of ordered orbits

The idea now is to calculate the rotation number of an ordered orbit. We will do this by showing that each ordered orbit induces a homeomorphism of the circle. The rotation number of the orbit, then, is just the rotation number of the induced circle map. To define this homeomorphism, we will have to show that every orbit can be realized as the graph of a Lipschitz function from a subset of S^1 to $[0,1]$. Essentially, because we have a twist diffeomorphism, the y coordinates of the orbit cannot jump around too quickly. The twist map itself satisifes a Lipschitz condition which forbids this from happening:

Lemma 4.3.1. *Let f be a twist map with lift F. Suppose $(x_i, y_i) = F^i(x_0, y_0)$ and $(x'_i, y'_i) = F^i(x'_0, y'_0)$. Then if $x'_i > x_i$ for $i = -1, 0, 1$, there exists $M > 0$ such that $|y'_0 - y_0| < M|x'_0 - x_0|$.*

Proof: First we assume $y'_0 < y_0$. Then by the twist condition, for any Y_1, Y_2 such that $Y_2 > Y_1$, there exists c bounded away from zero such that $Y_2 > Y_1$,

$$\frac{F_1(x'_0, Y_2) - F_1(x'_0, Y_1)}{Y_2 - Y_1} > c$$

Applying this to our case we have

$$F_1(x'_0, y_0) - x'_1 > c(y_0 - y'_0) \tag{4.1}$$

Now using the differentiability of f, we have that there exists L such that $F_1(x'_0, y_0) - F_1(x_0, y_0) < L(x'_0 - x_0)$. L is bounded on the annulus. Since $x'_1 > x_1 = F_1(x_0, y_0)$,

$$F_1(x'_0, y_0) - x'_1 < L(x'_0 - x_0) \tag{4.2}$$

Let $M = L/c$ and combine inequalities (1) and (2) to obtain the lemma. The same argument with f^1 in place of f works for the $y'_0 > y_0$ case.

We can immediately apply this lemma. Given an ordered orbit O consisting of pairs (x_i, y_i), there must exist a function $\psi : S^1 \to [0,1]$ such that O is contained in the graph of ψ. (Observe $\psi(x_i)$ is just y_i.) Now if we apply the previous lemma to (x_l, y_l) and (x_m, y_m) for $l = i-1, i, i+1$ and $m = j-1, j, j+1$, we obtain that ψ is a Lipschitz function. Let $\hat{O}$ be the set obtained by deleting from O all the y coordinates; we call $\hat{O}$ the projection of O onto S^1. Note that $\hat{O}$ is a closed subset of S^1. Then we have a map $g : \hat{O} \to \hat{O}$ defined by

$$g(\theta) = f_1(\theta, \psi(\theta))$$

The inverse map $g^{-1} : \hat{O} \to \hat{O}$ is defined by

$$g^{-1}(\theta) = f_1^{-1}(\theta, \psi(\theta))$$

Because f is a diffeomorphism and ψ is Lipschitz on $\hat{O}$, g is a homeomorphism of $\hat{O}$. Furthermore, g preserves the cyclic order of points of $\hat{O}$. We will extend g to a homeomorphism of S^1.

Denote the closure of $\hat{O}$ by K. Given a sequence x_n in $\hat{O}$ converging to $x \in K \backslash \hat{O}$, define $g(x) = \lim_{n \to \infty} g(x_n)$. If we do the same for g^{-1} we have a homeomorphism of K. The map can then be extended linearly to a homeomorphism of S^1—for example, if $x_1, x_2 \in K$ but $(x_1, x_2) \notin K$, then for $x \in (x_1, x_2)$, let $g(x) = [g(x_1)(x - x_1) + g(x_2)(x_2 - x)] \pmod 1$. Geometrically, we are simply filling in the gaps on the circle that are left after drawing in all the points of K.

Corollary 4.3.1. *The circle map g described above is the* induced homeomorphism of an ordered orbit*. Every ordered orbit induces such a homeomorphism.*

We are now in a position to calculate the rotation number of an ordered orbit:

Definition 4.3.1. Given a ordered orbit O, the rotation number $\rho(O)$ is defined to be the rotation number of the induced homeomorphism g.

This definition agrees with the informal definition given in the last section:

Lemma 4.3.2. *The rotation number of an ordered orbit $O = (\theta_i, r_i)_{i \in \mathbb{Z}}$ is*

$$\lim_{i \to \infty} \frac{\hat{\theta}_i}{i}$$

where $\pi \circ \hat{\theta}_i = \theta_i$.

Proof: Note that $g^i(\theta_0) = {f_1}^i(\theta_0, \psi(\theta_0)) = \pi_1 \circ f^i(\theta_0, r_0) = \theta_i$. Then the definition of the rotation number for a circle map applied to $\rho(g)$ yields the lemma.

4.4 Rotation number of ordered sets

In fact we can do even better than the results obtained in the last section. An ordered orbit is merely one example of a more general concept, an ordered set. We will define an ordered set as follows:

Definition 4.4.1. An *ordered set* T for a twist map f is a closed invariant set of f which satisfies

1. For $(x_1, y_1), (x_2, y_2) \in A$, $\hat{x}_1 < \hat{x}_2$ implies $F_1(\hat{x}_1, y_1) < F_1(\hat{x}_2, y_2)$ where $\pi(\hat{x}_i) = x_i$.

2. For $(x_1, y_1), (x_2, y_2) \in A$, $x_1 = x_2$ implies $y_1 = y_2$.

An ordered set induces a circle map in the same way as an ordered orbit. In fact, we can follow the argumentation of the last section to obtain an induced homeomorphism. The rotation number of this homeomorphism, then, is just the rotation number of an orbit that is sitting in the ordered set. Note that any orbit that is contained in an ordered set must be an ordered orbit, and hence the rotation number of the orbit exists. In this way we can calculate the rotation number of an ordered set.

Now note that a minimal ordered set is precisely what we called an Aubry-Mather set in the first chapter. It follows that any Aubry-Mather set is the subset of a graph of a Lipschitz function from S^1 to $[0, 1]$, that an Aubry-Mather set induces a homeomorphism of the circle, and also that we can calculate the rotation number of an Aubry-Mather set. These three properties will become useful later, when we consider ordered sets of irrational rotation number. For now, we will take a look at the rational case.

4.5 Ordered orbits of rational rotation number

Armed with a definition of rotation number that works for ordered orbits, we can try to find ordered orbits with particular rotation numbers. We found in the last chapter that a circle map with even one periodic orbit must have a rational rotation number. Consider a periodic ordered orbit of a twist map. By our definition, it must have rational rotation number. Just as for circle maps, we can ask the reverse question: given a rational number p/q in the twist interval, is there a periodic ordered orbit with rotation number p/q?

The answer is that for an area-preserving twist map which is a diffeomorphism of the annulus, there are two such orbits. Besides being ordered periodic orbits, they satisfy the condition that they can be parameterized according to their cyclic ordering after projection from the annulus to the circle. This parameterization takes a particular form, and in this case we call the orbit a Birkhoff periodic orbit:

Definition 4.5.1. Let f be a twist map with a lift denoted by F. A *Birkhoff periodic point of type* (p, q) is a point $w \in A = S^1 \times [0, 1]$ such that for a lift $z \in S = \mathbb{R} \times [0, 1]$, there exists a sequence $\{(x_n, y_n)\}_{n \in \mathbb{Z}}$ in S that satisfies the following properties:

1. $(x_0, y_0) = z$,

2. $x_{n+1} > x_n (n \in \mathbb{N})$,

3. $(x_{n+q}, y_{n+q}) = (x_n + 1, y_n)$,

4. $(x_{n+p}, y_{n+p}) = F(x_n, y_n)$.

The orbit $\omega(w)$ is then called a *Birkhoff periodic orbit of type* (p, q). [8]

Here we will only show the existence of one of the Birkhoff periodic orbits of type (p, q) since that is all we need for later purposes:

Theorem 4.5.1. *Let f be a twist map with a lift F. For any rational p/q in the twist interval $[\rho_0, \rho_1]$, f has a Birkhoff periodic orbit of type (p, q)*

Following the argument of [7], the proof of the theorem involves two stages. In the first stage, we set up a space of non-decreasing maps which correspond to the x-coordinates of candidate Birkhoff orbits. On this space, we can define a Lagrangian functional that takes a non-decreasing map and gives us a real number. In the second stage, we show that if the Lagrangian of a certain map $L(\phi)$ is a local minimum for L, then a certain periodicity condition is satisfied. If this condition is satisfied, we can, in a straightforward way, use ϕ to write down a Birkhoff periodic orbit of f. The difficult task is showing that a local minimum of L corresponds to a periodicity condition.

Proof: Start with the space $\mathcal{S}_{p,q}$ consisting of all nondecreasing real-valued functions ϕ on $\mathbb{Z}$ that satisfy

1. $\phi(n + q) = \phi(n) + 1$ and

2. $\phi(n) + F_1(\phi(n), 0) \leq \phi(n + p) \leq \phi(n) + F_1(\phi(n), 1)$

Let $\Phi_{p,q} = \mathcal{S}_{p,q}/\sim$, where $\sim$ is the equivalence relation that identifies two maps in $\mathcal{S}_{p,q}$ if they differ by a fixed integer translation.

We can think of $\phi(n)$ as the projection onto S^1 of a candidate Birkhoff periodic orbit. The first condition corresponds to item 3 in the Birkhoff periodic orbit definition, while the second corresponds to item 4: $(x_{n+p}, y_{n+p}) = F(x_n, y_n)$. The second requirement ensures that it is possible for $\phi(n + p)$ to be the x-coordinate of $F(\phi(n))$.

Consider $f : \Phi_{p,q} \to \mathbb{R}^q/\mathbb{Z}$ given by $f(\phi) = (\phi(0), ..., \phi(q - 1))$. In fact f is an embedding, giving us a natural topology on f. It follows from condition (2) above that $f(\Phi_{p,q})$ is a closed and bounded subset of $\mathbb{R}^q/\mathbb{Z}$. Hence $\Phi_{p,q}$ is compact. Furthermore we claim that $\Phi_{p,q} \neq \emptyset$.

Suppose we have $x' \in [F_1(x, 0), F_1(x, 1)]$ for some real x, x'. Then by the twist condition, the image of the interval $I = \{x\} \times [0, 1]$ intersects the interval $I' = \{x'\} \times [0, 1]$ in only one point—recall that the twist condition states that for a fixed x coordinate, $F_1(x, y)$ is a strictly increasing function of y. Since there is always only one point of intersection, we can write it as $(x', h(x, x'))$, defining the function h in the process. We obtain a three-sided region bounded by $\mathbb{R} \times \{0\}$, I', and the curve $F(I)$. Since this three-sided region is determined only by x and x', we denote it by $T(x, x')$. Because f is area-preserving, the lift of the area that f preserves is an area on $S = \mathbb{R} \times [0, 1]$ preserved by F. We can calculate the area of the triangular region $T(x, x')$; denote it by $H(x, x')$. H is the classical *generating function* for the twist map f.

Now we can define the Lagrangian $L_{p,q} : \Phi_{p,q} \to \mathbb{R}$ given by

$$L_{p,q}(\phi) = \sum_{n=0}^{q-1} H(\phi(n), \phi(n+p))$$

Thus ends the first stage of the proof. In order to continue, we must introduce the periodicity condition that will be satisfied when the Lagrangian hits a local minimum.

Let $h_1(n) = h(\phi(n-p), \phi(n))$, $\psi_1(n) = (\phi(n), h_1(n))$, and $\psi_2(n) = F^{-1} \circ \psi_1(n+p)$.

For the moment let $\psi_2(n) = (a, b)$. Then $F(a, b) = (\phi(n+p), h_1(n+p)) = (\phi(n+p), h(\phi(n), \phi(n+p)))$. This is just the point of intersection of $F(\phi(n) \times [0, 1])$ with $\phi(n+p) \times [0, 1]$. Therefore, $a = \phi(n)$. We define $h_2(n) = b$. In other words, $h_2(n)$ is defined by the equation $\psi_2(n) = (\phi(n), h_2(n))$.

We will now take a quick detour to state the periodicity condition and show that it gives us what we want:

Lemma 4.5.1. *If the* periodicity condition $h_1(n) = h_2(n)$ *is satisfied for every* n *then the map* $\psi_1 : \mathbb{Z} \to \mathbb{R}$ *defines a Birkhoff periodic orbit of type* (p, q).

Proof: Define a sequence of points by $(x_n, y_n) = \psi_1(n)$. Then the periodicity condition implies immediately that $F(\psi_1(n)) = \psi_1(n+p)$ and condition (4) in the Birkhoff definition is satisfied for $\{(x_n, y_n)\}$. The third condition is trivially satisfied. The only thing left to check is that $\phi(n)$ is strictly monotone.

Suppose it is not and there is some integer n for which $\phi(n) = \phi(n+1)$ but $h_1(n) \neq h_1(n+1)$. Now consider $(\phi(n-p), h_1(n-p)) = F^{-1}(\phi(n), h_1(n))$ and $(\phi(n-p+1), h_1(n-p+1)) = F^{-1}(\phi(n), h_1(n+1))$. Since $\phi(n)$ is nondecreasing, $\phi(n-p) \leq \phi(n-p+1)$. Then the reverse twist condition for F^{-1} implies that $h_1(n) > h_1(n+1)$. Then using this with $\phi(n) = \phi(n+1)$, the twist condition on F implies that $\phi(n+p) > \phi(n+p+1)$. Hence ϕ is not non-decreasing and we have a contradiction, proving the lemma.

We are now in a position to begin the second phase of the proof of the main theorem:

Lemma 4.5.2. *If the Lagrangian $L_{p,q}$ has a local minimum at ϕ, then with that ϕ, the periodicity condition $h_1(n) = h_2(n)$ is satisfied for all integers n.*

Proof: Suppose that for some n, $h_1(n) \neq h_2(n)$. The first case we consider is when $\phi(n-1) < \phi(n) < \phi(n+1)$. We assume for now that

$$0 \leq h_2(n) < h_1(n) \leq 1$$

We will construct a function which yields a smaller value of the Lagrangian than the alleged minimum. By definition, $F(\phi(n), h_2(n)) = \psi_1(n+p) = (\phi(n+p), h_1(n+p))$. Then from the assumption we conclude that $h_1(n+p) < 1$. Choose $\epsilon > 0$ and define

$$\tilde{\phi}_\epsilon(m) = \begin{cases} \phi(m) & \text{if } m \neq n \pmod{q}, \\ \phi(m) - \epsilon & \text{if } m = n \pmod{q}. \end{cases}$$

Choose $\epsilon > 0$ sufficiently small so that the assumption and the bound on $h_1(n+p)$ ensure that $\tilde{\phi}_\epsilon \in \Phi_{p,q}$. Now our task is to show that $L_{p,q}(\tilde{\phi}_\epsilon) < L_{p,q}$. We will do this using an area argument.

First we make the following definitions:

$$\begin{aligned} I_0 &= \{\phi(n-p)\} \times [0,1] \\ I_1 &= \{\phi(n)\} \times [0,1] \\ I_2 &= \{\phi(n+p)\} \times [0,1] \\ \tilde{I}_1 &= \{\tilde{\phi}_\epsilon(n)\} \times [0,1] \end{aligned}$$

Let A_1 be the four-sided region bounded by I_1, $F(I_0)$, $\tilde{I}_1$, and $\mathbb{R} \times 0$.

Let A_2 be the four-sided region bounded by $F(I_1)$, I_2, $F(\tilde{I}_1)$ and $\mathbb{R} \times 0$.

Note that three sides of $F^{-1}(A_2)$ coincide with A_1. If we want to show that $F^{-1}(A_2) \subset A_1$, the only sides we have to worry about are $F^{-1}(I_2)$ and $F(I_0)$.

If ϵ is chosen small enough, then $\tilde{I}_1$ is sufficiently close to I_1 to guarantee that all of the y-coordinates of the segment $F^{-1}(I_2) \cap [\tilde{\phi}_\epsilon, \phi(n)] \times [0,1]$ are less than the y-coordinates of the segment $F(I_0) \cap [\tilde{\phi}_\epsilon, \phi(n)] \times [0,1]$. Then $F^{-1}(A_2) \subset A_1$.

Denoting the preserved area function on $\mathbb{R}\times[0,1]$ by μ, we have $\mu(A_1) > \mu(F^{-1}(A_2)) = \mu(A_2)$.

Then applying the definition of $L_{p,q}$ yields

$$\begin{aligned} L_{p,q}(\phi) - L_{p,q}(\tilde{\phi}_\epsilon) &= H(\phi(n), \phi(n+p)) + H(\phi(n-p), \phi(n)) \\ &\quad - H(\phi(n) - \epsilon, \phi(n+p)) - H(\phi(n-p), \phi(n) - \epsilon) \\ &= \mu(A_1) - \mu(A_2) > 0 \end{aligned}$$

Now ϕ is no longer a local minimum; this contradiction implies the lemma, for the particular case we have considered. If we had assumed instead that

$$0 \geq h_2(n) > h_1(n) \geq 1$$

then we would be able to run the same argument with success, except now we would have to define $\tilde{\phi}_\epsilon(n) = \phi(n) + \epsilon$ and show that $A_1 \subset F^{-1}(A_2)$. The structure of the argument is identical.

We considered a case where $\phi(m)$ was strictly monotonic for $m = n-1, n, n+1$. To finish the proof of the lemma, we must generalize. Once again suppose $h_1(n) \neq h_2(n)$ for some integer n. This time, we suppose that there is some integer k such that

$$\phi(n-1) < \phi(n) = \phi(n+1) = \cdots = \phi(n+k) < \phi(n+k-1)$$

Using the twist condition we can show that there are only three possibilities to consider: either $1 \geq h_1(n) > h_2(n) \geq 0$, $1 \geq h_2(n+k) \geq h_1(n+k) \geq 0$, or $h_1(n+l) = h_2(n+l)$ for $l = 0, 1, ..., k$. The third case implies the lemma. As for the first two cases, they correspond to the two different assumptions we made in the more special monotonic-ϕ case above. They are simply perturbations to the left and to the right of $\phi(n)$ and $\phi(n+k)$ respectively—the basic argument is, once again, the same.

4.6 Ordered orbits of irrational rotation number

Having shown that for a twist map f, every rational number p/q in the twist interval of f corresponds to a Birkhoff periodic orbit of type (p,q), we can continue the analogy with circle maps still further: does an irrational number in the twist interval correspond to any particular orbits in the annulus? More specifically, is there an ordered set with irrational number α for each irrational α in the twist interval?

This question was answered affirmatively by S. Aubry and J. Mather, who, working independently, published extensively on this subject in the early 1980's. We will return to the methods they used to prove the main theorem, the Aubry-Mather theorem, after first examining a straightforward method of deducing the theorem given the existence of Birkhoff periodic orbits. This method is due to Katok [7].

Katok's idea is, in a nutshell, to approximate an irration number α in the twist interval by a sequence of rationals in the twist interval. For each of these rationals, there is an associated Birkhoff periodic orbit. We have a sequence of closed invariant sets, each of which is a Birkhoff periodic orbit, and hence we can take the limit of this sequence in the Hausdorff topology. The limit is a closed invariant set, in fact an Aubry-Mather set with rotation number α.

In order to present a detailed version of this proof, a few preliminaries have to be dealt with:

Definition 4.6.1. Consider the space H consisting of closed subsets of a metric space (X, d). We assign a metric to H, called the *Hausdorff metric*, as follows:

$$d(A, B) = \sup_{x \in A}\{d(x, B)\} + \sup_{x \in B}\{d(A, y)\}$$

where A and B are closed subsets of X. (Recall that for $A \subset X$, $d(x, A) = \inf_{y \in A}\{d(x, y)\}$).

Any homeomorphism f of a compact metric space X induces a homeomorphism of H. A closed set is an invariant set of f if and only if it is a fixed point of the induced homeomorphism; since the set of fixed points of a continuous function is closed, we find that the set of closed invariant sets of f is a closed set with respect to the Hausdorff metric.

We are now in a position to prove the Aubry-Mather theorem, Theorem 1.3.1.

Proof: Given an irrational number α in the twist interval of a twist map f, write down a sequence p_n/q_n of rationals (each in lowest terms) which converges to α. (Take each p_n/q_n to be in the twist interval.) Then for each rational p_n/q_n, we are guaranteed the existence of a Birkhoff periodic orbit O_n of type (p_n, q_n). As established in section 4.3, each O_n is contained in the graph of a Lipschitz function $\phi_n : S^1 \to [0, 1]$. Based on the argument in the proof of that lemma, the set of all the Lipschitz constants is a set of positive numbers bounded away from zero. If we take the infimum of that set, we have a positive number which can serve as Lipschitz constant for all the functions at once. Therefore, we have an equicontinuous family of Lipschitz functions and we can apply the Arzela-Ascoli theorem. The theorem implies that $\phi_n \to \phi$ where ϕ is Lipschitz.

We have a sequence of Birkhoff periodic orbits w_n, each of which is a closed invariant set of f. Denote the accumulation point of the w_n, in the Hausdorff metric, by the set I. Because the set of closed invariant sets of f is closed in the Hausdorff metric, I must be a closed invariant set of f. Furthermore, I is a subset of the graph of ϕ, since the graph of ϕ_n contains w_n for each n. Finally, f preserves the cyclic ordering of I since it preserves the cyclic ordering on all the w_n.

Now we show that the rotation number of I is indeed α. Let f_n be the circle map induced by w_n. Then $\rho(f_n) = p_n/q_n$. Since I is an ordered invariant set, we have an induced circle map f_α. From the induced circle map construction and the construction of I from the w_n we have $f_n \to f_\alpha$ uniformly. Now by the following lemma, we have $\rho(f_\alpha) = \lim_{n\to\infty} p_n/q_n = \alpha$ as desired.

Finally, we note that f_α is a circle map with irrational rotation number. Hence its minimal set is either the entire circle or an invariant Cantor subset of the circle.

The only thing left to show is the following lemma:

Lemma 4.6.1. *The rotation number $\rho(\cdot)$ is continuous in the C^0 topology.*

Proof: For a circle map f, write $\rho(f) = \rho$. Choose r/s and p/q in $\mathbb{Q}$ such that $r/s < \rho < p/q$. Lift f to F such that $-1 < F^q(x) \leq x + p$ for some $x \in \mathbb{R}$. Then for all $x \in \mathbb{R}$, $F^q(x) < x + p$—if not, then $\rho = p/q$, a contradiction. We saw earlier that $F^q - \text{id}$ is periodic and continuous—hence it attains its maximum M on $[-1, 0]$. Then $(F^q - \text{id})(x) \leq M < p$ for all $x \in \mathbb{R}$, so there exists $\delta > 0$ such that for all x, $(F^q - \text{id})(x) < p - \delta$. This means that a sufficiently small perturbation G of F satisfies $(G^q - \text{id})(x) < p$ for all x. Hence the rotation number of the circle map whose lift is G is less than p/q. We can run a similar argument to show that the rotation number of this circle map is greater than r/s as well, proving the lemma.

4.7 Connection with Denjoy

In the last chapter we mentioned the Denjoy counterexample, a C^1 diffeomorphism of the circle which is not topologically conjugate to a rotation. Suppose for a twist map f we obtain an Aubry-Mather set I which is a Cantor subset of an invariant circle of f. Then the restriction of the twist map to the invariant circle is essentially a circle map whose invariant set is a Cantor subset of the circle. This means that the circle map cannot be topologically conjugate to a rotation—hence it is a Denjoy counterexample.

4.8 Other proofs

There are several other strategies for proving the Aubry-Mather theorem. We will briefly discuss three other strategies.

4.8.1 The proof of G. R. Hall

Hall's proof [11] differs primarily in its construction of periodic ordered orbits for area-preserving twist maps. The concept of a Birkhoff periodic point is not needed in this approach. We give a brief outline.

First, Hall proves that each rational number p/q in the twist interval corresponds to two distinct periodic orbits of rotation number p/q. However, these periodic orbits are not guaranteed to be Birkhoff orbits, or even ordered orbits. The proof of the existence of periodic orbits, furthermore, is not based on an the extremal of any function—Hall takes a topological approach. The idea is to take the set of all p/q-periodic points of a twist map's lift F, denoted by Σ, and then perform a series of topological operations on Σ. Let $S = \mathbb{R} \times [0, 1]$. Then let U_1 be the component of $S \backslash \Sigma$ that contains $\mathbb{R} \times \{0\}$. Let V be the component of $S \backslash \bar{U_1}$ which contains $\mathbb{R} \times \{1\}$, where the bar denotes closure. Finally, we take the boundary of the complement of the closure of V. Call this set Γ. Hall proves that every point in $F^{-1}(\Gamma) \cap \Gamma$ is a p/q-periodic point for F.

Next, through a series of careful lemmas that rely on the topology of the plane, Hall shows that to every p/q-periodic orbit there is an associated ordered p/q-periodic orbit.

The final step is similar to Katok's approach: Hall takes a sequence of ordered periodic orbits of rotation number p_n/q_n converging to some irrational α in the twist interval. He takes from each orbit one p_n/q_n-periodic point and constructs a sequence z_n of periodic points (which come from ordered orbits). A subsequence of these points must converge to a point z_α. Using a simple lemma, Hall shows that the orbit of z_α must be an ordered orbit with rotation number equal to α.

One advantage to Hall's approach is that, because his results are rooted in the two-dimensionality of the annulus, his proofs can be generalized to twist maps which do not preserve area but instead satisfy weaker topological conditions.

4.8.2 The proof of J. N. Mather

Mather [10] proves the existence of quasi-periodic orbits without going through the intermediate step of showing the existence of periodic orbits. The main feature of his proof is an Euler-Lagrange condition which is a generalization of the argument used above to show the existence of Birkhoff periodic points. Given an irrational number α in the twist interval, Mather shows that when a cleverly defined functional $F_\alpha(\phi)$ takes its maximum on the space of all weakly order preserving mappings from $\mathbb{R}$ to $\mathbb{R}$, an Euler-Lagrange equation $V(\phi, t) = 0$ is satisfied for all real t. From this condition it is easily shown that M_ϕ, the closure of the set of $(\phi(t), \eta(t))$ such that ϕ is continuous at t, is either the entire real line or a Cantor set invariant under the lift of the twist map. (Here $\eta(t)$ is a function related to $\phi(t)$.) Hence the projection of M_ϕ gives us our Aubry-Mather set.

The great advantage to Mather's variational approach is that it is not difficult to extend his arguments to produce existence proofs of dynamical behavior besides the quasi-periodic Aubry-Mather behavior. With Katok's approach, it is possible to show the three behaviors associated with rational rotation numbers: the Birkhoff periodic orbit, the orbit that is asymptotic in forward and backward iterations to a single periodic orbit, and the orbit that is asymptotic in forward and backward iterations to two periodic orbits. But it is not possible to show the three kinds of irrational rotation number behavior: an orbit dense in the circle, an orbit dense in a Cantor set, and an orbit homoclinic to a Cantor set. Katok's approach can be used only to show the first two kinds of behavior for irrational rotation number orbits. With Mather's approach, the third kind of behavior can be described as well.

Another advantage to Mather's approach is that, because his methods do not rely on any Lipschitz conditions, he is able to state all of the results for area-preserving twist homeomorphisms of the annulus. The maps do not have to be differentiable.

4.8.3 The proof of Katznelson and Ornstein

Y. Katznelson and D. S. Ornstein [9] have the most novel approach to proving the existence of Aubry-Mather sets: they examine the images of curves in the annulus and using various "trimming operators," they actually trim the curves down to size until they are one-sided-invariant graphs of functions which are either continous or contain a countable number of jump discontinuities. In the former case, the graph is an invariant curve of the twist map, while in the latter, the graph contains a minimal invariant set, which is precisely the Aubry-Mather set we are after.

Their discussion centers on the concept of a (left) pseudograph, which is obtained by drawing in the vertical line segments on the graph of a function which has only positive jump discontinuities. The idea is to take a function ψ, and create a set that contains the graph of ψ, all the points $(\theta, \psi(\theta-))$ and $(\theta, \psi(\theta+))$, and the vertical line segments connecting $(\theta, \psi(\theta-))$ and $(\theta, \psi(\theta+))$ at every point of discontinuity θ of ψ.

By examining the action of a twist map f on these pseudographs, and inventing appropriate trimming recipes for pseudographs, they are able to show the existence of invariant sets of all rotation numbers, including Birkhoff periodic orbits and Aubry-Mather sets. The main idea of trimming is to take a curve $\gamma(t) = (\theta(t), r(t))$, find a maximal closed subset E of the circle on which $\theta(t)$ is monotone, and then delete the parts of γ which are not in E and replace them with vertical segments connecting the endpoints. This results in what is called a proper curve. The $\limsup$ of an infinite sequence of maps consisting of trimming, then twisting, trimming, then twisting, etc. applied to a (left) proper curve is a (left-proper) pseudograph. If these trimming operations are proper, i.e. if they preserve the area of the region bounded by the pseudograph and the inner circle of the annulus, then in fact the $\limsup$ is a curve invariant under the inverse of the twist map. From here it is an easy step to obtain the actual Aubry-Mather sets and show they have the right rotation number.

The method of Katznelson and Ornstein has numerous advantages: no Lipschitz condition is required, no complicated and careful variational arguments are required, and the methods can be extended to construct many different kinds of orbits besides those contained in Aubry-Mather sets.

4.9 The three-body problem revisited

What do Aubry-Mather sets tell us about the dynamics of the three-body problem? We now know that the Poincaré map f, in a neighborhood of the elliptic fixed point p, has Aubry-Mather sets of all rotation numbers. The rational rotation numbers correspond to periodic orbits of the Poincaré map, which also correspond to periodic trajectories of the three bodies. Since there are infinitely many rational numbers in the twist interval of f, there are infinitely many such periodic orbits.

As for the irrational numbers in the twist interval, we know that f has invariant sets with those rotation numbers. Each set is either an invariant circle or a Cantor subset of the circle. In the former case, if the invariant circle is conjugate to an irrational rotation, the orbits on the circle are dense subsets of the circle. This corresponds to quasiperiodic motion of the three bodies—their configurations always return to the invariant circle whenever $Q_2 = 0$ and the Poincaré slice is reached, but their configuration never returns to a point on the slice that the trajectory has already passed through.

If the invariant circle is not conjugate to an irrational rotation, or then there must be an invariant Cantor set sitting inside the invariant circle. The restriction of the Poincaré map to the circle is a Denjoy counterexample. If we start with a point in the Cantor set, then the orbit of the point will fill the Cantor set densely—hence we witness quasiperiodic motion of the three bodies again, but this time, they do not come infinitesimally close to every point of the invariant circle. If we start with a point outside the Cantor set, then we will never end up in the Cantor set.

Finally, if there is no invariant circle of a certain irrational rotation number, the closed invariant ordered set with that rotation number must project in $1-1$ correspondence onto a Cantor subset of the circle. The orbits in this case can be complicated, including both orbits on the Cantor set and orbits that are asymptotic to the minimal Aubry-Mather set. In either case, if we plot the orbit of a point in the invariant set on a sheet of paper representing the Poincaré slice, we end up with "Cantor dust" that is spread out on a loop in phase space surrounding the elliptic fixed point. It will appear as though an invariant circle near the plotted points has not only been deformed but has had infinitely many arcs cut out of it. The motions of the three bodies will once again trace out a complicated quasiperiodic trajectory.

One important point about these quasiperiodic trajectories is that they are *not* chaotic. The motion is restricted to a very small area of phase space. Once we know the configuration of the system at a particular time t_0, if the configuration lies in an invariant ordered subset of the Poincaré slice which has irrational rotation number, even though the motion will not be periodic, we can generally predict what will happen. The fact that we can do all of this without ever writing down the solutions of our differential equations shows is just one small illustration of the power of the methods we have employed.

Appendix A

Poisson bracket calculations

The six planar isosceles initial conditions guarantee that the system will always lie in an isosceles triangle.

One way to show this is to recast the problem as a nonlinear Hamiltonian system. Using Poisson brackets, we will be able to show very easily that the planar isosceles conditions are invariant under the time-evolution of our Hamiltonian system. This will involve using the fact that the solution $\mathbf{y}(t)$ of a system of differential equations $\dot{\mathbf{y}} = \mathbf{h}(\mathbf{y})$ is a real analytic function if $\mathbf{h}$ is a real analytic function.

To recast the problem in the Hamiltonian formalism, we introduce the following choice of canonical coordinates:

$$\begin{array}{ll} q_1 = x_1 & p_1 = m_1\dot{x}_1 = m_1 u_1 \\ q_2 = x_2 & p_2 = m_2\dot{x}_2 = m_2 u_2 \\ q_3 = x_3 & p_3 = m_3\dot{x}_3 = m_3 u_3 \\ q_4 = y_1 & p_4 = m_1\dot{y}_1 = m_1 v_1 \\ q_5 = y_2 & p_5 = m_2\dot{y}_2 = m_2 v_2 \\ q_6 = y_3 & p_6 = m_3\dot{y}_3 = m_3 v_3 \end{array}$$

The energy integral for the original system is the kinetic energy plus the potential energy. In the original coordinates, this is

$$h = \frac{1}{2}\sum_{i=1}^{3} m_i(\dot{x}_i^2 + \dot{y}_i^2) + U$$

where U is the potential energy from (1).

Substituting our canonical variables into the above expression for h, we (naively) obtain an expression for the Hamiltonian function $H : O \to \mathbb{R}$ where $O \subset \mathbb{R}^{12}$:

$$H = \sum_{i=1}^{3} \frac{{p_i}^2 + {p_{i+3}}^2}{2m_i} - \sum_{1 \le k < l \le 3} \frac{G m_k m_l}{\sqrt{(q_k - q_l)^2 + (q_{k+3} - q_{l+3})^2}} \tag{A.1}$$

Because the gravitational force is conservative, we have an autonomous (nonlinear) Hamiltonian system with a 12-dimensional phase space. We remark that the whole reason we are trying to show that the g_i identically vanish for all t is to reduce the dimension of this monstrous phase space to a more manageable number.

Note that the Hamiltonian function is defined for all $\mathbf{z} = (\mathbf{q}, \mathbf{p}) \in \mathbb{R}^{12}$ *except* when $q_1 = q_2$, $q_1 = q_3$, $q_2 = q_3$, $q_4 = q_5$, $q_4 = q_6$, or $q_5 = q_6$. These six conditions correspond geometrically to six different 11-dimensional closed sets; their union is again a closed set, so the phase space O is an open subset of $\mathbb{R}^{12}$. Also note that H is smooth, indeed analytic, on O.

To assist us in our task of writing the equations of motion in a simple and elegant way, we now introduce the 12-dimensional vector z, the 12×12 matrix J, and the gradient of H by:

$$\mathbf{z} = \begin{pmatrix} q_1 \\ \vdots \\ q_6 \\ p_1 \\ \vdots \\ p_6 \end{pmatrix}, \quad J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \quad \boldsymbol{\nabla_z H} = \begin{pmatrix} \frac{\partial H}{\partial z_1} \\ \vdots \\ \frac{\partial H}{\partial z_{12}} \end{pmatrix} \tag{A.2}$$

In the definition above, I is simply the 6×6 identity matrix. Now the equations of motion for the system can be summarized neatly by:

$$\dot{\mathbf{z}} = J \boldsymbol{\nabla_z H}(\mathbf{z}) \tag{A.3}$$

A straightforward exercise shows that this is equivalent to the standard statement of Hamilton's equations: $\dot{q}_i = H_{p_i}$, $\dot{p}_i = -H_{q_i}$.

We now introduce the Poisson bracket:

$$\{F, G\} = (\boldsymbol{\nabla_z F})^T J \boldsymbol{\nabla_z G} \tag{A.4}$$

Here F and G are smooth functions from $O \subset \mathbb{R}^{12}$ to $\mathbb{R}$. The Poisson bracket results in such a function as well; that is, if we define $H := \{F, G\}$, then $H : O \to \mathbb{R}$ is smooth.

Now the elegance of the Hamiltonian formalism becomes clear. Let $g : O \to \mathbb{R}$ be a smooth function, and let H be the Hamiltonian function as before. Then we note that

$$\begin{aligned}
\{g,\, H\} &= (\boldsymbol{\nabla}_{\boldsymbol{z}}\boldsymbol{g})^T J \boldsymbol{\nabla}_{\boldsymbol{z}}\boldsymbol{H} \\
&= (\boldsymbol{\nabla}_{\boldsymbol{z}}\boldsymbol{g})^T \dot{\mathbf{z}} \\
&= \left(\frac{\partial g}{\partial z_1}, \,\cdots,\, \frac{\partial g}{\partial z_{12}}\right) \begin{pmatrix} \frac{\partial z_1}{\partial t} \\ \vdots \\ \frac{\partial z_{12}}{\partial t} \end{pmatrix} \\
&= \sum_{i=1}^{12} \left(\frac{\partial g}{\partial z_i}\frac{\partial z_i}{\partial t}\right) = \frac{dg}{dt}
\end{aligned}$$

Next, we recast the planar isosceles conditions, using six functions on phase space, $g_i : O \to \mathbb{R}$ for $i = 1, 2, ..., 6$:

$$\begin{aligned}
g_1(\mathbf{z}) &= q_1 + q_2 \\
g_2(\mathbf{z}) &= q_4 - q_5 \\
g_3(\mathbf{z}) &= q_3 \\
g_4(\mathbf{z}) &= p_1 + p_2 \\
g_5(\mathbf{z}) &= p_4 - p_5 \\
g_6(\mathbf{z}) &= p_3
\end{aligned}$$

Hence, the isosceles triangle initial conditions amount to stipulating that

$$g_i \circ \mathbf{z}(0) = 0 \quad \text{for} \quad i = 1, 2, ..., 6$$

for a given trajectory $\mathbf{z}(t) : \mathbb{R} \to O$, where $\mathbf{z}(t)$ satisfies (4).

Since $\{g_i,\, H\} = \frac{dg_i}{dt}$ for any function g on phase space, we have

$$\left.\frac{dg_i}{dt}\right|_{\mathbf{z}(0)} = \{g_i,\, H\} \circ \mathbf{z}(0) = 0, \tag{A.5}$$

with detailed calculations to follow. The first calculations are short:

$$\frac{dg_1}{dt} = \frac{d}{dt}(q_1 + q_2) = \left(\frac{p_1}{m_1} + \frac{p_2}{m_2}\right) = \frac{1}{m} g_4, \text{ and therefore}$$

$$\left.\frac{dg_1}{dt}\right|_{\mathbf{z}(0)} = \frac{1}{m} g_4 \circ \mathbf{z}(0) = 0$$

Similarly,

$$\left.\frac{dg_2}{dt}\right|_{\mathbf{z}(0)} = \frac{1}{m} g_5 \circ \mathbf{z}(0) = 0$$

and

$$\left.\frac{dg_3}{dt}\right|_{\mathbf{z}(0)} = \frac{1}{m} g_6 \circ \mathbf{z}(0) = 0$$

The remaining calculations require actual evaluation of the Poisson brackets. To evaluate $\{F, G\}$ for a given F and G, we need an expression in terms of coordinates. For our problem,

$$\{F, G\} = (\boldsymbol{\nabla}_{\boldsymbol{z}} \boldsymbol{F})^T J \boldsymbol{\nabla}_{\boldsymbol{z}} \boldsymbol{G} = \left(\frac{\partial F}{\partial z_1}, \; \cdots, \; \frac{\partial F}{\partial z_{12}}\right) \; J \; \begin{pmatrix} \frac{\partial G}{\partial z_1} \\ \vdots \\ \frac{\partial G}{\partial z_{12}} \end{pmatrix}$$

This is equivalent to

$$\begin{aligned}
\{F, G\} &= \left(\frac{\partial F}{\partial q_1}, \; \cdots, \; \frac{\partial F}{\partial q_6}, \; \frac{\partial F}{\partial p_1}, \; \cdots, \; \frac{\partial F}{\partial p_6}\right) \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial G}{\partial q_1} \\ \vdots \\ \frac{\partial G}{\partial q_6} \\ \\ \frac{\partial G}{\partial p_1} \\ \vdots \\ \frac{\partial G}{\partial p_6} \end{pmatrix} \\
&= \left(\frac{\partial F}{\partial q_1}, \; \cdots, \; \frac{\partial F}{\partial q_6}, \; \frac{\partial F}{\partial p_1}, \; \cdots, \; \frac{\partial F}{\partial p_6}\right) \begin{pmatrix} \frac{\partial G}{\partial p_1} \\ \vdots \\ \frac{\partial G}{\partial p_6} \\ \\ -\frac{\partial G}{\partial q_1} \\ \vdots \\ -\frac{\partial G}{\partial q_6} \end{pmatrix} \\
&= \sum_{i=1}^{6} \left(\frac{\partial F}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q_i}\right)
\end{aligned}$$

Now we can calculate $\frac{dg_4}{dt}$:

$$\begin{aligned}
\{g_4,\, H\} &= \{p_1 + p_2,\, H\} = -\frac{\partial H}{\partial q_1} - \frac{\partial H}{\partial q_2} \\
\frac{\partial H}{\partial q_1} &= \frac{Gm_1m_2(q_1 - q_2)}{\left[(q_1 - q_2)^2 + (q_4 - q_5)^2\right]^{3/2}} + \frac{Gm_1m_3(q_1 - q_3)}{\left[(q_1 - q_3)^2 + (q_4 - q_6)^2\right]^{3/2}}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial H}{\partial q_2} &= -\frac{Gm_1m_2(q_1 - q_2)}{\left[(q_1 - q_2)^2 + (q_4 - q_5)^2\right]^{3/2}} + \frac{Gm_2m_3(q_2 - q_3)}{\left[(q_2 - q_3)^2 + (q_5 - q_6)^2\right]^{3/2}} \\
\{g_4,\, H\} &= -\frac{Gm_1m_3(q_1 - q_3)}{\left[(q_1 - q_3)^2 + (q_4 - q_6)^2\right]^{3/2}} - \frac{Gm_2m_3(q_2 - q_3)}{\left[(q_2 - q_3)^2 + (q_5 - q_6)^2\right]^{3/2}}
\end{aligned}$$

Finally,

$$\left.\frac{dg_4}{dt}\right|_{\mathbf{z}(0)} = \{g_4,\, H\} \circ \mathbf{z}(0) = -\frac{Gmm_3(q_1 + q_2)}{\left[{q_1}^2 + (q_4 - q_6)^2\right]^{3/2}} = 0 \qquad \text{(A.6)}$$

Next we calculate $\frac{dg_5}{dt}$:

$$\begin{aligned}
\{g_5,\, H\} &= \{p_4 - p_5,\, H\} = -\frac{\partial H}{\partial q_4} + \frac{\partial H}{\partial q_5} \\
\frac{\partial H}{\partial q_4} &= \frac{Gm_1m_2(q_4 - q_5)}{\left[(q_1 - q_2)^2 + (q_4 - q_5)^2\right]^{3/2}} + \frac{Gm_1m_3(q_4 - q_6)}{\left[(q_1 - q_3)^2 + (q_4 - q_6)^2\right]^{3/2}} \\
\frac{\partial H}{\partial q_5} &= -\frac{Gm_1m_2(q_4 - q_5)}{\left[(q_1 - q_2)^2 + (q_4 - q_5)^2\right]^{3/2}} + \frac{Gm_2m_3(q_5 - q_6)}{\left[(q_2 - q_3)^2 + (q_5 - q_6)^2\right]^{3/2}}
\end{aligned}$$

$$\begin{aligned}
\{g_5,\, H\} = &-\frac{2Gm_1m_2(q_4 - q_5)}{\left[(q_1 - q_2)^2 + (q_4 - q_5)^2\right]^{3/2}} - \frac{Gm_1m_3(q_4 - q_6)}{\left[(q_1 - q_3)^2 + (q_4 - q_6)^2\right]^{3/2}} \\
&+ \frac{Gm_2m_3(q_5 - q_6)}{\left[(q_2 - q_3)^2 + (q_5 - q_6)^2\right]^{3/2}}
\end{aligned}$$

Finally,

$$\left.\frac{dg_5}{dt}\right|_{\mathbf{z}(0)} = \{g_5, H\} \circ \mathbf{z}(0)$$

$$= 0 - \frac{Gmm_3(q_4 - q_6)}{\left[{q_1}^2 + (q_4 - q_6)^2\right]^{3/2}} + \frac{Gmm_3(q_5 - q_6)}{\left[{q_2}^2 + (q_5 - q_6)^2\right]^{3/2}}$$

$$= \frac{Gmm_3}{\left[{q_1}^2 + (q_4 - q_6)^2\right]^{3/2}} (-q_4 + q_6 + q_5 - q_6) = 0$$

as required. The last calculation we will carry out is $\frac{dg_6}{dt}$:

$$\{g_6, H\} = \{p_3, H\} = -\frac{\partial H}{\partial q_3}$$

$$= \frac{Gm_1m_3(q_1 - q_3)}{\left[(q_1 - q_3)^2 + (q_4 - q_6)^2\right]^{3/2}} + \frac{Gm_2m_3(q_2 - q_3)}{\left[(q_2 - q_3)^2 + (q_5 - q_6)^2\right]^{3/2}}$$

Therefore,

$$\left.\frac{dg_6}{dt}\right|_{\mathbf{z}(0)} = \{g_6, H\} \circ \mathbf{z}(0) = \frac{Gmm_3(q_1 + q_2)}{\left[{q_1}^2 + (q_4 - q_6)^2\right]^{3/2}} = 0$$

Thus we have shown that for each i, $\frac{dg_i}{dt}$ evaluated at $\mathbf{z}(0)$ gives zero. We remark that using similar tricks, we can show that $\frac{d^n g_i}{dt^n}$ evaluated at $\mathbf{z}(0)$ gives zero for all n. (In fact, we have successfully employed Mathematica scripts to evaluate the Poisson brackets to verify this claim.) It is merely a matter of noting that when successive p_i and q_i derivatives are taken of the above brackets, the new terms that appear in the numerator will always sum to zero when the planar isosceles conditions at $\mathbf{z}(0)$ are substituted.

Consider again the trajectory $\mathbf{z}(t)$. We know it must satisfy (4). Since H is analytic on O, we have a system of the form

$$\dot{\mathbf{z}} = F(\mathbf{z})$$

where F is an analytic function on O. Now we can apply the following standard theorem [4]:

Theorem A.0.1. *Given the first-order system of differential equations,*

$$\frac{d\mathbf{x}}{dt} = X(\mathbf{x}, t),$$

if $X(\mathbf{x}, t)$ is an analytic real function of the real variables $x_1, ..., x_n$ and t, then every solution of the above system is analytic.

We conclude that $\mathbf{z}(t)$ must be an analytic function of t. Since each g_i is an analytic function on O, each $(g_i \circ \mathbf{z})$ must be an analytic function of t. Expanding about $t = 0$, we have

$$(g_i \circ \mathbf{z})(a) = \sum_{i=0}^{\infty} \frac{[d^n(g_i \circ \mathbf{z})/dt^n]_{t=0}\, a^n}{n!} = \sum_{i=0}^{\infty} \frac{[d^n g_i/dt^n]_{\mathbf{z}=\mathbf{z}(0)}\, \dot{\mathbf{z}}(0)}{n!}$$

By the remark above, each term in the series is zero. Hence $(g_i \circ \mathbf{z})(a)$ is zero for all a such that $|a| < R(i, 0)$, where $R(i, 0) > 0$ is the radius of convergence for the power series of $(g_i \circ \mathbf{z})$ about $t = 0$. Note that there is nothing special about $t = 0$. We merely stipulated that the planar isosceles triangle conditions were to hold just at $t = 0$, and now we know they hold for all $t \in [-a, a]$. Running the same argument over again at the endpoints of this closed interval, we can extend the result for all $t \in \mathbb{R}$. ■

Appendix B

The Poincaré map is a diffeomorphism

Since the vector field X_K is smooth on M, the flow ψ^t is a one-parameter family of diffeomorphisms. We will use this fact to show that f is a diffeomorphism on S. Let g be the function defined by $g(t, \mathbf{z}) = \pi_3 \circ \psi^t(\mathbf{z})$, where π_3 denotes projection onto the third component, and $\mathbf{z} = (Q_1, P_1, Q_2, P_2) \in M$ such that $(Q_1, P_1) \in S$. Write $g(T, \mathbf{z}) = 0$. We wish to solve this equation for T, since T satisfies $Q_2(T) = 0$. (The initial conditions for the four variables, including $Q_2(0) = 0$, are given by $\mathbf{z}$.) We claim that $\partial g/\partial t \neq 0$. Since the velocity of the third mass can never be zero in this problem, the only way our claim could be false is if the third mass starts at $Q_2 = 0$ with zero velocity. This exception has been excluded by our choice of section S. By the implicit function theorem, there is a smooth solution $r(\mathbf{z})$ such that $g(r(\mathbf{z}), \mathbf{z}) = 0$. Since $r(\mathbf{z})$ is smooth, f must be smooth as well. The same argument can be used to show that f^{-1} is smooth, except that now we must define $r(\mathbf{z}) = T$ to be the greatest such $T < 0$ with $Q_2(T) = 0$ given initial conditions specified by $\mathbf{z}$. Hence f is a diffeomorphism of S.

I wish to express my thanks to Professor Daniel L. Goroff and Matt Leingang, from whom I received generous assistance in the form of questions, answers, discussions, encouragement, and patience. I am very grateful.

Bibliography

[1] Arnold, V.I. (1989). *Mathematical Methods of Classical Mechanics.* Second Edition, Graduate Texts in Mathematics **60**, Springer-Verlag.

[2] Arrowsmith, D.K. and Place, C.M. (1990). *An Introduction to Dynamical Systems.* Cambridge University Press.

[3] Bernstein, D. (1985). Birkhoff periodic orbits for twist maps with the graph intersection property. *Ergodic Theory and Dynamical Systems.* **5**, 531-537.

[4] Birkhoff, G. and Rota, G-C. (1969). *Ordinary Differential Equations.* Second Edition, Blaisdell.

[5] Birkhoff, G. D. (1927). *Dynamical Systems.* American Mathematical Society.

[6] Boyland, P.L. and Hall, G.R. (1987). Invariant Circles and the Order Structure of Periodic Orbits in Monotone Twist Maps. *Topology.* **26**, 21-35.

[7] Katok, A. (1982). Some remarks on Birkhoff and Mather twist map theorems. *Ergodic Theory and Dynamical Systems.* **2**, 185-194.

[8] Katok, A. and Hasselblatt, B. (1995). *Introduction to the Modern Theory of Dynamical Systems.* Encyclopedia of Mathematics and its Applications, **54**, Cambridge University Press.

[9] Katznelson, Y. and Ornstein, D. S. (1997). Twist Maps and Aubry-Mather Sets. *Contemporary Mathematics.* **211**, 343-357.

[10] Mather, J. N. (1982). Existence of Quasi-Periodic Orbits for Twist Homeomorphisms of the Annulus. *Topology.* **21**, 457-467.

[11] Meyer, K. R. and Hall, G. R. (1992). *Introduction to Hamiltonian Dynamical Systems and the N-Body Problem.* Applied Mathematical Sciences **90**, Springer-Verlag.

[12] Moser, J. (1973). *Stable and Random Motions in Dynamical Systems with Special Emphasis on Celestial Mechanics.* Princeton University Press.

[13] Zare, K. and Chesley, S. (1998). *Chaos.* **8**, 475-494.