UNDERSTANDING THE KAM THEOREM

JAMES ONG

Abstract. In the study of two-dimensional dynamical systems, one invariably stumbles upon the KAM theorem. More often than not, the results and corollaries of the theorem are quoted, but the theorem itself is not. A full statement of the theorem will often be stated in a form that requires a comprehensive knowledge of nomenclature used in previous sections. This report aims to gather together enough of a background for the reader to understand the KAM theorem as restricted to $\mathbb{R}^2$.

Contents

1. Introduction

Celestial mechanics has historically been an area of interest for mathematicians and physicists alike. One of the key problems, the stability problem, has traditionally been concerned with the form of the solar system in the far future. Its results relate to whether the planets will continue in their orbits around the sun or leave the solar system; it should predict collisions and take into account the forces that each body in the solar system exerts on everything else.

The stability problem has been solved a number of times by using various approximations to reduce the number of factors considered. However, to get a more accurate solution, methods are required that take into account the subtle perturbations that 'small body' approximations ignore. The KAM theorem gives some stability properties to tori that are perturbed by 'sufficiently small' perturbations, giving some insight into the problem. See Arnol'd [2, pp125-] or Moser [14, pp65-71] for a discussion of the stability of planetary motions using the KAM theorem.

The problem with understanding the KAM theorem arises from the amount of background knowledge required to appreciate the conditions imposed upon perturbations and to motivate the persistence of tori with a sufficient irrationality condition. A large amount of the background can easily be motivated by the use of one-dimensional maps, since the study of one-dimensional maps is one of the easiest ways of being introduced to the area of dynamical systems. Although they are inherently simple compared to higher dimensional maps, they retain many fundamental features that are important to note, like periodicity and attractors.

One-dimensional circle maps are the simplest way to be introduced to important properties of tori required for the KAM theorem, as well as properties of general dynamical systems like stability and periodicity. An overview of properties of circle maps will thus form the early sections of my report. Analogues in two dimensions will be illustrated, leading to a statement of the KAM theorem.

2. Useful Theorems

We will state two theorems that will be used in later sections to help us prove various results.

Theorem 2.1 (Intermediate Value Theorem). *Let I be an interval in $\mathbb{R}$ and $f : I \to \mathbb{R}$ be continuous on I. If $a, b \in I$ and if $k \in \mathbb{R}$ satisfies $f(a) < k < f(b)$, then there exists a point $c \in (a, b)$ such that $f(c) = k$.*

Proof. See Bartle and Sherbert [5, p156]. $\square$

Theorem 2.2 (Implicit Function Theorem). *Suppose $F : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$, $(x, \lambda) \to F(x, \lambda)$ is a C^1 function, where $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$ and*

$$F(x_0, \lambda_0) = 0 \quad \textit{and} \quad \frac{\partial F}{\partial \lambda}(x_0, \lambda_0) \neq 0.$$

Then there exists an neighbourhood U of x_0, a neighbourhood V of λ_0, and a unique C^1 function $z = \phi(x)$, $\phi : U \to V$ satisfying

$$\phi(x_0) = \lambda_0$$

and

$$F(x, \phi(x)) = 0$$

for all $x \in U$.

Proof. See Marsden and Tromba [12, pp280–]. $\square$

3. Basic concepts

In this section, we will define a number of useful concepts that will be used throughout the report.

Definition (S^1). The circle S^1 is defined to be $\mathbb{R}/\mathbb{Z}$, or equivalently the interval [0,1] with the endpoints identified.

Definition (Homeomorphism). Let $(X, \mathcal{T})$ and $(Y, \mathcal{U})$ be topological spaces. A map $f : X \to Y$ is said to be a homeomorphism provided that it is a bijection, and both f and f^{-1} are continuous.

Definition (Diffeomorphism). A map f is said to be a diffeomorphism provided that it is a homeomorphism, and both f and f^{-1} are of the class C^1.

The nth iterate of the map f will be denoted f^n. The identity mapping $x \mapsto x$ will be denoted Id.

3.1. Orientation.

In later sections, we will be working mainly with orientation preserving homeomorphisms.

Definition (Orientation on $\mathbb{R}^n$). Take a basis $\{v_i\}$ of $\mathbb{R}^n$, and define the coordinate transformation $A : \mathbb{R}^n \to \mathbb{R}^n$ as $Ae_i = v_i$, where $\{e_i\}$ is the standard set of orthonormal basis vectors

$$\begin{aligned} e_1 &= (1, 0, 0, \ldots, 0) \\ e_2 &= (0, 1, 0, \ldots, 0) \\ &\vdots \\ e_n &= (0, 0, 0, \ldots, 1). \end{aligned}$$

We take the orientation of $\{v_i\}$ to be positive if

$$\det A > 0,$$

Definition (Projection). The projection $\Pi : \mathbb{R} \to S^1$ is defined to be $x \mapsto x + \mathbb{Z}$.

Definition (Orientation on S^1). The orientation on S^1 is defined as the orientation induced from the real line via Π.

Definition (Orientation preserving). Let $U \subset \mathbb{R}^n$ be open. A differentiable mapping $M : U \to \mathbb{R}^n$ is orientation preserving if

$$\det DM(x) > 0$$

for all $x \in U$, where DM is the derivative of M.

3.2. **Lift of a function.** When studying functions defined on S^1, it is often easier to lift them to $\mathbb{R}$ to prove a result. Most often, we will use a standard lift which is continuous.

Definition (Lift). A lift $F : \mathbb{R} \to \mathbb{R}$ of $f : S^1 \to S^1$ is defined as

$$\Pi(F(x)) = f(\Pi(x)).$$

Lemma 3.1. *Let $f : S^1 \to S^1$ be an orientation preserving homeomorphism with lift F continuous. Then*

$$F(x+1) = F(x) + 1.$$

Proof. We know that

$$f(\Pi(x)) = f(\Pi(x+1)).$$

Using the definition of lift, we get

$$\Pi(F(x)) = \Pi(F(x+1))$$

and therefore

$$F(x+1) = F(x) + k(x)$$

where $k(x)$ can take only integer values. Since F is continuous, $k(x)$ must be a constant integer value k.

Suppose $k \geq 2$, so $F(x+1) \geq F(x) + 2$. By the Intermediate Value Theorem, there must be two values $x_1, x_2 \in [x, x+1)$ such that $F(x_2) = F(x_1) + 1$. Thus $\Pi(x_1) \neq \Pi(x_2)$ but $f(x_1) = f(x_2)$. However, this cannot be true since f is a homeomorphism.

Suppose now that $k \leq 0$. Then $F(x+1) \leq F(x)$ which cannot occur since F is a homeomorphism and orientation preserving (or equivalently strictly monotone increasing).

Therefore $k = 1$, so $F(x+1) = F(x) + 1$. $\square$

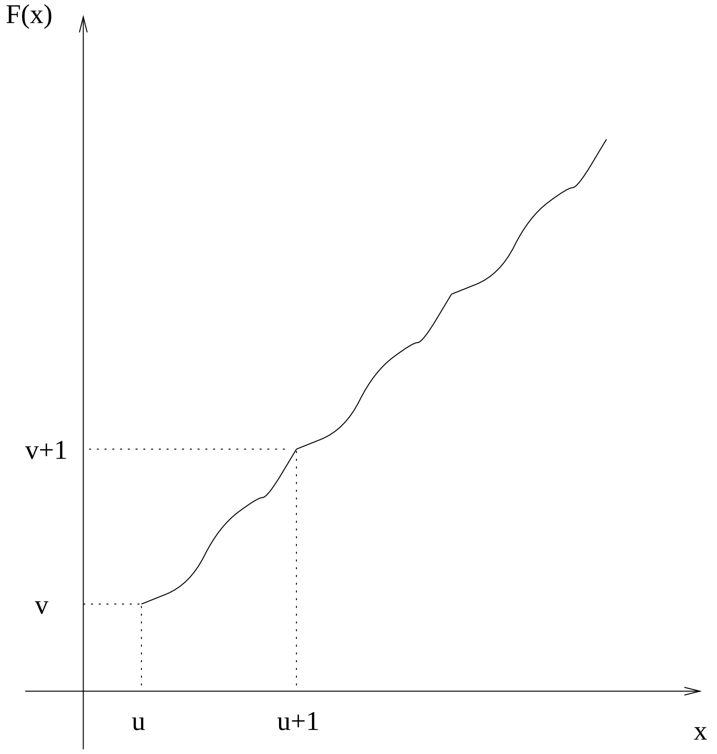

FIGURE 1. The standard lift

We will call $F : \mathbb{R} \to \mathbb{R}$ a standard lift of $f : S^1 \to S^1$ if it has the property above that $F(x+1) = F(x)+1$. This makes F a monotone increasing function which then gives us invertibility, causing F to be a homeomorphism. Note that orientation reversing mappings can be dealt with in a similar fashion, defining $F(x+1) = F(x)-1$. Other general mappings will not be invertible, and so not be homeomorphisms.

One way of understanding the standard lift is to cut the circle at one point, straighten it out and identify it with an interval of unit length. Then we have a function $\tilde{f}$ mapping domain $[u, u+1]$ to the interval $[v, v+1]$ for some $u, v \in \mathbb{R}$, satisfying $\tilde{f}(x+1) = \tilde{f}(x)+1$. Now take this function defined on a unit interval and extend it to $\mathbb{R}$ by copying the unit square and matching endpoints (see Figure 1).

Lemma 3.2. *Let $F : \mathbb{R} \to \mathbb{R}$ be a standard lift of $f : S^1 \to S^1$ such that $F(x+1) = F(x)+1$. Then $F^q(x+1) = F^q(x)+1$.*

Proof. We will prove this by induction. The claim is true for $q = 1$. Assume that it is true for $q = k$.

$$\begin{aligned} F^{k+1}(x+1) &= F\left(F^k(x+1)\right) \\ &= F\left(F^k(x)+1\right) \\ &= F^{k+1}(x)+1 \end{aligned}$$

Thus the claim is true for all q. $\square$

Invariant sets and limit sets of a dynamical system are essential features that help to define behaviour of orbits. The next two sections will introduce these concepts.

3.3. **Invariance.** In the following section, we will take a set U in some metric space X. We will denote its closure as $\overline{U}$ and its boundary as δU.

Definition (Orbit)**.** The orbit of a point is the set of all forward and backward iterates of the point under a mapping.

Definition (Invariant set)**.** An invariant set $U \subseteq X$ under the mapping $f : X \to X$ is defined by $f(U) = U$.

Lemma 3.3. *The closure of an invariant set is also invariant under a continuous bijective mapping.*

Proof. Let X be a metric space, $f : X \to X$ be a continuous bijective mapping and $U \subseteq X$ an invariant set. Any element x in the boundary of U has $d(x, U) = 0$ and thus $d(f(x), f(U)) = d(f(x), U) = 0$ by continuity. Therefore, $f(x)$ is also in $\overline{U}$. Because the mapping is bijective, then $f(\overline{U}) = \overline{U}$. $\square$

Lemma 3.4. *If U is a closed invariant set in a metric space X under a continuous bijective mapping, then δU is a closed invariant set.*

Proof. δU is closed, since it is the intersection of the two closed sets $\overline{U}$ and $\overline{X \backslash U}$. $X \backslash U$ must be an open invariant set, since X is invariant. Then the closure $\overline{X \backslash U}$ must be a closed invariant set due to continuity. Since

$$\delta U = \overline{U} \cap \overline{X \backslash U} = \left(\overline{X \backslash U}\right) \backslash (X \backslash U)$$

then δU must be invariant. $\square$

3.4. **Limit set.**

Definition (Limit point). A point y is said to be an ω_x-limit point of the mapping $f : X \to X$ if there is an infinite sequence $\{m_i\}$, $m_i \in \mathbb{Z}$ such that

$$\lim_{i\to\infty} f^{m_i}(x) = y$$

where $x, y \in X$. A point z is said to be an α_x-limit point of the mapping $f : X \to X$ if there is an infinite sequence $\{m_i\}$, $m_i \in \mathbb{Z}$ such that

$$\lim_{i\to\infty} (f^{-1})^{m_i}(x) = z$$

where $x, y \in X$.

Definition (Limit set). $L_f^\omega(x)$ is defined as the set of all ω_x-limit points of f, and $L_f^\alpha(x)$ is likewise defined as the set of all α_x-limit points of f. $L_f^\omega(x)$ is called the ω-limit set, while $L_f^\alpha(x)$ is called the α-limit set.

4. Circle homeomorphisms

The aim of this section is to introduce the concept of rotation number as applied to circle homeomorphisms. This concept will turn out to be useful to explain behaviour of curves under perturbation. Indeed, classification of homeomorphisms based on the rationality of the rotation number helps to give an insight into the persistence of KAM tori.

Section 4.1 has been adapted from de Melo and van Strien [6, pp16-17]. Much of the content of sections 4.2, 4.3, 4.4, 4.5 has been adapted from Katok and Hasselblatt [9, pp26-27,387-396].

4.1. **Periodic points in orientation preserving homeomorphisms.** In this section, we will see that the presence of periodic points is intimately related to asymptotic behaviour of orbits. Indeed, both the α- and ω-limit sets for orientation preserving homeomorphisms in S^1 are periodic orbits.

Lemma 4.1. *Let U be a closed interval and $f : U \to U$ be an orientation preserving continuous injective map. Then either all $x \in U$ are fixed or the limit sets $L_f^\omega(x)$ and $L_f^\alpha(x)$ are adjacent fixed points.*

Proof. Let U_{fixed} be the set of fixed points in U. If $U_{\text{fixed}} = U$ then there is nothing to prove. Otherwise, take some $y \in U \backslash U_{\text{fixed}}$, and let (a, b) be the maximal open interval of $U \backslash U_{\text{fixed}}$ containing y. Consider the case where $f(y) > y$. Then

$$f^2(y) = f(f(y)) > f(y)$$

since f is monotone increasing. Similarly

$$f^k(y) = f(f^{k-1}(y)) > f^{k-1}(y).$$

Thus $f^n(y)$ is monotone increasing, bounded above by b and converges to some $\bar{y} = \lim_{n\to\infty}\{f^n(y)\}$. Since f is continuous,

$$f(\bar{y}) = f(\lim_{n\to\infty} f^n(y)) = \lim_{n\to\infty} f^{n+1}(y) = \bar{y}.$$

Thus $\bar{y} \in U_{\text{fixed}}$, and so $\bar{y} = b$. Similarly, if $f(y) < y$, then iterates $f^n(y)$ are monotone decreasing and converge to a.

Now if f is monotone increasing then f^{-1} is monotone decreasing, and similarly if f is monotone decreasing then f^{-1} is monotone increasing. Thus we can use the same arguments as above to show that $L_f^\omega(x)$ and $L_f^\alpha(x)$ must be the adjacent fixed points a and b. $\square$

Theorem 4.2. *Let $f : S^1 \to S^1$ be an orientation preserving homeomorphism with periodic points $y, f(y), \ldots, f^{q-1}(y)$, with $f^q(y) = y$. Then for all $x \in S^1$, $L_f^\omega(x) = \{f^r(\bar{x}); r = 0, 1, \ldots, q-1\}$ where $f^q(\bar{x}) = \bar{x}$, for some q-periodic point $\bar{x} \in S^1$.*

Proof. f has a q-cycle of periodic points, so f^q must have fixed point y. Since f is orientation preserving, then f^q restricted to $U = S^1\backslash\{y\}$ is monotone increasing. Using Lemma 4.1, $L_{f^q}^\omega(x) = \bar{x}$ where $\bar{x}$ is a fixed point of f^q (or equivalently a q-periodic point of f). Now consider iterates $f^{nq+r}(x)$, $r = 1, \ldots, q-1$. For each value of r,

$$\begin{aligned}\lim_{n\to\infty} f^{nq+r}(x) &= \lim_{n\to\infty} f^r(f^{nq}(x)) \\ &= f^r\left(\lim_{n\to\infty} f^{nq}(x)\right) \\ &= f^r(\bar{x})\end{aligned}$$

Thus $L_f^\omega(x) = \{f^r(\bar{x}); r = 0, 1, \ldots, q-1\}$. $\square$

Since f is a homeomorphism in this case, $f(S^1) = S^1$, and we can apply Theorem 4.2 to f^{-1} to show that $L_f^\alpha(x) = \{f^r(\bar{x}^*); r = 0, 1, \ldots, q-1\}$ for some q-periodic point $\bar{x}^* \in S^1$.

We know that the set $L_{\text{per}} = L_f^\omega(S^1) = L_f^\alpha(S^1)$ is the set of periodic points of f. Using the two results above, if the q-cycles $L_f^\omega(S^1\backslash L_{\text{per}})$ and $L_f^\alpha(S^1\backslash L_{\text{per}})$ have empty intersection, then $L_f^\omega(S^1\backslash L_{\text{per}})$ is stable, while $L_f^\alpha(S^1\backslash L_{\text{per}})$ is unstable. Thus there exists one stable q-cycle and one unstable q-cycle.

4.2. **Rotations on the circle.** One of the simplest orientation preserving homeomorphisms we can consider is a pure rotation. Iterates under a rotation are either periodic, if we rotate by a rational amount, or fill the circle densely, if we rotate by an irrational amount.

Definition (Rotation). A rotation $\mathbf{R}_\alpha : S^1 \to S^1$ is defined on as $\mathbf{R}_\alpha x = \Pi(x+\alpha)$.

Note that the rotation mapping is both orientation preserving and length preserving, since it is the projection of a linear mapping. This mapping can be thought of as a rotation in some positive direction by angle $2\pi\alpha$.

4.2.1. *Rotation by a rational α.* When α is rational, we can write it as $\alpha = \frac{p}{q}$, where we can choose $p, q \in \mathbb{Z}$ to be relatively prime.

Lemma 4.3. *Rotation by a rational amount $\frac{p}{q}$ is a periodic mapping of period q.*

Proof.

$$\begin{aligned} \mathbf{R}_\alpha^q x &= \Pi\left(x + q\left(\frac{p}{q}\right)\right) \\ &= \Pi(x+p) \\ &= x. \end{aligned}$$

□

4.2.2. *Rotation by an irrational α.*

Theorem 4.4. *For irrational α, the orbit of a point is dense in S^1.*

Proof. We know that the orbit O of a point must be an invariant set. The closure of an invariant set is also invariant under a continuous bijective mapping by Lemma 3.3. Thus $\overline{O}$ is closed and invariant under rotation. The complement $C = S^1 \backslash \overline{O}$ must be open and invariant. Choose C to be nonempty, and take a connected component C_0 of C of maximum length. Each of the intervals $f^k(C_0)$ must have the same length because the rotation mapping preserves length. Thus the intervals $\{f^n(C_0); n \in \mathbb{Z}\}$ cannot all be disjoint. However, because we have initially taken a component of maximum length, one of the intervals $f^k(C_0)$ must map onto itself under f^m for some finite $m \in \mathbb{Z}$.

There must be a fixed point when mapping an interval (a, b) onto itself via a continuous function f (see below). This means that one of the intervals $f^k(C_0)$ has a fixed point under f^m. However, this would imply that we are rotating by a rational number. Therefore, there cannot exist a nonempty connected component C_0 of nonzero length.

However, since C_0 is open, it must be empty. Therefore O is dense on S^1.

We now want to prove our assumption that there is a fixed point when mapping an interval onto itself via a continuous function. We will apply the Intermediate Value Theorem to the function $f(x)-x$. If we assume that $f(x)-x$ does not have a zero at either $x=a$ or $x=b$, then $f(x)-x$ must be positive at $x=a$ and negative at $x=b$. This means that there must be a zero of $f(x)-x$ somewhere in (a,b). This implies that there is a fixed point of the mapping $f(x)=x$. $\square$

4.3. **Rotation number.** Let $f: S^1 \to S^1$ be an orientation preserving homeomorphism, $F: \mathbb{R} \to \mathbb{R}$ a standard lift of f and $\Pi: \mathbb{R} \to S^1$ the standard projection. Then we can define the rotation number as

$$\rho(f) = \Pi\left(\lim_{|n|\to\infty} \frac{F^n(x)-x}{n}\right).$$

Note that this definition requires that $\rho(f)$ does not depend on the initial x chosen, and that the limit exists for all x.

Lemma 4.5. *$\rho(f)$ is independent of x and exists for all $x \in S^1$.*

Proof. (i) Existence: For this proof, we will use the fact that for any monotone increasing function g, $g(x)-x$ varies in an interval of size at most 1 (see below).

This implies that for the lift F, $F^n(x)-x$ varies in an interval of size at most 1, since F^n is monotone increasing. For some $k \in \mathbb{Z}$ and $x_0 \in \mathbb{R}$

$$\begin{aligned}
f^{kn}(x_0) - x_0 &= \sum_{i=0}^{k-1} f^{(i+1)n}(x_0) - f^{in}(x_0) \\
&= \sum_{i=0}^{k-1} f^n(x_i) - x_i \\
&\qquad \text{where } x_i = f^{in}(x_0) \\
&= k(f^n(x_0) - x_0) + \epsilon(k,n,x_0) \\
&\qquad \text{where } |\epsilon(k,n,x_0)| \le (k-1) \\
\frac{f^{kn}(x_0) - (x_0)}{kn} &= \frac{f^n(x_0) - x_0}{n} + \delta(k,n,x_0) \\
&\qquad \text{where } \delta(k,n,x_0) = \frac{\epsilon(k,n,x_0)}{kn}
\end{aligned}$$

Note that $|\delta(k,n,x_0)| < \frac{1}{n}$.

$$\begin{aligned} & \left|\frac{f^n(x_0)-x_0}{n} - \frac{f^k(x_0)-x_0}{k}\right| \\ \leq & \left|\frac{f^n(x_0)-x_0}{n} - \frac{f^{kn}(x_0)-x_0}{kn}\right| + \left|\frac{f^{kn}(x_0)-x_0}{kn} - \frac{f^k(x_0)-x_0}{k}\right| \\ \leq & \frac{1}{n} + \frac{1}{k} \end{aligned}$$

This makes $\frac{f^n(x_0)-x_0}{n}$ a Cauchy sequence, so the limit must exist.

We now want to prove our assumption that $F(x) - x$ varies in an interval of size at most 1. We know that $F(x+1) = F(x) + 1$, so we can consider the interval $[0,1]$ since this describes an entire period of the function F. We know that F is monotone increasing, and maps $[0,1]$ onto $[F(0), F(1)]$. Take any two points $(a, F(a))$ and $(b, F(b))$, with $0 \leq a < b \leq 1$ and thus $F(0) \leq F(a) \leq F(b) \leq F(1)$. We will consider the four cases:

1. $a \leq f(a) - f(0), b \leq f(b) - f(0)$ for all a and b. Thus for all x,

$$f(0) \leq f(x) - x.$$

Since f is monotone increasing,

$$f(x) \leq f(0) + 1$$

with equality when $x = 1$, and thus

$$f(x) - x < f(0) + 1.$$

Put these results together to give

$$f(0) \leq f(x) - x < f(0) + 1.$$

2. $a \geq f(a) - f(0), b \geq f(b) - f(0)$ for all a and b. Thus for all x,

$$f(x) - x \leq f(0).$$

Since f is monotone increasing,

$$f(x) \geq f(0)$$

with equality when $x = 0$, and thus

$$f(x) - x > f(0) - 1.$$

Put these results together to give

$$f(0) - 1 < f(x) - x \leq f(0).$$

3. $a \leq F(a) - F(0), b \geq F(b) - F(0)$ for some a and b. $F(a) - a \geq F(0)$, while $F(b) - b \leq F(0)$. To find the overall fluctuation in $F(x) - x$, we thus consider

$$(F(a) - a) - (F(b) - b) = F(a) - F(b) + b - a \leq 1$$

since $F(a) - F(b) \leq 0$ and $b - a \leq 1$.

4. $a \geq F(a) - F(0), b \leq F(b) - F(0)$ for some a and b. $F(a) - a \leq F(0)$, while $F(b) - b \geq F(0)$. To find the overall fluctuation in $F(x) - x$, we thus consider

$$(F(b) - b) - (F(a) - a) = F(b) - F(a) + a - b \leq 1$$

since $F(b) - F(a) \leq 1$ and $a - b \leq 0$.

(ii) Uniqueness: Take any two different values of x, say x_1 and x_2.

$$\begin{aligned}
& \left| \left(\lim_{|n| \to \infty} \frac{F^n(x_1) - x_1}{n} \right) - \left(\lim_{|n| \to \infty} \frac{F^n(x_2) - x_2}{n} \right) \right| \\
= & \left| \lim_{|n| \to \infty} \left(\frac{F^n(x_1) - x_1}{n} - \frac{F^n(x_2) - x_2}{n} \right) \right| \\
\leq & \left| \lim_{|n| \to \infty} \frac{1}{n} \right| \\
& \quad \text{since } F^n(x) - x \text{ varies in an interval of size at most 1} \\
= & \ 0
\end{aligned}$$

$\square$

Lemma 4.6. *The rotation number of a pure rotation* $\mathbf{R}_\alpha$ *is* $\Pi(\alpha)$.

Proof. The lift $\hat{\mathbf{R}}_\alpha$ of $\mathbf{R}_\alpha$ maps $x \mapsto x + \alpha$.

$$\begin{aligned}
\rho(\mathbf{R}_\alpha) &= \Pi \left(\lim_{|n| \to \infty} \frac{\hat{\mathbf{R}}_\alpha^n(x) - x}{n} \right) \\
&= \Pi \left(\lim_{|n| \to \infty} \frac{x + n\alpha - x}{n} \right) \\
&= \Pi \left(\lim_{|n| \to \infty} \frac{n\alpha}{n} \right) \\
&= \Pi(\alpha)
\end{aligned}$$

$\square$

In some sense, the rotation number gives us some idea of the 'average' rotation of a mapping. The next two theorems show us that homeomorphic mappings of a circle have some properties in common with a pure rotation with the same rotation number.

Theorem 4.7. *Let $f : S^1 \to S^1$ be an orientation preserving homeomorphism. Then $\rho(f) \in \mathbb{Q}$ if and only if f has a periodic point.*

Proof. ($\Rightarrow$) $\rho(f) = \frac{p}{q} \in \mathbb{Q}$. We expect that this homeomorphism will have a q-periodic point, analogous to the rotation with the same rotation number. This is equivalent to saying that f^q has a fixed point.

$$\begin{aligned}
\rho(f^k) &= \Pi\left(\lim_{|n|\to\infty} \frac{(F^k)^n(x) - x}{n}\right) \\
&\qquad \text{where } F \text{ is a standard lift of } f \\
&= \Pi\left(k \lim_{|kn|\to\infty} \frac{(F^{kn}(x) - x}{kn}\right) \\
&= \Pi\left(k\rho(f)\right)
\end{aligned}$$

From this, $\rho(f^q) = \Pi\left(q(\frac{p}{q})\right) = 0$. Thus, it is sufficient to show that any map with rotation number 0 has a fixed point.

Suppose that map g has no fixed point, and take a standard lift G with $G(0) \in [0, 1)$ for convenience. Note that $G(x) - x \in \mathbb{R}\backslash\mathbb{Z}$, since g has no fixed point. Thus $G(x) - x \in (0, 1)$ by the Intermediate Value Theorem. We know that $G - Id$ is continuous for all x, so it achieves its maxima and minima. Thus we can find a $0 < \delta \leq \frac{1}{2}$ such that $\delta \leq G(x) - x \leq 1 - \delta$ for all x. If we rewrite $G^n(x) - x$ as

$$\sum_{i=0}^{n-1} G^{i+1}(x) - G^i(x)$$

we can see that

$$\begin{aligned}
& n\delta \leq G^n(x) - x \leq n(1 - \delta) \\
\Rightarrow \quad & \delta \leq \frac{G^n(x) - x}{n} \leq 1 - \delta
\end{aligned}$$

Thus $\rho(g) \neq 0$, proving the contrapositive of what is required.

($\Leftarrow$) f has a periodic point $x = a$ of period q. Then for a standard lift F, $F^q(a) = a + p$ for some $p \in \mathbb{Z}$. Since rotation number is independent of the point on S^1 used from Lemma 4.5, we will use a to calculate it.

$$\begin{aligned}
\frac{F^{nq}(a)-a}{nq} &= \frac{1}{nq}\sum_{i=0}^{n-1}\left(F^{(i+1)q}(a)-F^{iq}(a)\right) \\
&= \frac{1}{nq}\sum_{i=0}^{n-1}\{(a+(i+1)p)-(a+ip)\} \\
&= \frac{np}{nq} \\
&= \frac{p}{q} \\
\rho(f) &= \Pi\left(\lim_{|nq|\to\infty}\frac{F^{nq}(a)-a}{nq}\right)=\frac{p}{q}
\end{aligned}$$

□

Theorem 4.8. *Let $f : S^1 \to S^1$ be an orientation preserving homeomorphism with rational rotation number. Then all periodic orbits have the same period.*

Proof. Take $\rho(f) = \frac{p}{q}$, with $p, q \in \mathbb{Z}$ relatively prime. Take a periodic point $x = c$, where for any standard lift F for f, $F^r(c) = c + s$, where $r, s \in \mathbb{Z}$. Then

$$\begin{aligned}
\frac{p}{q} &= \Pi\left(\lim_{|n|\to\infty}\frac{F^{nr}(c)-c}{nr}\right) \\
&= \Pi\left(\lim_{|n|\to\infty}\frac{ns}{nr}\right) \\
&= \Pi\left(\frac{s}{r}\right)
\end{aligned}$$

Choose a lift F such that $\frac{s}{r} = \frac{p}{q}$. Then $s = mp$ and $r = mq$ for some $m \in \mathbb{Z}$ and so $F^{mq}(c) = c + mp$.

Now we know that $F^k(c-p) = F^k(c) - p$ from Lemma 3.2, and that $F(a) \geq F(b)$ if $a > b$ by monotonicity. If $F^q(c) - p > c$ then

$$\begin{aligned}
F^{mq}(c)-mp &= F^{mq}(c-p)-(m-1)p \\
&= F^{(m-1)q}\left(F^q(c-p)\right)-(m-1)p \\
&= F^{(m-1)q}\left(F^q(c)-p\right)-(m-1)p \\
&\geq F^{(m-1)q}(c)-(m-1)p \\
&\geq \cdots \\
&\geq F^q(c)-p \\
&> c
\end{aligned}$$

giving us a contradiction. Similarly, it cannot be true that $F^q(c) - p < c$. Thus $F^q(c) - p = c$, and so the orbit beginning at $x = c$ must have period q. □

4.4. **Effect of perturbation on rotation number.** In later sections, we will be concerned persistence of invariant circles under perturbation. This is intimately related to the effect that the perturbation has on the rotation number of the homeomorphism acting on the invariant circle.

Definition (Ordering of homeomorphisms on S^1)**.** Take two homeomorphisms $f_1 : S^1 \to S^1$ and $f_2 : S^1 \to S^1$, and lift them to $F_1 : \mathbb{R} \to \mathbb{R}$ and $F_2 : \mathbb{R} \to \mathbb{R}$ with the same standard lift. Define $f_1 \prec f_2$ if $F_1(x) < F_2(x)$ for all $x \in \mathbb{R}$.

The next two theorems will show us that rational rotation numbers are stable under small perturbations, while irrational rotation numbers are unstable.

Theorem 4.9. *Let $f : S^1 \to S^1$ be an orientation preserving homeomorphism with rational rotation number $\rho(f) = \frac{p}{q}$ and some non-periodic points. Then all sufficiently nearby perturbations $\tilde{f}$ have rotation number $\frac{p}{q}$.*

Proof. For any standard lift F of f, $F^q - Id - p$ is not identically zero since f has some non-periodic points, but has at least one zero since $\rho(f) = \frac{p}{q}$. If $F^q - Id - p$ has at least one zero after perturbation, $\rho(f)$ must remain unchanged (from Theorem 4.8).

If $F^q(x) - x - p > 0$ for any x, then for any sufficiently small perturbation $\tilde{f}$ with standard lift $\tilde{F}$, $\tilde{F}^q(x) - x - p > 0$, and so $\rho(\tilde{f}) \geq \frac{p}{q}$. Similarly, if $F^q(x) - x - p < 0$ for any x, then for any sufficiently small perturbation $\tilde{f}$ with lift $\tilde{F}$, $\tilde{F}^q(x) - x - p < 0$, and so $\rho(\tilde{f}) \leq \frac{p}{q}$.

Thus the only time when a zero can fail to persist under a small perturbation is when $F^q - Id - p$ has a stationary maximum or minimum at that zero. □

Note on the proof:
The scenario where the zero fails to persist is more generally described in dynamical systems by bifurcation theory, specifically relating to saddle-node bifurcations. This property occurs in many other problems in dynamical systems, and will invariably signal a change in the nature of the dynamics being studied.

Theorem 4.10. *If $f_1 \prec f_2$ and $\rho(f_1) \in \mathbb{R} \backslash \mathbb{Q}$ then $\rho(f_1) < \rho(f_2)$.*

Proof. Use the same standard lift λ to take lifts F_1 and F_2 of f_1 and f_2. By definition, $F_2(x) - F_1(x) > 0$ for all x. By continuity and periodicity, $F_2(x) - F_1(x) > \delta$ for some $\delta > 0$ and all x. Take $\frac{p}{q} \in \mathbb{Q}$ such that $\frac{p}{q} - \frac{\delta}{q} < \rho(f_1) < \frac{p}{q}$.

Assume for the moment that there exists $x = a$ such that $F_1^q(a) > p - \delta$. Then

$$\begin{aligned} F_2^q(a) &= F_2(F_2^{q-1}(a)) \\ &> F_1(F_2^{q-1}(a)) + \delta \\ &> F_1(F_1^{q-1}(a)) + \delta \\ &= F_1^q(a) + \delta \\ &> a + p. \end{aligned}$$

By the Intermediate Value Theorem, since $F_2^q(a) - a > p$, then either $F_2^q(x) - x > p$ for all x or $F_2^q(b) - b = p$ for some $b \in \mathbb{R}$. Thus $\rho(f_1) < \frac{p}{q} \leq \rho(f_2)$.

Now we want to prove the assumption above that there exists $x = a$ such that $F_1^q(a) > p - \delta$. Assume that there does not exist such a value a. Then $F_1^q(x) \leq p - \delta$ for all x.

$$\begin{aligned} \rho(f_1) &= \Pi\left(\lim_{|nq|\to\infty} \frac{F_1^{nq}(x) - x}{nq}\right) \\ &= \Pi\left(\lim_{|n|\to\infty} \frac{F_1^{nq}(x) - x}{nq}\right) \\ &\leq \Pi\left(\lim_{|n|\to\infty} \frac{n(p-\delta) - x}{nq}\right) \\ &= \lim_{|n|\to\infty} \frac{n(p-\delta)}{nq} \\ &= \frac{p}{q} - \frac{\delta}{q} \end{aligned}$$

This contradicts the assumption that $\frac{p}{q} - \frac{\delta}{q} < \rho(F_1) < \frac{p}{q}$. Thus, such a value a must exist. $\square$

4.5. **Behaviour of limit sets.** We will study asymptotic behaviour of orientation preserving homeomorphisms on S^1 to motivate the presence of invariant circles, periodic orbits and Cantor sets that we observe later.

Definition (Homoclinic orbit). An orbit O is homoclinic under a mapping f to an invariant set U if O is disjoint from U and $L_f^\alpha(x) = L_f^\omega(x) = U$ for all $x \in O$.

Definition (Heteroclinic orbit). An orbit O is heteroclinic under a mapping f to two disjoint invariant sets U_1, U_2 if O is disjoint from both of them and $L_f^\alpha(x) = U_1$, $L_f^\omega(x) = U_2$ for all $x \in O$.

Theorem 4.11. *Let $f : S^1 \to S^1$ be an orientation preserving homeomorphism with rotation number $\rho(f) = \frac{p}{q} \in \mathbb{Q}$. Then*

(i) If f has exactly one periodic orbit then every other point is heteroclinic under f^q to two points on the periodic orbit. These points are different if $q > 1$.

(ii) If f has more than one periodic orbit then each nonperiodic point is heteroclinic under f^q to two points on different periodic orbits.

Proof. Take a standard lift F of f, and consider $G = F^q - p$ restricted to an interval $[z, z+1]$. Lemma 4.1 tells us that $L_G^\alpha(x)$ and $L_G^\omega(x)$ are adjacent fixed points of G for all $x \in \mathbb{R}$.

If f has exactly one periodic orbit, then adjacent fixed points of G will project to adjacent q-periodic points of f if $q > 1$, or to the same fixed point if $q = 1$.

Let f have more than one periodic orbit, and let adjacent fixed points of G be a, b. If a and b project to the same periodic orbit, then $\Pi(a) = \bar{x}$ and $\Pi(b) = f^k(\bar{x})$, and since $\bigcup_{n=0}^{q-1} f^{nk}\Pi(a,b)$ covers $S^1 \backslash \{f^n(\bar{x}); n = 0, \ldots, q-1\}$, there must be no periodic points apart from $\{f^n(\bar{x}); n = 0, \ldots, q-1\}$. Thus a and b must project to different periodic orbits. □

Figure 2 shows an example of each type predicted by the theorem. These correspond to projections of the mappings:

$$\begin{aligned} (a) f(x) &= x + \frac{1}{2} + \epsilon \sin(2\pi x) \\ (b) g(x) &= x + \frac{1}{2} + \epsilon \sin(4\pi x) \end{aligned}$$

for ϵ small enough to keep the mappings as homeomorphisms.

It is easy to solve for the limit sets of (a) exactly:
Fixed points (in $[0,1)$) are $\{0, \frac{1}{2}\}$. In the interval $(0, \frac{1}{2})$, $f(x) > x$, so the α-limit set is 0, while the ω-limit set is $\frac{1}{2}$. In the interval $(\frac{1}{2}, 1)$, $f(x) < x$, so the α-limit set is $\frac{1}{2}$, while the ω-limit set is $\Pi(1) = 0$.

We can do a similar analysis for $g^2(x)$ to find the limit sets of (b) exactly - in [0,1), the α-limit set is $\{0, \frac{1}{2}\}$ and the ω-limit set is $\{\frac{1}{4}, \frac{3}{4}\}$.

The following lemma will be useful in studying the behaviour of orbits of homeomorphisms with irrational rotation number.

Lemma 4.12. *Let $f : S^1 \to S^1$ be an orientation preserving homeomorphism, with rotation number $\rho(f) \in \mathbb{R}\backslash\mathbb{Q}$ and $U \subset S^1$ a closed*

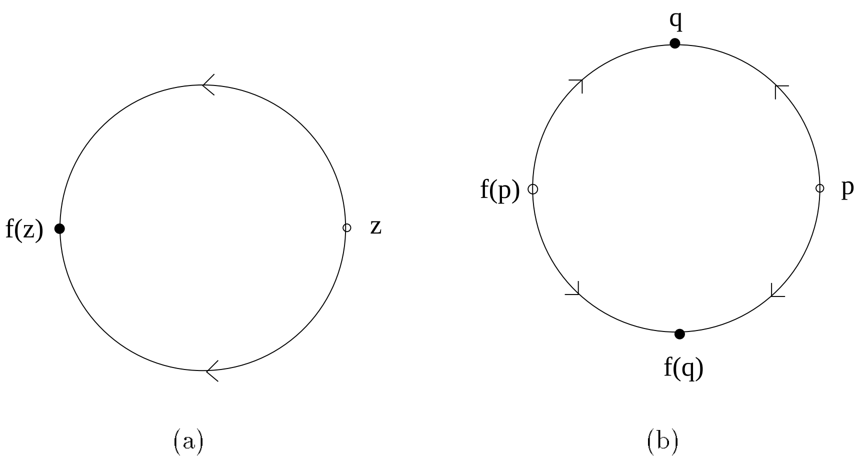

FIGURE 2. Examples of limit sets

interval with endpoints $f^m(x)$ and $f^k(x)$, for some $x \in S^1$, $m, k \in \mathbb{N}$, $m \neq k$. Then at least one positive iterate $f^n(x)$, $n \in \mathbb{N}$ will be in U.

Proof. We want to show that $f^n(x) \in U$, or equivalently that $x \in (f^{-1})^n(U)$ for some $x \in S^1$, $n \in \mathbb{N}$. Thus, it is sufficient to show that $S^1 = \bigcup_{r \in \mathbb{N}} f^{-r}(U)$.

Let $U_k = f^{-k(n-m)}(U)$. U_k and U_{k+1} share a common endpoint since

$$f^{-k(n-m)}(f^m(x)) = f^{-(k+1)(n-m)}(f^n(x)),$$

and thus the intervals U_k are all contiguous. Thus if $S^1 \neq \bigcup_{r \in \mathbb{N}} U_k$, then the sequence of endpoints converges to some $z \in S^1$. However, that would imply that z is periodic, since

$$\begin{aligned}
z &= \lim_{k \to \infty} f^{-k(n-m)} f^m(x) \\
&= \lim_{k \to \infty} f^{-(k+1)(n-m)} f^m(x) \\
&= \lim_{k \to \infty} f^{(n-m)} f^{-k(n-m)} f^m(x) \\
&= f^{(n-m)} \lim_{k \to \infty} f^{-k(n-m)} f^m(x) \\
&= f^{(n-m)}(z),
\end{aligned}$$

which contradicts the assumption $\rho(f) \in \mathbb{R} \backslash \mathbb{Q}$. □

Definition (Cantor set). A set is a Cantor set if it is perfect (closed with no isolated points) and nowhere dense.

Theorem 4.13. *Let $f : S^1 \to S^1$ be an orientation preserving homeomorphism with rotation number $\rho(f) \in \mathbb{R}\backslash\mathbb{Q}$. Then $L^\omega_f(x)$ is independent of x and is either dense in S^1 or a Cantor set.*

Proof. (i) Independence of x: We want to show that $L^\omega_f(x) = L^\omega_f(y)$ for all $x, y \in S^1$. Let $a \in L^\omega_f(x)$, and choose a sequence l_n such that $f^{l_n}(x) \to z$. If $y \in S^1$ then by Lemma 4.12 there exist $k_n \in \mathbb{N}$ such that $f^{k_n}(y) \in U_n$ where U_n has endpoints $f^{l_n}(x)$ and $f^{l_{n+1}}(x)$. Thus $\lim_{n\to\infty} f^{k_n}(y) = z$, and so $z \in L^\omega_f(y)$.

Thus $L^\omega_f(x) \subseteq L^\omega_f(y)$ for all $x, y \in S^1$, and by symmetry $L^\omega_f(x) = L^\omega_f(y)$ for all $x, y \in S^1$.

(ii) Dense in S^1 or a Cantor set: Consider a nonempty closed f-invariant set $A \subseteq S^1$. For every $x \in A$, $\{f^k(x); k \in \mathbb{Z}\} \subseteq A$ since A is invariant and $L^\omega_f(x) \subseteq A$ since A is closed. Thus $\emptyset$ and $L^\omega_f(x)$ are the only closed invariant subsets of $L^\omega_f(x)$. Since the boundary of a closed invariant set is a closed invariant set by Lemma 3.4, then $\delta L^\omega_f(x) = \emptyset$ or $\delta L^\omega_f(x) = L^\omega_f(x)$.

If $\delta L^\omega_f(x) = \emptyset$ then $L^\omega_f(x) = S^1$.

If $\delta L^\omega_f(x) = L^\omega_f(x)$ then $\mathrm{int}(L^\omega_f(x)) = \emptyset$ which implies that $L^\omega_f(x)$ is nowhere dense. Now take any element $a \in L^\omega_f(x)$. Then there is a sequence k_n such that $\lim_{n\to\infty} f^{k_n} = x$, and since $\rho(f) \in \mathbb{R}\backslash\mathbb{Q}$, $f^{k_n}(x) \neq x$ for all n. Thus a cannot be an isolated point, and so $L^\omega_f(x)$ has no isolated points. □

5. Perturbed integrable dynamical systems in $\mathbb{R}^2$

In this section, we will investigate behaviour of integrable and almost integrable dynamical systems. Along the way, we will see that the dynamics of the system can sometimes be reduced to a set of circle homeomorphisms, which we already know about. We will introduce the concepts of symplectic structure and Hamiltonian system, and look at special properties associated with these. The twist homeomorphism will be introduced to allow discussion of the Poincaré-Birkhoff theorem and KAM theorem, which help us to pinpoint features that must be present in almost integrable systems.

5.1. **Integrable systems, symplectic structure and Hamiltonian systems.** Integrable systems are probably the most obvious introduction to two-dimensional systems, since it is possible to specify the behaviour of these systems analytically. This occurs due to the lack of any 'chaotic' regions where orbits mix and tangle, which lead to a sensitive dependence on initial conditions.

Definition (Symplectic vector space)**.** A vector space V is symplectic if it has a bilinear form ω which is antisymmetric and nondegenerate, i.e. for $u, v \in V$,

$$\omega(u,v) = -\omega(v,u)$$

and for every $u \neq 0$, there is a v satisfying $\omega(u,v) \neq 0$. We call the symplectic space (V, ω).

Definition (Symplectic diffeomorphism)**.** Let $\varphi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ in the symplectic space $(\mathbb{R}^{2n}, \omega)$. φ is symplectic if

$$\omega(D\varphi u, D\varphi v) = \omega(u,v).$$

We can find a bilinear form defined as $\omega(u,v) = \langle Ju, v\rangle$ for $u, v \in \mathbb{R}^{2n}$ where $\langle u, v\rangle$ is the inner product of u and v, $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$ and I is the $n \times n$ identity matrix. Then

$$\langle JD\varphi u, D\varphi v\rangle = \langle Ju, v\rangle.$$

Rewriting this, we get

$$\langle D\varphi^T JD\varphi u, v\rangle = \langle Ju, v\rangle$$

Since this is true for all u, v, this is equivalent to

$$[D\varphi]^T JD\varphi = J$$

Definition (Hamiltonian system)**.** Take a smooth function $H : \mathbb{R}^{2n} \to \mathbb{R}$ in symplectic space $(\mathbb{R}^{2n}, \omega)$. We define the Hamiltonian vector field X_H as

$$-dH(x)v = \omega(X_H(x), v).$$

Then the Hamiltonian flow is defined as

$$\frac{d\varphi^t(q,p)}{dt} = X_H.$$

Using the bilinear form defined previously,

$$-\langle \nabla H(x), v\rangle = \langle JX_H(x), v\rangle.$$

Thus

$$X_H(x) = J\nabla H(x).$$

Let $q_1, \ldots, q_n; p_1, \ldots, p_n$ (or equivalently (q,p)) be a coordinate system in $\mathbb{R}^{2n}$ and $H(q,p)$ a smooth function. Then, we can express the equations of the Hamiltonian system (known as the canonical equations) as

$$\begin{aligned} \frac{dq}{dt} &= \frac{\partial H(q,p)}{\partial p} \\ \frac{dp}{dt} &= -\frac{\partial H(q,p)}{\partial q}. \end{aligned}$$

The canonical equations define the Hamiltonian flow φ^t as

$$X_H = \frac{d\varphi^t(q,p)}{dt} = \begin{bmatrix} \frac{\partial H(\varphi^t(q,p))}{\partial p} \\ -\frac{\partial H(\varphi^t(q,p))}{\partial q} \end{bmatrix} = J\nabla_{(q,p)} H(\varphi^t)$$

where $\varphi^0(q,p) = (q,p)$.

Definition (Integral)**.** A map $M : U \to U$ has an integral if there is a function $I : U \to \mathbb{R}$ such that $I(M(x)) = I(x)$ for all $x \in U$ where $U \subseteq \mathbb{R}^n$.

Definition (Integrable)**.** A Hamiltonian system with n degrees of freedom is integrable if it has n integrals $I_1, I_2, \ldots, I_n$, where the integrals satisfy

$$\frac{\partial I_i}{\partial q}\frac{\partial I_j}{\partial p} - \frac{\partial I_i}{\partial p}\frac{\partial I_j}{\partial q} \equiv 0$$

for all $i, j \in \{1, 2, \ldots, n\}$.

Definition (Volume preserving)**.** A diffeomorphism $M : U \to \mathbb{R}^n$ is volume preserving if

$$|\det\ DM| = 1$$

for all $x \in U$ where $U \subseteq \mathbb{R}^n$.

Note that when we work in $\mathbb{R}^2$, we refer to this property as area preserving instead of volume preserving.

Lemma 5.1. *Symplectic diffeomorphisms are volume preserving.*

Proof. Take a symplectic diffeomorphism $f : U \to \mathbb{R}^{2n}$. By definition,

$$[Df]^T J Df = J$$

where $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$ Since $\det\ J = 1$ and $\det\ Df = \det\ [Df]^T$, then $|\det\ Df| = 1$. Thus f is volume preserving. □

Lemma 5.2. *In two dimensions, a diffeomorphism* $M(x,y) = (f(x,y), g(x,y))$ *is symplectic if and only if it is area preserving and orientation preserving.*

Proof.

$$\begin{aligned} DfJ[Df]^T &= \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial g}{\partial x} \\ \frac{\partial f}{\partial y} & \frac{\partial g}{\partial y} \end{bmatrix} \\ &= \begin{bmatrix} 0 & \frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x} \\ \frac{\partial f}{\partial y}\frac{\partial g}{\partial x} - \frac{\partial f}{\partial x}\frac{\partial g}{\partial y} & 0 \end{bmatrix} \end{aligned}$$

This means that $\frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x} = 1$ if and only if f is symplectic, and so f must be both area preserving and orientation preserving. □

Theorem 5.3 (Liouville Theorem)**.** *Hamiltonian flows are symplectic.*

Proof. Let the Hamiltonian flow be denoted φ^t. From the definition, we have

$$\frac{d\varphi^t(q,p)}{dt} = J\nabla_{(q,p)}H$$

where $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$. Since H is smooth,

$$\begin{aligned} \frac{d}{dt}D\varphi^t(q,p) &= D\frac{d\varphi^t(q,p)}{dt} \\ &= DJ\nabla_{(q,p)}H \\ &= JH''(\varphi^t(q,p))D\varphi^t(q,p). \end{aligned}$$

Note that $J^T = -J$, $JJ = -I$, and $H'' = (H'')^T$.

$$\begin{aligned} \frac{d}{dt}((D\varphi^t)^T JD\varphi^t) &= \frac{d}{dt}(D\varphi^t)^T JD\varphi^t + (D\varphi^t)^T J\frac{d}{dt}(D\varphi^t) \\ &= (D\varphi^t)^T(H'')^T J^T JD\varphi^t + (D\varphi^t)^T JJH''D\varphi^t \\ &= (D\varphi^t)^T H''D\varphi^t - (D\varphi^t)^T H''D\varphi^t \\ &= 0. \end{aligned}$$

Thus $(D\varphi^t)^T JD\varphi^t$ is constant for all t. However, we know that $\varphi^0(q,p) = (q,p)$, so $D\varphi^0 = I$. Thus

$$(D\varphi^t)^T JD\varphi^t = (D\varphi^0)^T JD\varphi^0 = J.$$

This means that the Hamiltonian flow φ^t is symplectic. □

Corollary 5.4. *Hamiltonian flows are volume preserving.*

Lemma 5.5. *Hamiltonian flows in $\mathbb{R}^2$ are integrable.*

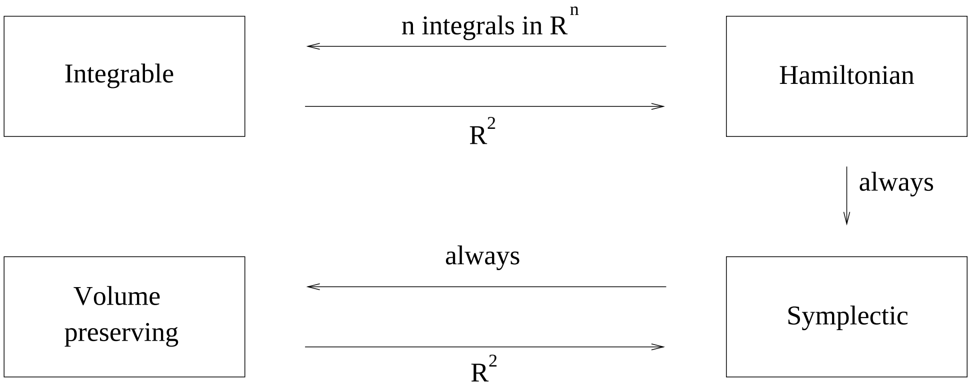

FIGURE 3. Relationships between important concepts

Proof.

$$\begin{aligned}\frac{dH}{dt} &= \frac{\partial H}{\partial p}\frac{dp}{dt} + \frac{\partial H}{\partial q}\frac{dq}{dt} \\ &= \frac{dq}{dt}\frac{dp}{dt} - \frac{dp}{dt}\frac{dq}{dt} = 0\end{aligned}$$

□

Figure 3 shows the relationships between the concepts of integrable, Hamiltonian, symplectic and volume preserving that we have derived.

In two dimensions, any compact level set $I(x, y) = c$ can be transformed into 'action-angle' variables (θ, r) such that

$$\begin{aligned}\theta &\mapsto \theta + \omega(r) \\ r &\mapsto r\end{aligned}$$

where θ is taken mod 2π. This result follows directly from the following theorem.

Theorem 5.6 (Liouville-Arnol'd Theorem). *Suppose that we have two functions $f(q, p)$ and $g(q, p)$ defined on $\mathbb{R}^2$ such that*

$$\frac{\partial f}{\partial q}\frac{\partial g}{\partial p} - \frac{\partial f}{\partial p}\frac{\partial g}{\partial q} \equiv 0.$$

Consider a level set $L = \{x; f(x) = k_1, g(x) = k_2\}$. Assume that the differentials Df and Dg are pointwise linearly independent on L. Then

(i) L is a smooth manifold, invariant under any Hamiltonian flow with Hamiltonian function $H = f$ or $H = g$.

(ii) If L is compact and connected, then it is diffeomorphic to S^1.

(iii) The Hamiltonian flow determines a conditionally periodic motion on L. This means that L is mapped to a rotation under a diffeomorphism to S^1.

(iv) The canonical equations with Hamiltonian function H can be integrated by quadrature.

Proof. See Arnol'd [3, pp272-]. □

In effect, this means that closed curves in the plane can be transformed into circles with rotation number $\Pi\left(\frac{\omega(r)}{2\pi}\right)$.

5.2. **The twist map.** We will use the twist map in the statement of the Aubry-Mather theorem, Poincaré-Birkhoff theorem and KAM theorem.

Definition (Twist homeomorphism)**.** Let $A = [a, b] \times S^1$ be an annulus in $\mathbb{R}^2$. A homeomorphism $\tilde{\mathbf{T}} : A \to A$ defined as

$$(x, y) \mapsto (f(x, y), g(x, y))$$

is a twist homeomorphism if it preserves orientation, maintains the boundary components of A, and $f(x, y)$ is a strictly monotone function of y for every $x \in \mathbb{R}$.

In the following sections, we will use a special example of the twist homeomorphism, which we will call $\mathbf{T}$, defined as

$$\mathbf{T} : A \to A, \quad (\theta, r) \mapsto (\theta + \alpha(r), r)$$

where $\alpha'(r) \neq 0$.

The twist map $\mathbf{T}$ leaves each circle with radius $r \in [a, b]$ invariant. On each of these circles, the twist map simply causes a pure rotation by $\alpha(r)$. Thus the dynamics in the annulus are totally described by circle mappings. Figure 4 shows an example when $\alpha'(r) > 0$.

We will define the perturbed area preserving twist map $\mathbf{T}_\epsilon$ as

$$\mathbf{T}_\epsilon : A \to \mathbb{R}^2, \quad (\theta, r) \mapsto (\theta + \alpha(r) + f(\theta, r, \epsilon), r + g(\theta, r, \epsilon))$$

where $a \leq r \leq b$, $\alpha'(r) \neq 0$ and f, g are 2π-periodic in θ with $f(\theta, r, 0) = g(\theta, r, 0) \equiv 0$. $\mathbf{T}_\epsilon$ is defined on an annulus $a \leq r \leq b$, $a < b$ and is area preserving for all values of ϵ. As a consequence, $\mathbf{T}_\epsilon$ must map a closed curve onto a closed curve. Note that $\mathbf{T}_\epsilon$ does not necessarily maintain the boundary components of the annulus A.

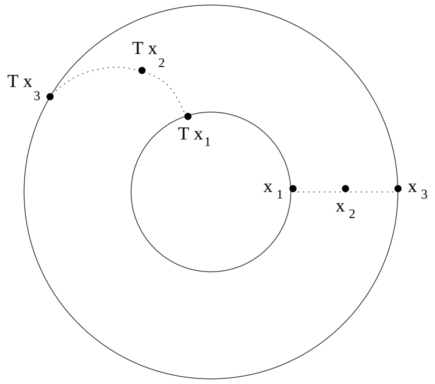

FIGURE 4. A twist homeomorphism on an annulus

5.3. **Aubry-Mather theorem.** The Aubry-Mather theorem tells us the limit sets of twist homeomorphisms.

Theorem 5.7 (Aubry-Mather). *Let* $\mathbf{T}$ *be an area preserving twist homeomorphism of the annulus with rotation numbers in the interval* $[a, b]$. *Then, for every* $\rho \in [a, b]$, $\mathbf{T}$ *has a Mather set with rotation number* ρ, *where Mather sets can be as follows:*

(i) If ρ *is rational, then the Mather set is a Birkhoff periodic orbit (see Arrowsmith and Place* [4, p336]*).*

(ii) If ρ *is irrational, then the Mather set can either be* S^1 *or a Cantor set.*

This theorem puts together the results from Theorems 4.11 and 4.13.

5.4. **Poincaré-Birkhoff theorem.** This theorem tells us that the perturbed twist map will maintain a number of periodic orbits for each rational rotation present in the unperturbed twist map.

This version of the theorem has been adapted from Arrowsmith and Place [4, p309].

Theorem 5.8 (Poincaré-Birkhoff). *Given any rational number* $\frac{p}{q}$, *between* $\frac{\alpha(a)}{2\pi}$ *and* $\frac{\alpha(b)}{2\pi}$, *there are at least* $2q$ *fixed points* (θ_0, r_0) *of* $\mathbf{T}_\epsilon^q : (\theta, r) \mapsto (\theta_q, r_q)$ *satisfying*

$$(\theta_q, r_q) = (\theta + 2\pi p, r),$$

where $\mathbf{T}_\epsilon$ *is the perturbed twist map of an annulus, provided that* ϵ *is sufficiently small.*

Proof. Take $\mathbf{T}$ as the twist map with $\frac{\alpha(r)}{2\pi} = \frac{p}{q}$, and perturb it to the map $\mathbf{T}_\epsilon$ given. Then for $\mathbf{T}_\epsilon^q$ at some $\bar{r}$ near r_0 and a fixed value $\bar{\theta}$

$$(\theta_q, r_q) = \Big(\bar{\theta} + q\alpha(\bar{r}) + \tilde{f}(\bar{\theta}, \bar{r}, \epsilon), \bar{r} + \tilde{g}(\bar{\theta}, \bar{r}, \epsilon)\Big) \quad \text{(see (a) below)}$$

We want

$$(\theta_q, r_q) = \left(\bar{\theta} + 2\pi p, \bar{r}\right)$$

or equivalently

$$\chi_{\bar{\theta}}(\epsilon, \bar{r}) = q\alpha(\bar{r}) - 2\pi p + \tilde{f}(\bar{\theta}, \bar{r}, \epsilon) = 0$$

We can use the Implicit Function Theorem for $\chi_{\bar{\theta}}$, since

$$\chi_{\bar{\theta}}(0, r_0) = 0 \quad \text{and} \quad \frac{\partial \chi_{\bar{\theta}}}{\partial r}(0, r_0) = q\alpha'(r_0) \neq 0.$$

Thus there must exist an open interval U about 0 and V about r_0 and a unique function $\phi_{\bar{\theta}}(\epsilon)$, $\phi_{\bar{\theta}} : U \to V$ such that

$$\phi_{\bar{\theta}}(0) = r_0 \quad \text{and} \quad \chi_{\bar{\theta}}(\epsilon, \phi_{\bar{\theta}}(\epsilon)) = 0$$

for all $\epsilon \in U$. Thus, on the curve

$$\Gamma : \bar{r}(\theta, \epsilon) = \phi_{\bar{\theta}}(\epsilon),$$

$\chi \equiv 0$, or equivalently

$$\theta_q = \bar{\theta} + 2\pi p.$$

This means that each point on the curve Γ is mapped radially under $\mathbf{T}_\epsilon^q$. The curve Γ is continuous by the restrictions on f and g in the definition of $\mathbf{T}_\epsilon$ (and thus on $\tilde{f}$ and $\tilde{g}$). Because the perturbed twist map is area preserving, then Γ and $\mathbf{T}_\epsilon^q \Gamma$ must intersect at least twice. This means that there must be at least two fixed points of $\mathbf{T}_\epsilon^q$ (see (b) below). Moreover, if $\hat{x}$ is one of these fixed points, then so are $\mathbf{T}_\epsilon^k(\hat{x})$ for $k = 1, \ldots, q-1$ since

$$\mathbf{T}_\epsilon^q \mathbf{T}_\epsilon^k(\hat{x}) = \mathbf{T}_\epsilon^k \mathbf{T}_\epsilon^q(\hat{x}) = \mathbf{T}_\epsilon^k(\hat{x}).$$

Thus $\mathbf{T}_\epsilon^q$ must have at least $2q$ fixed points.

(a) Take any point (θ_k, r_k) and map it to (θ_{k+1}, r_{k+1}) via $\mathbf{T}_\epsilon$. Thus

$$\begin{aligned} \theta_1 &= \theta_0 + \alpha(r_0) + f(\theta_0, r_0, \epsilon) \\ r_1 &= r_0 + g(\theta_0, r_0, \epsilon). \end{aligned}$$

Applying $\mathbf{T}_\epsilon$ again, we get

$$\begin{aligned}
\theta_2 &= \theta_1 + \alpha(r_1) + f(\theta_1, r_1, \epsilon) \\
&= \theta_0 + \alpha(r_0) + \alpha(r_1) + f(\theta_0, r_0, \epsilon) + f(\theta_1, r_1, \epsilon) \\
&= \theta_0 + 2\alpha(r_0) + f_2(\theta_0, r_0, \epsilon) \\
&\quad \text{where } f_2(\theta_0, r_0, \epsilon) = \alpha(r_1) - \alpha(r_0) + f(\theta_0, r_0, \epsilon) + f(\theta_1, r_1, \epsilon) \\
r_2 &= r_1 + g(\theta_1, r_1, \epsilon) \\
&= r_0 + g(\theta_0, r_0, \epsilon) + g(\theta_1, r_1, \epsilon) \\
&= r_0 + g_2(\theta_0, r_0, \epsilon) \\
&\quad \text{where } g_2(\theta_0, r_0, \epsilon) = g(\theta_0, r_0, \epsilon) + g(\theta_1, r_1, \epsilon).
\end{aligned}$$

If we continue with this process and define

$$\begin{aligned}
f_1(\theta_0, r_0, \epsilon) &= f(\theta_0, r_0, \epsilon) \\
g_1(\theta_0, r_0, \epsilon) &= g(\theta_0, r_0, \epsilon)
\end{aligned}$$

we end up with

$$\begin{aligned}
\theta_k &= \theta_0 + k\alpha(r_0) + f_k(\theta_0, r_0, \epsilon) \\
r_k &= r_0 + g_k(\theta_0, r_0, \epsilon)
\end{aligned}$$

where

$$\begin{aligned}
f_k(\theta_0, r_0, \epsilon) &= \alpha(r_{k-1}) - \alpha(r_0) + f_{k-1}(\theta_0, r_0, \epsilon) + f(\theta_{k-1}, r_{k-1}, \epsilon) \\
g_k(\theta_0, r_0, \epsilon) &= g_{k-1}(\theta_0, r_0, \epsilon) + g(\theta_{k-1}, r_{k-1}, \epsilon).
\end{aligned}$$

(b) Take any two closed curves with radial distance functions $f(\theta)$ and $g(\theta)$ as measured from some point enclosed by both curves, where $\theta \in [0, 2\pi)$. Let the curves enclose the same area,

$$\text{i.e.} \quad \int_0^{2\pi} f(\theta)d\theta = \int_0^{2\pi} g(\theta)d\theta.$$

Take lifts $F : \mathbb{R} \to \mathbb{R}$, $G : \mathbb{R} \to \mathbb{R}$ for $f : S^1 \to \mathbb{R}$, $g : S^1 \to \mathbb{R}$ such that $F(\theta + 2\pi z) = F(\theta)$ and $G(\theta + 2\pi z) = G(\theta)$ for all $z \in \mathbb{Z}$. Define $H(\theta) = F(\theta) - G(\theta)$. We want to show that $H(\theta) = 0$ for two distinct $\theta \in [0, 2\pi)$ by applying the Intermediate Value Theorem, since this is equivalent to saying that the two curves intersect at two points.

Take the case where $f(\theta) > g(\theta)$ for all θ. Obviously

$$\int_0^{2\pi} f(\theta)d\theta > \int_0^{2\pi} g(\theta)d\theta,$$

so this cannot occur. Similarly, it cannot be true that $f(\theta) < g(\theta)$ for all θ.

There must exist the situation $F(a) < G(a)$ and $F(b) > G(b)$ for some $a, b \in [0, 2\pi)$. Arbitrarily choose $a < b$ (if $a > b$, then just swap F and G). Applying the Intermediate Value Theorem to $H(a)$ and $H(b)$, there must be some $c \in (a, b)$ such that $H(c) = 0$. However, it is also true that $F(a+1) < G(a+1)$, so there must also be some $d \in (b, a+1)$ such that $H(d) = 0$. Since c, d are distinct and are both in the interval $(a, a+1)$, then there must be two distinct $\theta \in [0, 2\pi)$ such that $H(\theta) = 0$. □

The Poincaré-Birkhoff theorem predicts fixed points that exist after mapping q times with the perturbed twist map. Arbitrarily taking $\alpha'(r) > 0$, points inside Γ have $\alpha(r) < 0$, while points outside Γ have $\alpha(r) > 0$. If we assume that curves Γ and $\mathbf{T}_\epsilon^q\Gamma$ pass through each other at their points of intersection (the fixed points of $\mathbf{T}_\epsilon^q$), then these points are alternately elliptic (locally stable) and hyperbolic (locally unstable). It happens that the elliptic points form a stable q-cycle, while the hyperbolic points form an unstable q-cycle. This is analogous to the situation with circle maps which we would expect, since for small ϵ, the curves Γ and $\mathbf{T}_\epsilon^q\Gamma$ look similar to circles with periodic points, with behaviour described in Theorem 4.2.

5.5. **KAM Theorem.** The KAM Theorem derives its name from the initials of the three eminent mathematicians Kolmogorov, Arnol'd and Moser who contributed most of the initial content to the theorem. It predicts the persistence of invariant curves under small perturbations.

5.5.1. *Statement of the theorem.* Here are two forms of the theorem, stated for use in $\mathbb{R}^2$:

Theorem 5.9 (Kolmogorov, Arnol'd)**.** *Suppose that the perturbed Hamiltonian function $H(r, \theta, \epsilon)$ is analytic on the annulus $A : [a, b] \times S^1$, where $a, b \in \mathbb{R}$ and H has period 2π in the circular coordinate θ. Let $H(r, \theta, \epsilon) = H_0(r) + H_1(r, \theta, \epsilon)$ where $H_0(r) = H(r, \theta, 0)$ and*

$$\frac{\partial^2 H_0}{\partial r^2} \neq 0$$

in A. Then if $|H_1|$ is sufficiently small, there exist invariant curves

$$\theta = \xi + F(\xi, \epsilon), \;\; r = c + G(\xi, \epsilon)$$

where F, G are analytic functions of period 2π in ξ. The motion of the invariant curves is given by

$$\dot{\xi} = \omega$$

where

$$\omega = \left.\frac{\partial H_0}{\partial r}\right|_{r=c}$$

for some c in some set X dense in $[a, b]$.

Proof. See Arnol'd [1]. □

Theorem 5.10 (Moser Twist Theorem). *Take a mapping $\mathbf{T}_\epsilon$ such that for any curve $\Gamma : r = \phi(\theta) = \phi(\theta + 2\pi)$, Γ always intersects $\mathbf{T}_\epsilon\Gamma$, and $\alpha, f, g \in C^k$, for some k, in some annulus $A : a \le r \le b$.*

Given any rotation ω such that $\alpha(a) \le \omega \le \alpha(b)$ and ω is incommensurable with 2π, there exists a continuously differentiable closed invariant curve

$$\theta = \xi + F(\xi, \epsilon), \; r = c + G(\xi, \epsilon)$$

in A provided that ϵ is sufficiently small, where $a < c < b$, and F, G are small perturbations of period 2π. This curve satisfies

$$\left|\frac{w}{2\pi} - \frac{p}{q}\right| \ge \frac{\gamma(\epsilon)}{|q|^\mu}$$

for all $p, q \in \mathbb{Z}$. In addition, the induced mapping of the curve is given by $\xi \to \xi + \omega$.

Proof. See Siegel and Moser [15, pp225-243] for the full details of the proof. A brief outline of the proof will be provided below.

Assume that we have an annulus A as above, and choose an ω satisfying the condition

$$\left|\frac{\omega}{2\pi} - \frac{p}{q}\right| \ge \frac{c}{|q|^\mu}$$

for all $p, q \in \mathbb{Z}$ and some c, μ. Also assume that all curves Γ and $\mathbf{T}_\epsilon\Gamma$ intersect.

We will change the mapping

$$\mathbf{T}_\epsilon : (\theta, r) \mapsto (\theta + \alpha(r) + f(\theta, r, \epsilon), r + g(\theta, r, \epsilon))$$

to the linear mapping

$$\mathbf{M} : (x, y) \mapsto (x + \gamma y + \tilde{f}(x, y, \epsilon), y + \tilde{g}(x, y, \epsilon))$$

where $x = \theta$, $y = \frac{\alpha(r)}{\gamma}$, $\gamma = |\alpha(b) - \alpha(a)|$.

Now we will use a result (without proof) that will reduce the size of the perturbation after a coordinate transformation:

Under the assumptions for $\mathbf{M}$ above, and taking $\gamma = 1$, there exists a coordinate transformation

$$\mathbf{U} : (x, y) \mapsto (\xi + u(\xi, \eta, \epsilon), \eta + v(\xi, \eta, \epsilon))$$

with u, v real analytic in a domain D, such that the transformed map

$$\mathbf{U}^{-1}\mathbf{M}\mathbf{U} : (\xi, \eta) \mapsto (\xi + \eta + \phi(\xi, \eta, \epsilon), \eta + \psi(\xi, \eta, \epsilon))$$

has ϕ, ψ real analytic functions well defined in $D_1 \subset D$.

Apply this result successively to form a sequence of mappings

$$\mathbf{M}_{n+1} = \mathbf{U}_n^{-1}\mathbf{M}_n\mathbf{U}_n$$

where $\mathbf{M}_0 = \mathbf{M}$. If we define $\mathbf{V}_n = \mathbf{U}_0\mathbf{U}_1 \ldots \mathbf{U}_n$, then we can write

$$\mathbf{V}_n : (x, y) \mapsto (\xi + p_n(\xi, \eta, \epsilon), \eta + q_n(\xi, \eta, \epsilon))$$

where p_n, q_n are analytic functions in D_n, where $D_n \subset D_{n-1} \subset \ldots \subset D_1 \subset D$. This means that

$$\mathbf{M}_{n+1} = \mathbf{V}_n^{-1}\mathbf{M}_0\mathbf{V}_n.$$

As $n \to \infty$, $|\eta - \omega| \to 0$. This means that the dynamics on the curve can be transformed into the twist mapping

$$(\xi, \eta) \mapsto (\xi + \eta, \eta)$$

where $\eta = \omega$.

Denote the limit of p_n by $u(\xi)$ and the limit of q_n by $v(\xi) - \omega$. We need to check that $u(\xi)$, $v(\xi)$ describe the invariant curve required. We know that $\mathbf{V}_n\mathbf{M}_{n+1} = \mathbf{M}_0\mathbf{V}_n$. Consider this in the limit $n \to \infty$. We then have the relations

$$\begin{aligned} \xi + \omega + u(\xi + \omega) &= \xi + u(\xi) + v(\xi) + \tilde{f}(\xi + u, v, \epsilon) \\ v(\xi + \omega) &= v(\xi) + \tilde{g}(\xi + u, v, \epsilon) \end{aligned}$$

which say that the curve

$$x = \xi + u(\xi), \ y = v(\xi)$$

is invariant under $\mathbf{M}_0$ with an induced mapping given by $\xi \mapsto \xi + \omega$. □

Theorem 5.9 has been adapted from Kolmogorov [10] and Arnol'd [1]. Theorem 5.10 has been adapted from Moser [13]. The second form is the one most commonly referred to as the KAM theorem. Katok [8, p53] mentions that the perturbations need to be at least C^3.

5.5.2. *The small denominator problem.* In the statement of the KAM theorem, we have the assumption that

$$\left| \frac{\omega}{2\pi} - \frac{p}{q} \right| \geq \frac{\gamma(\epsilon)}{|q|^\mu}.$$

where $\mu \geq 2$ (Siegel and Moser [15, p231]). This is often referred to as the sufficient irrationality condition. Arrowsmith and Place [4, pp331-332] review results regarding the relationship between the value of μ and the minimum amount of differentiability required for perturbations. They state that a less stringent condition on sufficient irrationality implies that more differentiability is required in the perturbations. In the following section, we will outline why this condition is required.

In the proof of the KAM theorem, we require $\mathbf{U}$ to transform $\mathbf{M}$ into

$$\mathbf{T} : (\xi, \eta) \mapsto (\xi + \eta, \eta)$$

and thus if

$$\mathbf{U}^{-1}\mathbf{M}\mathbf{U} = \mathbf{T}$$

then

$$\mathbf{M}\mathbf{U} = \mathbf{U}\mathbf{T}.$$

Writing this out in full, we get

$$\begin{aligned} \mathbf{M}\mathbf{U}(x, y) &= \mathbf{U}\mathbf{T}(\xi, \eta) \\ \mathbf{M}(\xi + u(\xi, \eta, \epsilon), \eta + v(\xi, \eta, \epsilon)) &= \mathbf{U}(\xi + \eta, \eta, \epsilon) \end{aligned}$$

$$(\xi + u(\xi, \eta, \epsilon) + \eta + v(\xi, \eta, \epsilon) + \tilde{f}(x, y, \epsilon), \eta + v(\xi, \eta, \epsilon) + \tilde{g}(x, y, \epsilon))$$

$$= (\xi + \eta + u(\xi + \eta, \eta, \epsilon), \eta + v(\xi + \eta, \eta, \epsilon)).$$

Thus u, v satisfy the equations

$$\begin{aligned} u(\xi + \eta, \eta, \epsilon) &= u(\xi, \eta, \epsilon) + v(\xi, \eta, \epsilon) + \tilde{f}(\xi + u(\xi, \eta, \epsilon), \eta + v(\xi, \eta, \epsilon), \epsilon) \\ v(\xi + \eta, \eta, \epsilon) &= v(\xi, \eta, \epsilon) + \tilde{g}(\xi + u(\xi, \eta, \epsilon), \eta + v(\xi, \eta, \epsilon), \epsilon). \end{aligned}$$

To solve these, we need to linearise them, and thus we replace them with linear approximations

$$\begin{aligned} u(\xi + \omega, \eta, \epsilon) &= u(\xi, \eta, \epsilon) + v(\xi, \eta, \epsilon) + \tilde{f}(\xi, \eta, \epsilon) \\ v(\xi + \omega, \eta, \epsilon) &= v(\xi, \eta, \epsilon) + \tilde{g}(\xi, \eta, \epsilon) - g^*(\eta, \epsilon) \end{aligned}$$

where the change from η to ω accounts for the approximation in the angle term ξ, and we choose $g^*(\eta, \epsilon)$ to be a linear term to account for average error in the radius term η.

The standard way to solve these linear equations is to expand each of the terms by using Fourier series, and then to equate each coefficient of ϵ^n. For higher order terms to equate, we find that we need the irrationality assumption for ω, since otherwise the Fourier series expansions fail to converge due to the denominators being arbitrarily close to zero. For more detail, see Siegel and Moser [15, pp235-237].

5.5.3. *Invariant circles.* The KAM theorem concerns the persistence of invariant circles under small perturbations. It predicts the persistence of circles with a sufficiently irrational rotation number if the perturbation is sufficiently differentiable. We know that there will be an infinite number of invariant circles for small perturbations, since most irrational rotation numbers will be sufficiently irrational to persist. As perturbation strength increases, the number of invariant circles decreases until eventually there are no invariant circles left (this is easily seen with the help of Arnol'd tongues [6, p70]). The presence of invariant circles is important in governing the dynamics in the plane, since they divide the plane into two distinct parts. Iterates encased between closed invariant circles cannot escape from the enclosed region, maintaining an ordered framework that limits chaotic behaviour.

We can now use results from previous sections to discuss why we might have expected this theorem to hold true. We have seen that rational rotation numbers are locally stable under perturbation, while irrational rotation numbers are unstable. We have also looked at ω-limit sets of homeomorphisms, and discovered that if the homeomorphism is not a pure rotation, only a homeomorphism with irrational rotation number can create an ω-limit set dense in S^1. Taking these results together, we can see that we would only expect invariant curves to arise when the original unperturbed rotation was sufficiently irrational such that after perturbation the rotation number remained irrational.

5.5.4. *Applications of the KAM theorem to the stability problem.* We can think of the stability problem of the solar system as a collection of bodies held in their orbits by gravitational attraction to each other. It turns out that if we consider these orbits as the perturbed curves considered in the KAM theorem, we can predict that most orbits will be stable, although there will some orbits that may not be stable, corresponding to a rotation number that is not sufficiently irrational.

6. Concluding remarks

The study of one-dimensional circle homeomorphisms has given us insight and motivation for the more difficult and involved theory present in higher dimensions. This arises from the fact that the dynamics in higher dimensions will often incorporate lower dimensional dynamics since we can consider dynamics restricted to level sets which are of lower dimension. This insight can be useful to explain and predict behaviour, and indeed may give the tools required to prove theorems in higher dimensions.

References

[1] V. I. Arnol'd. Proof of a theorem of A. N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian. *Russ. Math. Surv.*, 18(5):9–36, 1963.

[2] V. I. Arnol'd. Small denominators and problems of stability of motion in classical and celestial mechanics. *Russ. Math. Surv.*, 18(6):85–191, 1963.

[3] V. I. Arnol'd. *Mathematical methods of classical mechanics.* Springer-Verlag, New York, 2nd edition, 1989.

[4] D. K. Arrowsmith and C. M. Place. *An introduction to dynamical systems.* Cambridge University Press, Melbourne, 1st edition, 1990.

[5] R. G. Bartle and D. R. Sherbert. *Introduction to real analysis.* John Wiley & Sons, USA, 2nd edition, 1992.

[6] W. de Melo and S. van Strien. *One-dimensional dynamics*, volume 25 of *A series of modern surveys in mathematics.* Springer-Verlag, Heidelberg, 1st edition, 1993.

[7] B-L. Hao. *Chaos II.* World Scientific Publishing, Singapore, 1990.

[8] A. Katok. Periodic and quasi-periodic orbits for twist maps. *Springer Lect. Notes Phys.*, 179:47–65, 1983. Reprinted in MacKay and Meiss pp376–394.

[9] A. Katok and B. Hasselblatt. *Introduction to the modern theory of dynamical systems*, volume 54 of *Encyclopedia of mathematics and its applications.* Cambridge University Press, Melbourne, 1st edition, 1995.

[10] A. N. Kolmogorov. Preservation of conditionally periodic movements with small change in the Hamilton function. *Akad. Nauk. SSSR Doklady*, 98:527–, 1954. Reprinted in Hao pp107-112.

[11] R. S. MacKay and J. D. Meiss. *Hamiltonian dynamical systems.* J W Arrowsmith, Bristol, 1987.

[12] J. E. Marsden and A. J. Tromba. *Vector calculus.* W. H. Freeman, USA, 3rd edition, 1988.

[13] J. K. Moser. *Stable and random motions in dynamical systems: with special emphasis on celestial mechanics*, volume 18 of *Annals of mathematics studies.* Princeton University Press, New Jersey, 1st edition, 1973.

[14] J. K. Moser. Is the solar system stable? *Math Intelligencer*, 1:65–71, 1978.

[15] C. L. Siegel and J. K. Moser. *Lectures on celestial mechanics*, volume 187 of *Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen.* Springer-Verlag, Heidelberg, 1971.