MASTER'S THESIS

Master of Science in Advanced Mathematics and Mathematical Engineering

Title: Instabilities in Hamiltonian Systems

Author: Francesc Granell i Yuste

Advisor: Marcel Guàrdia Munárriz and Maria Teresa Martínez-Seara Alonso

Department: Mathematics Department

Academic year: 2016-2017

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat de Matemàtiques i Estadística

MASTER'S DEGREE THESIS

Instabilities in Hamiltonian Systems

Author:
Francesc Granell i Yuste

Thesis advisors:
Marcel Guàrdia Munárriz
Maria Teresa Martínez-Seara Alonso

MSc in Advanced Mathematics and Mathematical Engineering
Facultat de Matemàtiques i Estadística
Universitat Politècnica de Catalunya
June 2017

Als meus pares

Abstract

In 1964, V. I. Arnol'd proved the existence of nearly-integrable Hamiltonian systems which have global instabilities (global chaotic behaviour). This phenomenon is nowadays termed under the name *Arnol'd diffusion*. One of the key ideas that he used is to "travel" along invariant manifolds of the Hamiltonian system. The purpose of this project is to understand the Arnol'd instability mechanism and study new ones using different invariant objects.

Keywords: Arnol'd diffusion, Melnikov function, Near-integrable Hamiltonian system, splitting of separatrices.

Contents

List of Symbols

$\mathbb{A}^k(\mathbb{V})$	Vector space of $k-$forms over $\mathbb{V}$
$\mathcal{B}$	Set of Brjuno numbers
$[F, G]$	Lie bracket of the functions F and G
$\{F, G\}$	Poisson bracket of the functions F and G
$\mathcal{D}$	Set of Diophantine numbers
$\mathcal{D}(\tau, \gamma)$	Set of Diophantine numbers of class (τ, γ)
F, G	Smooth functions from $\mathbb{R}^{2n+1}$ to $\mathbb{R}$
$\mathbb{F}$	Field
$\phi(t; t_0, z_0)$	Flow of a dynamical system satisfying $\phi(t_0; t_0, z_0) = z_0$
$gl(m, \mathbb{F})$	Set of $m \times m$ matrices over $\mathbb{F}$
$Gl(m, \mathbb{F})$	Set of non-singular $m \times m$ matrices over $\mathbb{F}$
H	Hamiltonian function from $\mathbb{R}^{2n+1}$ to $\mathbb{R}$
J	$(2n \times 2n)-$matrix defined by $J = \begin{pmatrix} 0 & \mathrm{Id}_n \\ -\mathrm{Id}_n & 0 \end{pmatrix}$
M, N	Complex manifolds
$\mu(A)$	Lebesgue measure of a set A
Ω	Standard symplectic structure in $\mathbb{R}^{2n}$
ω	Non-degenerate alternating bilinear form on a vector space
$\mathcal{R}_\theta$	Rigid rotation of angle $\theta \in \mathbb{R}$
$sp(m, \mathbb{F})$	Set of Hamiltonian or infinitesimally symplectic matrices over $\mathbb{F}$
$Sp(m, \mathbb{F})$	General linear group (i.e. set of symplectic matrices over $\mathbb{F}$)
S_ρ	Horizontal strip of width 2ρ
$\mathbb{T}$	One-dimensional torus (sometimes also $\mathbb{R}/\mathbb{Z}$)
$\mathbb{T}^n$	$n-$dimensional torus
$\mathcal{T}_\omega$	Torus with fixed frequency ω
A^T	Transpose of the matrix A
$\mathcal{U}, \mathcal{V}, \mathcal{W}$	Open subsets of $\mathbb{R}^{2n}$
$\mathbb{V}$	$m-$dimensional vector space over a field $\mathbb{F}$
$\mathbb{V}^k$	Cartesian product of the vector space $\mathbb{V}$ with itself k times
$W^{s,u}(\Lambda)$	Stable and unstable (respectively) invariant manifolds of the set Λ

Introduction

> *"Il peut arriver que des petites différences dans les conditions initiales en engendrent de très grandes dans les phénomènes finaux ; une petite erreur sur les premières produirait une erreur énorme sur les derniers. La prédiction devient impossible et nous avons le phénomène fortuit."*
>
> – Henri Poincaré, *Science et Méthode*

Dynamical Systems

The first idea one could have of a *dynamical system* is an object which is used in order to model a physical phenomenon whose state changes over time. Such models are used in financial and economic forecasting, environmental modelling, medical diagnosis, industrial equipment diagnosis, and a host of other applications. From a mathematical point of view, a continuous dynamical system is a differential equation

$$\dot{x} = f(x),$$

where $f\colon \mathcal{U} \subset M \to N$ is a smooth function and M and N are manifolds. If the system can be solved, it can be *integrated* to find its solutions. Then, given an initial point it is possible to determine all its future positions. Such a collection of points is known as a *trajectory* or *orbit* of the dynamical system, which is given by the *flux* (also *semigroup*) $\phi(t; t_0, z_0)$.

Finding an orbit may require sophisticated mathematical techniques and can be accomplished only for a small class of dynamical systems. This leads to the search of techniques that help understand the behaviour of a system from a qualitative point of view, even when the solutions cannot be computed. The collection of such techniques define a field which is nowadays referred to as *qualitative theory* of dynamical systems, and started with the work of Henry Poincaré in *New Methods of Celestial Mechanics* (1899), which is considered the first oeuvre in the subject. The methods developed therein laid the basis for the local and global analysis of nonlinear differential equations, including the use of first-return (Poincaré) maps, stability theory for fixed points and periodic orbits and stable and unstable manifolds. More strikingly, using the example of a periodically-perturbed pendulum, Poincaré showed that mechanical systems with $n \geq 2$ degrees of freedom may not be integrable, due to the presence of homoclinic orbits.

Hamiltonian Systems

In many cases, dynamical systems have a special structure which makes them somewhat rigid, but also gives them a bunch of very useful properties. We refer to *Hamiltonian systems*, and they can be introduced in a natural way by means of the well-known Newton's Second Law, which describes the relation between the forces that are exerted onto a body whose mass is m and its acceleration, namely

$$ma = \sum_{i \in I} F_i.$$

This law gives rise to a system of second-order differential equations in $\mathbb{R}^n$ and so to a system of first-order equations in $\mathbb{R}^{2n}$. When the forces are "derived" from a potential function (in a way that needs to be specified), the system is called *conservative.* Under this condition, the equations of motion of the mechanical system have many special properties, which can be easily formalised in the setting of Hamiltonian systems.

A Hamiltonian system is a system of $2n$ differential equations of the form

$$\begin{cases} \dot{q} = \dfrac{\partial H}{\partial p}(q,p,t), \\ \dot{p} = -\dfrac{\partial H}{\partial q}(q,p,t), \end{cases}$$

where

$$\begin{array}{rrcl} H: & \mathcal{U} = \mathring{\mathcal{U}} \subset \mathbb{R}^{2n+1} & \longrightarrow & \mathbb{R} \\ & (q,p,t) & \longmapsto & H(q,p,t) \end{array}$$

is a smooth real-valued map called the *Hamiltonian.* The system is then said to have n *degrees of freedom*, and the vectors $q = (q_1, q_2, \ldots, q_n)$ and $p = (p_1, p_2, \ldots, p_n)$ are called its *positions* and *momenta*, respectively. Alternatively, q and p are called the *actions* and *angles*, respectively. The choice of either of these two terminologies depends on the context of the problem. The terms *position* and *momentum* relate to a Classical Mechanics context, while the terms *action* and *angle* are more used in Celestial Mechanics. These and other basic definitions and results in Hamiltonian systems can be found in Chapter 1.

Given a Hamiltonian system, a *first integral* or *conserved quantity* is a smooth function $F\colon \mathcal{U} \to \mathbb{R}$ that is constant along the solutions of the system, that is

$$F(\phi(t, z_0)) = F(z_0) = \text{constant}.$$

The level surfaces $F^{-1}(c) \subset \mathbb{R}^{2n}$, where c is a constant, are invariant sets. In general, they are $(2n-1)-$manifolds. In the (very special) case when there are $2n-1$ independent integrals $F_1, F_2, \ldots, F_{2n-1}$, then fixing them defines a *solution curve* in $\mathbb{R}^{2n}$, and in the classical sense, the problem is said to be *integrated.*

There is a whole theory that focuses on the integrability of Hamiltonian systems, which we concentrate in Chapter 2. We say that a system of differential equations in $\mathbb{R}^n$ is *integrable by quadratures* if its solutions can be computed by a finite number of *algebraic* operations (including inversion of functions) and quadratures (i.e. computation of integrals of known functions). As we shall see, all 1−degree-of-freedom Hamiltonian systems are integrable in this sense. For Hamiltonian systems with more degrees of freedom, the integrability is subject to the existence of a "large enough" set of first integrals. Liouville-Arnol'd's Theorem (see Theorem 2.1.4) states that if these conserved quantities satisfy certain conditions, there exists a submanifold which is invariant under the flow associated to any Hamiltonian defined by one of the integrals. Moreover, it gives a topological classification of the submanifold.

Hamiltonian systems which satisfy Liouville-Arnol'd's Theorem are called *completely integrable*. A particular case occurs when the invariant submanifold is compact and connex. Then, it is diffeomorphic to an $n-$dimensional torus

$$\mathbb{T}^n = \{(\varphi_1, \varphi_2, \dots, \varphi_n) \colon \varphi_i = \mathbb{R}/2\pi\mathbb{Z}\}, \qquad \varphi_i = \varphi_i^0 + t\omega_i,$$

where n is the number of first integrals $F_1, F_2, \dots, F_n \colon \mathcal{U} \subset \mathbb{R}^{2n} \to \mathbb{R}$. We say that the motion is *quasiperiodic* with frequency vector $\omega = (\omega_1, \omega_2, \dots, \omega_n)$ (see Definition 2.1.1). Moreover, ω is called

- *non-resonant* or *rationally-independent* if $\omega \cdot k \neq 0$, for any $k \in \mathbb{Z}^n \setminus \{0\}$,
- *resonant* or *rationally-dependent* if there exists $k^* \in \mathbb{Z}^n \setminus \{0\}$ such that $\omega \cdot k^* = 0$.

An important consequence of Liouville-Arnol'd's Theorem is that it gives a system of coordinates which is particularly suitable for studying completely-integrable systems. These are the so-called *action-angle coordinates*. The actions are usually denoted by I and the angles by θ (also φ or ϕ). In Section 2.1.2.1 we shall see that any completely-integrable Hamiltonian system in action-angle coordinates can be written as $H(\varphi, I) = H(I)$, so that H does not depend on the angles.

Near-Integrable Hamiltonian Systems and Arnol'd Diffusion

Sometimes we consider Hamiltonian systems of the form

$$H(\theta, I) = H_0(I) + \varepsilon H_1(\theta, I),$$

for $\varepsilon \ll 1$, $\theta \in \mathbb{T}^n$ and $I \in V \subset \mathbb{R}^n$. These systems, which are treated in more detail in Chapter 3, result from a small perturbation of an integrable Hamiltonian system. Even though the perturbation can be very small, it can cause significant changes in the dynamics of the complete system, as is evidenced by its equations of motion, namely

$$\begin{cases} \dot{\theta} = \partial_I H_0(I) + \varepsilon \partial_I H_1(\theta, I), \\ \dot{I} = -\varepsilon \partial_\theta H_1(\theta, I). \end{cases}$$

If $\varepsilon = 0$, the actions are constants of motion and orbits are confined to $\mathbb{T}^n = \{I = I_0\}$. For $\varepsilon > 0$, we have $\dot{I} = \varepsilon$, hence the actions change much more slowly than the actions. This brings up two questions:

- when we let $0 < \varepsilon \ll 1$, are there still some invariant $n-$dimensional tori, even if slightly deformed? In the literature, such tori are said to *persist* under the perturbation, and even if in principle some tori persist, we expect some others to disappear, for otherwise the whole system would still be integrable;

- which of these tori $I = I_0$ persist and which break down, and what is the new dynamics that they originate?

The answer to the first question is affirmative. Moreover, persistent tori correspond to those that have "sufficiently irrational" frequencies (see Definition 2.1.2). However, this "sufficiently irrational" condition becomes increasingly difficult to satisfy as the number of degrees of freedom grows, so that other tori are destroyed, allowing for the actions to drift indefinitely. This drift in the actions is known as *Arnol'd Diffusion*, and we present it in the third chapter of this text. We provide and study the example that Arnol'd devised to prove this phenomenon. The system that he chose is a 5−degrees-of-freedom Hamiltonian

$$H(I_1, I_2, \phi_1, \phi_2, t; \varepsilon, \mu) = H_0(I_1, I_2) + H_1(\phi_1, \phi_2; t), \quad \text{for } \varepsilon, \mu > 0,$$

where

$$H_0(I_1, I_2) = \frac{1}{2}(I_1^2 + I_2^2), \quad \text{and} \quad H_1(\phi_1, \phi_2, t; \varepsilon, \mu) = \varepsilon(\cos\phi_1 - 1)(1 + \mu(\sin\phi_2 + \cos t)).$$

When $\varepsilon > 0$ and $\mu = 0$, this system possesses invariant tori that have coincident stable and unstable manifolds, thus forming a homoclinic orbit (or *separatrix*). When we let $\mu > 0$, the perturbation causes a break in the homoclinic orbit, so that the phenomenon of *splitting of the separatrix* appears. The key point to prove diffusion is to see that this splitting makes the stable manifold of each torus intersect transversally the unstable manifold of a different torus, so that the existence of a chain of intersecting manifolds is established. This allows to prove that there are orbits that travel along different tori by closely following these intersecting manifolds.

Diffusion in More General Near-Integrable Systems

After Arnol'd's paper was published, the main cause of diffusion in Hamiltonian systems was believed to be the splitting of separatrices. This arises an almost instinctive interest in studying the splitting in more general near-integrable Hamiltonian systems. In Chapter 4, we deal with the dynamics on a torus originated by a high-frequency perturbation of the pendulum. The size of the splitting of such a perturbation is given up to order one by the Melnikov function. In [1] the value of the splitting is shown to be exponentially small with respect to ε provided that the perturbation's amplitude is small enough with respect to ε. We give a similar result even when the perturbation exists in a strip whose width is logarithmic with respect to ε. Namely, we consider a high-frequency perturbation of the pendulum described by the Hamiltonian function

$$\frac{\omega \cdot I}{\varepsilon} + h(x, y, \theta, \varepsilon),$$

where

$$\omega \cdot I = \omega_1 I_1 + \omega_2 I_2, \quad \text{and} \quad h(x, y, \theta, \varepsilon) = \frac{y^2}{2} + \cos x + \varepsilon^p m(\theta_1, \theta_2)\cos x,$$

where the function m is assumed to be a $2\pi-$periodic function of two variables θ_1 and θ_2. We assume that ε is a small positive parameter and p is a positive parameter. We also assume that the frequency is of the form ω/ε, where

$$\omega = (1, \gamma), \quad \text{and} \quad \gamma = \frac{1+\sqrt{5}}{2}.$$

The Hamiltonian function given above can be considered as a singular perturbation of the pendulum

$$h_0 = \frac{y^2}{2} + \cos x,$$

which has a saddle point $(0, 0)$ and a homoclinic trajectory. For $p > 3$ and small $\varepsilon > 0$, the invariant manifolds split. As is well known, the Melnikov function gives a first-order approximation of the difference between the values of the unperturbed pendulum energy h_0 on the stable and unstable manifolds. We provide an estimate of its size which shows that the splitting occurs. Indeed, taking m to be analytic in a strip $\{|\mathrm{Im}(\theta_1)| < r_1\} \times \{|\mathrm{Im}(\theta_2)| < r_1\}$, where

$$r_i = b_i \log \frac{1}{\varepsilon}, \qquad \text{for } i = 1, 2,$$

we obtain the following

Proposition (Estimate for the size of the Melnikov coefficients)**.** *The maximum of the modulus of the Melnikov function*

$$\max_{(\theta_1,\theta_2)\in\mathbb{T}^2} |M(\theta_1, \theta_2; \varepsilon)|,$$

taken on real arguments, can be bounded from above and from below by terms of the form

$$const\ \varepsilon^{p-1} \exp\left(-\sqrt{\frac{-\log\varepsilon}{\varepsilon}}c(\log(-\varepsilon\log\varepsilon))\right)$$

with different $\varepsilon-$independent constants, where the function c in the exponent is defined by

$$c(\delta) = C_0 \cosh\left(\frac{\delta-\delta_0}{2}\right), \quad \textit{for } \delta \in [\delta_0 - \log\gamma, \delta_0 + \log\gamma],$$

where

$$C_0 = \sqrt{2\pi C_F(\gamma b_1 + b_2)}, \quad C_F = \frac{1}{\gamma+\gamma^{-1}}, \quad \delta_0 = \log\varepsilon^* \quad \textit{and} \quad \varepsilon^* = \frac{\pi(\gamma+\gamma^{-1})}{2\gamma^2(b_1\gamma+b_2)},$$

and continued by $2\log\gamma-$periodicity onto the whole real axis. The function is piecewise analytic and continuous.

Structure of the text

Let us end this introduction by explaining how we have structured this text. The first two chapters review the mathematical background needed, and fix the notation used in the remainder of the text. The third one gives a detailed description of the phenomenon of Arnol'd diffusion based on the original paper [2]. The last chapter contains some computations to see if the splitting occurs in a particular case of a near-integrable system.

1

Preliminaries on Hamiltonian Systems

In this chapter we give a very shallow overview of Hamiltonian systems. It begins with the definition of a Hamiltonian system of ordinary differential equations and gives some basic results about such systems. The second section is devoted to studying the *Poisson bracket* along with some of its most important properties. We finally dedicate the third part of the chapter to Symplectic Hamiltonian Systems, focusing on the linear case in the first place and moving on to the general case afterwards. Although this discussion is far from complete and sometimes informal, it is intended to provide a general overview of the subject, which will help follow the subsequent chapters. All the explanation is based on [10, Ch. 1].

1.1 Hamilton's Equations

The well-known Newton's second law gives rise to systems of second-order differential equations in $\mathbb{R}^n$ and so to a system of first-order equations in $\mathbb{R}^{2n}$, that is an even-dimensional space. If the forces are derived from a potential function, the equations of motion of the mechanical system have many special properties, most of which follow from the fact that the equations of motion can be written as a *Hamiltonian system*. The Hamiltonian formalism is the natural mathematical structure in which to develop the theory of conservative mechanical systems.

A Hamiltonian system is a system of $2n$ differential equations of the form

$$\begin{cases} \dot{q} = \dfrac{\partial H}{\partial p}(q,p,t), \\ \dot{p} = -\dfrac{\partial H}{\partial q}(q,p,t), \end{cases} \tag{1.1}$$

where

$$\begin{array}{rrcl} H: & \mathcal{U} = \mathring{\mathcal{U}} \subset \mathbb{R}^{2n+1} & \longrightarrow & \mathbb{R} \\ & (q,p,t) & \longmapsto & H(q,p,t) \end{array} \tag{1.2}$$

is a smooth real-valued map called the *Hamiltonian*. The system is then said to have n *degrees of freedom*, and the vectors $q = (q_1, q_2, \dots, q_n)$ and $p = (p_1, p_2, \dots, p_n)$ are called its *positions* and *momenta*, respectively. Alternatively, q and p are called the *actions* and *angles*, respectively. The choice of either of these two terminologies depends on the context of the problem. The terms *position* and *momentum* relate to a Classical Mechanics context, while the terms *action* and *angle* are more used in Celestial Mechanics.

We introduce the $2n$ vector z, the $2n \times 2n$ skew symmetric matrix and the gradient of H by

$$z = \begin{pmatrix} q \\ p \end{pmatrix}, \quad J \equiv J_n = \begin{pmatrix} 0 & \mathrm{Id}_n \\ -\mathrm{Id}_n & 0 \end{pmatrix} \quad \text{and} \quad \nabla H \equiv \nabla_z H = \begin{pmatrix} \dfrac{\partial H}{\partial z_1} \\ \dfrac{\partial H}{\partial z_2} \\ \vdots \\ \dfrac{\partial H}{\partial z_{2n}} \end{pmatrix},$$

where 0 and Id_n denote the $n \times n$ zero and identity matrix, respectively. The particular case J_2 is sometimes denoted by K. In this notation, System (1.1) can be written as

$$\dot{z} = J \nabla H(t, z). \tag{1.3}$$

One of the basic results from the general theory of ordinary differential equations is the existence and uniqueness theorem. This theorem states that for each $(t_0, z_0) \in \mathcal{U}$, there is a unique solution $z = \phi(t; t_0, z_0)$ of System (1.3) defined for t near t_0 that satisfies the initial condition $\phi(t_0; t_0, z_0) = z_0$. The map ϕ is defined on an open neighbourhood Q of $(t_0; t_0, z_0) \in \mathbb{R}^{2n+2}$ into $\mathbb{R}^{2n}$. The function $\phi(t; t_0, z_0)$ is smooth in all its displayed arguments, and so ϕ is $\mathcal{C}^\infty$ if the equations are $\mathcal{C}^\infty$, and it is analytic if the equations are analytic.

When H is independent of t, so that $H\colon \mathcal{U} = \mathring{\mathcal{U}} \subset \mathbb{R}^{2n} \to \mathbb{R}$, the differential equations (1.1) can be written as

$$\begin{cases} \dot{q} = \dfrac{\partial H}{\partial p}(q, p), \\ \dot{p} = -\dfrac{\partial H}{\partial q}(q, p), \end{cases} \tag{1.4}$$

and are autonomous. In this case, the Hamiltonian system is called *conservative*. It follows that $\phi(t - t_0; 0, z_0) = \phi(t; t_0, z_0)$ holds, because both sides satisfy Equation (1.3) and the same initial conditions. Usually, the t_0 dependence is dropped and only $\phi(t, z_0)$ is considered, where $\phi(t, z_0)$ is the solution of Equation (1.3) satisfying $\phi(0, z_0) = z_0$. The solutions are pictured as parameterised curves in $\mathcal{U} \subset \mathbb{R}^{2n}$, and the set $\mathcal{U}$ is called the *phase space*. By the existence and uniqueness theorem, there is a unique curve through each point in $\mathcal{U}$, and by the uniqueness theorem, two solution curves cannot cross in $\mathcal{U}$.

A *first integral* (also *conserved quantity*) for the system in (1.3) is a smooth function $F\colon \mathcal{U} \to \mathbb{R}$ that is constant along the solutions of (1.3), that is

$$F(\phi(t, z_0)) = F(z_0) = \text{constant}.$$

The level surfaces $F^{-1}(c) \subset \mathbb{R}^{2n}$, where c is a constant, are invariant sets. In general, they are $(2n-1)-$manifolds. In the (very special) case when there are $2n-1$ independent integrals $F_1, F_2, \ldots, F_{2n-1}$, then fixing them defines a *solution curve* in $\mathbb{R}^{2n}$, and in the classical sense, the problem is said to be *integrated.*

A non-autonomous Hamiltonian system can always be made autonomous in the following way. We consider a new Hamiltonian $\widetilde{H}\colon \widetilde{\mathcal{U}} \to \mathbb{R}$ with $n+1$ degrees of freedom, where the *extended phase space* $\widetilde{\mathcal{U}} \subset \mathbb{R}^{2n+2}$ is an open set. This Hamiltonian has time as another variable (which we now call s) and there is a new variable A, so that

$$\widetilde{H}(q,s,p,A) := A + H(q,p,s). \tag{1.5}$$

The equations of this new Hamiltonian system are

$$\begin{cases} \dot{q} = \dfrac{\partial H}{\partial p}(q,p,s), \\ \dot{s} = 1, \\ \dot{p} = -\dfrac{\partial H}{\partial q}(q,p,s), \\ \dot{A} = 0. \end{cases} \tag{1.6}$$

The orbits of H in the energy level $\{(q,p,t) \in \mathcal{U} \mid \widetilde{H}(q,p,t) = 0\}$ are equivalent to the dynamics of $\widetilde{H}$ in the extended phase space. We denote the flow in $\widetilde{\mathcal{U}}$ as $\widetilde{\phi}(t,q,s,p,A)$, so that

$$\widetilde{\phi}(0,q,s,p,A) = (q,s,p,A).$$

1.2 The Poisson Bracket

Many of the special properties of Hamiltonian systems are formulated in terms of the *Poisson bracket* operator, hence it plays a central role in the theory of Hamiltonian systems. Let $F, G, H\colon \mathcal{U} = \mathring{\mathcal{U}} \subset \mathbb{R}^{2n+1} \to \mathbb{R}$ be smooth functions, and define the Poisson bracket of F and G by

$$\{F,G\} := \nabla F^T J \nabla G = \sum_{i=1}^{n} \left(\frac{\partial F}{\partial q_i}(t,q,p)\frac{\partial G}{\partial p_i}(t,q,p) - \frac{\partial F}{\partial p_i}(t,q,p)\frac{\partial G}{\partial q_i}(t,q,p) \right). \tag{1.7}$$

Clearly, $\{F,G\}$ is a smooth function from $\mathcal{U}$ to $\mathbb{R}$ as well, which defines a bilinear, skew symmetric form. A tedious calculation shows that *Jacobi's Identity* is satisfied:

$$\{F,\{G,H\}\} + \{G,\{H,F\}\} + \{H,\{F,G\}\} = 0. \tag{1.8}$$

The Poisson bracket is useful to see how a function evolves along the trajectories of a Hamiltonian system. Consider a function F as above. We have

$$\begin{aligned}\frac{dF}{dt}(\phi(t;t_0,z_0),t) &= (D_zF)(\phi(t;t_0,z_0),t)\cdot\frac{d\phi}{dt}(t;t_0,z_0)+\frac{\partial F}{\partial t}(\phi(t;t_0,z_0),t)\\ &= (\nabla F)^T(\phi(t;t_0,z_0),t)\cdot J\cdot\nabla H(\phi(t;t_0,z_0),t)+\frac{\partial F}{\partial t}(\phi(t;t_0,z_0),t)\\ &= \{F,G\}(\phi(t;t_0,z_0),t)+\frac{\partial F}{\partial t}(\phi(t;t_0,z_0),t).\end{aligned}$$

In particular,

$$\frac{dH}{dt}(\phi(t;t_0,z_0),t)=\frac{\partial H}{\partial t}(\phi(t;t_0,z_0),t).$$

Moreover, if $H = H(q,p)$, then

$$\frac{dH}{dt}(\phi(t;t_0,z_0))=0.$$

In general, F is a first integral for System (1.3) if, and only if

$$\{F,H\}(\phi(t;t_0,z_0),t)=0.$$

In many examples, the Hamiltonian H is the total energy of a physical system, in which case it remains constant along the trajectories. In the conservative case when H is independent of t, a critical point $z^* = (q^*,p^*) \in \mathcal{U} \subset \mathbb{R}^{2n}$ of H (i.e. $\nabla H(z^*) = 0$), is an equilibrium point of the system of differential equations (1.4), that is $z(t) = z^*$.

Definition 1.2.1 (Lyapunov stability)**.** A critical point z^* of System (1.4) is *stable in the sense of Lyapunov* (or *Lyapunov-stable*) if for any $\varepsilon > 0$ there exists $\delta > 0$ such that if $\|z_0 - z^*\| < \delta$, then

$$\|\phi(t,z_0)-z^*\|<\varepsilon,$$

for any $t \in \mathbb{R}$.

Theorem 1.2.1 (Dirichlet)**.** *If z^* is a strict local maximum or minimum of H, then z^* is stable in the sense of Lyapunov.*

Proof. Without loss of generality, assume that $z^* = 0$ and $H(0) = 0$. We shall do the case where z^* is a strict local minimum. Then, there exists $\eta > 0$ such that $H(z) > 0$ for any η satisfying $0 < \|z\| < \eta$. Let $\varepsilon > 0$ be fixed. Let $k = \min\{\varepsilon, \eta\}$ and $m = \min\{H(z) \mid \|z\| < 1\}$, which exist since the unit ball is a compact set. Since $H(0) = 0$ and H is continuous, there exists $\delta > 0$ such that $H(z) < m$ for any z satisfying $\|z\| < \delta$, we have $|H(z)| < m$. Now we have that for any z_0 whose norm is strictly less than δ and for any $t \in \mathbb{R}$,

$$H(\phi(t,z_0))=H(z_0)<m.$$

Then, $z^* = 0$ is stable, for otherwise there would exist z_0 with $\|z_0\| < \delta$ and there would exist $t^* \in \mathbb{R}$ such that

$$\|\phi(t^*,z_0)\|=k<\varepsilon,$$

so $|H(\phi(t^*, z_0))| \geq m$, which is a contradiction. □

1.3 Symplectic Hamiltonian Systems

1.3.1 The Linear Case

In this section, we study Hamiltonian systems that are linear differential equations. Many of the basic facts about Hamiltonian systems and symplectic geometry are easy to understand in this simple context.

Let $gl(m, \mathbb{F})$ denote the set of all $m \times m$ matrices with entries in the field $\mathbb{F}$ and $Gl(m, \mathbb{F})$ the set of all nonsingular $m \times m$ matrices with entries in $\mathbb{F}$. The set $Gl(m, \mathbb{F})$ is a group under matrix multiplication, and is called the *general linear group*. The matrices $I = I_m$ and $0 = 0_m$ denote the $m \times m$ identity and zero matrices, respectively. In general, the subscript is clear from the context.

In this context, a special role is played by the $2n \times 2n$ matrix

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}.$$

Note that J is an orthogonal skew-symmetric matrix, that is

$$J^{-1} = J^T = -J.$$

Let z be a coordinate vector in $\mathbb{R}^{2n}$, I an interval in $\mathbb{R}$ and $S\colon I \to gl(2n, \mathbb{R})$ be continuous and symmetric. A *linear Hamiltonian system* is a system of $2n$ ordinary differential equations

$$\dot{z} = J\frac{\partial H}{\partial z} = JS(t)z = A(t)z, \tag{1.9}$$

where

$$H = H(t, z) = \frac{1}{2}z^T S(t)z.$$

The Hamiltonian H is a quadratic form in the zs with coefficients that are continuous in $t \in I \subset \mathbb{R}$. If S, and hence H, is independent of t, then H is an integral for (1.9) as seen in Section 1.2.

Let $t_0 \in I$ be fixed. From the theory of differential equations, for each $z_0 \in \mathbb{R}^{2n}$, there exists a unique solutions $\phi(t; t_0, z_0)$ of (1.9) for all $t \in I$ that satisfies the initial condition $\phi(t_0; t_0, z_0) = z_0$. Let $Z(t, t_0)$ be the $2n \times 2n$ fundamental matrix solution of (1.9) that satisfies $Z(t_0, t_0) = I$. Then, $\phi(t; t_0, z_0) = Z(t, t_0)z_0$.

In the case where S and A are constant, we take $t_0 = 0$ and

$$Z(t) = e^{At} = \exp At = \sum_{i=1}^{\infty} \frac{A^n t^n}{n!}.$$

Definition 1.3.1 (Hamiltonian matrix)**.** A matrix $A \in gl(2n, \mathbb{F})$ is called *Hamiltonian* or *infinitessimally symplectic* if

$$A^T J + JA = 0. \tag{1.10}$$

The set of all $2n \times 2n$ Hamiltonian matrices is denoted by $sp(2n, \mathbb{F})$.

Theorem 1.3.1 (Characterisation of Hamiltonian matrices)**.** *Let* $A \in gl(2n, \mathbb{R})$ *and* $\alpha \in \mathbb{R}$*. The following conditions are equivalent:*

1. *A is Hamiltonian,*
2. *$A = JR$, where R is symmetric,*
3. *JA is symmetric.*

Moreover, if A and B are Hamiltonian, then so are A^T, αA, $A \pm B$ and $[AB] = AB - BA$.

Proof. See [10]. □

The function

$$\begin{array}{rrcl} [\cdot,\cdot]: & gl(m,\mathbb{F}) \times gl(m,\mathbb{F}) & \longrightarrow & gl(m,\mathbb{F}) \\ & (A,B) & \longmapsto & [A,B] = AB - BA \end{array}$$

that appears in Theorem 1.3.1 is called the *Lie product.* The second part of this theorem implies that the set of all $2n \times 2n$ Hamiltonian matrices $sp(2n, \mathbb{F})$ is a Lie algebra.

Definition 1.3.2 (Symplectic matrix with multiplier μ)**.** A $2n \times 2n$ matrix T is called *symplectic* with multiplier μ if

$$T^T JT = \mu J,$$

where μ is a nonzero constant. If $\mu = +1$, then T is simply called *symplectic.* The set of all $2n \times 2n$ symplectic matrices is denotes by $Sp(2n, \mathbb{R})$.

Theorem 1.3.2. *If T is symplectic with multiplier μ, then T is nonsingular and*

$$T^{-1} = \mu^{-1} JT^T J.$$

If T and R are symplectic with multiplier μ and ν, respectively, then T^T, T^{-1} and TR are symplectic with multipliers μ, μ^{-1} and $\mu\nu$, respectively.

This theorem implies that $Sp(2n, \mathbb{R})$ is a group, a subgroup of $Gl(2n, \mathbb{R})$. In the 2×2 case, a matrix is symplectic with multiplier μ if and only if it has determinant $+\mu$. Thus a 2×2 symplectic matrix defines a linear transformation which is orientation-preserving (and area-preserving, if $\mu = +1$).

Theorem 1.3.3. *The fundamental matrix solution $Z(t, t_0)$ of a linear Hamiltonian system* (1.9) *is symplectic for all $t, t_0 \in I$. Conversely, if $Z(t, t_0)$ is a continuously differential function of symplectic matrices, then Z is a matrix solution of a linear Hamiltonian system.*

Corollary 1.3.4. *The (constant) matrix A is Hamiltonian if and only if e^{At} is symplectic for all t.*

By changing variables by $z = T(t)u$ in System (1.9), we obtain

$$\dot{u} = (T^{-1}AT - T^{-1}\dot{T})u, \tag{1.11}$$

which, in general, is not a Hamiltonian equation.

Theorem 1.3.5. *If T is symplectic with multiplier μ^{-1}, then* (1.11) *is a Hamiltonian system with Hamiltonian given by*

$$H(t,u) = \frac{1}{2}u^T(\mu T^T S(t)T + R(t))u,$$

where

$$R(t) = JT^{-1}\dot{T}.$$

Conversely, if (1.11) *is Hamiltonian for every Hamiltonian system* (1.9)*, then U is symplectic with constant multiplier μ.*

This is an example of a change of variables that preserves the Hamiltonian character of the system of equations. The general problem of which changes of variables preserve the Hamiltonian character is the matter of study of *symplectic transformations* (see [10, Ch. 6]).

1.3.1.1 Symplectic Linear Spaces

In this section, we present a way to interpret the matrix J, which plays an important role in our context. We address the topic from the point of view of abstract linear algebra.

Let $\mathbb{V}$ be an $m-$dimensional vector space over the field $\mathbb{F}$ (here, $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$). A bilinear form is a mapping $B\colon \mathbb{V}\times\mathbb{V}\to\mathbb{F}$ that is linear in both variables. It is *skey-symmetric* or *alternating* if $B(u,v) = -B(v,u)$ for all $u,v\in\mathbb{V}$, and it is *non-degenerate* if $B(u,v) = 0$ for all $v\in\mathbb{V}$ implies $u = 0$.

Let B be a bilinear form and $\{e_i\}_{i=1}^m$ be a basis for $\mathbb{V}$. Given any vector $v\in\mathbb{V}$, we write $v = \sum \alpha_i e_i$ and define an isomorphism

$$\begin{array}{rccl} \Phi: & \mathbb{V} & \longrightarrow & \mathbb{F}^m \\ & v & \longmapsto & a = (\alpha_1,\ldots,\alpha_m). \end{array}$$

Define $s_{ij} = B(e_i,e_j)$ and S to be the $m\times m$ matrix $S = (s_{ij})$, the matrix of B in the basis $\{e_i\}_{i=1}^m$. Let $\Phi(u) = b = (\beta_1,\ldots,\beta_m)$. Then,

$$B(u,v) = \sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_i\beta_j B(e_i,e_j) = b^T S a.$$

So in the coordinates defined by the basis $\{e_i\}_{i=1}^m$, the bilinear form is just $b^T Sa$, where S is the matrix $(B(e_i,e_j))_{ij}$. If B is alternating, then S is skew-symmetric, and if B is non-degenerate, then S is nonsingular and conversely.

Theorem 1.3.6. *Let S be any skew-symmetric matrix. Then, there exists a nonsingular matrix Q such that*

$$R = QSQ^T = \operatorname{diag}(K,K,\ldots,K,0,\ldots,0),$$

where

$$K = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Or, given an alternating form B, there is a basis for $\mathbb{V}$ such that the matrix of B in this basis is R.

Note that the rank of a skew-symmetric matrix is always even. Thus, every non-degenerate, alternating bilinear form is defined on an even dimensional space.

A *symplectic (linear) space* is a pair $(\mathbb{V}, \omega)$, where $\mathbb{V}$ is a $2n-$dimensional vector space over $\mathbb{F}$ and ω is a non-degenerate alternating bilinear form on $\mathbb{V}$. The form ω is called the *symplectic form* or the *symplectic inner product.* Throughout the rest of this section, we shall assume that $\mathbb{V}$ is a symplectic space with symplectic form ω. The standard example is $\mathbb{F}^{2n}$ and $\omega(x, y) = x^T Jy$. In this example, we shall write $\{x, y\} = x^T Jy$ and denote the space by $(\mathbb{F}^{2n}, J)$ or simply by $\mathbb{F}^{2n}$, if no confusion can arise.

A symplectic basis for $\mathbb{V}$ is a basis $v_1, \ldots, v_{2n}$ for $\mathbb{V}$ such that $\omega(v_i, v_j) = J_{ij}$, that is, the $(i, j)-$th entry of J. A *symplectic basis* is a basis so that the matrix of ω is just J. The standard basis $\{e_i\}_{i=1}^{2n}$, where e_i is 1 in the $i-$th position and zero elsewhere, is a symplectic basis for $(\mathbb{F}^{2n}, J)$. Given two simplectic spaces $(\mathbb{V}_i, \omega_i)$, for $i = 1, 2$, a *(symplectic) isomorphism* is a linear isomorphism

$$\begin{array}{rrcl} L: & \mathbb{V}_1 & \longrightarrow & \mathbb{V}_2 \\ & x & \longmapsto & L(x) \end{array}$$

such that $\omega_2(L(x), L(y)) = \omega_1(x, y)$ for all $x, y \in \mathbb{V}_1$. In other words, L preserves the symplectic form. In this case, we say that the two spaces are *symplectically isomporphic* or *symplectomorphic.*

Corollary 1.3.7. *Let* $(\mathbb{V}, \omega)$ *be a symplectic space of dimension* $2n$. *Then,* $\mathbb{V}$ *has a symplectic basis. Moreover,* $(\mathbb{V}, \omega)$ *is symplectically isomorphic to* $(\mathbb{F}^{2n}, J)$*, that is all symplectic spaces of dimension* $2n$ *are isomorphic.*

In view of these results, it is clear that the study of symplectic linear spaces is really the study of one canonical example, say $(\mathbb{F}^{2n}, J)$. Or put another way, J is just the coefficients matrix of the symplectic form in a symplectic basis.

1.3.2 General Case

Differential forms play an important part in the theory of Hamiltonian systems. It gives a natural higher-dimensional generalisation of the results of classical vector calculus. We give a brief introduction with some, but not all, proofs and refer the reader to [4] for another informal introduction but a more complete discussion with many applications, or to [9] or [6] for a more complete mathematical discussion. What is presented here is not meant to be a complete development but simply an introduction to a few results that are used sparingly later.

In this section, we introduce and use the notation of classical differential geometry by using superscripts and subscripts to differentiate between a vector and its dual. This convention helps sort out the multitude of different types of vectors encountered.

1.3.2.1 Exterior Algebra

Let $\mathbb{V}$ be a vector space of dimension m over the real number $\mathbb{R}$. Let $\mathbb{V}^k$ denote k copies of $\mathbb{V}$, that is $\mathbb{V}^k = \mathbb{V} \times \overset{k)}{\cdots} \times \mathbb{V}$. A function

$$\begin{array}{rrcl} \phi : & \mathbb{V}^k & \longrightarrow & \mathbb{R} \\ & (a_1, a_2, \ldots, a_k) & \longmapsto & \phi(a_1, a_2, \cdots, a_k) \end{array}$$

is called *k-multilinear* if it is linear in each argument. We sometimes refer to a $1-$multilinear map as a *covector* or $1-$*form*. A $k-$multilinear function ϕ is *skey-symmetric* or *alternating* if interchanging any two arguments changes its sign. Thus ϕ is zero if two of its arguments are the same. We call an alternating $k-$multilinear function a $k-$*linear form* or a $k-$*form*. Let $\mathbb{A}^k = \mathbb{A}^k(\mathbb{V})$ be the space of all $k-$forms for $k \geq 1$. It is easy to verify that $\mathbb{A}^k$ is a vector space when using the usual definition of addition of functions and multiplication of functions by a scalar.

If ψ is a $2-$multilinear function, then ϕ defined by

$$\phi(a, b) = \frac{\psi(a, b) - \psi(b, a)}{2}$$

is alternating and is sometimes called the *alternating part* of ψ. This construction is generalised as follows. Let $\mathfrak{S}_k$ be the set of all permutations of the k numbers $1, 2, \cdots, k$ and let

$$\begin{array}{rccl} \mathrm{sgn}: & \mathfrak{S}_k & \longrightarrow & \{+1, -1\} \\ & \sigma & \longmapsto & \mathrm{sgn}(\sigma) \end{array}$$

be the function that assigns $+1$ to an even permutation and -1 to an odd permutation. So if ϕ is alternating,

$$\phi(a_{\sigma(1), \ldots, a_{\sigma(k)}}) = \mathrm{sgn}(\sigma)\phi(a_1, \ldots, a_k).$$

If ψ is a $k-$multilinear function, then ϕ defined by

$$\phi(a_1, \ldots, a_k) = \frac{1}{k!} \sum_{\sigma \in \mathfrak{S}_k} \mathrm{sgn}(\sigma)(a_{\sigma(1), \ldots, a_{\sigma(k)}})$$

is alternating. We write $\phi = \mathrm{alt}(\psi)$. If ψ is already alternating, then $\psi = \mathrm{alt}(\psi)$. If $\alpha \in \mathbb{A}^k$ and $\beta \in \mathbb{A}^r$, then *exterior product* is the operator defined by

$$\begin{array}{rrcl} \wedge : & \mathbb{A}^k \times \mathbb{A}^r & \longrightarrow & \mathbb{A}^{k+r} \\ & (\alpha, \beta) & \longmapsto & \alpha \wedge \beta = \dfrac{(k+r)!}{k!\ r!} \mathrm{alt}(\alpha\beta). \end{array}$$

Proposition 1.3.8. *For all $\alpha \in \mathbb{A}^k$, $\beta, \delta \in \mathbb{A}^r$ and $\gamma \in \mathbb{A}^s$,*

1. $\alpha \wedge (\beta + \delta) = \alpha \wedge \beta + \alpha \wedge \delta$,
2. $\alpha \wedge (\beta \wedge \gamma) = (\alpha \wedge \beta) \wedge \gamma$,
3. $\alpha \wedge \beta = (-1)^k r \beta \wedge \alpha$.

Proposition 1.3.9. *The space of alternate $k-$forms $\mathbb{A}^k$ has dimension $\binom{m}{k}$. In particular, a basis for $\mathbb{A}^k$ is*

$$\{f^{i_1} \wedge f^{i_2} \wedge \cdots \wedge f^{i_k}\}_{1 \leq i_1 < i_2 < \cdots < i_k \leq m}.$$

In particular, the dimension of $\mathbb{V}^m$ is 1, and the space has as a basis the single element $f^1 \wedge \cdots \wedge f^m$.

Proposition 1.3.10. *Let $g^1, \ldots, g^r \in \mathbb{V}^*$. Then, $g^1, \ldots, g^r$ are linearly independent if and only if $g^1 \wedge \cdots \wedge g^r \neq 0$.*

A linear map $L \colon \mathbb{V} \to \mathbb{V}$ induces a linear map $L_k \colon \mathbb{A}^k \to \mathbb{A}^k$ by the formula

$$L_k\phi(a_1, \ldots, a_k) = \phi(La_1, \ldots, La_k).$$

If M is another linear map of $\mathbb{V}$ onto itself, then $(LM)_k = M_k L_k$, because

$$\begin{aligned}(LM)_k\phi(a_1, \ldots, a_k) &= \phi(LMa_1, \ldots, LMa_k) \\ &= L_k\phi(Ma_1, \ldots, Ma_k) \\ &= M_k L_k \phi(a_1, \ldots, a_k).\end{aligned}$$

Recall that $\mathbb{A}^1 = \mathbb{V}^*$ is the *dual space*, and $L_1 = L^*$ is called the *dual map*. If $\mathbb{V} = \mathbb{R}^m$ (column vectors), then we can identify the dual space $\mathbb{V}^*$ with $\mathbb{R}^m$ by the isomorphism

$$\begin{array}{rccl}\Phi: & \mathbb{R}^m & \longrightarrow & \mathbb{V}^* \\ & \hat{f} & \longmapsto & f := \langle \hat{f}^T, \cdot \rangle.\end{array}$$

In this case, L is and $m \times m$ matrix, and Lx is the usual matrix product. The product $L_1 f$ is defined by

$$L_1 f(x) = f(Lx) = \hat{f}^T Lx = \hat{f}^T Lx = (L^T \hat{f})^T x.$$

Therefore, the matrix representation of L_1 is the transpose of L, that is $L_1(f) = L^T \hat{f}$. The matrix representation of L_k is discussed in [4].

By Proposition 1.3.9, $\dim \mathbb{A}^m = 1$, and so every element in $\mathbb{A}^m$ is a scalar multiple of a single element. Since L_m is a linear map, there is a constant ℓ such that $L_m f = \ell f$, for all $f \in \mathbb{A}^m$. Define the determinant of L to be this constant ℓ, and denote it by $\det(L)$, so

$$L_m = \det(L) f,$$

for all $f \in \mathbb{A}^m$.

Proposition 1.3.11. *Let $L, M \colon \mathbb{V} \to \mathbb{V}$ be linear. Then,*

1. $\det(LM) = \det(L)\det(M)$,
2. $\det(\mathrm{Id}) = 1$, *where* $\mathrm{Id} \colon \mathbb{V} \to \mathbb{V}$ *is the identity map,*
3. *L is invertible if and only if* $\det(L) \neq 0$, *in which case,* $\det(L^{-1}) = \det(L)^{-1}$.

Let $\mathbb{V} = \mathbb{R}^m$ and let $\{e_i\}_{i=1}^m$ be the standard basis of $\mathbb{R}^m$. Let L be the matrix (L_i^j), so that

$$Le_i = \sum_{j=1}^{m} L_i^j e_j.$$

Let ϕ be a nonzero element of $\mathbb{A}^m$. Then, from the definition that we have given for the determinant, we have

$$\begin{aligned}
\det(L)\phi(e_1, \ldots, e_m) &= L_m \phi(e_1, \ldots, e_m) \\
&= \phi(Le_1, \ldots, Le_m) \\
&= \sum_{j_1=1}^{m} \cdots \sum_{j_m=1}^{m} L_1^{j_1} \cdots L_m^{j_m} \phi(e_{j_1}, \ldots, e_{j_m} \\
&= \sum_{\sigma \in \mathfrak{S}_m} \operatorname{sgn}(\sigma) L_1^{\sigma(1)} \cdots L_m^{\sigma(m)} \phi(e_1, \ldots, e_m).
\end{aligned}$$

In the second-to-last sum above, the only nonzero terms are the ones with distinct es. Thus, the sum over the nonzero terms is the sum ove all permutations of the es. From the above,

$$\det(L) = \sum_{\sigma \in \mathfrak{S}_m} \operatorname{sgn}(\sigma) L_1^{\sigma(1)} \cdots L_m^{\sigma(m)},$$

which is one of the classical formulae for the determinant of a matrix.

1.3.2.2 The Symplectic Form

In this section, let $(\mathbb{V}, \omega)$ be a symplectic space of dimension $2n$ (see Section 1.3.1.1). Recall that a symplectic form ω on a vector space $\mathbb{V}$ is a nondegenerate, alternating bilinear form on $\mathbb{V}$, and the pair $(\mathbb{V}, \omega)$ is called a symplectic space.

Theorem 1.3.12. *There exitsts a basis $f^1, \ldots, f^{2n}$ for $\mathbb{V}^*$ such that*

$$\omega = \sum_{i=1}^{n} f^i \wedge f^{n+i}.$$

The basis $\{f^1, \ldots, f^{2n}\}$ is a symplectic basis for the dual space $\mathbb{V}^*$. By the above,

$$\omega^n = \omega \wedge \overset{n)}{\cdots} \wedge \omega = \pm n! f^1 \wedge f^2 \wedge \cdots \wedge f^{2n},$$

where the sign is $(-1)^n$. Thus, ω^n is a nonzero element of $\mathbb{A}^{2n}$. Because a symplectic linear transformation preserves ω, it preserves ω^n and, therefore, its determinant is $+1$.

Corollary 1.3.13. *The determinant of a symplectic linear transformation is $+1$.*

1.3.2.3 Tangent and Cotangent Vectors

This section deals with the concepts of tangent vectors, which arise from the linearisation of a curve at a point, and cotangent vectors, which arise from the linearisation of a function at a point. The analysis of these two objects is fundamental, since much of analysis reduces to studying maps from an interval in $\mathbb{R}$ into an open set of a vector space and vice-versa.

Let $\mathcal{O}$ be an open set in an $m-$dimensional vector space $\mathbb{V}$ over $\mathbb{R}$. Let $\{e_i\}_{i=1}^m$ be a basis for $\mathbb{V}$, and $f^1, \ldots, f^m$ the dual basis. Let $x = (x^1, \ldots, x^m)$ be coordinates in $\mathbb{V}$ and also in $\mathbb{V}^*$ relative to the basis $\{e_i\}_{i=1}^m$ and to the dual basis $f^1, \ldots, f^m$, respectively. Let $I = (-1, 1) \subset \mathbb{R}$, and let $t \in \mathbb{R}$. A *tangent vector* at $p \in \mathcal{O}$ can be thought of as the tangent vector to a curve through p. Two curves that have the same tangent vector at a point p are called *equivalent*. Tangency, in fact, is an equivalence relation, and tangent vectors to $\mathcal{O}$ at p are the equivalence classes of this relation, which we denote by $\{g\}$. The set of all tangent vectors to $\mathcal{O}$ at p is called the *tangent space* to $\mathcal{O}$ at p and is denoted by $T_p\mathcal{O}$, and it can be made into a vector space by using the coordinates

$$\frac{dg}{dt}(0) = \left(\frac{dg^1}{dt}(0), \ldots, \frac{dg^m}{dt}(0)\right) = (\gamma^1, \ldots, \gamma^m),$$

which are the coordinates relative to the x coordinates. The curve

$$\begin{array}{rccc} \xi_i : & \mathbb{R} & \longrightarrow & \mathbb{V} \\ & t & \longmapsto & p + te_i \end{array}$$

satisfies

$$\frac{d\xi_i}{dt}(0) = e_i = (0, \ldots, \overset{i}{\overbrace{1}}, 0, \ldots, 0)$$

in the x coordinates. The tangent vector consisting of all curves equivalent to ξ_i at p is denoted by the partial derivative with respect to x^i, so that the vectors

$$\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^m}$$

form a basis for $T_p\mathcal{O}$. Therefore, a vector $v_p \in T_p\mathcal{O}$ can be written as

$$v_p = \gamma^1 \frac{\partial}{\partial x^1} + \cdots + \gamma^m \frac{\partial}{\partial x^m}.$$

In classical tensor notation, one uses Einstein's summation convention, understanding that a repeated index, one as a superscript and one as a subscript has to be summed over from 1 to m, so that

$$v_p = \gamma^i \frac{\partial}{\partial x^i}.$$

A *cotangent vector* or *covector* at p can be thought as the differential of a function at p. As in the case of tangent vectors, two smooth functions that have the same differential at p are said to be equivalent, and this defines an equivalence relation. Each equivalence class is now called the cotangent vector or covector to $\mathcal{O}$ at p, and we denote it by $\{h\}$. In the x coordinates, we write

$$Dh(p) = \left(\frac{\partial h(p)}{\partial x^1}, \ldots, \frac{\partial h(p)}{\partial x^m}\right) = (\eta_1, \ldots, \eta_m).$$

The set of all covectors at p is called the *cotangent space* to $\mathcal{O}$ at p and is denoted by $T_p^*\mathcal{O}$. It can be made into a vector space by using the coordinate representation given above. The function

$$x^i : \mathcal{O} \to \mathbb{R}$$

defines a cotangent vector at p, which is $(0, \dots, \underbrace{1}_{i}, \dots, 0)$. The covector consisting of all functions equivalent to x^i at p is denoted by dx^i. The covectors $dx^1, \dots, dx^m$ form a basis of $T_p^*\mathcal{O}$. A typical covector $v^p \in T_p^*\mathcal{O}$ can be written as

$$\eta_1 dx^1 + \cdots + \eta_m dx^m = \eta_i dx^i.$$

The two constructions that we have just explained are clearly parallel. In fact, they are dual. Let g and h be as above, so $h \circ g : I \subset \mathbb{R} \to \mathbb{R}$. By the chain rule,

$$D(h \circ g)(0)(1) = Dh(p) \circ Dg(0)(1),$$

which is a real number. Therefore, $Dh(p)$ is a linear functional on tangents to curves. In coordinates, if

$$\{g\} = v_p = \frac{dg^i}{dt}(0)\frac{\partial}{\partial x_i} = \gamma^i \frac{\partial}{\partial x_i} \quad \text{and} \quad \{h\} = v^p = \frac{\partial h}{\partial x_i}(p) dx^i = \eta_i dx^i,$$

then

$$v^p(v_p) = D(h \circ g)(0)(1) = \frac{dg^i}{dt}(0)\frac{\partial h}{\partial x_i}(p) = \gamma^i \eta_i.$$

Let $p \in \mathcal{O}$ and denote by $\mathbb{A}_p^k\mathcal{O}$ the space of $k-$forms on the tangent space $T_p\mathcal{O}$. A *$k-$differential form* or *$k-$form* on $\mathcal{O}$ is a smooth choice of a $k-$linear form in $\mathbb{A}_p^k\mathcal{O}$ for all $p \in \mathcal{O}$. That is, a $k-$form can be written as

$$F = \sum_{1 \le i_1 < \cdots < i_k \le m} f_{i_1, \dots, i_k}(x_1, \dots, x_m) dx^{i_1} \wedge \cdots \wedge dx^{i_k}, \tag{1.12}$$

where the functions $f_{i_1,\dots,i_k} : \mathcal{O} \to \mathbb{R}$ are smooth. Because $\mathbb{A}_p^0\mathcal{O} = \mathbb{R}$, $0-$forms are simply smooth functions, and because $\mathbb{A}_p^1\mathcal{O} = T_p^*\mathcal{O}$, $1-$forms are covector fields. Given a $0-$form F, dF is a $1-$form. The natural generalisation is the *exterior derivative* operator d which converts a $k-$form F as given in (1.12) into a $(k+1)-$form dF by the formula

$$dF = \sum_{j=1} m \sum_{1 \le i_1 < \cdots < i_k \le m} \frac{\partial f_{i_1,\dots,i_k}}{dx^j} dx^j \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}.$$

Lemma 1.3.14. *Let F and G be smooth forms defined on an open set $\mathcal{O}$. Then,*

1. *$d(F + G) = dF + dG$,*
2. *$d(F \wedge G) = dF \wedge G + (-1)^{\deg(F)} F \wedge dG$,*
3. *$d(dF) = 0$, for all F,*
4. *if F is a function, then dF agrees with the standard definition of the differential of F,*

5. *the operator d is uniquely defined by the properties given above.*

A $k-$form F is *closed* if $dF = 0$. A $k-$form is *exact* if there is a $(k-1)-$form G such that $F = dG$. We have the following result. Part 3 of Lemma 1.3.14 says that an exact form is closed. A partial converse is also true.

Theorem 1.3.15 (Poincaré's Lemma)**.** *Let $\mathcal{O}$ be a ball in $\mathbb{R}^m$ and F a $k-$form such that $dF = 0$. Then there is a $(k-1)-$form g on $\mathcal{O}$ such that $F = dg$.*

Remark 1.3.1. Poincaré's Lemma contains two classical theorems. If F is a vector field defined on a ball in $\mathbb{R}^3$ with $\nabla \times F = 0$, then there is a smooth function g such that $F = \nabla g$, and if F is a smooth vector field defined on a ball such that $\nabla \cdot F = 0$, then there is a smooth vector field g such that $F = \nabla \times g$.

1.3.2.4 The Symplectic Form and Darboux's Theorem

Let $\mathcal{O}$ be an open set in $\mathbb{R}^{2n}$. A *symplectic structure* or *symplectic form* on $\mathcal{O}$ is a closed non-degenerate $2-$form. The standard symplectic structure in $\mathbb{R}^{2n}$ is

$$\Omega = \sum_{i=1}^{n} dz^i \wedge dz^{i+n} = \sum_{i=1}^{n} dq^i \wedge dp^i, \tag{1.13}$$

where $z = (z^1, \ldots, z^{2n}) = (q^1, \ldots, q^n, p^1, \ldots, p^n)$ are coordinates in $\mathbb{R}^{2n}$. The coefficient matrix of Ω is just J. By Corollary 1.3.7, there is a linear change of coordinates so that the coefficients matrix of a non-degenerate $2-$form is J at one point. A much more powerful result is the following.

Theorem 1.3.16 (Darboux's Theorem)**.** *If F is a symplectic structure on an open ball in $\mathbb{R}^{2n}$, then there exists a coordinate system z such that F in this coordinate system is the standard symplectic structure Ω.*

For the proof, see [6]. A coordinate system for which a symplectic structure is Ω is called a *symplectic coordinate (for this form)*. A *symplectic transformation* ϕ is one that preserves the form Ω or preserves the coefficients matrix J, that is

$$D\phi^T J D\phi = J.$$

The standard symplectic form given in (1.13) is closed, since $d\Omega = 0$. In fact, it is also exact, because

$$\Omega = d\alpha,$$

where

$$\alpha = \sum_{i=1}^{n} q^i dp^i = qdp.$$

In short, Ω is a closed non-degenerate $2-$form. By Darboux's Theorem, for any closed, non-degenerate $2-$form, there are local coordinates such that in these coordinates, the $2-$form is given by (1.13). This says that J is simply the coefficients matrix of a closed, non-degenerate $2-$form in Darboux coordinates.

The definition of symplectic structure can be generalised to differentiable manifolds, which we include for the sake of completeness.

Definition 1.3.3 (Symplectic manifold)**.** Let M^{2n} be an even-dimensional differentiable manifold. A *symplectic structure* on M^{2n} is a closed differential $2-$form ω^2 on M^{2n} which is non-degenerate, that is, for any nonzero tangent vector $\xi \in T_pM$, there exists another tangent vector $\eta \in T_pM$ such that

$$\omega^2(\xi, \eta) \neq 0.$$

The pair (M^{2n}, ω^2) is called a *symplectic manifold*.

2

Integrability of Hamiltonian Systems

Differential equations, including Hamiltonian equations, are customarily classified into integrable and non-integrable. In this chapter, we give a brief survey of integrability of Hamiltonian systems. We start with the definition of *integrability by quadratures* before studying complete integrability and introducing action-angle coordinates. We then switch to *near-integrable Hamiltonian systems.* In order to fully understand the phenomena arising from these objects, we include a whole section on rational approximation of real numbers, and finally we give an introduction to the techniques of KAM theory, which allows us to understand the dynamics of these systems and which is of greatest importance in Chapter 3. For a deeper insight on the integrability of Hamiltonian systems, see [11].

2.1 Integrable Hamiltonian Systems

2.1.1 Quadratures

The expression *integration by quadratures* of a system of differential equations in $\mathbb{R}^n$ is the search for its solutions by a finite number of "algebraic" operations (including inversion of functions) and "quadratures" (i.e. calculation of integrals of known functions). The following result connects the integration by quadratures of Hamiltonian systems with the existence of a sufficiently rich set of first integrals. Notice that even if this is a local definition, it has global implications.

2.1.1.1 Hamiltonian Systems with One Degree of Freedom

All $1-$degree-of-freedom Hamiltonian systems are integrable in the sense described above. Indeed, consider $H(q,p)$ for $(q,p) \in \mathcal{U} \subset \mathbb{R}^2$. Take $(q_0,p_0) \in \mathcal{U}$. If $\nabla H(q_0,p_0) = 0$, then $q(t) = q_0$ and $p(t) = p_0$, for any $t \in \mathbb{R}$. Otherwise, asume without loss of generality that $\partial_q H(q_0,p_0)$, and call $h = H(q_0,p_0)$. Then, by the Implicit Function Theorem, $H(q,p) = h$ defines a curve containing (q_0, p_0) which can be locally parameterised as $q = Q_h(p)$, for p close to p_0. Then, Hamilton's equations are

$$\begin{cases} \dot{q} = \partial_p H(q,p), \\ \dot{p} = -\partial_q H(Q_h(p), p). \end{cases}$$

The second equation is a one-dimensional separable differential equation. Therefore,

$$-\int_{p_0}^{p} \frac{d\widetilde{p}}{\partial_q H(Q_h(\widetilde{p}), \widetilde{p})} = \int_{t_0}^{t} d\widetilde{t}.$$

The momentum $p(t)$ is then obtained by integrating and inverting a function. The solution to the whole system is given by

$$\begin{cases} p(t), \\ q(t) = Q_h(p(t)). \end{cases}$$

2.1.1.2 Hamiltonian Systems with Two Degrees of Freedom

Let $X_1, X_2 \colon \mathcal{U} \subset \mathbb{R}^{2n} \to \mathbb{R}$. The *Lie bracket* of X_1 and X_2 is the operator defined by the commutator

$$[X_1, X_2](x) = DX_2(x)X_1(x) - DX_1(x)X_2(x).$$

The following lemma will be needed to give the main result.

Lemma 2.1.1. *If two vector fields X_1 and X_2 satisfy that they are linearly independent in $\mathcal{U} \subset \mathbb{R}^2$ and $[X_1, X_2] = \lambda X_1$, for some $\lambda \in \mathbb{R}$, then $\dot{x} = X_1(x)$ can be integrated by quadratures locally in $\mathcal{U}$.*

Proof. The idea is to construct a first integral. Recall that F is a first integral of $\dot{x} = X_1(x)$ if and only if

$$\frac{d}{dt}\left(F(\phi^t_{X_1}(x)\right)\Big|_{t=0} = 0.$$

By the cain rule, this is equivalent to

$$\left(DF(\phi^t_{X_1}(x))\frac{d\phi^t_{X_1}(x)}{dt}\right)\Bigg|_{t=0},$$

that is,

$$DF(x)X_1(x) = 0.$$

Call $X_1 = (a_{11}, a_{12})^T$. We need to solve

$$a_{11}\partial_{x_1}F + a_{12}\partial_{x_2}F = 0. \tag{2.1}$$

We add another equation in order to be able to solve (2.1). Let us choose suitable coefficients a_{21}, a_{22} and a_{23}, and let $A = (a_{ij})^2_{i,j=1}$, so that we obtain the system

$$A\begin{pmatrix} \partial_{x_1}F \\ \partial_{x_2}F \end{pmatrix} = \begin{pmatrix} 0 \\ a_{23} \end{pmatrix}.$$

Assuming that A is regular, we have

$$\begin{pmatrix} \partial_{x_1} F \\ \partial_{x_2} F \end{pmatrix} = b := A^{-1} \begin{pmatrix} 0 \\ a_{23} \end{pmatrix}.$$

By Poincaré's Lemma (see Theorem 1.3.15), we have

$$\partial_{x_2} b_1 = \partial_{x_2} b_2. \tag{2.2}$$

Since F cannot be constant, $\nabla F \neq 0$, so $a_{23} \neq 0$. We can assume, without loss of generality, that $a_{23} = 1$. We now have to choose a_{21} and a_{22} satisfying

$$a_{21} \partial_{x_1} F + a_{22} \partial_{x_2} F = 1$$

such that (2.2) holds. Some calculations show that this is true if and only if

$$\det([X_1, X_2], X_1) = 0,$$

which is satisfied, since by hypothesis we have $[X_1, X_2] = \lambda X_1$, for some $\lambda \in \mathbb{R}$. □

Theorem 2.1.2. *Consider a $\mathcal{C}^2$ Hamiltonian $H = F_1 \colon \mathcal{U} \subset \mathbb{R}^4 \to \mathbb{R}$. Assume that there exists a $\mathcal{C}^2$ function $F_2 \colon \mathcal{U} \to \mathbb{R}$ such that*

1. *there exists $c \in \mathbb{R}$ such that $\{F_1, F_2\} = cF_1$ on $\mathcal{U}$,*
2. *the functions F_1 and F_2 are linearly independent on the manifold given by*
$$M_f := \{x \in \mathcal{U} \mid (F_1(x), F_2(x) = f \subset \mathbb{R}^2\},$$
3. *on M_f, either $c = 0$ or $F_1 = 0$.*

Then, the solutions of $\dot{x} = J\nabla H(x)$ on M_f can be found by quadratures.

Proof. Observe that, since F_1 and F_2 are linearly independent on the manifold M_f, it is moreover a submanifold $M_f \subset \mathcal{U} \subset \mathbb{R}^4$, which is invariant under the flows of F_1 and F_2.

Indeed, let $X_i = J\nabla F_i$, for $i = 1, 2$. Then M_f is invariant under the flows of X_i if and only if $DF_i X_j = 0$, for $i, j = 1, 2$. This is equivalent to having

$$DF_i J \nabla F_j = \{F_i, F_j\} = 0.$$

As a consequence, M_f is invariant under the flows of $X_i = J\nabla F_i$ or, equivalently, X_1 and X_2 are tangent to M_f.

Let us now check the conditions of Lemma 2.1.1. The functions F_1 and F_2 are linearly independent, that is, ∇F_1 and ∇F_2 are linearly independent as vectors, because they are the product of the invertible matrix J by the X_1 and X_2, which are linearly independent. In order to show the second hypothesis, we shall use the equality

$$[J\nabla F_1, J\nabla F_2] = J\nabla\{F_1, F_2\},$$

which can be checked with some tedious computations. Then, we have

$$[X_1, X_2] = J\nabla\{F_1, F_2\} = J\nabla(cF_1) = cJ\nabla F_1 = cX_1.$$

Therefore, we can apply Lemma 2.1.1, so the system $\dot{x} = J\nabla F_1(x) = J\nabla H(x)$, restricted to M_f, can be integrated by quadratures. □

2.1.1.3 Higher-Dimensional Hamiltonian Sytems

Theorem 2.1.3. *Consider a Hamiltonian* $H = F_1 \colon \mathcal{U} \subset \mathbb{R}^{2n} \to \mathbb{R}$*, and assume that the functions* $F_1, F_2, \ldots, F_n \colon \mathcal{U} \to \mathbb{R}$ *are linearly independent functions in involution, that is,* $\{F_i, F_j\} = 0$*, for any* $i, j = 1, 2, \ldots, n$*. Then,*

1. *the sets* $M_f = \{x \in \mathcal{U} \mid F_i(x) = f_i,\ \text{for any } i = 1, 2, \ldots, n\}$ *are invariant* $n-$*dimensional submanifolds under the flow of* $\dot{x} = J\nabla H(x)$*,*
2. *the trajectories of each* M_f *can be computed by quadratures.*

2.1.2 Complete Integrability

Theorem 2.1.4 (Liouville-Arnol'd)**.** *Consider* $F_1, F_2, \ldots, F_n \colon \mathcal{U} \subset \mathbb{R}^{2n} \to \mathbb{R}$ *functions in involutions and functionally independent. Then,*

1. *for any* $c = (c_1, c_2, \ldots, c_n) \in F_1(\mathcal{U}) \times F_2(\mathcal{U}) \times \cdots \times F_n(\mathcal{U})$*, the set*

$$M_c := \{z \in \mathbb{R}^n \colon F_i(z_i) = c_i,\ i = 1, 2, \ldots, n\}$$

   is a submanifold that is invariant by the flow associated to any Hamiltonian defined by F_i*,*

2. *if* M_c *is connected and compact, then* $M_c \simeq \mathbb{T}^n$*, where* $\mathbb{T}^n$ *is the* $n-$*dimensional torus,*
3. *if* M_c *is connected but not compact, but the flows associated to the Hamiltonians* F_i*s are complete, then* $M_c \simeq \mathbb{T}^k \times \mathbb{R}^{n-k}$*, for some* $0 \le k \le n-1$*,*
4. *the flow of any* $x_{F_i} = J\nabla F_i$ *on* M_c *is "of translation type", namely, in suitable coordinates* $\varphi \in \mathbb{T}^k$ *and* $y \in \mathbb{R}^{n-k}$*,*

$$\phi_t(\phi, y) = (\varphi + \omega(c)t \mod 2\pi, y + \nu(c)t),$$

   for some $\omega(c) \in \mathbb{R}^k$ *and* $\nu(c) \in \mathbb{R}^{n-k}$*.*

In the statement of Theorem 2.1.4 we use the letter ω to refer to a real number. See the List of Symbols for other uses of the same terminology.

Hamiltonian systems satisfying Liouville-Arnol'd's Theorem are called *completely integrable.* The most interesting case is when $M_c \simeq \mathbb{T}^n$,where

$$\mathbb{T}^n = \{(\varphi_1, \varphi_2, \ldots, \varphi_n) \colon \varphi_i = \mathbb{R}/2\pi\mathbb{Z}\} \quad \text{and} \quad \varphi_i = \varphi_i^0 + t\omega_i.$$

Definition 2.1.1 (Quasiperiodic motion)**.** If $M_c \simeq \mathbb{T}^n$ in Liouville-Arnol'd's Theorem, we say that the motion is *quasiperiodic* with frequency vector $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$.

Definition 2.1.2. Using the notation above, we say that ω is

- *non-resonant* or *rationally-independent* if $\omega \cdot k \neq 0$, for any $k \in \mathbb{Z}^n \setminus \{0\}$,
- *resonant* or *rationally-dependent* if there exists $k^* \in \mathbb{Z}^n \setminus \{0\}$ such that $\omega \cdot k^* = 0$.

In the first case, the orbits in M_c are dense for any $n \in \mathbb{N}$.

Remark 2.1.1. The theorem is not constructive, so the frequencies $\omega(c)$ and $\nu(c)$ are unkown.

Example 2.1.1 ($M_c \simeq \mathbb{T}^2$)**.**
In the non-resonant case, we have $\omega_1 k_1 + \omega_2 k_2 \neq 0$, for any $k_1, k_2 \in \mathbb{Z} \setminus \{0\}$. Equivalently,

$$\frac{\omega_1}{\omega_2} \neq -\frac{k_1}{k_2},$$

that is to say $\omega_1/\omega_2 \in \mathbb{R} \setminus \mathbb{Q}$. On the other hand, if ω is resonant, then all orbits are periodic.

Example 2.1.2 ($M_c \simeq \mathbb{T}^3$)**.**
Suppose $\omega = (\sqrt{2}, \sqrt{3}, 0)$. Then, the motion in $\mathbb{T}^3$ is given by

$$\varphi_1 = \varphi_1^0 + \sqrt{2}t, \quad \varphi_2 = \varphi_2^0 + \sqrt{3}t, \quad \text{and} \quad \varphi_3 = \varphi_3^0.$$

Therefore,

$$\mathbb{T}^3 = \sqcup_{\varphi_3 \in \mathbb{T}^1} \{(\varphi_1, \varphi_2, \varphi_3) \in \mathbb{T}^3 \colon (\varphi_1, \varphi_2) \in \mathbb{T}^2\}.$$

In non-resonant tori, we can say even more. Orbits are not only dense, but they also "spread uniformly".

Theorem 2.1.5 (H. Weyl)**.** *Let $f \colon \mathbb{T}^n \to \mathbb{C}$ be a Riemann-integrable map and $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$ be non-resonant. Then, for any $\varphi_0 \in \mathbb{T}^n$, the* time average *of the map f*

$$f^*(\varphi) := \lim_{T \to +\infty} \frac{1}{T} \int_0^T f(\varphi_0 + t\omega)\, dt$$

exists. Moreover, it coincides with the spatial average *of f*

$$\bar{f} := \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} f(\varphi)\, d\varphi.$$

Namely, for any $\varphi_0 \in \mathbb{T}^n$,

$$f^*(\varphi_0) = \bar{f}.$$

Corollary 2.1.6. *Let $D \subset \mathbb{T}^n$ be a rectangle and let $f = \chi_D$. Then,*

$$\lim_{T \to +\infty} \frac{\mu(t \in [0, T] \colon \varphi_0 + t\omega \in D)}{T} = \frac{\mu(D)}{(2\pi)^n}.$$

Remark 2.1.2. Corollary 2.1.6 is also valid for any $D \subset \mathbb{T}^n$ such that its characteristic function χ_D is Riemann-integrable.

2.1.2.1 Action-Angle Coordinates

Liouville-Arnol'd's Theorem (see Theorem 2.1.4) gives a coordinates system to study completely integrable systems. When M_c is connected and compact, these are called *action-angle coordinates*. More precisely,

1. there exists a neighbourhood of M_c which is diffeomorphic to $\mathbb{T}^n \times D$, where $D \subset \mathbb{R}^n$ is a small open ball,

2. on $\mathbb{T}^n \times D$ we can define new coordinates (by means of a symplectic transformation) $\varphi \in \mathbb{T}^n$ (angles) and $I \subset D$ (actions) such that the Is only depend on the first integrals $F_1, F_2, \ldots, F_n$. As a consequence, these Is are also first integrals. Indeed, the Hamiltonian system is given by

$$\begin{cases} \dot{\varphi} = \frac{\partial H}{\partial I}(\varphi, I), \\ \dot{I} = -\frac{\partial H}{\partial \varphi}(\varphi, I) = 0, \end{cases}$$

so H does not depend on φ and so $H(\varphi, I) = H(I)$.

Therefore, a completely integrable Hamiltonian system in action-angle coordinates has a $\varphi-$independent Hamiltonian. The motion in $\mathbb{T}^n \times D$ is given by

$$\begin{cases} \varphi = \varphi_0 + \partial_I H(I_0)t, \\ I = I_0 \end{cases}$$

for $(\varphi, I) \in \mathbb{T}^n \times D \subset \mathbb{T}^n \times \mathbb{R}^n$. Each set

$$\mathbb{T}^n_{I_0} := \{I = I_0\}$$

is an $n-$dimensional torus whose dynamics is quasi-periodic with frequencies $\omega(I_0) = \partial_I H(I_0)$.

2.2 Near-integrable Hamiltonian Systems

We now aim to study what happens when we slightly modify a completely integrable Hamiltonian System. Let $H_0 \colon \mathcal{U} \to \mathbb{R}^{2n} \to \mathbb{R}$ be the corresponding Hamiltonian. Consider

$$H(\theta, I) = H_0(I) + \varepsilon H_1(\theta, I), \tag{2.3}$$

for $\varepsilon \ll 1$, $\theta \in \mathbb{T}^n$ and $I \in V \subset \mathbb{R}^n$. The equations of motion are

$$\begin{cases} \dot{\theta} = \partial_I H_0(I) + \varepsilon \partial_I H_1(\theta, I), \\ \dot{I} = -\varepsilon \partial_\theta H_1(\theta, I). \end{cases}$$

For $\varepsilon = 0$, the actions are constants of motion and orbits are confined to $\mathbb{T}^n = \{I = I_0\}$. For $\varepsilon > 0$, we have $\dot{I} \neq 0$. In fact, $\dot{I} = \varepsilon$ (i.e. the actions change very slowly), so the angles move much faster than the actions.

In the light of this observation, one may consider several question. When we let $0 < \varepsilon \ll 1$, are there still some invariant $n-$dimensional tori, even if slightly deformed? In the literature, such tori are said to *persist* under the perturbation. In principle, even if some tori persist, we expect some others to disappear, for otherwise the whole system would still be integrable. The next thing we ask ourselves is which of these tori $I = I_0$ persist and which break down, and what is the new dynamics that they originate.

When $\varepsilon = 0$, the phase space is foliated by $n-$dimensional tori whose dynamics is given by

$$\theta = \theta_0 + \omega t.$$

The tori are called *resonant* if there exists $k \in \mathbb{Z}^n \setminus \{0\}$ such that $\omega \cdot k = 0$, and orbits are dense in tori of lower dimension. In the case that $\omega \cdot k \neq 0$, for any $k \in \mathbb{Z}^n \setminus \{0\}$, the tori are called *non-resonant*. In this setting, orbits are dense.

For $\varepsilon > 0$, some tori persist and others disappear. The main tool to study when each case occurs is *KAM Theory*, which we discuss in Section 2.2.2. The remainder of this section is devoted to studying Diophantine approximation.

2.2.1 Diophantine Approximation of Real Numbers

This section deals with the approximation of real numbers by irrational numbers. Let us begin with the following result.

Theorem 2.2.1 (Dirichlet's Approximation Theorem)**.** *Let $\omega \in \mathbb{R}$ and $N \in \mathbb{N}$. Then, there exist $p, q \in \mathbb{Z}$, where $1 \leq q \leq N$ such that*

$$|q \cdot \omega - p| < \frac{1}{N}. \tag{2.4}$$

As a consequence, the inequality

$$\left|\omega - \frac{p}{q}\right| < \frac{1}{q^2}$$

is satisfied by infinitely many $p, q \in \mathbb{Z}$.

Proof. We need to find $p, q \in \mathbb{Z}$ such that (2.4) is satisfied. We consider the real numbers

$$q \cdot \omega \mod 1, \qquad \text{for } q = 0, 1, \ldots, N,$$

so we have $N+1$ numbers in the interval $[0, 1)$. Let us divide it into N intervals as

$$I_k = \left[\frac{k}{N}, \frac{k+1}{N}\right), \qquad \text{for } k = 0, 1, \ldots, N-1.$$

By the Pigeonhole's Principle, there is one I_k containing two numbers, say

$$q_1\omega \mod 1 \qquad \text{and} \qquad q_2\omega \mod 1.$$

That is, there exist $k^*, q_1, q_2, p_1, p_2 \in \mathbb{Z}$ such that

$$q_1\omega + p_1, q_2\omega + p_2 \in I_{k^*}.$$

We can assume without loss of generality that $q_2 \geq q_1$. Then,

$$|(q_2\omega - p_2) - (q_1\omega - p_1)| < \frac{1}{N}.$$

Therefore, taking $q = q_2 - q_1$ and $p = p_1 - p_2$, we obtain

$$|q\omega - p| < \frac{1}{N} < \frac{1}{q},$$

since $q < N$. This is equivalent to

$$\left|\omega - \frac{p}{q}\right| < \frac{1}{q^2},$$

which has infinitely many solutions. □

Definition 2.2.1 (Diophantine number)**.** A number $\omega \in \mathbb{R}$ is *Diophantine of class* (τ, γ) or *belongs to* $\mathcal{D}(\tau, \gamma)$ there exist $\gamma > 0$ and $\tau \geq 0$ such that

$$\left|\omega - \frac{p}{q}\right| \geq \frac{\gamma}{q^{2+\tau}}$$

for any $p, q \in \mathbb{Z}$, where $q \neq 0$.

The set of all Diophantine numbers is often denoted by

$$\mathcal{D} = \cup_{\substack{\gamma>0 \\ \tau\geq 0}} \mathcal{D}(\tau, \gamma)$$

and is often referred to as the set of "badly approximable" numbers by rationals.

The intuition about what this set looks like is not straightforward. The first thing to check is that $\mathcal{D}$ is not empty. Indeed, the golden number

$$\frac{1+\sqrt{5}}{2}$$

belongs to $\mathcal{D}(\tau, \gamma)$ with $\tau \geq 0$ and $\gamma \geq 1/\sqrt{5}$. The following result gives a property on the measure of $\mathcal{D}$.

Lemma 2.2.2. *Let* $\tau > 0$ *and* $\gamma \geq 0$*. Then, there exists a function* $m(\tau)$ *such that*

$$\mu(D_{\tau,\gamma} \cap [0, 1]) \geq 1 - m(\tau)\gamma.$$

Proof. We have

$$\begin{aligned}
\mu([0,1] \setminus \mathcal{D}_{\tau,\gamma}) &= \mu\left(\bigcup_{\substack{q\in\mathbb{N} \\ 0<p\leq q}} \left\{\omega \in [0,1] \colon \left|\omega - \frac{p}{q}\right| < \frac{\gamma}{|q|^{2+\tau}}\right\}\right) \\
&= \sum_{\substack{q\in\mathbb{N} \\ 0<p\leq q}} \mu\left(\left\{\omega \in [0,1] \colon \left|\omega - \frac{p}{q}\right| < \frac{\gamma}{|q|^{2+\tau}}\right\}\right) \\
&= \sum_{\substack{q\in\mathbb{N} \\ 0<p\leq q}} \frac{2\gamma}{|q|^{2+\tau}} \\
&= \sum_{q=1}^{\infty}\sum_{p=1}^{\infty} \frac{2\gamma}{|q|^{2+\tau}} \\
&= \gamma \sum_{q=1}^{\infty} \frac{2q}{|q|^{2+\tau}}.
\end{aligned}$$

Therefore, by defining

$$m(\tau) := \sum_{q=1}^{\infty} \frac{2q}{|q|^{2+\tau}}$$

and taking into account that for $\tau > 0$, the series converges, we get the desired result. □

Notice that for $\tau = 0$, the series diverges, and $\mathcal{D}_{0,\gamma}$ has zero measure even if it is non-empty.

From a topological point of view, the set $\mathcal{D}$ is a *Cantor set.* The following proposition gives the necessary properties.

Proposition 2.2.3 ($\mathcal{D}$ is a Cantor set)**.** *Let $\gamma > 0$ and $\tau \geq 0$. Then,*

1. $\mathcal{D}_{\tau,\gamma}$ *is* nowhere dense*, that is,* $\overline{\mathcal{D}_{\tau,\gamma}}$ *has empty interior;*
2. $\mathcal{D}_{\tau,\gamma}$ *and* $\mathcal{D}$ *are* perfect sets*, that is, they are closed and have no isolated points;*
3. $\mathcal{D}_{\tau,\gamma}$ *is* totally disconnected*, that is, all of its connected components are points.*

The complement of $\mathcal{D}$ also has interesting properties. We say that a number $\omega \in \mathbb{R} \setminus \mathbb{Q}$ is a *Liouville number* if for any $n \in \mathbb{N}$, there exist $p, q \in \mathbb{Z}$, where $q > 1$ such that

$$0 < \left|\omega - \frac{p}{q}\right| < \frac{1}{q^n}.$$

Therefore, Liouville numbers are very well approximated by rational numbers. As before, we this set is non-empty, since it contains at least the so-called *Liouville number*

$$\sum_{n=1}^{\infty} 10^{-n!}.$$

Moreover, the set of Liouville numbers has zero measure, and from the point of view of Topology, it is a $G_\delta-$*set*, that is, a countable intersection of open dense sets.

So far, we have studied the rational approximation of real numbers. This can be generalised to higher dimensions.

Definition 2.2.2 (Diophantine vector)**.** A vector $\omega \in \mathbb{R}^n$ is *Diophantine of class* (τ, γ) or *belongs to* $\mathcal{D}_{\tau,\gamma}$ if there exist $\gamma > 0$ and $\tau \geq 0$ such that

$$|k \cdot \omega| \geq \frac{\gamma}{\|k\|^{n-1+\tau}},$$

for any $k \in \mathbb{Z}^n \setminus \{0\}$.

An analogous result to Lemma 2.2.2 is valid for Diophantine vectors.

Lemma 2.2.4. *For $\tau > 0$, there exists a function $m(\tau, n)$ such that*

$$\mu\left(\mathcal{D}_{\tau,\gamma} \cap [0,1]^n\right) \geq 1 - m(\tau, n)\gamma.$$

An immediate consequence of this lemma is the following result.

Proposition 2.2.5. *The set $\mathcal{D} = \cup_{\substack{\gamma>0 \\ \tau\geq 0}} \mathcal{D}_{\tau,\gamma}$ has full measure.*

2.2.2 KAM Theory

In nature one often encounters systems that differ from integrable ones by small perturbations. Thus, for example, the problem of the motion of planets around the Sun can be regarded as a perturbation of the integrable problem of the motion of non-interacting point masses around a fixed centre of attraction. Methods for studying such systems are grouped in what is known as *perturbation theory*. These methods often enable us to describe the perturbed motion almost as completely as the unperturbed motion. The justification of various methods in perturbation theory is rather difficult.

The *Kolmogorov-Arnol'd-Moser Theory* (or simply *KAM Theory*) is a perturbation theory for non-resonant quasiperiodic motions (see Definition 2.1.2) of Hamiltonian and related systems that works over infinite time intervals.

2.2.2.1 Invariant Tori of the Perturbed System

Let us begin this section by recalling some of the basic notions underlying integrable systems. Consider an unperturbed integrable Hamiltonian system with Hamiltonian $H_0(I)$. Its phase space is foliated by invariant tori $I = \text{const}$. The motion on each torus is conditionally-periodic with frequency vector $\omega(I) = \partial_I H_0$. A torus on which the frequencies are rationally-independent is said to be *non-resonant*. Each phase trajectory on such a torus fills it densely (and is called a *winding* of the torus). The remaining tori $I = \text{const}$ are said to be *resonant*. They are foliated by tori of smaller dimension. The unperturbed system is said to be *non-degenerate* if the frequencies are functionally independent, that is

$$\det(\partial_I \omega) = \det(\partial_I^2 H_0) \neq 0.$$

In a non-degenerate system, the non-resonant tori form a dense set of full measure. The resonant tori forma set of measure zero which, however, is also dense. Moreover, the sets of resonant tori with any number of independent frequencies from 1 to $n-1$ are each dense; in particular, the set of tori on which all phase trajectories are closed is dense.

In order to state the next result, we need to introduce some notation. Let $\sigma > 0$ be fixed and define

$$\mathbb{T}^n_\sigma := \{\theta \in \mathbb{C}^n/(2\pi\mathbb{Z})^n \mid |\mathrm{Im}(\theta)| \leq \sigma\}$$

and

$$\mathcal{U}_\sigma := \{I \in \mathbb{C}^n \mid \mathrm{Re}(I) \in \mathcal{U},\ |\mathrm{Im}(\theta)| \leq \sigma\}.$$

The norm of a function f in the product space $\mathbb{T}^n_\sigma \times \mathcal{U}_\sigma$ is defined as

$$\|f\|_{\mathbb{T}^n_\sigma \times \mathcal{U}_\sigma} := \sup_{(\theta,I)\in\mathbb{T}^n_\sigma \times \mathcal{U}_\sigma} |f(\theta, I)|\,.$$

Sometimes, we also work in the restricted spaces $\mathbb{T}^n_{\sigma-\rho}$ and $\mathcal{U}_{\sigma-\rho}$, for which the above definitions are generalised in the obvious way $\rho \in (0, \sigma)$. Consider now the perturbed system with Hamiltonian

$$H(I, \theta, \varepsilon) = H_0(I) + \varepsilon H_1(I, \theta, \varepsilon),$$

where $\theta \in \mathbb{T}^n$, $I \in \mathcal{U} \subset \mathbb{R}^n$, which is analytic in $\mathbb{T}^n_\sigma \times \mathcal{U}_\sigma \subset \mathbb{C}^n \times \mathbb{C}^n$ for some $\sigma > 0$. Assume that $\partial_I H_0(I) = \omega$, for $\omega \in \mathcal{D}_{\tau,\gamma}$, where $\gamma > 0$ and $\tau \geq 0$.

The subsequent theorem describes the fate of non-resonant tori under perturbation (see [7]).

Theorem 2.2.6 (Kolmogorov, 1954). *With the assumptions above, if the unperturbed system is non-degenerate, then there exists $\rho \in (0, \sigma)$ and a symplectic transformation*

$$\begin{array}{rrcl} \Phi : & \mathbb{T}^n_{\sigma-\rho} \times \mathcal{U}_{\sigma-\rho} & \longrightarrow & \mathbb{T}^n_\sigma \times \mathcal{U}_\sigma \\ & (r, \phi) & \longmapsto & \Phi(r, \phi) = (I, \theta) \end{array}$$

such that

$$H \circ \Phi(r, \phi) = a + \omega r + P(r, \phi, \varepsilon),$$

where $P(0, 0, \varepsilon) = 0$ and $\partial_r P(0, \phi, \varepsilon) = 0$.

As an immediate consequence of this theorem, we have that Hamilton's equations can now be written as

$$\begin{cases} \dot{\phi} = \omega + \partial_r P(r, \phi, \varepsilon), \\ \dot{r} = -\partial_\phi P(r, \phi, \varepsilon). \end{cases}$$

Therefore, if $r = 0$, then $\partial_r P(0, \phi, \varepsilon) = 0$, so that the $n-$dimensional torus $\{r = 0\}$ is invariant. Moreover, we have $\dot{\phi} = \omega$. In other words, after the change of coordinates given by Φ, the motion on the invariant torus $\{r = 0\}$ is the same as the one on the unperturbed torus, namely a rotation of frequency ω.

Remark 2.2.1. In the norm of the supremum, the transformation Φ is $\varepsilon-$*close to the identity*, that is

$$\|\Phi - \mathrm{Id}\,\|_{\mathbb{T}^n_{\sigma-\rho} \times \mathcal{U}_{\sigma-\rho}} = \mathcal{O}(\varepsilon).$$

Theorem 2.2.6 gives a first approach to how many of the tori that are present in the unperturbed system persist. Indeed, since the result holds for any $\omega \in \mathcal{D}_{\tau,\gamma}$ and

$$\frac{\mu(\mathcal{D}_{\tau,\gamma} \cap [0,1]^n)}{\mu([0,1]^n)} - 1 = \mathcal{O}(\gamma),$$

the union of sets $S_{\tau,\gamma}$ of persistent tori with Diophantine frequency $\omega \in \mathcal{D}_{\tau,\gamma}$ satisfies

$$\frac{\mu\left(\cup_{\substack{\gamma>0 \\ \tau \geq 0}} S_{\tau,\gamma}\right)}{\mu(\mathbb{T}_\sigma \times \mathcal{U}_\sigma)} - 1 = \mathcal{O}(\gamma).$$

This means that for γ small, every point except for a set of measure $\mathcal{O}(\gamma)$ belongs to one of the invariant tori given by Kolmogorov's Theorem.

In the sixties, Arnol'd proved that the measure of the set of points not belonging to invariant tori is of order $\mathcal{O}(\sqrt{\varepsilon})$. Around the same time, Moser gave an improved statement asking for H to be only $\mathcal{D}^k$, with $k > 2n$.

So far we have described the persistence of tori when we let $\varepsilon > 0$. The stability given by KAM Theory shows that, under suitable regularity and non-degeneracy assumptions, most (in the sense of measure theory) of the tori above persist (even if slightly deformed) under small Hamiltonian perturbations. In other words, there are a lot of bounded (stable) motions close to the unperturbed ones. We now wish to understand the behaviour of the rest of orbits of the system, that is in the complement of the set of persistent tori. We shall see in the subsequent pages that the number of degrees of freedom plays a crucial role in the stability of these solutions.

Systems with One Degree of Freedom If $n = 1$, the perturbed Hamiltonian system is also integrable, although not in the classical sense of Liouville-Arnol'd, for we don't have global action-angle variables. If we consider

$$H(I, \phi) = H_0(I) + \varepsilon H_1(I, \phi),$$

with $(I, \phi) \in \mathbb{R} \times \mathbb{T}$, then H is a first integral. In the lever curves $H(I, \phi) = h$ we see the structure of KAM Theorem.

Example 2.2.1. Consider the Hamiltonian for the pendulum given by

$$H(I, \phi) = \frac{1}{2}I^2 + \varepsilon V(\phi) = \frac{1}{2}I^2 + \varepsilon(\cos\phi - 1).$$

The equations that describe the dynamical system are

$$\begin{cases} \dot{\phi} = I, \\ \dot{I} = -\varepsilon V'(\phi) = \varepsilon \sin\phi. \end{cases}$$

The level curves have the equation (see Figure 2.1)

$$\frac{1}{2}I^2 + \varepsilon(\cos\phi - 1) = h.$$

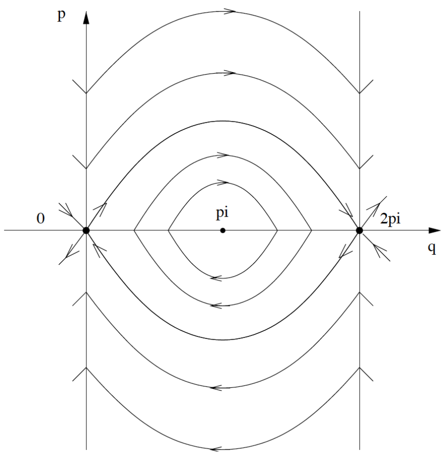

Figure 2.1: Phase portrait of the orbits of an ordinary pendulum.

Therefore, when $h > 0$ is far from zero, the level curves of H are close to the lever curves of H_0. Indeed, they have the form

$$H = I_0 + \mathcal{O}(\varepsilon),$$

with $I_0 = \sqrt{2h}$. The point $(0,0)$ is a saddle and its stable and unstable manifolds coincide along the *separatrix* (which is a homoclinic orbit)

$$H(I,\phi) = \frac{1}{2}I^2 + \varepsilon(\cos\phi - 1) = 0.$$

We observe that the tori close to $I = 0$ have disappeared. The energy level $h = 0$ contains an equilibrium point of saddle type and its separatrices (i.e. the stable and unstable manifolds, which coincide due to the integrability of H).

When $h < 0$ inside the separatrices of the saddle we have tori of different topology (contractible to a point). Summing up, there is stability in the case of one degree of freedom.

Systems with One and a Half Degrees of Freedom We now consider

$$H(I,\phi,t;\varepsilon) = H_0(I) + \varepsilon H_1(I,\phi,t;\varepsilon),$$

with $(I,\phi,t) \in \mathbb{R}\times\mathbb{T}^2$. We add time as a variable, obtaining the equations

$$\begin{cases} \dot{\phi} = \dfrac{\partial H}{\partial I}(I,\phi,s;\varepsilon), \\ \dot{I} = -\dfrac{\partial H}{\partial \phi}(I,\phi,s;\varepsilon), \\ \dot{s} = 1 \end{cases}$$

in the 3−dimensional extended phase space $(I,\phi,s) \in \mathbb{R}\times\mathbb{T}^2$. Denote the flow as $\phi(t,I,\phi,s;\varepsilon)$ (so that $\widetilde{\phi}(0,I,\phi,s;\varepsilon) = (I,\phi,s)$).

For $\varepsilon = 0$, the Hamiltonian is $H_0(I)$ and the system is integrable. The 3−dimensional space is foliated by 2−dimensional tori

$$\mathcal{T}_{I_0} = \{(I,\phi,s) \in \mathbb{R}\times\mathbb{T}^2 \mid I = I_0\}$$

and the flow in $\mathcal{T}_{I_0}$ is a rotation with frequency $\widetilde{\omega}_0(I_0) = (\omega_0(I_0),1)$, where $\omega_0(I_0) = \nabla H_0(I_0)$, so that

$$\{\widetilde{\phi}(t,I,\phi,s;0) \mid (I,\phi,s) \in \mathcal{T}_{I_0}\} = (I_0, \phi + \omega(I_0)t, s+t).$$

When $\varepsilon > 0$, KAM Theory gives tori $\mathcal{T}_{I_0,\varepsilon}$ close to $\mathcal{T}_{I_0}$ for the Hamiltonian system. We know that the invariant tori cover the whole space $\mathbb{R}\times\mathbb{T}^2$ except for a set of measure $\sqrt{\varepsilon}$, and that the invariant curves cover the whole space $\mathbb{R}\times\mathbb{T}^2$ except for a set of measure $\sqrt{\varepsilon}$. There appears, however, a phenomenon known as *splitting of separatrices*, which means that the in the zones between KAM tori do not coincide. This phenomenon makes the system look chaotic, although only locally. In fact, the tori (curves) are barriers to the unstable motion (see Figure 2.2).

Higher-Dimensional Systems The previous argument does not work for perturbations of Hamiltonian systems with more than two degrees of freedom (neither for periodic perturbations of systems of two degrees of freedom). Suppose now that $\theta \in \mathbb{T}^n$ and $I \in \mathcal{U} \subset \mathbb{R}^n$. Then, the energy level has dimension

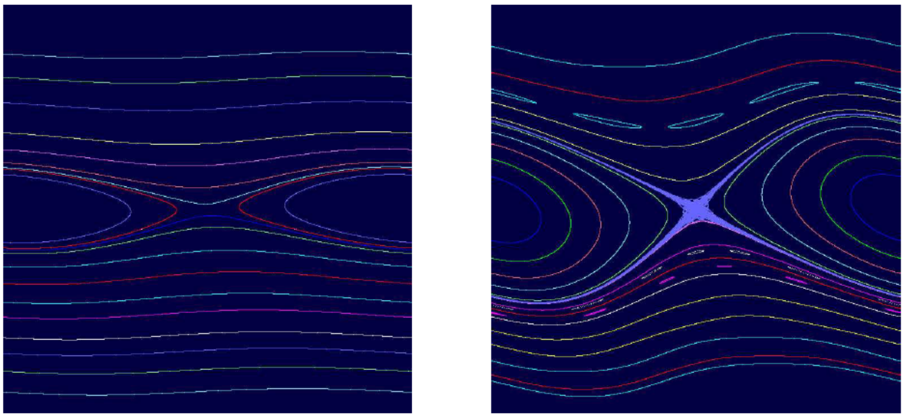

Figure 2.2: The perturbed system (on the right) shows how some of the orbits of the unperturbed one (on the left) are preserved, even if deformed, while others disappear. In the latter case, the KAM curves (i.e. persistent tori) are barriers to unstable motion, although there is chaos locally due to the splitting of separatrices.

$2n-1$ and the invariant tori have dimension n, so in principle, orbits may drift in actions. That is, there can be orbits $(\theta(t), I(t))$ such that there exists $T>0$ such that

$$|I(T)-I(0)|>1.$$

This drift in actions in usually called *Arnol'd diffusion*, and we devote Chapter 3 to giving a well-known example of this phenomenon in a Hamiltonian system with two and a half degrees of freedom.

Diffusion of Slow Variables and its Exponential Estimate In generic systems, the average velocity of the diffusion is exponentially small. Consider a Hamiltonian

$$H(\theta, I)=H_0(I)+\varepsilon H_1(\theta, I) \tag{2.5}$$

that is analytic in $\theta \in \mathbb{T}^n$ and $I \in \mathcal{U} \subset \mathbb{R}^n$. Assume that H_0 is strictly convex, that is there exists $M>0$ such that

$$\left|v^T \partial_I^2 H_0(I) v\right| \geq M\|v\|^2,$$

for any $v \in \mathbb{R}^n$.

Remark 2.2.2. Notice that strict convexity implies that the map $I \mapsto \omega(I)=\partial_I H_0(I)$ is a global diffeomorphism.

Theorem 2.2.7 (Nekhoroshev, 1977)**.** *With the above hypotheses, there exists $\varepsilon_0>0$ such that for any $\varepsilon \in (0, \varepsilon_0)$, any orbit $(\theta(t), I(t))$ of (2.5) satisfies*

$$|I(t)-I(0)| \leq C_1 \varepsilon^{\frac{1}{2n}}, \qquad \textit{for any } |t| \leq C_2 e^{C_3 / \sqrt[2n]{\varepsilon}}$$

where C_1, C_2 and C_3 are positive ε−independent constants.

3

Arnol'd Diffusion

In Chapter 2, we have seen how to deal with the persistence of quasiperiodic motions (see Definition 2.1.1) under small perturbations of conservative dynamical systems. Indeed, integrable Hamiltonian systems have quasiperiodic trajectories of the coordinates. If such a system is subjected to a weak nonlinear perturbation, some invariant tori are deformed and survive. They meet the non-resonance condition, that is they have "sufficiently irrational" frequencies (see Definition 2.1.2). However, this non-resonance condition becomes increasingly difficult to satisfy for systems with more degrees of freedom. In fact, we know that in this case, other tori are destroyed, allowing for the actions to drift indefinitely, thus giving rise to what is know as *Arnol'd Diffusion* (see Section 2.2.2.1).

The purpose of this chapter is to give the famous example that Arnol'd himself provided to show that diffusion occurs in higher-dimensional near-integrable systems. One of the reasons why Arnol'd's original paper is unusual is his employing a slightly special vocabulary in which, for instance, the usual terms *stable* and *unstable* manifolds are replaced by *arriving* and *departing* whiskers. A torus with an arriving and a departing whisker is then referred to as a *whiskered torus*. We shall try to use the more modern nomenclature for the stable and unstable manifolds, which is the most accepted one in Dynamical Systems, but we keep the expression "whiskered torus" in order to avoid the longer option "torus with stable and unstable manifolds". The original exposition as Arnol'd presented it can be found in [2]. Another reference is [3].

3.1 The Intuition Behind Arnol'd's Model

The particular example that Arnol'd conceived was a $5-$degrees-of-freedom Hamiltonian that presented instability, namely he proved the following

Theorem 3.1.1 (Arnol'd, 1964). *Let be a Hamiltonian system described by*

$$H(I_1, I_2, \phi_1, \phi_2, t; \varepsilon, \mu) = H_0(I_1, I_2) + H_1(\phi_1, \phi_2; t), \quad \text{for } \varepsilon, \mu > 0, \qquad (3.1)$$

where

$$H_0(I_1, I_2) = \frac{1}{2}(I_1^2 + I_2^2), \quad \textit{and} \quad H_1(\phi_1, \phi_2, t; \varepsilon, \mu) = \varepsilon(\cos\phi_1 - 1)(1 + \mu(\sin\phi_2 + \cos t)).$$

Then, there exist orbits of Hamilton's equations satisfying

$$|I(T) - I(0)| > 1.$$

The Hamiltonian (3.1) has two degrees of freedom and is $2\pi-$periodic in time. The system of differential equations that it defines is

$$\begin{cases} \dot{\phi}_1 = I_1, \\ \dot{\phi}_2 = I_2, \\ \dot{I}_1 = \varepsilon \sin\phi_1(1 + \mu(\sin\phi_2 + \cos t)), \\ \dot{I}_2 = \varepsilon(1 - \cos\phi_1)\mu\cos\phi_2. \end{cases} \tag{3.2}$$

This system can be made autonomous, following the procedure outlined in Section 1.1 Indeed, in order to get rid of the time t, we introduce the variables $s := t$ and I_3, and consider the new Hamiltonian

$$K := H + I_3,$$

whose equations for the motion are

$$\begin{cases} \dot{\phi}_1 = I_1, \\ \dot{\phi}_2 = I_2, \\ \dot{s} = 1, \\ \dot{I}_1 = \varepsilon \sin\phi_1(1 + \mu(\sin\phi_2 + \cos t)), \\ \dot{I}_2 = \varepsilon(1 - \cos\phi_1)\mu\cos\phi_2, \\ \dot{I}_3 = -\partial_s K. \end{cases} \tag{3.3}$$

In order to study the orbits, we restrict to the energy level $\{K = h\}$. Therefore, the new action I_3 is determined by

$$I_3 = h - H(I_1, I_2, \phi_1, \phi_2, s)$$

and hence it is irrelevant in the new dynamics. As a consequence, if we denote by $\mathbb{T}^k$ the $k-$*dimensional torus*, that is the direct product of k circumferences, which can be described by the k *angular coordinates* $(\phi_1, \phi_2, \ldots, \phi_k) \mod 2\pi$, then the phase space of System (3.3) is the $5-$dimensional direct product of $\mathbb{R}^2$ with the three-dimensional torus $\mathbb{T}^3$.

The best way to understand how instability is manifested in this system is probably to look at a simple diagram. Notice that any such figure must be very schematic, since three dimensions simply are not enough to show everything that is going on. Figure 3.1 shows the phase space of System (3.2) in three cases of increasing complexity. The last case shows the fully developed *transition chain* (see Definition 3.3.3) of Arnol'd's mechanism, in which the unstable manifold of one *whiskered torus* intersects the stable manifold of another *whiskered torus*.

The unperturbed system is very simple, with a phase space entirely foliated by $2-$dimensional invariant tori. We then use ε and μ to perturb the system in

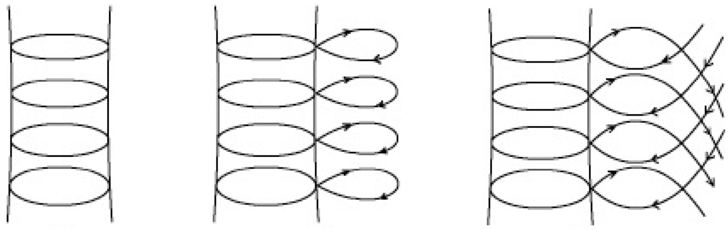

Figure 3.1: Schematic diagram of the invariant manifolds in Arnol'd's model in three cases. On the left, $\varepsilon = \mu = 0$; on the centre, $\varepsilon > 0$ and $\mu = 0$, and on the right, $\varepsilon > 0$ and $\mu > 0$.

two stages. First, we see that taking $\varepsilon > 0$ "switches on" the hyperbolicity of the system, that is to say, the tori develop homoclinic orbits (coincident stable and unstable whiskers). Finally, when we take both $\varepsilon > 0$ and $\mu > 0$, the homoclinic orbits break into separate stable and unstable invariant manifolds that intersect each other transversely (therefore breaking the integrability of the system). The truly interesting point is that not only do the stable and unstable manifolds from the same torus intersect each other, but they also intersect invariant manifold attached to nearby tori (see Theorem 3.3.3). This is the crucial fact that allows us to establish the existence of a *transition chain*, that is a set of invariant tori spread over a large expanse of phase space and linked by transverse intersection of their respective unstable and stable invariant manifolds (see Definition 3.3.3). Very near the chain are guiding channels through which unstable orbits of the system may travel large distances along resonances and through the thicket of invariant tori. This behaviour is often referred to as *shadowing*: the unstable orbits shadow orbits in one stable manifold and begin shadowing those in another as they move along the transition chain (see Lemma 3.3.8).

To be more explicit about the flow along the transition chain, consider System (3.2) with $\varepsilon > 0$ and $\mu > 0$ of the right magnitudes so that a transition chain is established (see Figure 3.2). We first imagine an orbit starting very close to an invariant torus of the chain. At first swept along by the flow on the torus, our orbit winds around the torus for a time until it is "picked up" by the flow of the unstable invariant manifolds emanating from the torus. The orbit then shadows an orbit in this unstable manifold until it reaches the vicinity of a point of intersection of this manifold with the stable manifolds of a second invariant torus. The orbit then switches to the stable manifold, following one of its orbits until it is drawn almost to the surface of the second torus. It then begins to wind around the second torus, and the process just described –the flow along one link of the chain– is repeated. Moving from link to link along the chain, our orbit may travel long distances from its initial location.

The process of switching from one invariant structure to the next may seem mysterious, and indeed, the precise way it occurs is highly sensitive to the orbit's initial conditions. Yet once the transition chain's existence is established, we know that such orbits must exist as a consequence of the Lambda Lemma (see Theorem 3.3.7). That the switching (precisely where and when it occurs) is so sensitive to initial conditions makes it seem like a random process, and soon after Arnol'd proposed this model, the instability in it was conjectured to be

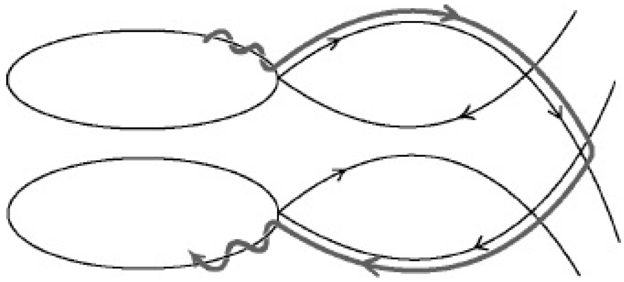

Figure 3.2: Schematic diagram of a transition chain in Arnol'd's model of instability, with an unstable orbit traversing one "link" in the chain.

a kind of diffusion. The terminology "Arnol'd diffusion" quickly became the standard name for the instability.

It is worth commenting on one of the sketchiest parts of the argument outlined in [2], namely the use of what Arnol'd calls *obstructing sets*, which in turn lies on the concept of *complementation*.

Definition 3.1.1 (Complementing manifold at a point)**.** Let M be a smooth submanifold of the space X. Let TM_x be the tangent plane to M at the point x. The manifold N *complements* M at the point $x \in M \cap N$ if

$$TM_x + TN_x = TX_x.$$

Definition 3.1.2 (Obstructing set)**.** With the same terminology as above, we say that the set Ω obstructs the manifold M at the point $x \in M$ if every manifold N which complements M at x is intersected by Ω.

Once the transverse intersection between the departing whisker of one torus and arriving whisker of another torus has been established, the concept of obstructing set is crucial in guaranteeing that trajectories moving very close alongside the departing whisker will transfer over so as to move closely alongside the arriving whisker of the next torus in the chain. In order to verify this process, a special versions of the so-called Lambda Lemma is needed (see Theorem 3.3.7).

In the remaining sections of this chapter we shall explain Arnol'd's proof in detail. The study of the system is done in three steps, corresponding to those in Figure 3.1.

3.2 The Unperturbed System

Let us fix $\varepsilon = 0$. The third and fourth equations in System (3.3) give $I_1 = \omega_1$ and $I_2 = \omega_2$, with ω_1, ω_2 constant values. These equations define a three-dimensional torus

$$\mathcal{T}_\omega = \{(I_1, I_2, \phi_1, \phi_2, s) \in \mathbb{R}^2 \times \mathbb{T}^3 \mid I_1 = \omega_1, I_2 = \omega_2\} \simeq \mathbb{T}^3,$$

where $\omega = (\omega_1, \omega_2, 1)$. The dynamics in this torus is described by the system

$$\begin{cases} \dot{\phi}_1 = \omega_1, \\ \dot{\phi}_2 = \omega_2, \\ \dot{s} = 1, \end{cases} \tag{3.4}$$

which can be explicitly solved to give

$$(\phi_1(t), \phi_2(t), s(t)) = (\phi_1^0, \phi_2^0, s^0) + (\omega_1, \omega_2, 1)t.$$

Therefore, the motion in $\mathcal{T}_\omega$ corresponds to a rigid rotation with frequency ω.

A way to find the instability that we seek is by seeing what occurs near surviving tori, that is, near resonance. For instance, if $\omega_2 \in \mathbb{Q}$, the equation $I_1 = \omega_1 = 0$ gives a four-dimensional manifold foliated by a family of resonant tori.

3.3 The Perturbed System

3.3.1 Two Uncoupled Subsystems

We now assume $\varepsilon > 0$ and $\mu = 0$. Then the original Hamiltonian

$$H(I_1, I_2, \phi_1, \phi_2, s; \varepsilon, \mu) = \frac{1}{2}(I_1^2 + I_2^2) + \varepsilon(\cos\phi_1 - 1) \tag{3.5}$$

can be split into two part, each one corresponding to one of the systems

$$\begin{cases} \dot{I}_1 = \varepsilon \sin\phi_1, \\ \dot{\phi}_1 = I_1, \end{cases} \quad \text{and} \quad \begin{cases} \dot{I}_2 = 0, \\ \dot{\phi}_2 = I_2. \end{cases} \tag{3.6}$$

The motion in the second subsystem is given by $(\phi_2(t), I_2(t)) = (\phi_2^0 + \omega_2 t, \omega_2)$. For the first system, the change of (ϕ_1, I_1) in time is that of an ordinary pendulum, which has a hyperbolic equilibrium point at $(0, 0)$. Moreover, the equation $I_2 = \omega_2$ defines a four-dimensional manifold containing several tori. Some are $3-$dimensional persistent tori which correspond to the rotation orbits in the pendulum and are described by

$$\frac{1}{2}I_1^2 + \varepsilon(\cos\phi_1 - 1) = c > 1, \quad I_2 = \omega_2, \quad \text{and} \quad (\phi_1, \phi_2, s) \in \mathbb{T}^3.$$

Other tori are destroyed, giving rise to $2-$dimensional whiskered tori of the form

$$\mathcal{T}_{\omega_2} = \{(I_1, I_2, \phi_1, \phi_2, s) \in \mathbb{R}^2 \times \mathbb{T}^3 \mid I_1 = 0, I_2 = \omega_2, \phi_1 = 0\} \simeq \mathbb{T}^2,$$

for any $\omega_2 \in \mathbb{T}$. The motion in these tori is given by

$$\phi_2 = \phi_2^0 + \omega_2 t, \qquad s = s_0 + t.$$

The three-dimensional arriving and departing whiskers (i.e. stable and unstable manifolds) form a homoclinic orbit (see Figure 3.3).

The following proposition gives an explicit expression for the whiskers of the $2-$dimensional whiskered tori. Before we state it, we need to give a definition.

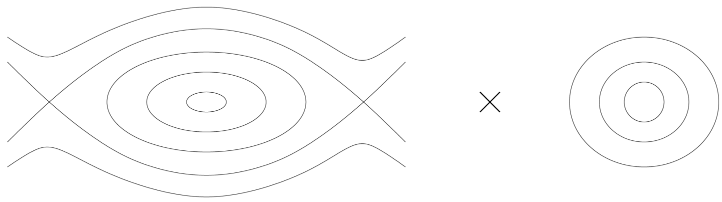

Figure 3.3: The phase portrait on the left is that of an ordinary pendulum. Persistent tori (given by KAM Theory) are the ones outside of the homoclinic orbit. The *gap* between surviving tori has size $\mathcal{O}(\sqrt{\varepsilon})$. The phase portrait on the right corresponds to an ordinary rotation.

Definition 3.3.1 (Higher-dimension stable and unstable invariant manifolds)**.** Let Λ be an invariant set contained in the phase space X. The *stable manifold* of Λ is the set

$$W^s(\Lambda) = \{p \in X \mid \text{ there exists } p_0^s \in \Lambda \text{ such that } \lim_{t\to+\infty} |\phi_t(p) - \phi_t(p_0^s| = 0\},$$

and the *unstable manifold* of Λ is the set

$$W^u(\Lambda) = \{p \in X \mid \text{ there exists } p_0^u \in \Lambda \text{ such that } \lim_{t\to-\infty} |\phi_t(p) - \phi_t(p_0^u| = 0\}.$$

Proposition 3.3.1 (Whiskered torus)**.** *The manifold*

$$T_{\omega_2} = \{(I_1, I_2, \phi_1, \phi_2, s) \in \mathbb{R}^2 \times \mathbb{T}^3 \mid I_1 = 0, I_2 = \omega_2, \phi_1 = 0\} \tag{3.7}$$

is a two-dimensional whiskered torus of System (3.3) *if* $\omega_2 \in \mathbb{R}\backslash\mathbb{Q}$*. The whiskers are* $3-$*dimensional and are defined by*

$$H^{(1)}(I_1, I_2, \phi_1, \phi_2, s; \varepsilon, \mu) = \frac{1}{2}I_1^2 + \varepsilon(\cos\phi_1 - 1) = 0 \tag{3.8}$$

and

$$H^{(2)}(I_1, I_2, \phi_1, \phi_2, s; \varepsilon, \mu) = \frac{1}{2}I_2^2 = \frac{1}{2}\omega_2^2, \tag{3.9}$$

which is equivalent to

$$I_1 = \pm 2\sqrt{\varepsilon}\sin\left(\frac{\phi_1}{2}\right) \quad \text{and} \quad I_2 = \omega_2. \tag{3.10}$$

Or, alternatively,

$$\begin{cases} I_1(t) = \pm 2\sqrt{\varepsilon}\operatorname{arccosh}(\sqrt{\varepsilon}(t - t^0)), \\ \phi_1(t) = \pm\operatorname{arccot}(-\sinh(\sqrt{\varepsilon}(t - t^0))), \\ \phi_2(t) = \phi_2^0 + \omega(t - t^0), \end{cases} \tag{3.11}$$

where $I_1(t^0) = \pm\sqrt{\varepsilon}$*,* $\phi_1(t^0) = \pm\pi$ *and* $\phi_2(t^0) = \phi_2^0$*.*

Proof. The fixed points are of the form $(0, \pi k)$, for $k \in \mathbb{Z}$. To see their stability, we compute

$$Df(I_1, \phi_1)|_{(0,\pi k)} = \begin{pmatrix} 0 & \varepsilon\cos\phi_1 \\ 1 & 0 \end{pmatrix}\Bigg|_{(0,\pi k)} = \begin{pmatrix} 0 & \varepsilon(-1)^k \\ 1 & 0 \end{pmatrix},$$

so that the eigenvalues are $\pm\sqrt{\varepsilon}$ for even k and $\pm i\sqrt{\varepsilon}$ for odd k.

In order to compute the homoclinic orbit, we notice that the points (I_1, ϕ_1) satisfy $H(I_1, \phi_1) = H(0, 0)$, so the orbit is given by $H^{(1)} = 0$ and $H^{(2)} = \frac{1}{2}\omega^2$, that is

$$I_1 = \pm\sqrt{2\varepsilon(1 - \cos\phi_1} = \pm\sqrt{4\varepsilon\sin^2\left(\frac{\phi_1}{2}\right)} = \pm 2\sqrt{\varepsilon}\sin\left(\frac{\phi_1}{2}\right), \quad \text{and} \quad I_2 = \omega.$$

Since $I_1 = \dot{\phi}_1$, we have

$$\pm\int \frac{1}{\sin\left(\frac{\phi_1}{2}\right)}\, d\phi_1 = 2\sqrt{\varepsilon}\int\, dt.$$

Imposing for instance $\phi_1(0) = \pi$, this can be solved to give

$$\phi_1(t) = 4\,\mathrm{arccot}\, e^{\mp 2\sqrt{\varepsilon}t}.$$

Then,

$$I_1(t) = \pm 2\sqrt{\varepsilon}\sin\left(\frac{\phi_1}{2}\right) = \pm 2\sqrt{\varepsilon}\,\mathrm{arccosh}(\sqrt{\varepsilon}t).$$

Thus, the whiskers satisfy

$$\lim_{t\to\infty} I_1(t) = 0, \qquad \lim_{t\to\infty}\phi_1(t) = 0, \qquad \lim_{t\to\infty}\phi_2(t) \in \{\phi_2^0 + \omega(t - t^0)\},$$

so that if

$$z_h(t) = (I_1(t), \omega_2, \phi_1(t), \phi_2^0 + \omega_2 t, s_0 + t) \qquad \text{and} \qquad z_{\omega_2} = (0, \omega_2, 0, \phi_2^0 + \omega_2 t, s^0 + t)$$

satisfy

$$\lim_{t\to\pm\infty} \|z_h(t) - z_{\omega_2}(t)\| = 0,$$

then, there is a homoclinic orbit. □

This proof gives, in addition, the existence of a homoclinic orbit for the whiskered torus.

3.3.2 Complete System

We now consider the whole system, that is, we assume $0 \ll \mu \ll \varepsilon \ll 1$. In [1,2] it is proved that for the majority of non-resonant conditions, the quantities $I_1(t)$ and $I_2(t)$ change little in the course of the whole infinite interval of time $(-\infty, \infty)$. This is in fact a consequence of KAM Theory (see Section 2.2.2). It turns out, however, that close to the resonant manifold $I_1 = 0$, there is a zone of nonstability.

Of course, for $\mu \neq 0$, the homoclinic orbit of the torus defined in (3.7) splits into two whiskers which intersect each other.

Proposition 3.3.2. *The manifold $\mathcal{T}_{\omega_2}$ is a whiskered torus of System* (3.3) *if μ is sufficiently small.*

Proof. The assertion can be proved by the standard method of contractive mappings. It is convenient to use to conical metric

$$\|f(x)\| = \max\left\{\frac{f(x)}{x}\right\}.$$

□

Definition 3.3.2 (Transition torus)**.** A *transition torus* T is a whiskered torus such that the images of an arbitrary neighbourhood of an arbitrary point ξ of one of its arriving whiskers obstruct the departing whisker at an arbitrary point η of the latter.

Definition 3.3.3 (Transition chain)**.** A *transition chain* is a collection of tori $T_1, T_2, \ldots, T_n, \ldots$ such that the departing whisker Y_s^+ of every preceding torus T_s complements the arriving whisker of the following torus Y_{s+1}^- at some point of their intersection $x_s \in Y_s^+ \cap Y_{s+1}^-$.

It turns out that the following lemma holds.

Lemma 3.3.3 (Existence of a transition chain)**.** *Assume $A < \omega < B$. Then, there exists $\kappa = \kappa(\varepsilon, \mu, A, B) > 0$ such that the departing whisker Y_ω^+ of the torus T_ω intersects the arriving whiskers Y_ω^-, of all tori $T_{\widetilde{\omega}}$ for which $|\omega - \widetilde{\omega}| \leq \kappa = \kappa(\varepsilon, \mu, A, B)$.*

Proof. The proof of this lemma requires certain calculations. The nonperturbed whiskers have the equations (3.8) and (3.9). Assume $\alpha > 0$ (for instance, $\alpha = \pi/2$). It is easy to see that for $|\phi_1| < 2\pi - \alpha$, the equations of the perturbed departing whisker Y_ω^+ can be written in the form

$$H^{(1)} = \Delta_1^+(\phi_1; \phi_2, t; \omega) \quad \text{and} \quad H^{(2)} = \frac{1}{2}\omega^2 + \Delta_2^+(\phi_1; \phi_2, t; \omega), \tag{3.12}$$

where the functions $\Delta_k^+ = \mathcal{O}(\mu)$ are $2\pi-$periodic with respect to ϕ_2 and t, and are equal to 0 for $\phi_1 = 0$. In exactly the same way, the arriving whisker Y_ω^-, for $|\phi_1 - 2\pi| < 2\pi - \alpha$ has equations

$$H^{(1)} = \Delta_1^-(\phi_1; \phi_2, t; \widetilde{\omega}) \quad \text{and} \quad H^{(2)} = \frac{1}{2}\widetilde{\omega}^2 + \Delta_2^+(\phi_1; \phi_2, t; \widetilde{\omega}). \tag{3.13}$$

We shall look for the intersection of the whiskers Y_ω^+ and $Y_{\widetilde{\omega}}^-$ on the plane $\phi_1 = \pi$. The statement of this lemma is an assertion concerning the solvability with respect to ϕ_2 and t of the system of equations

$$\begin{cases} \Delta_1^+(\pi; \phi_2, t; \omega) = \Delta_1^-(\pi; \phi_2, t; \widetilde{\omega}), \\ \frac{1}{2}\omega^2 + \Delta_2^+(\pi; \phi_2, t; \omega) = \frac{1}{2}\widetilde{\omega}^2 + \Delta_2^-(\pi; \phi_2, t; \widetilde{\omega}). \end{cases} \tag{3.14}$$

The solvability of this system can be deduced from an approximate expression for $\Delta_k^\pm$. Indeed, in accordance with Equations (3.12) and (3.13), the quantities $\Delta_k^\pm$ represent the increments of $H^{(k)}$ in the perturbed motion of System (3.2). Therefore, the derivatives of $H^{(k)}$ are exactly the Poisson bracket $\{H, H^{(k)}\}$.

Consequently, $\Delta_k^{\pm}$ is exactly equal to the following integrals extended over the perturbed trajectories:

$$\Delta_k^{\pm} = \int_{\mp\infty}^{0} \{H, H^{(k)}\}\ d(t - t^0), \tag{3.15}$$

where the Poisson bracket is integrated along the nonperturbed trajectory described in (3.11). We then derive the estimate

$$\Delta_k^{\pm} = \mu\delta_k^{\pm} + \mathcal{O}(\mu^2).$$

Taking now

$$\delta_k = \delta_k^{+}(\pi; \phi_2^0, t^0, \omega) - \delta_k^{-}(\pi; \phi_2^0, t^0, \omega) = \int_{-\infty}^{+\infty} \{H, H^{(k)}\}\ d(t - t^0),$$

it is obvious that the solvability of System (3.16) depends basically on the solvability with respect to ϕ_2^0 and t^0 of the ystem

$$\begin{cases} \delta_1 + \mathcal{O}(\mu) = 0, \\ \dfrac{1}{2}(\omega^2 - \widetilde{\omega}^2) + \mu\delta_2 + \mathcal{O}(\mu^2) = 0. \end{cases} \tag{3.16}$$

An easy calculation gives the result

$$\delta_1 = -2\varepsilon \int_{-\infty}^{+\infty} u \frac{\partial B}{\partial t}\ dt \quad \text{and} \quad \delta_2 = 2\varepsilon\omega \int_{-\infty}^{+\infty} u \frac{\partial B}{\partial \phi_2}\ dt, \tag{3.17}$$

where

$$u = \frac{1}{\cosh^2(\sqrt{\varepsilon}(t - t^0))}, \quad B = B(\phi_2, t) \quad \text{and} \quad \phi_2 = \phi_2^0 + \omega(t - t^0).$$

By Proposition 3.3.4, for $B = \sin(\phi_2) + \cos t$ we have

$$\delta_1 = \frac{2\pi \sin t^0}{\sinh\left(\dfrac{\pi}{2\sqrt{\varepsilon}}\right)} \quad \text{and} \quad \delta_2 = \frac{2\pi\omega \cos\phi_2^0}{\sinh\left(\dfrac{\omega\pi}{2\sqrt{\varepsilon}}\right)}.$$

The intersection between the two whiskers of the same torus can be found by imposing $\widetilde{\omega} = \omega$, where ω is fixed, so that we obtain

$$\begin{cases} \dfrac{2\pi \sin t^0}{\sinh\left(\dfrac{\pi}{2\sqrt{\varepsilon}}\right)} + \mathcal{O}(\mu) = 0, \\ \mu \dfrac{2\pi\omega \cos\phi_2^0}{\sinh\left(\dfrac{\omega\pi}{2\sqrt{\varepsilon}}\right)} + \mathcal{O}(\mu^2) = 0. \end{cases} \tag{3.18}$$

By the Implicit Function Theorem, we can find a solution of this system for $|\mu| < \mu_0$, where μ_0 is fixed.

We now suppose $\widetilde{\omega} \neq \omega$ in order to find the heteroclinic connection between two different tori of frequencies ω and $\widetilde{\omega}$. If we restrict to order $\mathcal{O}(\mu)$, the second equation in System (3.16) yields

$$\frac{\widetilde{\omega}^2}{2} = \frac{\omega^2}{2} + \mu \frac{2\pi\omega \cos\phi_2^0}{\sinh\left(\frac{\omega\pi}{2\sqrt{\varepsilon}}\right)}.$$

Therefore, the new frequency $\widetilde{\omega}$ satisfies

$$\frac{\widetilde{\omega}^2}{2} \in \left[\frac{\omega^2}{2} - \mu \frac{2\pi\omega \cos\phi_2^0}{\sinh\left(\frac{\omega\pi}{2\sqrt{\varepsilon}}\right)}, \frac{\omega^2}{2} + \mu \frac{2\pi\omega \cos\phi_2^0}{\sinh\left(\frac{\omega\pi}{2\sqrt{\varepsilon}}\right)}\right],$$

so we can assume that $\widetilde{\omega}$ is of the form

$$\widetilde{\omega} = a\mu + \omega, \tag{3.19}$$

where $a \in \mathbb{R}$. Then, the original system (3.16) becomes

$$\begin{cases} \dfrac{2\pi \sin t^0}{\sinh\left(\frac{\pi}{2\sqrt{\varepsilon}}\right)} + \mathcal{O}(\mu) = 0, \\ \mu \dfrac{2\pi\omega \cos\phi_2^0}{\sinh\left(\frac{\omega\pi}{2\sqrt{\varepsilon}}\right)} - a\mu(2\omega + a\mu) + \mathcal{O}(\mu^2) = 0. \end{cases} \tag{3.20}$$

Notice that we can simplify μ in the second equation in System (3.21). For $\mu = 0$, the second equation gives

$$\phi_2^0 = \arccos\left(\frac{1}{2\omega} 2a\omega \sinh\left(\frac{\omega\pi}{2\sqrt{\varepsilon}}\right)\right).$$

In order for ϕ_2^0 to be well defined, we need to impose a condition for the argument of the arcosine. This gives an interval for a so that the system has a solution. Namely,

$$a \in \left[-\sinh^{-1}\left(\frac{\omega\pi}{2\sqrt{\varepsilon}}\right), \sinh^{-1}\left(\frac{\omega\pi}{2\sqrt{\varepsilon}}\right)\right].$$

Let us fix a_0 belonging to this interval. In order to find the values of μ that allow the intersection between different tori to occur, we use the Implicit Function Theorem again. Hence we obtain a solution to the system

$$\begin{cases} \dfrac{2\pi \sin t^0}{\sinh\left(\frac{\pi}{2\sqrt{\varepsilon}}\right)} + \mathcal{O}(\mu) = 0, \\ \dfrac{2\pi\omega \cos\phi_2^0}{\sinh\left(\frac{\omega\pi}{2\sqrt{\varepsilon}}\right)} - a_0(2\omega + a_0\mu) + \mathcal{O}(\mu) = 0, \end{cases} \tag{3.21}$$

for μ satisfying $|\mu| < \widetilde{\mu}_0$. Therefore, Equation (3.19) gives an intersection between whiskers of different tori provided that

$$|\widetilde{\omega} - \omega| = |a_0 \mu| \leq \frac{\widetilde{\mu}_0}{\sinh\left(\frac{\omega\pi}{2\sqrt{\varepsilon}}\right)} \leq \kappa,$$

where $\kappa = \kappa(\varepsilon, \mu, A, B) > 0$. □

Lemma 3.3.4. *For $B = \sin(\phi_2) + \cos t$, the equations in* (3.17) *can be written as*

$$\delta_1 = \frac{2\pi \sin t^0}{\sinh\left(\frac{\pi}{2\sqrt{\varepsilon}}\right)} \quad \textit{and} \quad \delta_2 = \frac{2\pi\omega \cos \phi_2^0}{\sinh\left(\frac{\omega\pi}{2\sqrt{\varepsilon}}\right)}.$$

Proof. For the given B, we have

$$\begin{aligned}
\delta_1 &= -2\varepsilon \int_{-\infty}^{+\infty} \frac{-\sin t}{\cosh^2(\sqrt{\varepsilon}(t - t^0))}\, dt \\
&= 2\varepsilon \frac{1}{\sqrt{\varepsilon}} \int_{-\infty}^{+\infty} \frac{\sin\left(\frac{s}{\sqrt{\varepsilon}} + t^0\right)}{\cosh^2 s}\, ds \\
&= 2\varepsilon \frac{1}{\sqrt{\varepsilon}} \int_{-\infty}^{+\infty} \frac{\sin\left(\frac{s}{\sqrt{\varepsilon}}\right) \cos t^0}{\cosh^2 s}\, ds + 2\varepsilon \frac{1}{\sqrt{\varepsilon}} \int_{-\infty}^{+\infty} \frac{\cos\left(\frac{s}{\sqrt{\varepsilon}}\right) \sin t^0}{\cosh^2 s}\, ds \\
&= \sqrt{\varepsilon} \sin t^0 \int_{-\infty}^{+\infty} \frac{e^{\frac{is}{\sqrt{\varepsilon}}} + e^{\frac{-is}{\sqrt{\varepsilon}}}}{\cosh^2 s}\, ds \\
&= \sqrt{\varepsilon} \sin t^0 \int_{-\infty}^{+\infty} \frac{e^{\frac{is}{\sqrt{\varepsilon}}}}{\cosh^2 s}\, ds - \sqrt{\varepsilon} \sin t^0 \int_{+\infty}^{-\infty} \frac{e^{\frac{ir}{\sqrt{\varepsilon}}}}{\cosh^2(-r)}\, dr \\
&= 2\sqrt{\varepsilon} \sin t^0 \int_{-\infty}^{+\infty} \frac{e^{\frac{is}{\sqrt{\varepsilon}}}}{\cosh^2 s}\, ds, \qquad (3.22)
\end{aligned}$$

where we have used that $\cosh(-r) = \cosh(r)$. Let now the map f be defined as

$$f(s) = \frac{e^{\frac{is}{\sqrt{\varepsilon}}}}{\cosh^2 s}.$$

Then, the integral in (3.22) can be dealt with using the Residue Formula, according to which

$$\int_\gamma f(z)\, dz = 2\pi i \sum_{i=1}^{k} \mathrm{Res}(f, c_i),$$

where γ is a closed curve containing the k poles of the map f and $\mathrm{Res}(f, c_i)$ is the residue of f at the i−th pole c_i. Furthermore, each of the residues can be computed as

$$\mathrm{Res}(f, c_i) = \frac{1}{(n_i - 1)!} \lim_{s \to c_i} \frac{d^{n_i - 1}}{ds^{n_i - 1}} \left((s - c_i)^{n_i} f(s)\right),$$

where n_i is the order of the pole c_i.

In our case, the map f has one pole of order two at $c = \frac{\pi i}{2}$. Therefore,

$$\begin{aligned}
\mathrm{Res}(f,c) &= \lim_{s\to c} \frac{d}{ds}\left((s-c)^2 f(s)\right) \\
&= \lim_{s\to c} \frac{d}{ds}\left(\frac{(s-c)^2 e^{\frac{is}{\sqrt{\varepsilon}}}}{-(s-c)^2 - \frac{1}{3}(s-c)^4 - \cdots}\right) \\
&= \lim_{s\to c} \frac{d}{ds}\left(\frac{e^{\frac{is}{\sqrt{\varepsilon}}}}{-1 - \frac{1}{3}(s-c)^2 - \cdots}\right) \\
&= \lim_{s\to c} \frac{\frac{i}{\sqrt{\varepsilon}} e^{\frac{is}{\sqrt{\varepsilon}}}\left(-1-\frac{1}{3}(s-c)^2-\cdots\right) - e^{\frac{is}{\sqrt{\varepsilon}}}\left(-\frac{2}{3}(s-c)-\cdots\right)}{\left(-1-\frac{1}{3}(s-c)^2-\cdots\right)^2} \\
&= \frac{-i}{\sqrt{\varepsilon}} e^{\frac{ic}{\sqrt{\varepsilon}}} \\
&= \frac{-i}{\sqrt{\varepsilon}} e^{\frac{-\pi}{2\sqrt{\varepsilon}}}.
\end{aligned}$$

We now choose a suitable curve γ that allows us to compute the desired integral. In particular, let γ be the boundary of the closed rectangle of vertices $(R,0)$, $(R,\pi i)$, $(-R,\pi i)$ and $(-R,0)$, where $R>0$ is a constant, and let γ_i be each of the four segments that form γ, starting with the one lying on the $x-$axis and taking the counterclockwise orientation. Using the value obtained for the residue $\mathrm{Res}(f,c)$, we obtain

$$\sum_{i=1}^{4} \int_{\gamma_i} f(z)\ dz = 2\pi i \frac{-i}{\sqrt{\varepsilon}} e^{\frac{-\pi}{2\sqrt{\varepsilon}}} = \frac{2\pi}{\sqrt{\varepsilon}} e^{\frac{-\pi}{2\sqrt{\varepsilon}}}.$$

In particular, if we let R tend to infinity, we have

$$\begin{aligned}
\lim_{R\to+\infty} \int_{\gamma_2} f(z)\ dz &= \lim_{R\to+\infty} \int_{\gamma_4} f(z)\ dz = 0, \\
\lim_{R\to+\infty} \int_{\gamma_3} f(z)\ dz &= \int_{+\infty}^{-\infty} \frac{e^{i\frac{\pi i+\sigma}{\sqrt{\varepsilon}}}}{\cosh^2(\pi i+\sigma)}\ d\sigma \\
&= -\int_{-\infty}^{+\infty} \frac{e^{\frac{-\pi}{\sqrt{\varepsilon}}} e^{\frac{i\sigma}{\sqrt{\varepsilon}}}}{\cosh^2(\sigma)}\ d\sigma \\
&= -e^{\frac{-\pi}{\sqrt{\varepsilon}}} \lim_{R\to+\infty} \int_{\gamma_1} f(z)\ dz,
\end{aligned}$$

where we have used that

$$\cosh(\pi i+\sigma) = \frac{e^{\pi i+\sigma} e^{-\pi i-\sigma}}{2} = \frac{-e^{\sigma} + (-e^{\sigma})^{-1}}{2} = -\cosh(\sigma).$$

Therefore,

$$\left(1 - e^{\frac{-\pi}{\sqrt{\varepsilon}}}\right) \lim_{R\to+\infty} \int_{\gamma_1} f(z)\ dz = \frac{2\pi}{\sqrt{\varepsilon}} e^{\frac{-\pi}{2\sqrt{\varepsilon}}},$$

and so

$$\begin{aligned}\delta_1 &= 2\sqrt{\varepsilon}\sin t^0 \int_{-\infty}^{+\infty} \frac{e^{\frac{is}{\sqrt{\varepsilon}}}}{\cosh^2 s}\, ds = 2\sqrt{\varepsilon}\sin t^0 \lim_{R\to+\infty}\int_{\gamma_1} f(z)\, dz \\ &= 2\sqrt{\varepsilon}\sin t^0 \frac{2\pi}{\sqrt{\varepsilon}} \frac{e^{\frac{-\pi}{2\sqrt{\varepsilon}}}}{1-e^{\frac{-\pi}{\sqrt{\varepsilon}}}} = \frac{2\pi \sin t^0}{\sinh\left(\frac{\pi}{2\sqrt{\varepsilon}}\right)}.\end{aligned}$$

Similarly,

$$\begin{aligned}\delta_2 &= 2\varepsilon\omega \int_{-\infty}^{+\infty} \frac{\cos(\phi_2^0 + \omega(t-t^0))}{\cosh^2(\sqrt{\varepsilon}(t-t^0))}\, dt \\ &= 2\varepsilon\omega \int_{-\infty}^{+\infty} \frac{\cos\left(\phi_2^0 + \dfrac{\omega s}{\sqrt{\varepsilon}}\right)}{\cosh^2 s}\, dt \\ &= 2\sqrt{\varepsilon}\omega \int_{-\infty}^{+\infty} \frac{\cos\phi_2^0 \cos\left(\dfrac{\omega s}{\sqrt{\varepsilon}}\right)}{\cosh^2 s}\, ds - 2\sqrt{\varepsilon}\omega \int_{-\infty}^{+\infty} \frac{\sin\phi_2^0 \sin\left(\dfrac{\omega s}{\sqrt{\varepsilon}}\right)}{\cosh^2 s}\, ds \\ &= \sqrt{\varepsilon}\omega\cos\phi_2^0 \int_{-\infty}^{+\infty} \frac{e^{\frac{i\omega s}{\sqrt{\varepsilon}}} + e^{\frac{-i\omega s}{\sqrt{\varepsilon}}}}{\cosh^2 s}\, ds \\ &= \sqrt{\varepsilon}\omega\cos\phi_2^0 \int_{-\infty}^{+\infty} \frac{e^{\frac{i\omega s}{\sqrt{\varepsilon}}}}{\cosh^2 s}\, ds - \sqrt{\varepsilon}\omega\cos\phi_2^0 \int_{+\infty}^{-\infty} \frac{e^{\frac{i\omega r}{\sqrt{\varepsilon}}}}{\cosh^2 r}\, dr \\ &= 2\sqrt{\varepsilon}\omega\cos\phi_2^0 \int_{-\infty}^{+\infty} \frac{e^{\frac{i\omega s}{\sqrt{\varepsilon}}}}{\cosh^2 s}\, ds.\end{aligned}$$

This integral can be computed by following the exact same steps as before and taking into account that there is a factor ω multiplying the exponent of the numerator. Indeed, if we call $\tilde{f}$ the new integrand, then $\tilde{f}$ has the same pole of order two, and we obtain

$$\mathrm{Res}(\tilde{f}, c) = \frac{-\omega i}{\sqrt{\varepsilon}} e^{\frac{-\omega\pi}{2\sqrt{\varepsilon}}}.$$

Let us now take each γ_i as before. Then,

$$\begin{aligned}\lim_{R\to+\infty}\int_{\gamma_2} f(z)\, dz &= \lim_{R\to+\infty}\int_{\gamma_4} f(z)\, dz = 0, \\ \lim_{R\to+\infty}\int_{\gamma_3} f(z)\, dz &= \int_{+\infty}^{-\infty} \frac{e^{i\omega\frac{\pi i+\sigma}{\sqrt{\varepsilon}}}}{\cosh^2(\pi i+\sigma)}\, d\sigma \\ &= -\int_{-\infty}^{+\infty} \frac{e^{\frac{-\omega\pi}{\sqrt{\varepsilon}}} e^{\frac{i\omega\sigma}{\sqrt{\varepsilon}}}}{\cosh^2(\sigma)}\, d\sigma \\ &= -e^{\frac{-\omega\pi}{\sqrt{\varepsilon}}} \lim_{R\to+\infty}\int_{\gamma_1} f(z)\, dz.\end{aligned}$$

Therefore,

$$\left(1 - e^{\frac{-\omega\pi}{\sqrt{\varepsilon}}}\right) \lim_{R\to+\infty}\int_{\gamma_1} f(z)\, dz = 2\pi i\,\mathrm{Res}(\tilde{f}, c) = 2\pi i \frac{-\omega i}{\sqrt{\varepsilon}} e^{\frac{-\omega\pi}{2\sqrt{\varepsilon}}} = \frac{2\pi\omega}{\sqrt{\varepsilon}} e^{\frac{-\omega\pi}{2\sqrt{\varepsilon}}},$$

which gives

$$\int_{-\infty}^{+\infty} \frac{e^{\frac{i\omega s}{\sqrt{\varepsilon}}}}{\cosh^2 s}\, ds = \lim_{R\to+\infty} \int_{\gamma_1} f(z)\, dz = \frac{2\pi\omega}{\sqrt{\varepsilon}} e^{\frac{-\omega\pi}{2\sqrt{\varepsilon}}} \frac{1}{1-e^{\frac{-\omega\pi}{\sqrt{\varepsilon}}}} = \frac{\pi\omega}{\sqrt{\varepsilon}\sinh\left(\frac{\omega\pi}{2\sqrt{\varepsilon}}\right)}.$$

Hence, we finally obtain

$$\delta_2 = 2\sqrt{\varepsilon}\omega\cos\phi_2^0 \int_{-\infty}^{+\infty} \frac{e^{\frac{i\omega s}{\sqrt{\varepsilon}}}}{\cosh^2 s}\, ds = 2\sqrt{\varepsilon}\omega\cos\phi_2^0 \frac{\pi\omega}{\sqrt{\varepsilon}\sinh\left(\frac{\omega\pi}{2\sqrt{\varepsilon}}\right)} = \frac{2\pi\omega^2\cos\phi_2^0}{\sinh\left(\frac{\omega\pi}{2\sqrt{\varepsilon}}\right)}.$$

□

Assume $\omega_1 < A < B < \omega_s$. For the proof of instability (see Theorem 3.3.10), it is sufficient to construct a transition chain of tori $\mathcal{T}_{\omega_1}, \mathcal{T}_{\omega_2}, \ldots, \mathcal{T}_{\omega_s}$, as the one given by Lemma 3.3.3. Next we need a Lambda Lemma in order to prove the shadowing property. Such a lemma can be found in [5]. The authors study the existence and parameter dependence of invariant manifolds of lower-dimensional tori for families of maps, when the normal behaviour has both a hyperbolic and a central part. Once the existence of these invariant manifolds is established, a very general version of the lemma for (partially hyperbolic) tori is given. Due to the technicalities of the paper, we shall omit most of the details and certainly all of the proofs.

We consider maps $F_\nu(w)$ defined on $\mathcal{M}_\delta := \mathcal{B}_\delta^m \times \mathcal{B}_\delta^p \times \mathbb{T}^n \times \mathcal{B}_\delta^q$, where $\mathcal{M}_\delta \subset \mathcal{M} := \mathbb{R}^m \times \mathbb{R}^p \times \mathcal{T}^n \times \mathbb{R}^q$ and $\nu \in \mathcal{B}_\mu^s \subset \mathbb{R}^s$. Suppose that $F_\nu(w)$ is of the form

$$F_\nu(w) = (A_-(\theta)x, A_+(\theta)y, \theta + \omega(x,y,r), B(\theta)r) + f(x,y,\theta,r) \tag{3.23}$$

with some extra conditions. The stable manifold is then obtained for not necessarily invertible maps. The results for the unstable one readily follow from the stable one under suitable invertibility conditions. Existence of invariant stable and unstable manifolds $W^{s,u}$ of dimension $m+n$ and $p+n$, respectively, is obtained. These invariant manifolds are given by $W^s = \{(y,r) = \gamma^s(x,\theta)\}$ and $W^u = \{(x,r) = \gamma^u(y,\theta)\}$, for certain functions $\gamma^{s,u}$.

The proof of the lemma has two main parts. We first choose a normal form of F around the invariant torus.

Lemma 3.3.5. *Let $F\colon \mathcal{M}_\delta \to \mathcal{M}$ be a map of the form* (3.23) *satisfying certain hypotheses. If δ is small enough, there exists a change of variables C that is $\mathcal{C}^2-$close to the identity, defined in a neighbourhood of $\mathbb{T}_0$ such that $\widetilde{F} = C^{-1}\circ F\circ C$ takes the same form* (3.23)*, that is*

$$\widetilde{F}(w) = (\widetilde{A}_-(\theta)x, \widetilde{A}_+(\theta)y, \theta + \widetilde{\omega}(x,y,r), \widetilde{B}(\theta)r) + \widetilde{f}(w).$$

The result then follows from the following proposition. A concept that is used is that of a *$p-$dimensional disc*, by which we mean the image of a $p-$dimensional ball into $\mathcal{M}$ by a $\mathcal{C}^1$ map.

Proposition 3.3.6. *Let F satisfy certain hypotheses. Assume that $\omega_0 = \omega(0,0,0)$ is non-resonant. Let $p_0 \in W^u$ and Γ be a $\mathcal{C}^1$ $(p+q)-$dimensional manifold that* *is transverse to* W^s *on* q_0. *Then, there exists a* $p-$*dimensional disc contained in* Γ, *given by* $q(z) = (x(z), z, \theta(z), r(z))$, *for any* $z \in B_{\zeta_0} = \{\|z\| < \zeta_0\} \subset \mathbb{R}^p$, *where* $q(0) = q_0$, *such that for any* $\varepsilon > 0$ *there exists* $j \in \mathbb{N}$ *satisfying*

$$\{F^j(q(z)) \mid z \in B_{\zeta_0}\} \cap B(p_0, \varepsilon) \neq \emptyset.$$

Furthermore, there exists j_0 *such that* $D(F^j(q(z)))(\mathbb{R}^p)$ *is* $\varepsilon-$*close to a subspace of* TW^u *if* $j > j_0$.

Theorem 3.3.7 (Lambda lemma). *Let* F_ν *satisfy certain hypotheses. Let* ω_0 *be non-resonant. Let* γ *be a* $(p+q)-$*dimensinoal* $\mathcal{C}^1$ *manifold intersecting transversally* W^s *at* q_0. *Then,*

$$W^u = \overline{\cup_{n=0}^{\infty} F^n(\gamma)}.$$

Moreover, there exist $p-$*dimensional submanifolds* D *of* γ *such that if* D_n *is the connected component of* $F^n(D) \cup B(0, \delta)$ *containing* $F^n(q_0)$, *then for any* $\varepsilon > 0$ *there exists* n_0 *such that* TD_n *is* $\varepsilon-$*close to (a subset of)* TW^u *if* $n > n_0$.

The Lambda Lemma provides a way to see that arbitrary neighbourhoods of different tori are connected. Indeed, it is a consequence of

Lemma 3.3.8 (Shadowing Lemma). *Let* $\{\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_n, \dots\}$ *be a sequence of transition tori. Given* $\{\varepsilon_i\}_{i=1}^{\infty}$ *a sequence of strictly positive numbers, there exists a point* P *and an increasing sequence of numbers* $\{T_i\}_{i=1}^{\infty}$ *such that*

$$\phi_{T_i}(P) \in N_{\varepsilon_i}(\mathcal{T}_i),$$

where $N_{\varepsilon_i}(\mathcal{T}_i)$ *is a neighbourhood of size* ε_i *of the torus* $\mathcal{T}_i$.

Proof. Let $x \in W^s_{\mathcal{T}_i}$. We can find a closed ball B_1 centred at x such that

$$\phi_{T_1}(B_1) \subset N_{\varepsilon_i}(\mathcal{T}_1). \tag{3.24}$$

By the Lambda Lemma, we know that

$$W^s_{\mathcal{T}_2} \cap B_1 \neq \emptyset.$$

Hence, we can find a closed ball $B_2 \subset B_1$, centred at a point in $W^s_{\mathcal{T}_2}$ satisfying (3.24) and

$$\phi_{T_2}(B_2) \subset N_{\varepsilon_2}(\mathcal{T}_2).$$

Proceeding by induction, we can find a decreasing sequence of closed balls $B_1 \supset B_2 \supset \cdots \supset B_{i-1} \supset B_i$ such that

$$\phi_{T_j}(B_i) \subset N_{\varepsilon_j}(\mathcal{T}_j), \qquad \text{for any } i \leq j.$$

Since the balls are compact, their intersection is nonempty. Hence a point P in the intersection satisfies the required property. □

This shadowing property readily gives the following

Theorem 3.3.9 (Arbitrary neighbourhoods of different tori are connected). *Let* $T_1, T_2, \dots, T_n, \dots$ *be a transition chain. Then an arbitrary neighbourhood of the torus* T_1 *is connected with an arbitrary neighbourhood of the torus* T_s *by trajectories of the given dynamical system.*

The instability close to the resonant manifold then follows.

Theorem 3.3.10 (Nonstability close to the resonant manifold $I_1 = 0$)**.** *Assume* $0 < A < B$. *For every* $\varepsilon > 0$ *there exists* $\mu_0 > 0$ *such that, for* $0 < \mu < \mu_0$*, System 3.2 is nonstable, that is, there exists a trajectory which connects the region* $I_2 < A$ *with the region* $I_2 > B$.

4

Transversality of Manifolds for Near-Integrable Hamiltonian Systems with More General Perturbations

In this chapter, we consider nearly-integrable systems (see Chapter 2), which appear when one considers a perturbation of an integrable Hamiltonian

$$H(I, \varphi, \varepsilon) = H_0(I) + \varepsilon H_1(I, \varphi),$$

where ε is a small parameter, $I = (I_1, I_2, \ldots, I_n)$ and $\varphi = (\varphi_1, \varphi_2, \ldots, \varphi_n)$ (see (2.3)). Recall that the values of the actions I such that the unperturbed frequencies

$$\omega_k(I) = \frac{\partial H_0}{\partial I_k}$$

are rationally dependent are called *resonances* (see Definition 2.1.2). Such systems provide a realistic model for the motion near a resonance only in the case of two degrees of freedom. If one considers simple resonances of systems with more than two degrees of freedom, one can choose all the angles except one to be the fast variables.

Autonomous models with perturbations that depend on time in a quasiperiodic way appear in several problems of Celestial Mechanics. For instance, the motion of a spacecraft in the Earth-Moon system can be modelled assuming that the Earth and the Moon revolve in circles around their common centre of masses (this gives an autonomous model), and the main perturbations (difference between the circular and the real motion of the Moon, effect of the Sun, etc.) are modelled as a time-dependent quasiperiodic function.

After Arnol'd's paper [2] saw the light, the main cause of the stochastic behaviour in Hamiltonian systems was considered to be the phenomenon of the splitting of separatrices. Indeed, Arnol'd used this fact to show existence of instability for a near-integrable Hamiltonian system with a $2-$harmonic periodic perturbation. However, the main problem studying systems near a resonance is that the splitting is of separatrices is exponentially small with respect to ε. Namely, for a Hamiltonian of the form

$$H\left(x, y, \frac{t}{\varepsilon}\right) = H_0(x, y) + H_1\left(x, y, \frac{t}{\varepsilon}\right),$$

where the Hamiltonian system of H_0 has a saddle and an associated homoclinic orbit, and the perturbation of H_1 is a periodic function of time with zero mean value, Neishtadt's Theorem (see [8]) implies that the splitting can be bounded from above by $\mathcal{O}\left(\exp\left(-\frac{\text{const}}{\varepsilon}\right)\right)$. For this estimate to be valid, all the functions have to be real analytic in x and y, whereas $\mathcal{C}^1$−dependence on time is enough.

The aim of the present chapter is to study the size of the above-mentioned splitting in more general near-integrable Hamiltonian systems. Working on the most general case is extremely difficult, for it requires dealing with manifolds for infinitely many harmonics. In fact, there is no know result for such a system if one assumes analiticity. Hence we shall restrict to a general perturbation of a well-known Hamiltonian system. Namely, we study the dynamics on a torus originated by a high-frequency perturbation of the pendulum.

The size of the splitting of such a perturbation is given up to order one by the Melnikov function. In [1] the value of the splitting is shown to be exponentially small with respect to ε provided that the perturbation's amplitude is small enough with respect to ε. We give a similar result even when the perturbation exists in a strip whose width is logarithmic with respect to ε (see Section 4.2).

4.1 High-Frequency Perturbations of an Ordinary Pendulum

Consider a high-frequency perturbation of the pendulum described by the Hamiltonian function

$$\frac{\omega \cdot I}{\varepsilon} + h(x, y, \theta, \varepsilon), \tag{4.1}$$

where

$$\omega \cdot I = \omega_1 I_1 + \omega_2 I_2, \quad \text{and} \quad h(x, y, \theta, \varepsilon) = \frac{y^2}{2} + \cos x + \varepsilon^p m(\theta_1, \theta_2)\cos x, \tag{4.2}$$

with symplectic form $dx \wedge dy + d\theta_1 \wedge dI_1 + d\theta_2 \wedge dI_2$. We assume that ε is a small positive parameter and p is a positive parameter. We also assume that the frequency is of the form ω/ε, where

$$\omega = (1, \gamma), \quad \text{and} \quad \gamma = \frac{1 + \sqrt{5}}{2}.$$

The number γ is the golden mean number, which is the "most irrational" number. The reason why we choose this particular frequency is merely pragmatic. Indeed, as is well-known, γ is closely related to the *Fibonacci sequence*. This fact will allow us to obtain certain estimates that would be more difficult to infer otherwise (see Section 4.2.1).

The function m is assumed to be a $2\pi-$periodic function of two variables θ_1 and θ_2. Thus, it can be represented as a Fourier series

$$m(\theta_1, \theta_2) = \sum_{k_1, k_2 \in \mathbb{Z}} m_{k_1,k_2} e^{i(k_1\theta_1 + k_2\theta_2)}. \tag{4.3}$$

We assume that, for some positive numbers r_1 and r_2,

$$\sup_{k_1,k_2\in\mathbb{Z}} \left| m_{k_1,k_2} e^{r_1|k_1| + r_2|k_2|} \right| < \infty, \tag{4.4}$$

and that there are positive numbers a and k_0 such that

$$|m_{k_1,k_2}| \geq a e^{-r_1|k_1| - r_2|k_2|}, \tag{4.5}$$

for all $|k_1| / |k_2|$, which are continuous fraction convergents of γ with $|k_2| \geq k_0$. In fact, k_1 and k_2 are consecutive Fibonacci numbers, that is

$$k_1 = \pm F_{n+1} \quad \text{and} \quad k_2 = \mp F_n.$$

Fibonacci numbers are defined by the recurrence

$$F_0 = 1, \quad F_1 = 1 \quad \text{and} \quad F_{n+1} = F_n + F_{n-1}, \quad \text{for all} n \geq 1.$$

We call the corresponding terms in the perturbation to be *resonant* or *Fibonacci terms.*

The upper bound (4.4) implies that the function m is analytic on the strip $\{|\mathrm{Im}\,\theta_1| < r_1\} \times \{|\mathrm{Im}\,\theta_2| < r_2\}$. Equation (4.5) implies that this function cannot be prolonged analytically onto a larger strip. Let us choose $\alpha \in (0, 1]$. Estimate (4.4) implies that

$$|m(\theta_1, \theta_2)| \leq K\varepsilon^{-2\alpha} \tag{4.6}$$

on the strip

$$\{|\mathrm{Im}\,\theta_1| \leq r_1 - \varepsilon^{\alpha}\} \times \{|\mathrm{Im}\,\theta_2| \leq r_2 - \varepsilon^{\alpha}\}. \tag{4.7}$$

Formula (4.5) implies that the upper bound in (4.6) cannot be improved. The value of the splitting depends essentially on the width of these strips.

The Hamiltonian function defined in (4.1) can be considered as a singular perturbation of the pendulum

$$h_0 = \frac{y^2}{2} + \cos x. \tag{4.8}$$

The unperturbed system has a saddle point $(0, 0)$ and a homoclinic trajectory given by

$$x_0(t) = 4\arctan(e^t), \quad \text{and} \quad y_0(t) = \dot{x}_0(t). \tag{4.9}$$

The complete system given by Hamiltonian (4.1) has a whiskered torus $\mathcal{T}\colon (0, 0, \theta_1, \theta_2)$. The whiskers are three-dimensional hypersurfaces in the four-dimensional extended phase space $(x, y, \theta_1, \theta_2)$. These invariant manifolds are close to the unperturbed pendulum separatrix.

For $p > 3$ and small $\varepsilon > 0$, the invariant manifolds split. In fact, the value of the splitting is predicted by the Melnikov function

$$M(\theta_1, \theta_2; \varepsilon) = \int_{\infty}^{\infty} \{h_0, h\} \left(x_0(t), y_0(t), \theta_1 + \frac{t}{\varepsilon}, \theta_2 + \gamma\frac{t}{\varepsilon} \right) \, dt, \tag{4.10}$$

which gives a first-order approximation of the difference between the values of the unperturbed pendulum energy h_0 on the stable and unstable manifolds.

4.2 Size of the Melnikov function in a Strip of Logarithmic Size With Respect to ε

As is know, the size of the splitting of separatrices when perturbing a Hamiltonian system is given up to order one by the Melnikov function. Hence, an accurate estimate of its size becomes principal in order to show that the splitting occurs. Here we consider that the perturbative function m defined in (4.2) is analytic in a strip $\{|\mathrm{Im}(\theta_1)| < r_1\} \times \{|\mathrm{Im}(\theta_2)| < r_1\}$, where

$$r_i = b_i \log \frac{1}{\varepsilon}, \qquad \text{for } i = 1, 2.$$

The result that we obtain is the following

Proposition 4.2.1. *The maximum of the modulus of the Melnikov function*

$$\max_{(\theta_1,\theta_2)\in\mathbb{T}^2} |M(\theta_1, \theta_2; \varepsilon)|, \tag{4.11}$$

taken on real arguments, can be bounded from above and from below by terms of the form

$$const\ \varepsilon^{p-1} \exp\left(-\sqrt{\frac{-\log\varepsilon}{\varepsilon}}\, c(\log(-\varepsilon\log\varepsilon))\right)$$

with different $\varepsilon-$independent constants, where the function c in the exponent is defined by

$$c(\delta) = C_0 \cosh\left(\frac{\delta - \delta_0}{2}\right), \quad \textit{for } \delta \in [\delta_0 - \log\gamma, \delta_0 + \log\gamma], \tag{4.12}$$

where

$$C_0 = \sqrt{2\pi C_F(\gamma b_1 + b_2)}, \quad C_F = \frac{1}{\gamma + \gamma^{-1}}, \quad \delta_0 = \log\varepsilon^* \quad \textit{and} \quad \varepsilon^* = \frac{\pi(\gamma + \gamma^{-1})}{2\gamma^2(b_1\gamma + b_2)},$$

and continued by $2\log\gamma-$periodicity onto the whole real axis. The function is piecewise analytic and continuous.

We devote the rest of this chapter to proving this proposition. Section 4.2.1 presents some results on the approximation of the golden number γ by rational numbers and Section 4.2.2 gives upper bounds that are used to justify the size of the Fourier coefficients of the Melnikov function.

4.2.1 Rational Approximation of the Golden Number

The best approximation of the golden number is given in terms of Fibonacci numbers, which is defined by the recurrent formula

$$F_0 = 1, \qquad F_1 = 1 \qquad \text{and} \qquad F_{n+1} = F_n + F_{n-1},$$

for any $n > 1$. It is easy to check that

$$F_{n-1} = \frac{\gamma^n - (-1)^n\gamma^{-n}}{\gamma + \gamma^{-1}} \tag{4.13}$$

and

$$F_n - \gamma F_{n-1} = \frac{(-1)^n}{\gamma^n} = \frac{(-1)^n}{F_n + \gamma^{-1} F_{n-1}}.$$

For large values of n this implies

$$F_n - \gamma F_{n-1} = (-1)^n \frac{C_F}{F_{n-1}} + \mathcal{O}\left(\frac{1}{F_{n-1}^3}\right), \tag{4.14}$$

where

$$C_F = \frac{1}{\gamma + \gamma^{-1}}.$$

The estimate of the following lemma is not sharp, but is sufficient for our purposes. The proof can be found in [1].

Lemma 4.2.2. *If $N \in \mathbb{N}$ is not a Fibonacci number, then*

$$|k - \gamma N| > \frac{\gamma C_F}{N},$$

for any integer k.

4.2.2 Exponentially Small Upper Bounds

The proof of the following lemma is standard, and provides a tool to show Proposition 4.2.1.

Lemma 4.2.3. *Let $F(\theta_1 + s/\varepsilon, \theta_2 + \gamma s/\varepsilon)$ be a $2\pi-$periodic function of the variables θ_1 and θ_2 that is analytic in the product of strips $|\mathrm{Im}(\theta_1)| \le r_1$, $|\mathrm{Im}(\theta_1)| \le r_2$ and $|\mathrm{Im}(s)| \le \rho$, where $r_i = b_i \log \frac{1}{\varepsilon}$ for $i = 1, 2$. Assume that $|F| \le A$ for these values of the variables. Then,*

$$|F_{k_1,k_2}| \le A \exp\left(\log \varepsilon (|k_1|\, b_1 + |k_2|\, b_2) - \rho \frac{|k_1 + \gamma k_2|}{\varepsilon}\right).$$

Consider now the $2\log\gamma-$periodic function $c_{\rho,b_1,b_2}(\delta)$ defined on the interval $[\log\varepsilon^* - \log\gamma, \log\varepsilon^* + \log\gamma]$ by

$$c_{\rho,b_1,b_2}(\delta) = C_0 \cosh\left(\frac{\delta - \log\varepsilon^*}{2}\right), \tag{4.15}$$

where

$$C_0 = 2\sqrt{\frac{(\gamma b_1 + b_2)\rho}{\gamma + \gamma^{-1}}} \qquad \text{and} \qquad \varepsilon^* = \frac{\rho(\gamma + \gamma^{-1})}{\gamma^2(b_1\gamma + b_2)},$$

and continued by $2\log\gamma-$periodicity. The following lemma gives the exponentially small upper bound for the function F for real values of the variables.

Lemma 4.2.4. *Let F satisfy the condition of Lemma 4.2.3. If γ is the golden mean number*

$$\gamma = \frac{1 + \sqrt{5}}{2}$$

and the mean value of the function F is zero, then

$$|F(\theta_1,\theta_2)| \le const\, A \exp\left(-\sqrt{\frac{-\log\varepsilon}{\varepsilon}}c_{\rho,b_1,b_2}(\log(-\varepsilon\log\varepsilon))\right) \tag{4.16}$$

on the real values of its arguments. The constant depends continuously on b_1 and b_2 on $b_1 > 0$ and $b_2 > 0$.

Proof. If the arguments of the function F are real, then

$$\begin{aligned} |F(\theta_1,\theta_2)| &\le \sum_{|k_1|+|k_2|\neq 0} |F_{k_1,k_2}| \\ &\le A \sum_{|k_1|+|k_2|\neq 0} \exp\left(\log\varepsilon(|k_1|\,b_1 + |k_2|\,b_2) - \frac{\rho\,|k_1+\gamma k_2|}{\varepsilon}\right), \end{aligned} \tag{4.17}$$

due to the estimate in Lemma 4.2.3. In order to estimate the last sum in (4.17)we separate it into two parts. The first one contains non-resonant terms, that is all the terms such that $|k_1+\gamma k_2| \ge 1/2$. We then obtain the upper estimate

$$\begin{aligned} &\sum_{|k_1+\gamma k_2|\ge 1/2} \left(\log\varepsilon(|k_1|\,b_1 + |k_2|\,b_2) - \frac{\rho\,|k_1+\gamma k_2|}{\varepsilon}\right) \\ &< \exp\left(-\frac{\rho}{2\varepsilon}\right) \sum_{|k_1|+|k_2|\neq 0} \exp\left(\log\varepsilon(|k_1|\,b_1 + |k_2|\,b_2)\right) \\ &= \frac{2(e^{b_1\log\varepsilon} + e^{-b_2\log\varepsilon} - e^{(b_1+b_2)\log\varepsilon})e^{-\frac{\rho}{2\varepsilon}}}{(1-e^{b_1\log\varepsilon})(1-e^{b_2\log\varepsilon})}. \end{aligned} \tag{4.18}$$

For the resonant terms we have $|k_1+\gamma k_2| < 1/2$. Obviously, for every k_2 there exists exactly one integer $k_1 = k_1(k_2)$ such that this inequality holds. Since the coefficients of the sum (4.17) are even with respect to (k_1,k_2) we can assume that k_2 is positive and multiply the estimates by 2 in the end. The sum of the resonant terms with $k_2 \ge \varepsilon^{-1}$ can be estimated as

$$\begin{aligned} &\sum_{k_2\ge\varepsilon^{-1}} \exp\left(\log\varepsilon(|k_1|\,b_1 + |k_2|\,b_2) - \frac{\rho\,|k_1+\gamma k_2|}{\varepsilon}\right) \\ &\le \sum_{k_2\ge\varepsilon^{-1}} \exp(\log\varepsilon(|k_1|\,b_1 + |k_2|\,b_2)) \\ &\le \sum_{k_2\ge\varepsilon^{-1}} \exp\left(-\log\varepsilon\left(\frac{b_1}{2} - (\gamma b_1 + b_2)k_2\right)\right) \\ &\le \frac{e^{-\frac{b_1}{2}\log\varepsilon}e^{(\gamma b_1+b_2)\frac{\log\varepsilon}{\varepsilon}}}{1-e^{(\gamma b_1+b_2)\log\varepsilon}}. \end{aligned} \tag{4.19}$$

We now estimate the resonant terms with $1 \le k_2 < \varepsilon^{-1}$. The number of such terms is large, but finite. We shall show that all of them, except for at most 4, can be estimated by $\mathcal{O}(e^{-\sqrt{\frac{-\log\varepsilon}{\varepsilon}}C_1})$, where

$$C_1 > \max_\delta c_{\rho,r_1,r_2}(\delta).$$

Let B denote the following expressions from the exponent of the right-hand side of (4.17), obtained after substituting $|k_1| = \gamma k_2$:

$$B(k_2, \varepsilon) := -\log \varepsilon(\gamma b_1 + b_2)k_2\sqrt{\varepsilon} + \frac{\rho\,|k_1 + \gamma k_2|}{\sqrt{\varepsilon}}.$$

It is sufficient to provide an appropriate lower bound for this function.

If k_2 is not a Fibonacci number, then by Lemma 4.2.2, we obtain

$$\begin{aligned}
B(k_2, \varepsilon) &\geq -\log \varepsilon(\gamma b_1 + b_2)k_2\sqrt{\varepsilon} + \frac{\rho\gamma C_F}{k_2\sqrt{\varepsilon}} \\
&\geq 2\sqrt{-\log \varepsilon(\gamma b_1 + b_2)\rho\gamma C_F} \\
&= \gamma C_0\sqrt{-\log \varepsilon} \\
&\equiv C_1\sqrt{-\log \varepsilon}.
\end{aligned}$$

If k_2 is a Fibonacci number, then we use (4.14), so

$$|k_1 + \gamma k_2| = \frac{1}{|k_1 + \gamma^{-1}k_2|} \geq \frac{1}{\gamma k_2 + 1 + \gamma^{-1}k_2} = \frac{C_F}{k_2 + C_F}$$

and we obtain

$$B(k_2, \varepsilon) \geq -\log \varepsilon(\gamma b_1 + b_2)k_2\sqrt{\varepsilon} + \frac{\rho C_F}{(k_2 + C_F)\sqrt{\varepsilon}}.$$

Provided that ε is small, $0 < \varepsilon < \varepsilon_0$, there are two positive numbers K_1 and K_2 such that the right hand side of the last inequality is larger than C_1 for k_2 outside of the interval $(K_1/\sqrt{\varepsilon}, K_2/\sqrt{\varepsilon})$. Moreover, this interval contains at most two Fibonacci numbers, which means that the inequality

$$B(k_2, \varepsilon) \geq \sqrt{-\log \varepsilon}\, C_1 \tag{4.20}$$

holds for all except for at most two terms. For these exceptional terms, we have

$$B(k_2, \varepsilon) \geq -\log \varepsilon(\gamma b_1 + b_2)k_2\sqrt{\varepsilon} + \frac{\rho C_F}{k_2\sqrt{\varepsilon}} - \sqrt{\varepsilon}\frac{\rho C_F^2}{K_1(K_1 + C_F\sqrt{\varepsilon})},$$

and it is convenient to rewrite

$$B(k_2, \varepsilon) \geq \sqrt{-\log \varepsilon}\, C_0 \cosh\left(\log(k_2\sqrt{\varepsilon}) - \log\sqrt{\frac{\rho C_F}{\gamma r_1 + r_2}}\right) - \mathcal{O}(\sqrt{\varepsilon}).$$

The above $\mathcal{O}(\sqrt{\varepsilon})$ term affects only the constant in front of estimate (4.16), since the terms in the sum of (4.17) are of the form $\exp(-B(k_2, \varepsilon)/\sqrt{\varepsilon})$. Since k_2 is a Fibonacci number, we have $k_2 = F_n$ for some n and taking into account (4.13), we obtain

$$B(k_2, \varepsilon) \geq \sqrt{-\log \varepsilon}\, C_0 \cosh\left(\frac{1}{2}\log \varepsilon + n\log\gamma + \log\frac{\gamma + 1}{\gamma + 2} - \log(k_2\sqrt{\varepsilon}) - \log\sqrt{\frac{\rho C_F}{\gamma r_1 + r_2}}\right) - \mathcal{O}(\sqrt{\varepsilon}).$$

The envelope of this family of curves is the function $c_{\rho,b_1,b_2}(\delta)$ defined by equation (4.15). Thus, in the sum of the resonant terms there is one leading term which is exponentially larger than the others except in the neighbourhoods of $\varepsilon = \varepsilon^* \gamma^n$, when the index of the leading term changes, and there are two terms of the same order. Moreover, we have established that for all resonant terms with $k_2 < \varepsilon^{-1}$,

$$B(k_2, \varepsilon) \geq c_{\rho,b_1,b_2}(\log(-\varepsilon \log \varepsilon)) - \mathcal{O}(\sqrt{\varepsilon}).$$

Together with estimates (4.18), (4.19) and (4.20) completes the proof. □

4.2.3 Recapitulation

The subsequent proof has two main parts. The first one consists in computing the exact expression for the Fourier coefficients of the Melnikov function in (4.10). In order to bound these coefficients, we take into account the estimate in (4.4) for the coefficients of the perturbation m and notice that the coefficients of the Melnikov function with the biggest size correspond to those whose indices are consecutive Fibonacci numbers. The Fourier coefficients that are not related to Fibonacci numbers can be estimated to be exponentially small with respect to the Fibonacci ones for small values of ε. This can be shown by repeating the proof of Lemma 4.2.4.

Proof of Proposition 4.2.1. Taking into account the explicit formula (4.9) for $x_0(t)$ and $y_0(t)$ we easily obtain that

$$\begin{aligned} M(\theta_1, \theta_2; \varepsilon) &= \varepsilon^p \int_{-\infty}^{\infty} y_0(t) \sin(x_0(t)) m \left(\theta_1 + \frac{t}{\varepsilon}, \theta_2 + \gamma \frac{t}{\varepsilon} \right) \, dt \\ &= -\varepsilon^p \int_{-\infty}^{\infty} \frac{4 \sinh t}{\cosh^3 t} m \left(\theta_1 + \frac{t}{\varepsilon}, \theta_2 + \gamma \frac{t}{\varepsilon} \right) \, dt. \end{aligned}$$

Then, the Fourier coefficients of the Melnikov function are given by

$$M_{k_1,k_2}(\varepsilon) = \left(-\varepsilon^p \int_{-\infty}^{\infty} \frac{4 \sinh t}{\cosh^3 t} e^{i(k_1+\gamma k_2)\frac{t}{\varepsilon}} \, dt \right) m_{k_1,k_2}.$$

Computing the integral by residues we obtain

$$M_{k_1,k_2}(\varepsilon) = -\frac{2\pi i \varepsilon^p (k_1 + \gamma k_2)^2}{\varepsilon^2 \sinh\left(\frac{\pi(i_1+\gamma k_2)}{2\varepsilon}\right)} m_{k_1,k_2}. \tag{4.21}$$

Taking into account the bound in (4.4), we have that for ε small, the coefficients in (4.21) have the same order as

$$\text{const } \varepsilon^{p-2}(k_1 + \gamma k_2)^2 \exp\left(-r_1 |k_1| - r_2 |k_2| - \frac{\pi(k_1 + \gamma k_2)}{2\varepsilon} \right).$$

Since the convergents of γ are quotients between successive terms of the Fibonacci sequence, the most resonant terms correspond to $|k_1| = F_{n+1}$ and $|k_2| = F_n$, so that

$$-r_1 |k_1| - r_2 |k_2| = -r_1 F_{n+1} - r_2 F_n = -(r_1 \gamma + r_2) F_n.$$

Therefore, we can bound these coefficients from below and above by terms of the form

$$\text{const}\ \frac{\varepsilon^{p-2}}{F_n^2} \exp\left(-\frac{\pi C_F}{2\varepsilon F_n} - (r_1\gamma + r_2)F_n\right), \quad r_i = b_i \log\frac{1}{\varepsilon}. \tag{4.22}$$

Let us take $A = \dfrac{\pi C_F}{2}$ and $B = (b_1\gamma + b_2)$. Then, the argument of the exponential in (4.22) can be written as

$$\begin{aligned}
-\frac{\pi C_F}{2\varepsilon F_n} - (r_1\gamma + r_2)F_n &= -\sqrt{\frac{-AB\log\varepsilon}{\varepsilon}}\left(\sqrt{\frac{A}{-B\varepsilon\log\varepsilon}}\frac{1}{F_n} + \sqrt{\frac{-B\varepsilon\log\varepsilon}{A}}F_n\right) \\
&= -2\sqrt{\frac{-AB\log\varepsilon}{\varepsilon}}\cosh\left(\log\left(\sqrt{\frac{A}{-B\varepsilon\log\varepsilon}}\frac{1}{F_n}\right)\right) \\
&= -2\sqrt{\frac{-AB\log\varepsilon}{\varepsilon}}\cosh\left(\frac{1}{2}\log\frac{A}{B} - \frac{1}{2}\log\left(-\varepsilon\log\varepsilon F_n^2\right)\right).
\end{aligned}$$

Calling now

$$C_0 = 2\sqrt{AB} = \sqrt{2\pi C_F(b_1\gamma + b_2)} \quad \text{and} \quad \phi_0 = \frac{A}{B} = \frac{\pi C_F}{2(b_1\gamma + b_2)},$$

we finally obtain that the exponents in (4.22) can be written as

$$-C_0\sqrt{\frac{-\log\varepsilon}{\varepsilon}}\cosh\left(\frac{1}{2}\log\left(-\varepsilon\log\varepsilon F_n^2\right) - \frac{1}{2}\log\phi_0\right).$$

For a fixed $0 < \varepsilon \ll 1$, the largest coefficient corresponds to the minimal value of the hyperbolic cosine or, equivalently, minimises

$$\left|\log(-\varepsilon\log\varepsilon F_n^2) - \log\phi_0\right|.$$

This happens when F_n is closest to

$$F^*(\varepsilon) = \sqrt{\frac{\phi_0}{-\varepsilon\log\varepsilon}}. \tag{4.23}$$

Let us denote the closest number to $F^*(\varepsilon)$ by $F_{n(\varepsilon)}$. We have that the index of the biggest term in the Fourier series of the Melnikov function grows as

$$\sqrt{\frac{1}{-\varepsilon\log\varepsilon}},$$

and it changes when $F^*(\varepsilon)$ crosses the value given by (4.23). The largest terms correspond to

$$(k_1, k_2) = \pm(F_{n(\varepsilon)+1}, -F_{n(\varepsilon)}).$$

Except when $F^*(\varepsilon)$ lies exactly in the centre of an interval $[F_n, F_{n+1}]$, there is only one Fibonacci number closest to $F^*(\varepsilon)$ and then, the two terms

$$(F_{n(\varepsilon)+1}, -F_{n(\varepsilon)}) \qquad \text{and} \qquad (-F_{n(\varepsilon)+1}, F_{n(\varepsilon)})$$

dominate the Fourier series. Otherwise, there are four dominating terms.

We now use that $F_n = C_F(\gamma^{n+1} + (-1)^n \gamma^{-n-1})$ in order to write

$$\begin{aligned}\log(-\varepsilon \log \varepsilon F_n^2) &= \log(-\varepsilon \log \varepsilon) + 2 \log C_F + \log(\gamma^{n+1}(1 + (-1)^n \gamma^{-2(n+1)}))^2 \\ &= \log(-\varepsilon \log \varepsilon) + 2 \log C_F + 2(n+1) \log \gamma + 2 \log(1 + (-1)^n \gamma^{-2(n+1)}).\end{aligned} \tag{4.24}$$

When we simultaneously increase n by 1 and decrease $\log(-\varepsilon \log \varepsilon)$ by $2 \log \gamma$, the value of (4.24) is repeatedly reached, up to order $\mathcal{O}(\gamma^{-2n-2}) = \mathcal{O}(F_n^{-2}) = \mathcal{O}(\varepsilon \log \varepsilon)$. Thus we obtain that the Fourier coefficients $M_{k_1,k_2}(\varepsilon)$ given in (4.21) can be estimated from below and from above by

$$\text{const } \varepsilon^{p-1} \exp\left(-\sqrt{\frac{-\log \varepsilon}{\varepsilon}} c(\log(-\varepsilon \log \varepsilon)\right),$$

where the function c is the one defined in (4.12). □

Bibliography

[1] A. Jorba A. Delshams V. Gelfreich and T. Martínez-Seara. "Exponentially small splitting of separatrices under fast quasiperiodic forcing". In: *Communications in Mathematical Physics* 189.1 (1997), pp. 35–71.

[2] V. I. Arnol'd. "Instability of Dynamical Systems with Several Degrees of Freedom". In: *Soviet Mathematics-Dodlaky* 5 (1964), pp. 633–636.

[3] H. Scott Dumas. *The KAM Story. A Friendly Introduction to the Content, History, and Significance of Classical Kolmogorov-Arnold-Moser Theory.* World Scientific Pulbishing Co., 2014.

[4] H. Flanders. *Differential Forms with Applications to Physical Sciences,* New York: Academic Press, 1963.

[5] E. Fontich and P. Martín. "Differentiable invariant manifolds for partially hyperbolic tori and a Lambda lemma". In: *Nonlinearity* 13.5 (2000), pp. 1561–1593.

[6] *Foundation of Mechanics.* London.: Benjamin-Cummings.

[7] A. N. Kolmogorov. "On conservation of conditionally periodic motions under small perturbations of the Hamiltonian". In: *Doklady Akademii Nauk SSSR* 4 (1954), pp. 527–530.

[8] A. I. Neishtadt. "The separation of motions in systems with rapidly rotating phase". In: *Journal of Applied Mathematics and Mechanics* 48.2 (1984), pp. 133–139.

[9] M. Spivak. *Calculus on Manifolds.* New York: W. A. Benjamin, 1965.

[10] L. Sirovich S.S. Antman J.E. Marsden, ed. *Introduction to Hamiltonian Dynamical Systems and the N-Body Problem.* Second Edition. Applied Mathematical Sciences 90. Springer, 2009.

[11] V. V. Kozlov V. I. Arnol'd and A. I. Neishtadt. *Dynamical Systems.* Ed. by R. V. Gamkrelidze. Second Edition. Vol. III. Encyclopaedia of Mathematical Sciences. Springer-Verlag, 1993.