Integrability in Molecular Dynamics: Investigating the Jellinek-Berry Thermostat via KAM Theory and Normal Forms.

by

Alireza Sharifi

A thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements of the degree of

Master of Science

Department of Mathematics

University of Manitoba

Winnipeg

Abstract

In general, Hamiltonian Thermostats, as proposed in the model by Jellinek and Berry in the paper [JB88], are derived from the Hamiltonian $H = H(p, q)$ of a mechanical system using some elementary constructions. This thesis is dedicated to studying the properties of the Jellinek–Berry (JB) thermostat applied to an ideal gas.

An ideal gas is a mechanical system with zero potential energy. In an idealized gas, atoms or molecules do not interact, so they only possess kinetic energy. If, as is usually assumed, the kinetic energy is just the euclidean squared-norm of momenta, then the Hamiltonian is

$$H(p, q) = \frac{1}{2}|p|^2$$

where $q \in M$, $p \in T_q^*(M)$ and M is a flat manifold,. This Hamiltonian is completely integrable because p is constant along the solution (conservation of momentum). Two major findings are:

- If $G(s, p_s)$ is the internal energy of the JB thermostat, then the thermostatted Hamiltonian $F_\epsilon(q, p, s, p_s) = H_\epsilon(q, \alpha(s)p) + G(s, p_s)$ for some scalar function $\alpha(s)$, is completely integrable when H is the total energy of an ideal gas.

- If H_ϵ is a small, real-analytic perturbation of the ideal-gas Hamiltonian $H = H_0$ sufficient conditions are determined that imply the existence of positive-measure sets of invariant tori for F_ϵ.

Preface

The present thesis embodies an interlaced structure of reviewed and newly conducted research in the domain of Hamiltonian Thermostats. The approach towards unfolding the intricacies of the subject matter is dichotomous - the first half provides the essential theoretical underpinnings, while the second half offers a novel perspective with original contributions.

The initial chapters, chapters 1 to 4, serve as an exhaustive exposition of the preexisting knowledge in the field. These chapters not only provide a comprehensive examination of the pertinent literature but also function as a theoretical grounding for the subsequent sections of this work. The content herein is largely derived from established texts and notable research papers, the references for which have been diligently cited to maintain the academic integrity. The primary objective of these sections is to elucidate the fundamental concepts and tools which serve as the bedrock of the subject matter, thereby paving the way for the subsequent, more advanced discourse.

The latter half (chapters 5 and 6) of the thesis marks a transformation from existing theories to the exploration of uncharted territories within the realm of Hamiltonian Thermostats. This portion of the thesis is primarily composed of original research and novel contributions, developed by leveraging the foundational theories elucidated in the preceding chapters. Herein, the focus shifts towards harnessing the acquired knowledge to solve intricate problems, and explore potential applications. This divergence from the well-trodden path is aimed at enriching the corpus of knowledge in the field and proposing new directions for future research.

In synthesizing established literature with original ideas, this thesis aims to offer a comprehensive yet innovative approach to understanding Hamiltonian Thermostats. It is my sincere hope that this approach not only provides clarity to the reader but also stimulates further exploration in this fascinating field.

Introduction

This thesis delves into the generalized Nosé-Hoover Hamiltonian system proposed by Julius Jellinek and R. Stephen Berry[JB88] in 1988, focusing specifically on an ideal gas (zero potential energy) scenario. We commence our study with a comprehensive review of the concept of canonical transformation, a technique that facilitates the transition from given coordinates to action-angle coordinates.

The initial section of this thesis engages in a comprehensive exploration and proof of the Liouville-Arnol'd theorem, a critical instrument in determining the integrability of the Jellinek–Berry thermostat (henceforth, JB thermostat) around the equilibrium point in action-angle variables. Construction of action-angle variables is generally quite challenging, but we resort to the computation of a type of Taylor expansion, the Birkhoff normal form, to help avoid the difficulties. Consequently, this system is completely solvable.

In the ensuing section, we set out to demonstrate that the trajectories of the system in phase space are conditionally periodic, occupying invariant tori, leading to a dense infilling of the phase space by these tori.

As our investigation advances, we pivot our focus towards an essential concept - normal forms. The process of normalization to discover the Birkhoff normal form of a simple harmonic oscillator is examined. Upon the inclusion of quadratic terms and generalized cubic terms, I compute the Birkhoff normal form to order 2 and determine the lowest-order terms in the system's frequency. This intricate process involves a sequence of symplectic structure-preserving transformations through which we demonstrate the normal-form algorithm on a simple Hamiltonian that is the sum of a harmonic oscillator and an cubic potential energy term.

The renowned Kolmogorov–Arnol'd–Moser (KAM) theorem is then introduced to gain insights into the fundamental structure and qualitative behavior of trajectories in the phase space. I then analyze the JB-thermostated Hamiltonian, first by reduction to a single degree of freedom system, and then the unreduced (two degree of freedom) system, within each of which conditions are established. If the Nosé–Hoover Hamiltonian is perturbed by a smooth function under these conditions, the majority of invariant KAM tori will persist, ensuring the system's stability.

Acknowledgements

In the quest of knowledge, one is never alone, and I have been incredibly fortunate to embark on this journey under the guidance of my advisor, Dr. Leo T. Butler.

His profound knowledge and unwavering belief in my capabilities have been the beacon that lighted my path through the complex maze of Hamiltonian Thermostats. Every discussion with him was an enlightening experience, unraveling the intricacies of the subject and broadening my perspective. His uncompromising dedication to excellence in research have been nothing short of inspirational.

His patience and understanding have given me the freedom to explore, the courage to challenge, and the resilience to persevere. His mentorship has been far more than academic; he has imparted lessons of critical thinking, professional integrity, and scholarly passion that will remain with me throughout my life.

For his invaluable guidance, and for his generous support, I offer my deepest gratitude. Thank you, Dr. Leo T. Butler, for making this journey an enriching and rewarding experience. This work is a testament to your mentorship, and I am proud to be your student.

Contents

List of Figures

1

Theory of Manifolds

This chapter reviews background material that can be found in the following textbooks:

- Geometrical methods in the theory of ordinary differential equations, Vladimir Arnol'd. [Arn88].
- Introduction to Smooth Manifolds Book by John M. Lee [Lee13].
- Mathematical Methods of Classical Mechanics Textbook by Vladimir Arnol'd [Arn89].

1.1 Smooth Manifolds

We start off the chapter by reviewing the concept of smooth manifold as it is used throughout the following materials.

Recall that a topological space is a set with a distinguished collection of subsets, called open sets. This collection contains the whole set, the empty set; it is closed under finite intersection and arbitrary unions.

Definition 1.1.1. (Topological Manifold)

Suppose M is a topological space. We say that M is a **topological manifold** of dimension n or a topological n-manifold if it has the following properties:

1. M is a Hausdorff space.

2. M is second-countable: there exists a countable basis for the topology of M.

3. M is locally Euclidean of dimension n: each point of M has a neighborhood that is homeomorphic to an open subset of $\mathbb{R}^n$.

For the third property, more specifically it means around each point $p \in M$ we can find:

- an open subset $U \subset M$ containing p,

- an open subset $\tilde{U} \subset \mathbb{R}^n$,

- a homeomorphism $\phi : U \longrightarrow \tilde{U}$.

Each (U, ϕ) is called a chart.

On top of topological properties we can equip the manifold with a smooth structure in order to also talk about the notion of differentiability on manifolds.

Definition 1.1.2. (Transition maps)

Let M be a topological n-manifold. If (U, ϕ) and (V, ψ) are two charts such that $U \cap V \neq \varnothing$, the composite map $\psi \circ \phi^{-1} : \phi(U \cap V) \to \psi(U \cap V)$ is called the **transition map** from ϕ to ψ. A transition map basically is a change of coordinates.

Here is worth to mention that one of the most surprising results is that there are topological manifolds with nondiffeomorphic smooth structures.

Definition 1.1.3. (Smoothly compatible charts)

Two charts (U, ϕ) and (V, ψ) are said to be **smoothly compatible** if either $U \cap V = \varnothing$ or the transition map $\psi \circ \phi^{-1}$ is a diffeomorphism.

Definition 1.1.4. (Atlas)

We define an atlas for M to be a collection of charts whose domains cover M. An atlas A is called a **smooth atlas** if any two charts in A are smoothly compatible with each other. A smooth atlas A on M is maximal if it is not properly contained in any larger smooth atlas.

Now we can define the main concept of this chapter.

Definition 1.1.5. (Smooth Manifold)

If M is a topological manifold, a smooth structure on M is a maximal smooth atlas. A **smooth manifold** is a pair (M, A), where M is a topological manifold and A is a smooth structure on M.

1.2 Tangent and Cotangent bundle

The background material in this section can be found in *Introduction to Smooth Manifolds* by John M. Lee [Lee13].

Definition 1.2.1. Given a function $f : M \longrightarrow \mathbb{R}^k$ and a chart (U, ϕ) for M; the function $\hat{f} : \phi(U) \longrightarrow \mathbb{R}^k$ defined by $\hat{f}(x) = f \circ \phi^{-1}(x)$ is called the **coordinate representation of** f. By definition, f is smooth if and only if its coordinate representation is smooth in some smooth chart around each point.

It can be shown that smooth functions have smooth coordinate representations in every smooth chart.

Definition 1.2.2. We indicate $\mathbf{C}^\infty(\mathbf{M})$ as a set of all smooth maps $f : M \longrightarrow \mathbb{R}$.

Definition 1.2.3. (Tangent space to M)

Let M be a smooth manifold and let $p \in U$ be a point of M and (U, x^i) be a coordinate neighborhood around the point p. A linear map $v : C^\infty(M) \longrightarrow \mathbb{R}$ is called a derivation at p if it satisfies

$$v(fg)(p) = f(p)vg(p) + g(p)vf(p) \qquad \forall f, g \in C^{\infty}(M). \tag{1.2.1}$$

To avoid confusion in above equation, notice that $vf(a) = v_a(f)$ which is identified by an isomorphism. In fact, the map $v_a \mapsto D_v|_a$ defined by

$$D_v|_a f = D_v f(a) = \left.\frac{d}{dt}\right|_{t=0} f(a + tv) = \sum_{i=1}^{n} v^i \frac{\partial f}{\partial x^i}(a) \tag{1.2.2}$$

establishes an isomorphism between $\mathbb{R}^n_a$ onto the tangent space at a to $\mathbb{R}^n$.

Definition 1.2.4. The set of all derivations of $C^{\infty}(M)$ at p, denoted by T_pM, is called the **tangent space** to M at p. An element of T_pM is called a tangent vector at p.

Theorem 1.2.5. *If M is an n-dimensional smooth manifold, then for each $p \in M$ the tangent space T_pM is an n-dimensional vector space.*

To find a better understanding of tangent space, we restrict our attention to $\mathbb{R}^n$.

1.2.1 Coordinate computation of tangent vector

Theorem 1.2.6. *For any $a \in \mathbb{R}^n$, the n derivations*

$$\left.\frac{\partial}{\partial x^1}\right|_a, \ldots, \left.\frac{\partial}{\partial x^n}\right|_a \quad \text{defined by} \quad \left.\frac{\partial}{\partial x^i}\right|_a f = \frac{\partial f}{\partial x^i}(a) \tag{1.2.3}$$

form a basis for $T_a\mathbb{R}^n$.

Above results can be extended to a smooth manifold M. To this end, let (U, ϕ) be a coordinate chart and $\phi : U \to \tilde{U}$ be in particular a diffeomorphism from U to open subset $\tilde{U}$ in $\mathbb{R}^n$.

Therefore, the coordinate computation of tangent vector is as follows,

$$\partial_i|_p = (d\phi_p)^{-1}(\frac{\partial}{\partial x^i}\bigg|_{\phi(p)}) = d(\phi^{-1})_{\phi(p)}(\frac{\partial}{\partial x^i}\bigg|_{\phi(p)}). \tag{1.2.4}$$

Then by invoking the definition of derivation we can evaluate,

$$\frac{\partial}{\partial x^i}\bigg|_p f = \frac{\partial}{\partial x^i}\bigg|_{\phi(p)} (f \circ \phi^{-1}) \tag{1.2.5}$$

which is an euclidean computation.

1.2.2 Tangent Bundle

Given a smooth manifold M with or without boundary, we define the tangent bundle of M, denoted by TM, to be the disjoint union of the tangent spaces at all points of M,

$$TM = \bigsqcup_{p \in M} T_pM. \tag{1.2.6}$$

We usually write an element of this disjoint union as an ordered pair (p, v) with $p \in M$ and $v \in T_pM$. The tangent bundle comes equipped with a natural projection map $\pi : TM \longrightarrow M$, which sends each vector in T_pM to the point p at which it is tangent: $\pi(p, v) = p$.

Theorem 1.2.7. *For any smooth n-manifold M without boundary, the tangent bundle TM has a natural topology and smooth structure that make it into a 2n-dimensional smooth manifold. With respect to this structure, the projection $\pi : TM \longrightarrow M$ is smooth.*

1.2.3 Derivation of a map and rank theorem

Definition 1.2.8. Let M and N be smooth manifolds and $F : M \longrightarrow N$ a smooth map, for each $p \in M$ we define a map

$$dF_p : T_pM \longrightarrow T_{F(p)}N \tag{1.2.7}$$

called the **differential of** F at p and is defined as follows:

Given $v \in T_pM$, we let $dF_p(v)$ be the derivation of F at p that acts on $f \in C^\infty(N)$, then $dF_p(v)(f)$ is defined by the rule

$$dF_p(v)(f) = v(f \circ F). \tag{1.2.8}$$

The rank of F at p is the rank of dF at p.

Theorem 1.2.9 ([Lee13], page 83)**. *(Global rank theorem)***

Let M and N be smooth manifolds of dimensions m and n respectively, and suppose $F : M \longrightarrow N$ is a smooth map of constant rank r. Then for each $p \in M$ there exists smooth charts (U, ϕ) for M centered at p and (V, ψ) for N centered at $F(p)$ such that $F(U) \subset V$, in which F has a coordinate representation of the form

$$\tilde{F}(x_1, \ldots, x_r, x_{r+1}, \ldots, x_m) = (x_1, \ldots, x_r, 0, \ldots, 0). \tag{1.2.9}$$

In particular, if F is a smooth submersion, this becomes

$$\tilde{F}(x_1, \ldots, x_n, x_{n+1}, \ldots, x_m) = (x_1, \ldots, x_n) \tag{1.2.10}$$

and if F is a smooth immersion, it is

$$\tilde{F}(x_1, \ldots, x_m) = (x_1, \ldots, x_m, 0, \ldots, 0) \tag{1.2.11}$$

Remark 1.2.10. Let M and N be smooth manifolds, and suppose $F : M \longrightarrow N$ is a smooth map of constant rank, then one can conclude that if F is bijective then it is a smooth homeomorphism with a smooth inversion, i.e., a diffeomorphism.

1.3 Differential 1-Form df

Consider a manifold M of dimension n, and let $f : M \to \mathbb{R}$ be a function in $C^r(M)$, the set of all r-times continuously differentiable functions from M to $\mathbb{R}$. The derivative of f at a point $p \in M$, denoted by df_p, is a covariant 1-tensor (covector) or 1-form defined by:

$$df_p : T_pM \to \mathbb{R}. \tag{1.3.1}$$

Here, T_pM denotes the tangent space to M at p, and df_p is a linear map that sends a tangent vector $v \in T_pM$ to the directional derivative of f at p along v.

1.4 Differential 1-Forms and Cotangent bundle

For each $p \in M$, the set of all linear maps from T_pM to $\mathbb{R}$ is called the cotangent space at p, denoted by T_p^*M. A differential 1-form ω on M is a smooth assignment of a cotangent vector $\omega_p \in T_p^*M$ to each $p \in M$. The set of all differential 1-forms on M forms a real vector space, denoted by $\Omega^1(M)$. Also the disjoint union

$$T^*M = \coprod_{p \in M} T_p^*M \tag{1.4.1}$$

is called the **cotangent bundle of M**. It has a natural projection map $\pi : T^*M \longrightarrow M$ sending $\omega_p \in T_p^*M$ to $p \in M$. Given any smooth local coordinate $(U, (x^i))$ on M we denote $dx^i|_p$ the basis for T_p^*M dual to $\frac{\partial}{\partial x^i}\big|_p$. This defines n maps $(dx^1, \ldots, dx^n) : U \to (T^*M)|_U$ called **coordinate covector fields.** These one forms trivialize the cotangent bundle of M restricted to U.

Theorem 1.4.1. *Let M be a smooth n-manifold with or without boundary. With its standard projection map and the natural vector space structure on each fiber, the cotangent bundle T^*M has a unique topology and smooth structure making it into a smooth rank-n vector bundle over M for which all coordinate covector fields are smooth local sections.*

1.5 Higher-Order Differential Forms

For an integer $r \geqslant 1$, an r-form at a point $p \in M$ is a multi-linear, skew-symmetric real-valued function on the r-fold product of the tangent space at p. A differential r-form is a smooth section of the r^{th} exterior power of the cotangent bundle which has a natural topology and smooth structure (to be a smooth manifold), meaning that it assigns to each $p \in M$ an r-form ω_p. The set of all differential r-forms on M forms a real vector space, denoted by $\Omega^r(M)$.

A general differential r-form ω on M is a function that for each $p \in M$ takes r vectors $v_1, \ldots, v_r \in T_pM$ and returns a real number, $\omega_p(v_1, \ldots, v_r) \in \mathbb{R}$, such that the following conditions are satisfied:

1. **Multi-linearity:** For each $i = 1, \ldots, r$ and each pair of vectors $u, v \in T_pM$ and scalar $a \in \mathbb{R}$, we have

$$\begin{aligned} &\omega_p(v_1, \ldots, v_{i-1}, au + v, v_{i+1}, \ldots, v_r) \\ &\quad = a\omega_p(v_1, \ldots, v_{i-1}, u, v_{i+1}, \ldots, v_r) + \omega_p(v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_r) \end{aligned} \tag{1.5.1}$$

2. **Anti-symmetric:** For each $i = 1, \ldots, r$, $v_i \in T_pM$ and any permutation σ of v_i we have,

$$\omega(v_{\sigma(1)}, \ldots, v_{\sigma(i)}) = sign(\sigma)\omega(v_1, \ldots, v_r) \quad \text{for } \sigma \in S_r \tag{1.5.2}$$

Every differential k-form on the space T_xM with a given coordinate system $(x_1, \ldots, x_n)$ can be written uniquely in the form

$$\omega^k = \sum_{i_1 < \cdots < i_k} a_{i_1,\ldots,i_k}(x) dx_{i_1} \wedge \cdots \wedge dx_{i_k} \tag{1.5.3}$$

where $a_{i_1,\ldots,i_k}(x)$ are smooth function on $\mathbb{R}^n$.

1.6 Operations on Differential Forms

Given $\omega \in \Omega^p(M)$ and $\eta \in \Omega^q(M)$, their wedge product $\omega \wedge \eta$ is a $(p+q)$-form in $\Omega^{p+q}(M)$ defined by:

$$(\omega \wedge \eta)(v_1, ..., v_p, w_1, ..., w_q) = \frac{1}{(p!q!)} \sum_{\sigma \in S_{p+q}} \operatorname{sign}(\sigma) \omega(v_{\sigma(1)}, ..., v_{\sigma(p)}) \eta(w_{\sigma(p+1)}, ..., w_{\sigma(p+q)}). \tag{1.6.1}$$

The exterior derivative $d : \Omega^r(M) \to \Omega^{r+1}(M)$ is a linear operator defined by extending the following rules to general r-forms:

$$df(v_1) = v_1(f) \quad \text{for } f \in C^\infty(M), v_1 \in T_pM \quad \text{which for the case of } r = 0 \tag{1.6.2}$$

$$d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^r \omega \wedge d\eta \quad \text{for } \omega \in \Omega^r(M), \eta \in \Omega^q(M). \tag{1.6.3}$$

Obviously the case for $r = 0$ we have $\Omega^0(M) = C^\infty(M)$.

1.7 Cotangent Bundle and Natural projection

The background material in this section can be found in the *Mathematical Methods of Classical Mechanics* textbook by Vladimir Arnol'd [Arn89].

Let V be an n-dimensional differentiable manifold. A 1-form on the tangent space to V at a point x is called a cotangent vector to V at x. The set of all cotangent vectors to V at x forms an n-dimensional vector space, dual to the tangent space T_xV. We will denote this vector space of cotangent vectors by T_x^*V and call it the cotangent space to V at x. The union of the cotangent spaces to the manifold at all of its points is called the cotangent bundle of V and is denoted by T^*V. The set T^*V has a natural structure of a differentiable manifold of dimension $2n$. A point of T^*V is a 1-form on the tangent space to V at some point of V. If q is a choice of n local coordinates for points in V, then such a form is given by its n components p. Together, the $2n$ numbers p, q form a collection of local coordinates for points in T^*V. In particular, if $\phi(q) = (x_1, \ldots, x_n)$ is a coordinate chart on $U \subset M$ then $p \in T^*V$ has the expression $p = p_1 dx_1 + \cdots + p_n dx_n$. There is a natural projection $f : T^*V \longrightarrow V$ (sending every 1-form on T_x^*V to the point x). The projection f is differentiable and surjective. The pre-image of a point $x \in V$ under f is the cotangent space T_x^*V.

1.8 Stokes' Theorem

Stokes' theorem is a cornerstone of calculus in multivariate real and complex analysis. In the language of differential forms, it states:

$$\int_{\partial M} \omega = \int_M d\omega. \tag{1.8.1}$$

for a manifold M with boundary ∂M, and ω a differential form of degree $\dim(M)-1$. This theorem unifies the fundamental theorem of calculus, Green's theorem, Gauss's divergence theorem, and the classical Stokes' theorem.

1.9 Transversality

The background material in this section can be found in the book *Geometrical Methods in the Theory of Ordinary Differential Equations* by Vladimir Arnol'd [Arn88]. Another important concept is the notion of transversality in which the **Thom Transversality theorem** plays the main role.

Before the formal statement of the Thom transversality theorem, I introduce the notion of a k-jet and then state the theorem using this language.

1.9.1 K-jets

Definition 1.9.1. Two functions $f_1, f_2 : M \longrightarrow N$ are said to be k-tangent at a point x, when their Taylor series in any fixed coordinate system up to and including the k^{th} term are identical at that point.

Remark 1.9.2. a) The definition above is actually coordinate independent;

b) it defines an equivalence relation on the set of pointed, smooth maps $(M, x) \longrightarrow (N, y)$ whose equivalence classes can be identified with maps that have a same Taylor expansion up to the order k.

Definition 1.9.3. A k-jet of a smooth mapping f at a point x is an equivalence class of k-tangent mappings at x.

We agree on the following notation that

$$j_x^k(f) = \{f_1 : f_1 \text{ is } k\text{-tangent to } f \text{ at } x\}. \tag{1.9.1}$$

Definition 1.9.4. Suppose $f, g : M \longrightarrow N$ are smooth maps. Define the relation $f \sim_x g$ if and only if there is an open neighbourhood U containing x such that $f|_U = g|_U$. It is easy to see that $\sim_x$ is symmetric, reflexive and transitive. A **germ** of a map at $x \in M$ is an equivalence class of $\sim_x$.

Definition 1.9.5. Two mappings in two neighborhoods of one and the same point have a common **germ** at that point if they coincide in a third neighborhood of this point. (The third neighborhood can be smaller than the intersection of the first two neighborhoods).

Definition 1.9.6. The set of all k-jets of germs of C^k mappings of M into N is called the space of k-jets of mappings of M into N and denoted by

$$J^k(M,N) = \{\text{the space of } k\text{-jets of germs of } C^k \text{ mappings of } M \text{ into } N\}. \quad (1.9.2)$$

It is notable that the set $J^k(M,N)$ has a natural smooth manifold structure. Indeed, we choose coordinate systems in the neighborhood of a point of M and in the neighborhood of the image of this point in N under some mapping f. Then the k-jet of f and all nearby jets can be given by coordinates of the preimage and the collections of Taylor coefficients up to order k at this point. Thus, we have constructed a chart of the manifold $J^k(M,N)$ of jets in the neighborhood of the point which is the k-jet of f.

Definition 1.9.7. The group of k-jets of diffeomorphisms of M leaving a point x fixed is called the group of k-jets of local diffeomorphisms of the manifold M at the point x and denoted by $J^k_x(M)$.

We review the notion of transversality by the following definition.

1.9.2 Transversality

Definition 1.9.8. Two linear subspaces X and Y of a linear space L are said to be **transversal** if their sum is the whole space:

$$L = X + Y. \quad (1.9.3)$$

For example, two planes intersecting at a nonzero angle in the three dimensional space are transversal and two straight lines are not.

Definition 1.9.9. Let M and N be smooth manifolds and let Z be a smooth submanifold of N. A mapping $f : M \longrightarrow N$ is said to be **transversal to** Z at a point a of M if either $f(a)$ does not belong to Z or the tangent plane to Z at $f(a)$ and the image of the tangent plane to M at a are transversal, i.e.,

$$d_a f T_a M + T_{f(a)} Z = T_{f(a)} N. \tag{1.9.4}$$

A mapping $f : M \longrightarrow N$ is transversal to Z if it is transversal to Z all point of the M.

Theorem 1.9.10 ([Arn88], page 231)**.** ***(Weak transversality theorem)*** *Let M be a compact manifold and let Z be a compact submanifold in a manifold N. The mappings $f : M \longrightarrow N$ transversal to Z form an open, everywhere dense set in the space of all sufficiently smooth (C^r with $r \geqslant 1$) mappings $M \longrightarrow N$.*

Remark 1.9.11. With regard to the topology of $C^r(M, N)$, on a quick note, one could define a basis open set as follows. Take $f \in C^k(M, N)$, a compact set of $K \subset M$ and an open neighbourhood, W, $J^k(M, N) \supset W \supset j^k(f)(K)$ of the graph of the $k-$jet of f restricted to K. A neighbourhood $\mathcal{W} \subset C^k(M, N)$, $\mathcal{W} = \mathcal{W}(f, K, W)$ is the set of all $g \in C^k(M, N)$ such that $j^k(g)(K) \subset W$.

Remark 1.9.12. This theorem is called the weak transversality theorem. Its assertion means that a mapping not transversal to a fixed submanifold can be turned into a transversal mapping by a small perturbation. If, on the other hand, transversality is present, then it is preserved under C^r-small perturbations.

1.9.3 Thom Transversality Theorem

The Thom transversality theorem is a generalization of the weak transversality theorem, in which the role of the submanifold Z is played by a submanifold of a space of jets.

With every smooth mapping $f : M \longrightarrow N$, we associate its "k-jet extension" $\tilde{f} : M \longrightarrow J^k(M,N)$, $\tilde{f}(x) = j_x^k(f)$. (To a point x of M, there corresponds the k-jet of the mapping f at x.)

Theorem 1.9.13 ([Arn88], page 234)**.** *Let C be a submanifold of the space $J^k(M,N)$ of jets. The set of mappings $f : M \longrightarrow N$ whose k-jet extensions are transversal to C is an everywhere dense countable intersection of open sets in the space to all smooth mappings of M into N.*

1.10 Symplectic structures on manifolds

Definition 1.10.1. Let M^{2n} be an even-dimensional differentiable manifold. A **symplectic structure** on M^{2n} is a closed non-degenerate differential 2-form ω on M^{2n}

$$d\omega = 0 \quad \textit{and} \quad \forall \xi \neq 0, \exists \eta : \omega(\xi,\eta) \neq 0 \qquad (\xi,\eta \in T_xM). \tag{1.10.1}$$

The pair (M^{2n},ω) is called a symplectic manifold. An example of a symplectic manifold is $T^*\mathbb{R}^n \cong \mathbb{R}^{2n}$ with coordinates p_i and q_i with $\omega = \sum_{i=1}^n dp_i \wedge dq_i$.

The following theorem gives a complete allowance to do local calculations in any symplectic manifold.

Theorem 1.10.2 ([Arn89], page 230)**.** ***Darboux's theorem***

Let ω be a closed non-degenerate differential 2-form in a neighborhood of a point x in the space $\mathbb{R}^{2n}$. Then in some neighborhood of x one can choose a coordinate system

$(p_1, \ldots, p_n, q_1, \ldots, q_n)$ *such that* ω *has the standard form:*

$$\omega = \sum_{i=1}^{n} dp_i \wedge dq_i. \tag{1.10.2}$$

1.10.1 Symplectic structure of cotangent space

Now we spend a little bit of time to understand a natural structure of any cotangent bundle of an arbitrary smooth manifold. It turns out that it carries a natural symplectic structure. The following theorem is talking about the formation of such type of structure.

Theorem 1.10.3 ([Arn89], page 202)**.** *Let* V *be a smooth manifold. The cotangent bundle* T^*V *has a natural symplectic structure. In the local coordinates system* $(p, q) = (p_1, \ldots, p_n, q_1, \ldots, q_n)$*, this symplectic structure is given by the formula*

$$\omega = dp \wedge dq = dp_1 \wedge dq_1 + \cdots + dp_n \wedge dq_n. \tag{1.10.3}$$

The following proof is an adaptation of the proof that appears in [Arn89].

Proof. First, we define a distinguished 1-form on T^*V. Let $\xi \in T_p(T^*V)$ be a vector tangent to the cotangent bundle at the point $p \in T^*_xV$. The derivative $df : T(T^*V) \longrightarrow TV$ of the natural projection $f : T^*V \longrightarrow V$ takes ξ to a vector $df(\xi)$, tangent to V at x. We define a 1-form ω^1 on T^*V by the relation $\omega^1(\xi) = p(df(\xi))$. In the local coordinates described above, this form is $\omega^1 = pdq$. Then, obviously the the exterior derivative $d\omega$ is going to be a closed non-degenerate 2-form. □

2

Hamiltonian Mechanic; A classical approach to describe the motion of a mechanical system.

The foundational material presented in this chapter is derived from the seminal work in *Mathematical Methods of Classical Mechanics* by Vladimir Arnol'd [Arn89].

2.1 From Newton's equation to Hamiltonian formalism

Suppose that you have point-mass of mass m that is moving under the influence of a force field F that depends on the position q, the velocity $\frac{dq}{dt}$ and possibly t then the Newton's second law of motion yields the following differential equation:

$$m\frac{d^2q}{dt^2} = F(q, \frac{dq}{dt}, t). \tag{2.1.1}$$

It is interpreted as a sophisticated expression of Newton's second law, $F = ma$.

Equation 2.1.1 is particularly applicable to elementary systems and scenarios in which the forces are explicitly known.

2.1.1 Hamiltonian formalism

Assuming a conservative force field in 2.1.1the equation 2.1.1 can be rewritten to the following by considering the $F = \frac{d}{dt}(m\dot{q})$

$$\frac{d}{dt}(m\dot{q}) = -\frac{\partial U}{\partial q} \tag{2.1.2}$$

for a differentiable potential function U. This implies conservation of total energy. Since the system is isolated this is reasonable. [1].

Now by introducing the variable $p = m\dot{q}$ we rewrite 2.1.2 as the following system of equations

$$\begin{cases} \dot{p} = -\frac{\partial U}{\partial q} \\ \dot{q} = \frac{1}{m}p. \end{cases} \tag{2.1.3}$$

In fact, given a smooth scalar function

$$H(p,q) = \frac{1}{2m}p^2 + U(q), \tag{2.1.4}$$

called Hamiltonian, one can rewrite the 2.1.3 with respect to H as follows:

$$\begin{cases} \dot{p} = -\frac{\partial H}{\partial q} \\ \dot{q} = \frac{\partial H}{\partial p}. \end{cases} \tag{2.1.5}$$

Remark 2.1.1. In above setup, it can be seen that the Hamiltonian 2.1.4 (total

[1] Notice that this is consistent within the scope of total this thesis and not necessarily every Newtonian system is conservative.

energy) is a conserved quantity. Indeed,

$$\frac{dH}{dt} = \frac{1}{m}\dot{p}p + \frac{\partial U}{\partial q}\dot{q} = -\frac{\partial U}{\partial q}\frac{p}{m} + \frac{\partial U}{\partial q}\dot{q} = \dot{p}\dot{q} + (-\dot{p})\dot{q} = 0. \tag{2.1.6}$$

A good example to start with would be the simple harmonic oscillator.

Example 2.1.2. By Hooke's law $F = -kq$ and invoking equation 2.1.1 we get:

$$m\ddot{q} = -kq. \tag{2.1.7}$$

It is worth to note that one can we turn 2.1.1 into system of first order ODE by introducing a new variable $\dot{q} = p$, then we get:

$$\begin{cases} \dot{p} = -\frac{k}{m}q \\ \dot{q} = p. \end{cases} \tag{2.1.8}$$

Example 2.1.3. We borrow the simple harmonic oscillator again as an example: Hamiltonian represents the total energy

$$H(p,q) = \frac{1}{2m}p^2 + \frac{1}{2}kq^2 \tag{2.1.9}$$

where $p = m\dot{q}$. Invoking the Hamiltonian formalism,

$$\dot{p} = -\frac{\partial H}{\partial q} \quad , \quad \dot{q} = \frac{\partial H}{\partial p} \tag{2.1.10}$$

which leads to

$$\dot{p} = -kq \quad , \quad \dot{q} = \frac{p}{m} \tag{2.1.11}$$

then simply by substitution we get

$$m\ddot{q} = -kq. \tag{2.1.12}$$

Example 2.1.4 (Simple pendulum). The following is the Hamiltonian function and the equations of motion

$$H(q,p) = \frac{p^2}{2mL^2} + mgL(1-\cos(q)),$$
$$\begin{cases} \dot{p} = -\dfrac{\partial H}{\partial q} = -mgL\sin(q), \\ \dot{q} = \dfrac{\partial H}{\partial p} = \dfrac{p}{mL^2}. \end{cases} \tag{2.1.13}$$

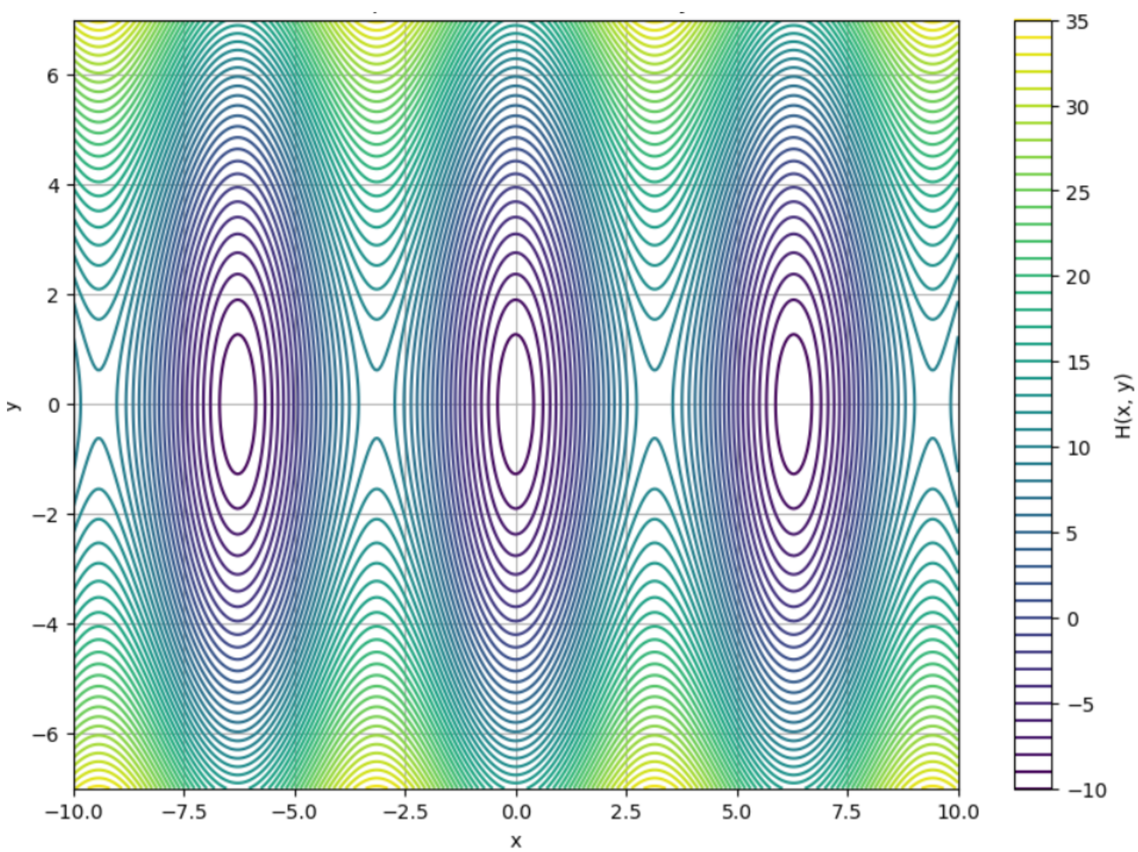


Figure 2.1: Phase space of simple pendulum for $m = L = 1$ 2.1.4.

2.2 Liouville's theorem

The equations 2.1.10 define vector field that attaches at each point of (p,q) of the phase space a vector $(-\partial H/\partial q, \partial H/\partial p)$. The phase flow is the one parameter group of transformations of phase space

$$g^t : (p(0), q(0)) \mapsto (p(t), q(t)) \tag{2.2.1}$$

in which $p(t)$ and $q(t)$ are the solution of Hamiltonian system of equations.

Theorem 2.2.1 ([Arn89] , page 69)**.** *[Liouville's theorem]*
The phase flow preserves volume. In other words, for any region D in phase space we have:

$$\textit{volume of } g^t D = \textit{volume of } D.$$

2.2.1 Hamiltonian vector field

From now on, we seek to find a modern correspondence of Hamiltonian mechanic in the language of symplectic geometry and derive underlying classical properties.

Definition 2.2.2. To each vector ξ, tangent to a symplectic manifold (M^{2n}, ω) at the point x, there exist a 1-form ω_x^1 associated to ξ on T_xM defined as follows

$$\omega_x^1(\eta) = \omega(\eta, \xi) \qquad \forall \eta \in T_xM. \tag{2.2.2}$$

Remark 2.2.3. The correspondence $\xi \longrightarrow \omega_\xi^1$ is an isomorphism between the $2n$-dimensional vector spaces of vectors and 1-forms.

Define J in local coordinate $(p, q) = (p_1, \ldots, p_n, q_1, \ldots, q_n)$ as follows:

$$\begin{aligned} J : T^*M &\longrightarrow TM \\ J(dp) &= \frac{\partial}{\partial q} \\ J(dq) &= -\frac{\partial}{\partial p}. \end{aligned} \tag{2.2.3}$$

We also declare that J defined above commutes with scalar multiplication by smooth function. It is readily seen that the matrix representation of J, according to 2.2.3, is

$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. It is notable that for $2n$-dimensional case $\mathbb{I}_n$ and $-\mathbb{I}_n$ are replaced by 1 and -1.

Let H be a smooth function on a symplectic manifold M^{2n}. Then, obviously, dH is a differential 1-form on M, and there is an associated vector field according to the isomorphism 2.2.3. The corresponding vector field for this differential is denoted by JdH.

Definition 2.2.4. The vector field JdH is called a **Hamiltonian vector field** and H is called the Hamiltonian function.

Now let's consider an example. Take a $2n$-dimensional manifold $M^{2n} = \mathbb{R}^{2n} = \{(p,q)\}$, where $(p,q) = (p_1,\ldots,p_n,q_1,\ldots,q_n)$ represents a point in the phase space. We can construct the canonical Hamiltonian equations by finding the corresponding Hamiltonian vector field:

$$\begin{aligned}(\dot{q},\dot{p}) = \dot{x} = JdH(x) &= J(\partial_p H dp + \partial_q H dq) \\ &= \partial_p H J(dp) + \partial_q H J(dq) \\ &= \partial_p H \frac{\partial}{\partial q} - \partial_q H \frac{\partial}{\partial p} \\ &\implies \begin{cases} \dot{p} = -\frac{\partial H}{\partial q} \\ \dot{q} = \frac{\partial H}{\partial p}.^2 \end{cases}\end{aligned} \tag{2.2.4}$$

The above result is the classical view of Hamiltonian formalism, however the goal of this chapter is to introduce the Poisson bracket to be able to see the Hamiltonian equations with a different perspective.

Remark 2.2.5. The isomorphism J and the equation 2.2.4 can also be obtained from the fact that $i_{X_H}\omega = -dH$ where X_H is the Hamiltonian vector field.

2.2.2 Conservation of symplectic form (Liouville's theorem)

Let (M^{2n}, ω) be a symplectic manifold and $H : M^{2n} \longrightarrow \mathbb{R}$ be a Hamiltonian function. Assume that the vector field JdH corresponding to H is complete then it gives a 1-parameter group of diffeomorphisms $g^t : M^{2n} \longrightarrow M^{2n}$,

$$\frac{d}{dt}\bigg|_{t=0} g^t(x) = JdH(x). \tag{2.2.5}$$

The group g^t is called Hamiltonian phase flow.

In order to define Lie derivative we need to **pull back**. If $f : M \longrightarrow N$ is a smooth map and $\alpha \in \Omega^k(N)$ is a smooth $k-$form, then $f^*\alpha \in \Omega^k(M)$ is a smooth $k-$form defined by

$$(f^*\alpha)_x(X) = \alpha_{f(x)}(d_x f(X))$$

the form $f^*\alpha$ is called the pullback of α.

Definition 2.2.6 (Lie derivative)**.** Suppose α is a smooth differential form and X is a smooth vector field with g_t being the flow, then the **Lie derivative** of α along X is defined as follows,

$$L_X\alpha = \lim_{t\to 0} \frac{g_t^*\alpha - \alpha}{t}. \tag{2.2.6}$$

Before the following theorem we recall the **Cartan's Formula**. It states that, on a smooth manifold M, for any smooth vector field X and any smooth differential form α,

$$L_X\alpha = i_X(d\alpha) + d(i_X\alpha). \tag{2.2.7}$$

Theorem 2.2.7. *[Arn89] A Hamiltonian phase flow preserves the symplectic structure:*

$$(g^t)^*\omega = \omega \tag{2.2.8}$$

This proof is an adaptation of the proof that appears in [Arn89].

Proof. [3] Note that the Hamiltonian vector field in coordinate (q, p) can be locally represented as $X_H = H_p \frac{\partial}{\partial q} - H_q \frac{\partial}{\partial p}$. Then invoking theorem 1.10.3 writing the ω in canonical coordinate we have

$$\omega = d(pdq) = \sum_i dp_i \wedge dq_i \tag{2.2.9}$$

which is obviously a closed form, i.e., $d\omega = 0$. Recalling the construction of Hamiltonian flow clearly we have $i_{X_H}\omega = -H_q dq - H_p dp = -dH$ which shows that $i_{X_H}\omega$ is exact. Therefore by Cartan's Formula 2.2.7,

$$L_{X_H}\omega = i_{X_H}dw + d(i_{X_H}\omega) = d(-dH) = -d^2H = 0 \tag{2.2.10}$$

hence by definition of Lie derivative we get:

$$(g^t)^*\omega = \omega \tag{2.2.11}$$

□

Remark 2.2.8. One can show that a form ω^k ($\omega^k = \underbrace{\omega \wedge \cdots \wedge \omega}_{k-times}$) is conserved along the Hamiltonian flow. Since pullback commutes with the wedge product, so if the Hamiltonian flow ϕ_t preserves ω, it preserves the $k-$fold wedge product, ω^k. In other words, not only ω (symplectic form) is conserved along the Hamiltonian flows, but also all the wedge powers of ω are conserved, resulting the Liouville's theorem that says the volume is conserved along the Hamiltonian flow.

Definition 2.2.9. A map $g : \mathbb{R}^{2n} \longrightarrow \mathbb{R}^{2n}$ is called **canonical map** if $g^*\omega = \omega$.

With this tool we can derive all the results that we already know about classical mechanics. For example we can prove that total energy is conserved. However,

[3] The proof provided is for $\mathbb{R}^{2n}$.

in this language it is equivalent to say the function H is a first integral of the Hamiltonian phase flow with Hamiltonian function H. This can be shown easily by evaluating dH over the Hamiltonian vector field $\eta = JdH$ that is:

$$dH(\eta) = \omega(\eta, JdH) = \omega(\eta, \eta) = 0. \tag{2.2.12}$$

2.3 Lie Algebra of Hamiltonian vector field, Lie bracket and Poisson bracket

Definition 2.3.1. A **Lie algebra** is a vector space L over a field $\mathbf{F} = \mathbb{R}$, that is equipped with a skew symmetric, bilinear map $[,] : L \times L \to L$ which satisfies

$$[x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0 \qquad \text{for each } x, y, z \in L.$$

This latter identity is called the Jacobi identity; the bilinear map $[,]$ is called the Lie bracket.

Example 2.3.2. An example of Lie algebra would be the space of $n \times n$ matrices by defining operation to be $[A, B] = AB - BA$. Note that, this works for any associative algebra

Example 2.3.3. The space of smooth vector fields on a smooth manifold in which the operation is the Lie bracket (which will be discussed in the next subsection).

Example 2.3.4. The space of smooth vector fields on a smooth manifold that preserve some fixed tensor (e.g. a k-form).

In the following subsection, we aim to construct a specific vector field out of two given vector fields, exploring its properties and its relation with the Poisson bracket of two functions.

2.3.1 Lie bracket in coordinates and Lie algebra of Hamiltonian fields

To this end we define an operator called the first-order differential operator L_A where A, ϕ^t are the vector field and its flow respectively. Note that this is the same as Lie derivative applied on 0-forms, i.e. smooth functions. For any smooth function $f : M \longrightarrow \mathbb{R}$ the value of $L_A f$ at a point x is:

$$(L_A f)(x) = \frac{d}{dt}\Big|_{t=0} f(\phi^t x) \tag{2.3.1}$$

which in coordinate $(x_1, \dots, x_n)$ is represented as:

$$L_A(f) = A_1 \frac{\partial f}{\partial x_1} + \cdots + A_n \frac{\partial f}{\partial x_n}. \tag{2.3.2}$$

Taking two smooth vector fields on M, A and B and computing the operator $L_B L_A - L_A L_B$ gives rise to another vector field which is obtained as follows:

$$(L_A L_B - L_B L_A) f = \sum_{i,j=1}^{n} (B_i \frac{\partial A_j}{\partial x_i} - A_i \frac{\partial B_j}{\partial x_i}) \frac{\partial f}{\partial x_j}. \tag{2.3.3}$$

Then we define the Lie bracket of two vector fields A and B in coordinate as follows:

$$[A, B]_j = \sum_{i}^{n} (B_i \frac{\partial A_j}{\partial x_i} - A_i \frac{\partial B_j}{\partial x_i}). \tag{2.3.4}$$

Then in coordinate by evaluating on function $f : M \longrightarrow \mathbb{R}$ we define

$$L_{[A,B]} f = \sum_{j=1}^{n} [A, B]_j \frac{\partial f}{\partial x_j}. \tag{2.3.5}$$

We skip to show to that the Lie bracket defied above possess the properties bilinearity, skew-symmetry that are clear. Also the Jacobi identity can be found in [Arn89].

Theorem 2.3.5 ([Arn89]). *The Lie bracket makes the vector space of vector fields on a manifold M into a Lie algebra.*

In addition to Lie algebra of Hamiltonian vector field we turn our attention to Lie Algebra of Hamiltonian functions which turns out to be useful in interpretation of Hamiltonian formalism and canonical transformations. Interestingly enough, there exists a homomorphism that establishes a structure preserving close relationship between the Lie algebra of Hamiltonian functions and Lie algebra of Hamiltonian vector fields.

Now we define the Poisson bracket of two maps.

Let (M^{2n}, ω) be a symplectic manifold. Consider a Hamiltonian $H : M^{2n} \longrightarrow \mathbb{R}$ and the flow of X_H, $g_H^t : M^{2n} \longrightarrow M^{2n}$ (which of course is a canonical transformation of phase space). Let $F : M^{2n} \longrightarrow \mathbb{R}$ be another function, then define:

$$(F, H)(x) = \left.\frac{d}{dt}\right|_{t=0} F(g_H^t(x)) \tag{2.3.6}$$

be the Poisson bracket of F, H shown by (F, H) [4] which is in fact the derivative of F in direction of flows of H.

Definition 2.3.6. A function F is a **first integral** of the phase flow with Hamiltonian function H if and only if its Poisson bracket with H is identically zero, i.e., $(F, H) \equiv 0$.

Corollary 1 ([Arn89], page 215)**.** The Poisson bracket of the functions F and H is

$$(F, H) = dF(JdH) \tag{2.3.7}$$

then again by applying established isomorphism between 1-forms and vector fields 2.2.3 we get

$$(F, H) = \omega(JdH, JdF). \tag{2.3.8}$$

[4]Note that the notation $(,)$ is utilized for Poisson bracket throughout the thesis.

2.3.2 Poisson bracket interpretation of Hamiltonian formalism

Before running into the next proposition it is important to attain a new interpretation for equations of motion which accounts for proving the form of Hamiltonian vector field is preserved under canonical transformation. We now try to reformulate everything in terms of Poisson bracket.

Suppose $F : (M^{2n}, \omega) \longrightarrow \mathbb{R}$ is a smooth map from symplectic manifold M. Then the time derivative of F along the Hamiltonian flow of H is

$$\dot{F} = L_{X_H} F = dF(JdH) = \omega(JdH, JdF) = (F, H) \tag{2.3.9}$$

as a result, in coordinate (q, p) we have $\dot{q} = (q, H)$ and $\dot{p} = (p, H)$.

In conclusion, Hamiltonian formalism can be converted as follows:

$$\begin{cases} \dot{q}_k = (q_k, H) \\ \dot{p}_k = (p_k, H). \end{cases} \tag{2.3.10}$$

2.3.3 Darboux coordinates and the correspondence between Hamiltonian vector field and Hamiltonian function

Theorem 2.3.7. *Let (M, ω) be a symplectic manifold with Poisson bracket $(,)$. Then, $(C^\infty(M), (,))$ is a Lie algebra. The proof is clear: skew-symmetry and bilinearity are obvious, while the Jacobi identity follows from the definition of a Poisson bracket and the fact that $d\omega = 0$.*

The results of this subsection show that the space of Hamiltonian functions equipped with Poisson operator form a Lie Algebra. The following corollary establishes neat and intimate connection between Poisson bracket of Hamiltonian fields and Hamiltonian functions.

Corollary 2 ([Arn89], page 217)**.** Let X_F and X_H be Hamiltonian vector fields with Hamiltonian functions F and H. Consider the Lie bracket $[X_F, X_H]$, this is again a Hamiltonian vector field with Hamiltonian function equal to the Poisson bracket of the Hamiltonian functions (F, H).

This proof is an adaptation of the proof that appears in [Arn89].

Proof. Set $(F, H) = D$. The Jacobi identity for a given map $G \in C^\infty(M)$ can be rewritten in the form

$$(G, D) = ((G, F), H) - ((G, H), F) \tag{2.3.11}$$

Comparing with the definition of the Lie derivative and the fact that $L_{X_{(F,G)}} H = (H, (F, G))$, we have

$$L_{X_D} = L_{X_H} L_{X_F} - L_{X_F} L_{X_H} \quad \text{where } L_{X_D} = L_{[X_F, X_H]} \tag{2.3.12}$$

as was to be shown. □

Theorem 2.3.8 ([Arn89], page 217)**.** *The first integrals of a Hamiltonian phase flow form a Lie subalgebra of the Lie algebra of all functions.*

2.3.4 A Brief Review of Achievements in Hamiltonian Vector Fields and Symplectic Structures

In synthesizing the achievements and conclusions we have reached to this point, we may succinctly summarize as follows:

A symplectic structure on a manifold is a closed nondegenerate differential 2-form. The phase space of a mechanical system has a natural symplectic structure. On a symplectic manifold, as on a Riemannian manifold, there is a natural isomorphism between vector fields and 1-forms. A vector field on a symplectic manifold corresponding to the differential of a function is called a Hamiltonian vector field. A complete vector field on a manifold determines a phase flow, i.e., a one-parameter group of diffeomorphisms. The phase flow of a Hamiltonian vector field on a symplectic manifold preserves the symplectic structure of phase space. Vector fields on a manifold form a Lie algebra. The Hamiltonian vector fields on a symplectic manifold also form a Lie algebra.

3

Generating functions

The background material in this chapter can be found in the textbook *"Mathematical Methods of Classical Mechanics"* by Vladimir Arnol'd [Arn89].

Generating functions are a powerful tool used in Hamiltonian dynamics to simplify the analysis of the equations of motion. In classical mechanics, Hamiltonian dynamics describes the motion of a system in terms of its energy and momentum. A generating function is a function that can be used to transform one set of canonical variables (such as position and momentum) into another set.

The use of generating functions can simplify the analysis of Hamiltonian systems, making it easier to solve for the trajectories of the system's components. In particular, they can be used to find canonical transformations that preserve the Hamiltonian form of the differential equations. These transformations are called symplectic transformations, and they preserve the structure of the phase space of the system.

A common use of generating functions is found in computing action-angle variables for an integrable Hamiltonian (see §4), or normal-form theory (see §5).

In summary, generating functions are a powerful tool in Hamiltonian dynamics, allowing for the simplification of the equations of motion and the calculation of integrals of motion. The use of generating functions has applications in classical and quantum mechanics and can be used to transform canonical variables, Lagrangian, and wave functions into new sets of variables.

3.1 The Hamilton-Jacobi method for integrating Hamilton's canonical equations

The idea is that under a canonical change of coordinates, the canonical equations of motion are preserved. If we are able to find such a transformation that simplifies the Hamiltonian to a form in which the canonical equations can be integrated, then we can integrate the original canonical equations. It turns out that this problem essentially boils down to solving a partial differential equation known as the Hamilton-Jacobi equation.

3.2 Generating functions

Consider the symplectic diffeomorphism $g : \mathbb{R}^{2n} \longrightarrow \mathbb{R}^{2n}$ that takes the canonical coordinates (p, q) to the new coordinates $(P(p, q), Q(p, q))$ where P and Q are $2n$ functions of $2n$ variables. It is readily observed that the 1-form $pdq - PdQ$ is a closed differential form. This can be deduced from the satisfaction of the condition $dp \wedge dq = dP \wedge dQ$. Since $\mathbb{R}^{2n}$ is contractible, then by Poincaré lemma,

$$pdq - PdQ = dS(q, Q). \tag{3.2.1}$$

We now assume that the following determinant is non-zero around the point (p_0, q_0),

$$\det(\frac{\partial(Q, q)}{\partial(p, q)}) = \det(\frac{\partial Q}{\partial p}) \neq 0 \tag{3.2.2}$$

therefore by the implicit function theorem we have:

$$S(q, Q) = S_1(q, p) \tag{3.2.3}$$

where S_1 is a function on the region of $\mathbb{R}^n_q \times \mathbb{R}^n_p$ of $2n$-dimensional coordinates spaces, whose points are denoted by q and p.

Definition 3.2.1. The function $S(q, Q)$ is called a **generating function** of the canonical transformation g.

From the equation 3.2.1 we drive the following:

$$\frac{\partial S(q, Q)}{\partial q} = p \quad and \quad \frac{\partial S(q, Q)}{\partial Q} = -P. \tag{3.2.4}$$

In the following theorem we prove that under a mild condition every function S gives a canonical transformation.

Theorem 3.2.2 ([Arn89], page 259)**.** *Let $W \subset \mathbb{R}^n \times \mathbb{R}^n$ be an open set around the point (q_0, Q_0), and $S(q, Q) : W \longrightarrow \mathbb{R}^{2n}$ be a smooth function. If*

$$\det(\frac{\partial^2 S}{\partial Q \partial q})\Big|_{(Q_0, q_0)} \neq 0 \tag{3.2.5}$$

then equations 3.2.4 locally defines a canonical transformation.

Proof. Consider the equation for the p coordinates,

$$\frac{\partial S(q, Q)}{\partial q} = p. \tag{3.2.6}$$

Condition 3.2.5 enables us to leverage the inverse function theorem thereby the equation 3.2.6 can be solved to determine a function $Q(p, q)$ in a neighborhood of the point

$$(q_0, p_0) = \frac{\partial S(q, Q)}{\partial q}\Big|_{(Q_0, q_0)} \tag{3.2.7}$$

with $Q(p_0, q_0) = Q_0$.

We now consider the function

$$P_1(q, Q) = -\frac{\partial}{\partial Q} S(q, Q), \tag{3.2.8}$$

and set

$$P(q, p) = P_1(q, Q(p, q)) \tag{3.2.9}$$

then the local map $g : \mathbb{R}^{2n} \longrightarrow \mathbb{R}^{2n}$ sending the point (p, q) to the point (P, Q) is canonical with generating function S since

$$pdq - PdQ = \frac{\partial S(q, Q)}{\partial q} dq + \frac{\partial S(q, Q)}{\partial Q} dQ. \tag{3.2.10}$$

□

3.2.1 Application of generating function

We can now apply the generating function to simplify the Hamiltonian. The process leads to a Hamilton-Jacobi equation by which we find our new coordinates. The idea is to find a canonical transformation by which the Hamiltonian turns to be independent of P variable. In this case the system is trivially integrable. That is to say, if $H = K(Q)$ then the canonical equations have the following forms,

$$\dot{Q} = 0 \qquad \dot{P} = \frac{\partial K}{\partial Q} \tag{3.2.11}$$

which can be integrated as follows

$$Q(t) = Q(0), \quad P(t) = P(0) + \int_0^t \left. \frac{\partial K}{\partial Q} \right|_{Q=Q(0)} dt. \tag{3.2.12}$$

So we now look for the generating function that makes that happen. Meaning that it transforms the $H(p, q)$ to the form $K(Q)$. By invoking the idea of generating functions, from 3.2.4 we obtain the condition

$$H(\frac{\partial S(Q,q)}{\partial q}, q) = K(Q) \tag{3.2.13}$$

where after differentiation we must substitute $q(P,Q)$ for q.

Definition 3.2.3 (Integrability)**.** Let (M,ω) be a symplectic manifold of dimension $2n$. A Hamiltonian $H \in C^\infty(M)$ is said to be **completely integrable** in the sense of Liouville if

a) there exist n functionally independent Hamiltonians $F_1, \ldots, F_n$ that Poisson commute with H;

b) the Hamiltonians $F_1, \ldots, F_n$ pairwise Poisson commute. In other words, for all i, j where $i, j = 1, \ldots, n$, we have $(F_i, F_j) \equiv 0$.

Theorem 3.2.4 ([Arn89], page 260)**. *Jacobi's theorem***
If a solution $S(Q,q)$ is found to the Hamilton-Jacobi equation 3.2.13 depending on n parameters Q_i such that $\det(\frac{\partial^2 S}{\partial Q \partial q}) \neq 0$ *then the canonical equations*

$$\dot{p} = -\frac{\partial H}{\partial q} \quad \textit{and} \quad \dot{q} = \frac{\partial H}{\partial p} \tag{3.2.14}$$

can be completely solved, that is to say the Hamiltonian H is completely integrable. The components $(Q_1, \ldots, Q_n)$ determined by the formulas $\frac{\partial S(Q,q)}{\partial q}$ are first integrals, i.e. $(F_1, \ldots, F_n)$.

Proof. Let us try to find a generating function to transform $H(p,q)$ to $K(Q)$. Consider the canonical transformation with generating function $S(q,Q)$. By 3.2.4 we have

$$p = \frac{\partial S}{\partial q}(q,Q) \tag{3.2.15}$$

from which we can determine $Q(p, q)$. Then, we have

$$H(p, q) = H(\frac{\partial S}{\partial q}(q, Q), q). \tag{3.2.16}$$

In order to find the Hamiltonian function in the new coordinates we must substitute into this expression (after differentiation) for q its expression in terms of P and Q. However, by 3.2.15 this expression does not depend on P at all, so we have simply

$$H(p, q) = K(Q). \tag{3.2.17}$$

which entails an integrable Hamiltonian system since the conjugate variable corresponding to Q does not exist in the new Hamiltonian which means Q is the first integral of the motion.

$\square$

3.3 The Generating function $S_2(P, q)$

Let $g : \mathbb{R}^{2n} \longrightarrow \mathbb{R}^{2n}$ be a canonical transformation with

$$g(p, q) = (P, Q) \tag{3.3.1}$$

This time we start off with different primitive, e.g.

$$pdq + QdP = dS_2(P, q). \tag{3.3.2}$$

By similar arguments as in the preceding section, this equation can also be used to create a canonical transformation. For this function $S_2(P, q)$ we find,

$$p = \frac{\partial S_2(P, q)}{\partial q} \quad and \quad Q = \frac{\partial S_2(P, q)}{\partial P}. \tag{3.3.3}$$

Remark 3.3.1. For some technical reasons, certain canonical transformations cannot be represented by a generating function. For example, in the case of the identity transformation where q and $Q = q$ are dependent, the identity transformation cannot be given by a generating function $S_1(q, Q)$ nor any other generating function in terms of (Q, q). This is why S_2 has been introduced. Also the generating function $S_2(P, q)$ is convenient also because there are no minus signs in the formulas 3.3.3. Also they are easy to remember if we remember that the generating function of the identity transformation is Pq.

Remark 3.3.2. The case for two degrees of freedom follows an analogous path. Let $(q_1, p_1, q_2, p_2) \equiv ((q_1, p_1), (q_2, p_2)) \in T^*\mathbb{R}^n_1 \times T^*\mathbb{R}^n_2$ and $(Q_1, P_1, Q_2, P_2) = \phi(q_1, p_1, q_2, p_2)$ be a canonical transformation. Then,

$$p_1 dq_1 - q_2 dp_2 + (Q_1 dP_1 - P_2 dQ_2) = dS, \tag{3.3.4}$$

where $S = S(q_1, p_2, P_1, Q_2)$. Then,

$$p_1 = \frac{\partial S}{\partial q_1}, \qquad q_2 = -\frac{\partial S}{\partial p_2} \tag{3.3.5}$$

$$Q_1 = \frac{\partial S}{\partial P_1}, \qquad P_2 = -\frac{\partial S}{\partial Q_2}. \tag{3.3.6}$$

4

Action-Angle variables

The background material in this section can be found in the textbook *Mathematical Methods of Classical Mechanics* by Vladimir Arnol'd [Arn89].

4.1 Introduction

In this chapter, we first illustrate the integration of a $2n$-dimensional canonical system of differential equations under certain topological conditions. The integration process in fact necessitates only the knowledge of n initial integrals with certain conditions, a statement affirmed by the Liouville-Arnol'd theorem. An overview of the proof for this theorem is also provided. The theorem's consequence instigates the introduction of action-angle variables within the phase space, incorporating a generating function. In section 4.2.6, we extend this conceptual framework to a Euclidean space of $2n$ dimensions. Beginning with the introduction of the action variable, which draws its inspiration from lower dimensions, we prove that the transition from canonical to action-angle coordinate is structure preserving.

Recall that a function F is a first integral of a system with Hamiltonian function H if and only if the Poisson bracket

$$(F, H) \equiv 0 \tag{4.1.1}$$

is identically equal to zero.

Definition 4.1.1. Two functions F_1 and F_2 are in **involution** (or they Poisson commute) if their Poisson bracket is equal to zero.

4.2 The Formulation of the Liouville-Arnol'd Theorem

Suppose that we are given n functions in involution 4.1.1 on a symplectic $2n$-dimensional manifold M,

$$F_1, \ldots, F_n \qquad (F_i, F_j) \equiv 0 \qquad i, j = 1, 2, \ldots, n. \tag{4.2.1}$$

Let $f = (f_1, \ldots, f_n)$ and then the common level set of functions F_i

$$M_f = \{x : F_i(x) = f_i, i = 1, \ldots, n\}. \tag{4.2.2}$$

In addition, we assume that the functions F_i are independent on M_f, i.e., the 1-forms dF_i, where $i = 1, 2, \ldots, n.$ are linearly independent at each point of M_f.

Theorem 4.2.1 ([Arn89], page 272)**.** *[Liouville-Arnol'd]*

Under above assumptions, i.e. Hamiltonian being integrable 3.2.3, the theorem asserts:

1. *M_f is a smooth manifold and is invariant under the phase flow of Hamiltonian function $H = h(F_1, \ldots, F_n)$.*

2. *If the manifold M_f, is compact and connected, then it is diffeomorphic to the n-dimensional torus $\mathbb{T}^n$,*

$$M_f \cong \mathbb{T}^n = (\mathbb{R}/\mathbb{Z})^n. \tag{4.2.3}$$

3. *The phase flow with Hamiltonian H determines a conditionally periodic motion on M_f, i.e., in angular coordinates $\phi = (\phi_1, \ldots, \phi_n)$ we have*

$$\frac{d\phi}{dt} = \omega, \qquad \omega = \omega(f). \tag{4.2.4}$$

4. *In an open neighborhood of M_f there exists action-angle coordinates (I, θ) in which the Hamiltonian H only depends on action variable I.*

Remark 4.2.2. The equations 4.2.4 can be trivially integrated.

The following corollary is immediate consequence of the Liouville-Arnol'd theorem which is the **key idea** of proving integrability of the Jellinek-Berry Hamiltonian system.

Remark 4.2.3. For example if in a canonical system with two degrees of freedom the extra, independent first integral (beside H) implies complete integrability due to the Liouville-Arnol'd theorem. Also a compact connected two-dimensional submanifold of the phase space is an invariant torus, and motion on it is conditionally periodic.

4.2.1 Proof of Liouville-Arnol'd theorem, statement 1

Lemma 1 ([Arn89], page 273)**.** On the n-dimensional manifold M_f, there exist n tangent vector fields which commute with one another and are linearly independent at every point. Also Hamiltonian vector field JdF_i induced by the first integral F_i is tangent to M_f.

Proof. The symplectic structure of phase space defines an operator J taking 1-forms to vector fields. JdF_i is the Hamiltonian vector field associated to F_i. We will show that the n vector fields JdF_i are tangent to M_f commute, and are independent. The independency of the JdF_i at every point of M_f, follows from the the fact that $J : T^*M \longrightarrow TM$ is a bundle isomorphism, if $dF_1, \ldots, dF_n \in T^*M_f$ are linearly independent, then $JdF_1, \ldots, JdF_n \in TM_f$ are linearly independent. The fields JdF_i commute with one another, since the Poisson brackets of their Hamiltonian functions (F_i, F_j) are identically zero. For the same reason, the derivative of the function F_i in the direction of the field JdF_j is equal to zero for any $i, j = 1, \ldots, n$. Indeed,

$$(F_i, F_j) = dF_i(JdF_j) = \omega(JdF_i, JdF_j) = 0 \tag{4.2.5}$$

as ω and Poisson bracket are skew-symmetric. Thus the fields JdF_i are tangent to M_f and Lemma 1 is proved.

□

Remark 4.2.4. The fact that M_f is a smooth submanifold is the consequence of above lemma that is the existence of n lineally independent tangent vector fields along M_f and submersion theorem.

Remark 4.2.5. By assumption M_f is compact so the flows g_i^t of the Hamiltonian vector field JdF_i are defined on M_f. From 4.2.5 we can conclude that M_f is invariant with respect to each of the n commuting phase flows g_i^t with Hamiltonian functions F_i, i.e.,

$$g_i^t g_j^s = g_j^s g_i^t \quad \forall s, t \in \mathbb{R}. \tag{4.2.6}$$

4.2.2 Proof of Liouville-Arnol'd theorem, statement 2

Remark 4.2.6. The symplectic form ω is zero on the tangent bundle restricted to a point x on the manifold M_f, i.e., $\forall x \in M_f \;\; \omega|_{T_x M_f} = 0$. In other words, the manifold

M_f is a **Lagrangian submanifold**. Indeed the n-vectors $JdF_i|_x$ form basis for the tangent plane to the manifold M_f at the point x.

It is worth mentioning that the following lemma proves the second statement of the **Liouville-Arnol'd** theorem.

Lemma 2 ([Arn89], page 274)**.** Let M^n be a compact connected differentiable n-dimensional manifold, on which we are given n pairwise commutative and linearly independent at each point vector fields. Then M^n is diffeomorphic to an n-dimensional torus.

We denote by g_i^t, $i = 1, \ldots, n$ the one-parameter groups of diffeomorphisms of M corresponding to the n given vector fields. Since the fields commute, the groups g_i^t and g_j^s commute. Therefore, we can define an action g of the commutative group $\mathbb{R}^n = \{t\}$ on the manifold M by setting

$$g^t : M \to M \qquad g^t = g_1^{t_1} \ldots g_n^{t_n} \qquad t = (t_1, \ldots, t_n) \in \mathbb{R}^n. \tag{4.2.7}$$

From the above definition it is clear that

$$g^{t+s} = g^t g^s \quad \text{where} \quad t, s \in \mathbb{R}^n. \tag{4.2.8}$$

Now fix a point x_0 and define the map:

$$G : \mathbb{R}^n \to M \quad G(t) = g^t x_0. \tag{4.2.9}$$

Lemma 3. The map G 4.2.9 of a sufficiently small neighborhood V of the point $0 \in \mathbb{R}^n$ gives a chart in a neighborhood of x_0, that is to say every point $x_0 \in M$ has a neighborhood U ($x_0 \in U \subset M$) such that G maps V diffeomorphically onto U.

Proof. By construction of the map G, if we take the derivative with respect to t_i and construct the Jacobian matrix at the point $0 \in \mathbb{R}^n$, then since the vector fields are linearly independent, inverse function theorem implies that G is a local diffeomorphism.

□

Lemma 4. The map $G : \mathbb{R}^n \to M$ is surjective.

Proof. Consider the x_0 in the preceding lemma and connect an arbitrary $x \in M$ to x_0 by a curve. Since M is compact the curve is compact, hence it can be covered by finite number of open subsets. Now define t as sum of shifts t_i corresponding to pieces of the curve.

□

Remark 4.2.7. It is quite clear that the map $G : \mathbb{R}^n \to M$ 4.2.9 can not be injective since M^n is compact and $\mathbb{R}^n$ is not.

Definition 4.2.8. The **stationary group** of the point x_0 is the set Γ of points $t \in \mathbb{R}^n$ for which $g^t x_0 = x_0$.

Two following lemmas indicate the Γ is a well-defined subgroup of $\mathbb{R}^n$ independent of the point x_0.

Lemma 5. In a sufficiently small neighborhood V of the point $0 \in \mathbb{R}^n$ there is no point of the stationary group other than t=0.

Lemma 6. In the neighborhood t + V of any point $t \in \Gamma \subset \mathbb{R}^n$ there is no point of the stationary group Γ other than t. Therefore, the points of Γ lie in $\mathbb{R}^n$ discretely.

Proof. It is derived directly from the fact that $G_x(t) = g^t(x)$ is a local diffeomorphism of $t = 0$ with a neighbourhood of $x \in M$ for all $x \in M$. □

Lemma 7 ([Arn89], page 276)**.** Let Γ be a discrete subgroup of $\mathbb{R}^n$. Then there exist k $(0 \leqslant k \leqslant n)$ linearly independent vectors $e_1, \ldots, e_k \in \Gamma$ such that Γ is exactly the set of all their integral linear combinations.

Considering above results, we jump into the proof of lemma 2 which in fact states that M_f is diffeomorphic to a torus $\mathbb{T}^n$.

Proof. Fix $x \in M_f$, let $\Gamma = \Gamma_x$ be the stationary subgroup of x under the action g and let $p : \mathbb{R}^n/\Gamma \to M_f$ be the map induced by the covering map $G_x : \mathbb{R}^n \to M_f$. By construction, $p(t + \Gamma) = G_x(t + \Gamma) = g^{t+\Gamma}(x) = g^t(g^\Gamma(x)) = g^t(x)$, so p is well-defined. Thus, p is surjective and a covering map, since G_x is. On the other hand, p is injective, since if $p(t_1 + \Gamma) = p(t_2 + \Gamma)$, then from the preceding, $g^{t_2 - t_1}(x) = x$ and so $t_2 - t_1 \in \Gamma$. This proves that p is an injective covering map, hence a diffeomorphism.

Remark 4.2.9. By lemmas 6, 7 and $\{e_i\}$ being the basis for Γ, one can establish an isomorphism $A : \Gamma \longrightarrow \mathbb{Z}^n$. The map A gives rise to a diffeomorphism $\tilde{A} : \mathbb{R}^n/\Gamma \longrightarrow \mathbb{R}^n/\mathbb{Z}^n$.

4.2.3 Proof of Liouville-Arnol'd theorem, statement 3

Remark 4.2.10. Returning to the assumption of the theorem 4.2.1 and by considering the fact that $\tilde{A} \circ p^{-1}$ is a diffeomorphism, one can establish an angular coordinate on M_f by pull back of the coordinate on $\mathbb{T}^n$ under $\tilde{A} \circ p^{-1} : M_f \to \mathbb{R}^n/\mathbb{Z}^n$.

$$\frac{d\phi_i}{dt} = \omega_i, \qquad \omega_i = \omega_i(f), \qquad \phi(t) = \phi(0) + \omega t. \tag{4.2.10}$$

In other words, motion on the invariant torus M_f is conditionally periodic.

□

4.2.4 Proof of Liouville–Arnol'd theorem, statement 4

In order to determine the Action-Angle variables under the hypotheses of Liouville–Arnol'd theorem, we generate a new symplectic coordinates (I, ϕ) such that, Hamiltonian depends only on I and the system becomes integrable.

Consider a ball U around $f \in \mathbb{R}^n$ that contains only regular values of the first-integral map $F = (F_1, \ldots, F_n)$, then by submersion theorem there exist a neighbourhood $W = F^{-1}(U) \subset M$ that is diffeomorphic to $\mathbb{T}^n \times U$ with coordinates (θ, g). We choose angular ϕ_i on M so that the phase flow with Hamiltonian function $H = F_1$ takes a special simple form:

$$\frac{d\phi}{dt} = \omega(g) \qquad \phi(t) = \phi(0) + \omega t. \tag{4.2.11}$$

Now the goal is to introduce the new canonical coordinates. We note that the variables (F, ϕ) are not, in general, canonical coordinates. It turns out that there are functions of F, which we will denote by $I = I(F)$ and $I = (I_1, \ldots, I_n)$ such that the variables (I, ϕ) are canonical coordinates. Now in the new coordinate the original symplectic structure ω is shown to be following:

$$\omega = dI \wedge d\phi = \sum dI_i \wedge d\phi_i. \tag{4.2.12}$$

In equation 4.2.12, the variable I is called the action variable, and together with the angle variables ϕ, they form the action-angle system of canonical coordinates in a neighborhood of M_f. The quantities I_i are first integrals of the system with the Hamiltonian function $H = F_1$. The variables F_i can be expressed in terms of I and, in particular $H = F_1 = H(I)$.

4.2.5 The case for $M = T^*\mathbb{R}^1$

Let's consider the case $n = 1$, in action-angle variables the differential equations of our flow take the form:

$$\frac{dI}{dt} = 0 \quad , \quad \frac{d\phi}{dt} = \omega(I). \tag{4.2.13}$$

Now we move into the construction of action-angle variable. In order to construct the action-angle variables, we will look for a canonical transformation $(p, q) \rightarrow (I, \phi)$ satisfying two following conditions:

$$I = I(h), \tag{4.2.14}$$

$$\oint_{M_h} d\phi = 2\pi \quad \text{where } M_h = M_h(H = h). \tag{4.2.15}$$

The tool that we leverage to construct the (I, ϕ) out of (p, q) is the Hamilton-Jacobi equation. So, we look for a generating function $S(I, q)$ that satisfies the following equations:

$$p = \frac{\partial S(I, q)}{\partial q} \qquad \phi = \frac{\partial S(I, q)}{\partial I} \qquad H(\frac{\partial S(I, q)}{\partial q}, q) = h(I) \tag{4.2.16}$$

we first assume that the function $h(I)$ is known and invertible, so that every curve M_h is determined by the value of I, i.e., $M_h = M_{h(I)}$. For the fixed value of I according to the equation 4.2.16,

$$dS|_{I=const} = pdq. \tag{4.2.17}$$

This relation defines a 1-form dS on the curve $M_{h(I)}$. Integrating this 1-form on the curve $M_{h(I)}$ we obtain (in a neighborhood of a point q_0) a function

$$S(I, q) = \int_{q_0}^{q} pdq. \tag{4.2.18}$$

This function will be the generating function of the desired symplectic transformation in a neighborhood of the point (I, q_0). So far, the first condition of action-angle variable (equation 4.2.14), by construction, is satisfied automatically. For the second condition, we analyze the behaviour of $S(I, q)$. We go a circuit along the curve $M_{h(I)}$

then the integral 4.2.17 becomes,

$$\Delta S(I) = \oint_{M_{h(I)}} pdq \tag{4.2.19}$$

which is equal to the area of $G = G(I) = \Delta S(I)$ enclosed by the curve $M_{h(I)}$. Therefore, the function S is a multiple-valued function on $M_{h(I)}$; meaning that it is determined up to addition of integral multiples of G. This term has no effect on the derivative $\frac{\partial S(I,q)}{\partial q}$ but it leads to multi-valuedness of $\phi = \frac{\partial S}{\partial I}$ this derivative is defined up to multiples of $G'(I)$, also this multi-valuedness is desired, since ϕ is meant to be an angle. More precisely, the Hamilton-Jacobi formulas 4.2.16 define a 1-form $d\phi$ on the curve $M_{h(I)}$.

The integral of $d\phi$ form on $M_{h(I)}$ is equal to $d\Delta S(I)/dI$, in fact:

$$\begin{aligned}
\oint_{M_{h(I)}} d\phi &= \oint_{M_{h(I)}} \frac{\partial \phi}{\partial I} dI + \oint_{M_{h(I)}} \frac{\partial \phi}{\partial q} dq \\
&= \oint_{M_{h(I)}} \frac{\partial \phi}{\partial I} dI + \oint_{M_{h(I)}} \frac{\partial}{\partial q} \frac{\partial S(I,q)}{\partial I} dq \\
&= 0 + \oint_{M_{h(I)}} \frac{\partial}{\partial I} \frac{\partial S(I,q)}{\partial q} dq \\
&= \frac{d}{dI} \oint_{M_{h(I)}} pdq = \frac{d}{dI} \Delta S(I).
\end{aligned} \tag{4.2.20}$$

In order to fulfill the second condition 4.2.15 we need that

$$2\pi = \frac{d}{dI} \Delta S(I) \tag{4.2.21}$$

then by taking integral from both side, we get

$$2\pi \Delta I = \Delta G, \text{ i.e. } 2\pi(I - I_0) = G(I) - G(I_0) \Rightarrow 2\pi I = G + c_0. \tag{4.2.22}$$

where $G = \oint_{M_h} pdq$ is the area bounded by the phase curve $H = h$.

Because the constant of integration c_0 are largely irrelevant, they are generally omitted.

Definition 4.2.11. The **action variable** of a system of one degree of freedom with Hamiltonian function $H(p, q)$ is the quantity $I(h) = (1/2\pi)G(h)$.

Let $\frac{dG}{dh} \neq 0$, then the inverse $I(h)$ of the function $h(I)$ is defined. We remark that if $G'(h) = 0$ then $\frac{dI}{dh} = 0$ or $\frac{dh}{dI} = \infty$. This implies the action variable I cannot be smoothly continued to the level $\{H = h\}$, and so that level must be singular for H. Similarly, if $\frac{dh}{dI} = 0$, then either the level set is entirely critical for H (contrary to hypothesis) or I does not extend smoothly to $\{H = h\}$. In either case, the h is a critical value. However, if $G'(h) \neq 0$ then by scaling things up, geometrically speaking, one can say that while the angle variable ϕ is tracing out the circuit $M_{h(I)=h}$, the value of I depends smoothly only on the corresponding energy level set $H = h$ which captures the area of the circuit.

Theorem 4.2.12 ([Arn89], page 282)**.** *Let* $S(I, q) = \int_{q_0}^{q} pdq\Big|_{H=h(I)}$*, then the equations of Hamilton-Jacobi equations 4.2.16 give a canonical transformation* $(p, q) \to (I, \phi)$ *satisfying the conditions of action angle variables 4.2.14 and 4.2.15.*

4.2.6 Action Angle variable in $\mathbb{R}^{2n}$

We turn now to systems with n degrees of freedom given in $\mathbb{R}^{2n} = (p, q)$ by a Hamiltonian function $H(p, q)$ and having n first integrals in involution $F_1, \ldots, F_n$. By a similar argument we immediately define n action variables I. We remark that the $I_i(f)$ is defined analogously to the 1-dim case.

Proposition 1. *Let* $\gamma_1, \ldots, \gamma_n$ *be a basis for the one-dimensional cycles on the torus* M_f *(the increase of the coordinate* ϕ_i *on the cycle* γ_j *is equal to* 2π *if* $i = j$ *and is* *zero if $i \neq j$). Then* ***action variable***

$$I_i(f) = \frac{1}{2\pi} \oint_{\gamma_i} pdq \tag{4.2.23}$$

does not depend on the choice of the curve γ_i.

Proof. Cycles γ and γ' are homologous on the torus M_f so that $\gamma' - \gamma$ bounds to a 2-cell σ. In fact, since the symplectic form $\omega = dp \wedge dq$ is zero on M_f then by Stokes formula,

$$\oint_{\gamma} - \oint_{\gamma'} pdq = \iint_{\sigma} dp \wedge dq = 0. \tag{4.2.24}$$

where $\partial\sigma = \gamma - \gamma'$. Note that the σ is the area stuck between the curves γ and γ'. □

We assume now that, for the given values f_i of the n integrals F_i the n quantities I_i are independent: $\det(\partial I/\partial f)|_f \neq 0$. Then in a neighborhood of the torus M_f, we can take the variables I, ϕ as coordinates.

4.2.7 Proof of the Preservation of Structural Integrity in the Transition to Action-Angle Coordinate Variables

Theorem 4.2.13 ([Arn89], page 283)**.** *The transformation $(p, q) \to (I, \phi)$ is canonical, i.e.,*

$$\sum dp_i \wedge dq_i = \sum dI_i \wedge d\phi_i. \tag{4.2.25}$$

Proof. Consider the differential 1-form pdq on M_f. Since the manifold M_f is Lagrangian this 1-form is closed on M_f. That is to say, its exterior derivative, which is the symplectic form, is zero on M_f.

Therefore, by Stokes' theorem the value of

$$S(q) = \int_{q_0}^{q} pdq \Big|_{M_f} \tag{4.2.26}$$

does not change under deformations of the path of integration.

With a similar argument to the case of 1 degree of freedom $S(q)$ is a multiple-valued function on M_f, with periods equal to

$$\Delta_i S = \int_{\gamma_i} dS = 2\pi I_i. \tag{4.2.27}$$

Now let q_0 be a point on M_f, in a neighborhood of which the n variables q are coordinates on M_f, such that the submanifold $M_f \subset \mathbb{R}^{2n}$ is given by n equations of the form $p = p(I, q)$ and q. In a simply connected neighborhood of the point q_0 a single-valued function is defined,

$$S(I, q) = \int_{q_0}^{q} p(I, q)dq \tag{4.2.28}$$

and we can use it as a generating function of the canonical transformation from (p, q) to (I, ϕ)

$$p = \frac{\partial S}{\partial q} \qquad \phi = \frac{\partial S}{\partial I}. \tag{4.2.29}$$

Now it is just left to show why is it canonical. To this end, in fact it suffices to show that $dp \wedge dq = dI \wedge d\phi$.

$$dp \wedge dq = (\frac{\partial}{\partial I}\frac{\partial S(I,q)}{\partial q}dI + \frac{\partial}{\partial q}\frac{\partial S(I,q)}{\partial q}dq) \wedge dq = \frac{\partial^2 S(I,q)}{\partial I \partial q} dI \wedge dq \tag{4.2.30}$$

$$d\phi \wedge dI = (\frac{\partial}{\partial I}\frac{\partial S(I,q)}{\partial I}dI + \frac{\partial^2 S(I,q)}{\partial I \partial q}dq) \wedge dI = \frac{\partial^2 S(I,q)}{\partial q \partial I} dq \wedge dI \tag{4.2.31}$$

by putting 4.2.30 and 4.2.31 together we find $dp \wedge dq = dI \wedge d\phi$. The coordinates ϕ will be multiple-valued with periods:

$$\Delta_i \phi_j = \Delta_i \frac{\partial S}{\partial I_j} = \frac{\partial}{\partial I_j} \Delta_i S = \frac{\partial}{\partial I_j} 2\pi I_i = 2\pi \delta_{ij} \tag{4.2.32}$$

as was to be shown.

We remark that the multi-valuedness of S does not destroy the new local canonical coordinates because they differ by a constant, i.e. $I' = I + c_0$ and $\phi' = \phi + c(I)$. so the corresponding generating function is going to be $S(I, \phi') = I \cdot \phi' + c_0 \cdot \phi' + \sigma(I, \phi')$. As a result, $\phi = \frac{\partial S}{\partial I} = \phi' + \frac{\partial \sigma}{\partial I}$ which implies $c(I) = -\frac{\partial \sigma}{\partial I}$. Moreover, $I' = \frac{\partial S}{\partial \phi'} = I + c_0$. Therefore, in the new coordinate we have

$$dI' \wedge d\phi' = (dI + dc_0) \wedge (d\phi + dc(I)) = dI \wedge d\phi. \tag{4.2.33}$$

□

We conclude this chapter by finding action-angle coordinate for simple harmonic oscillator.

Example 4.2.14. [Arn89]

Recall the Hamiltonian

$$H(x, y) = \frac{1}{2}(x^2 + \omega^2 y^2) \tag{4.2.34}$$

Note that the area-preserving change of $x = \sqrt{\omega} u, y = v/\sqrt{\omega}$ normalizes the Hamiltonian to the form $H(u, v) = \omega \frac{1}{2}(u^2 + v^2)$.

Now by transformation $u = \sqrt{I} sin(\theta), v = \sqrt{I} cos(\theta)$ the Hamiltonian becomes

$$H = \omega I. \tag{4.2.35}$$

Writing Hamiltonian formalism in the new coordinate,

$$\dot{I} = -\frac{\partial H}{\partial \phi} = 0 \quad and \quad \dot{\phi} = \frac{\partial H}{\partial I} = \omega. \tag{4.2.36}$$

The following figure demonstrates the phase portrait of original Hamiltonian. The phase space, in the new coordinate, the circles become horizontal straight lines.

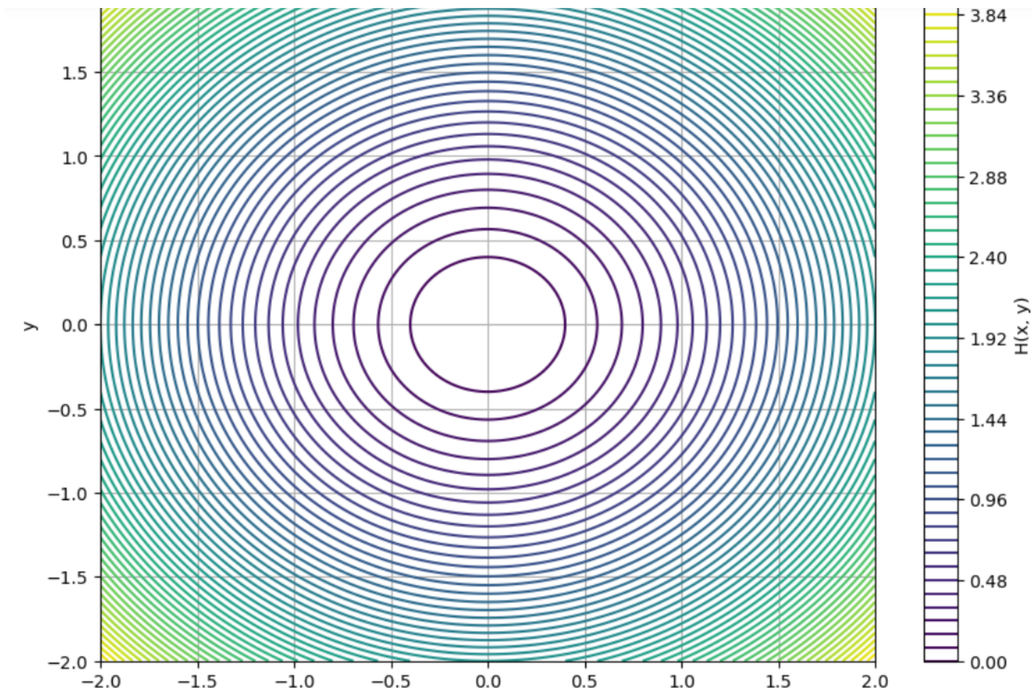


Figure 4.1: Phase space of simple harmonic oscillator for ω=1 and m=1

5

Normal forms

This chapter contains novel materials which are my own work. Also, some of the background material in this section can be found in:

- *Mathematical Methods of Classical Mechanics* textbook by Vladimir Arnol'd [Arn89].
- *Birkhoff Normal Form and Hamiltonian Equations* [Gre06].

When examining the behaviour of solutions to Hamilton's equations in the vicinity of an equilibrium position, relying solely on the linearized equation often proves adequate at least in the absence of further work. Indeed, due to Liouville's theorem, which asserts the preservation of phase space volume, asymptotically stable equilibrium positions cannot exist within Hamiltonian systems. This highlights the necessity of more nuanced analytical methods beyond mere linearization for thorough understanding and prediction of system behaviours.

5.0.1 Near-Equilibrium Stability and Classification of Equilibrium Points

This section is intended to give an informal overview of equilibrium stability.

Consider a general autonomous vector field

$$\dot{x} = f(x), \qquad x \in \mathbb{R}^n \tag{5.0.1}$$

where $f \in C^r, r \geqslant 1$. An **equilibrium solution** of 5.0.1 is a point $\bar{x} \in \mathbb{R}^n$ such that

$$f(\bar{x}) = 0. \tag{5.0.2}$$

Let $x(t)$ be any solution of 5.0.1. Roughly speaking, $\bar{x}$ is **stable** if solutions that start "close" to $\bar{x}$ at a given time remain close to $\bar{x}$ for all later times. It is **asymptotically stable** if nearby solutions converge to $\bar{x}$ as $t \to \infty$.

In the context of Hamiltonian vector fields, if the quadratic part of the Hamiltonian is positive definite, then the fixed point is Liapunov stable. This is because the positive definiteness of the quadratic part of the Hamiltonian ensures that the total energy has a local minimum at the fixed point.

When the quadratic part is not positive-definite nor negative definite (neutral), then the equilibrium is not stable. In fact, the positive-definite part gives rise to a *"center manifold"* where the dynamics can be quasi-periodic. But at the same time there is a unstable manifold.

Conventionally, by analogy of the general case we aim to reduce the Hamiltonian to lower order terms by implementing a canonical change of variables. However, within the purview of Hamiltonian systems, this process encounters certain limitations such as resonance of quadratic part with higher order terms. Indeed, while terms of power 3 can be removed, the presence of resonance prevents the elimination of terms of order for example 4. To navigate around this issue, we employ a suitably chosen canonical coordinate, such as action-angle variables.

5.0.2 Classification of Planar Systems

Definition 5.0.1 (Liapunov Stability)**.** $\bar{x}$ is said to be **stable (or Liapunov stable)** if, for a given $\epsilon > 0$, there exists a $\delta = \delta(\epsilon) > 0$ such that, for any other solution, $y(t)$ of 5.0.1 satisfying $y(0) \in B_\delta(\bar{x})$, then $y(t) \in B_\epsilon(\bar{x})$ for all $t > 0$.

Definition 5.0.2 (Asymptotic Stability)**.** $\bar{x}$ is said to be **asymptotically stable** if it is Liapunov stable, and if there exists a constant $\delta > 0$ such that for any other solution $y(t)$ of 5.0.1 satisfying $y(0) \in B_\delta(\bar{x})$, then $\lim_{t\to\infty} \|\bar{x} - y(t)\| = 0$.

Definition 5.0.3 (Hyperbolic Fixed Point)**.** Let $\bar{x}$ be a fixed point of 5.0.1. Then $\bar{x}$ is called a **hyperbolic fixed point** if none of the eigenvalues of $Df(\bar{x})$ have zero real part.

Theorem 5.0.4. *[Wig03] Consider the C^r autonomous vector field given by 5.0.1. Let $\bar{x}$ be a fixed point of this vector field. Assume that there exists a C^1 function $V : U \to \mathbb{R}$, defined on a neighborhood U of $\bar{x}$, satisfying the following properties:*

- $V(\bar{x}) = 0$*;*
- $V(x) > 0$ *for* $x \neq \bar{x}$*;*
- *The level sets of V are compact.*

Then, the following statements hold:

1. *If $\dot{V}(x) \leqslant 0$ for all $x \in U$, then $x = \bar{x}$ is stable.*
2. *If $\dot{V}(x) < 0$ for all $x \in U$, then $x = \bar{x}$ is asymptotically stable.*

Theorem 5.0.5. *For the planar system $\dot{x} = Ax$, $x \in \mathbb{R}^2$ the eigenvalues of A give qualitative properties about the behavior of the system around the origin.*

$$\lambda^2 - (tr(A))\lambda + \det A = 0. \tag{5.0.3}$$

Taking $tr(A) = T$ and $\det A = D$, the location of (T, D) relative to the parabola $T^2 - 4D$ in the TD-plane classifies the sort of trajectories around the origin as follows:

- *If $T^2 - 4D < 0$ and*

  1. *$T < 0$, then the origin is a spiral sink.*
  2. *$T > 0$, then the origin is a spiral source.*
  3. *$T = 0$, then the origin is a center.*

- *If $T^2 - 4D > 0$ and*

  1. *$D < 0$, then the origin is a saddle.*
  2. *$D > 0$ and $T < 0$, then origin is a sink.*
  3. *$T > 0$ and $D > 0$, then the origin is a source.*

5.0.3 Linearized Hamiltonian and spectral properties for one and two degrees of freedom

Assume that $\bar{x} = 0$ and write the Taylor expansion of the Hamiltonian H with one degree of freedom. Let's write it in a classical form:

$$H = \frac{1}{2}x^T Ax + O(3) \tag{5.0.4}$$

where A is the Hessian of H which is by definition symmetric $A^T = A$ and $x = (p, q) \in \mathbb{R}^n \times \mathbb{R}^n$. Also, in 5.0.4, $O(3)$ stands for the terms of order 3 and higher. Therefore, the equation for Hamiltonian is going to be:

$$\dot{x} = JAx = Bx + O(2) \tag{5.0.5}$$

where $JA = B$ and $J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$ where I is the $n \times n$ identity matrix. The matrix J is taken to be symplectic identity matrix that obviously is a skew-symmetric and orthogonal matrix.

Definition 5.0.6. The action of the Klein 4-group $G = \langle \sigma, \tau | \sigma^2 = \tau^2 = 1, \sigma\tau = \tau\sigma \rangle$ on $\mathbb{C}$:

$$\forall z \in \mathbb{C} : \tau(z) = \bar{z}, \quad \sigma(z) = -z.$$

Lemma 8. The stabilizer of $z \in \mathbb{R}\backslash\{0\}$ is generated by τ; the stabilizer of $z \in i\mathbb{R}\backslash\{0\}$ is generated by $\sigma\tau$; 0 is fixed by G and G acts freely on $\mathbb{C}\backslash(\mathbb{R} \cup i\mathbb{R})$.

Theorem 5.0.7 (Spectral Properties of Linear Hamiltonian)**.** *Let $B = JA$ be a $2n \times 2n$ Hamiltonian matrix.*

- *If $n = 1$, then the eigenvalues of B are $\pm\lambda \in \mathbb{R} \cup i\mathbb{R}$.*
- *If $n = 2$, either the eigenvalues of B come in two distinct pairs $\pm\lambda_1, \pm\lambda_2$ where $\lambda_1, \lambda_2 \in \mathbb{R} \cup i\mathbb{R}$; otherwise, there is an eigenvalue $\lambda \in \mathbb{C}\backslash(\mathbb{R} \cup i\mathbb{R})$ and the set of eigenvalues is the orbit $G \cdot \lambda$. In the first case, the characteristic polynomial $p_B(z)$ factors over $\mathbb{R}$ into a product of two quadratics of the form $z^2 + \lambda_i^2$ and $\mathbb{R}^4$ decomposes into the direct sum of two B-invariant symplectic subspaces; in the second case, when $\lambda = a + ib, ab \neq 0, p_B(z)$ factors over $\mathbb{R}$ into two quadratics $z^2 \pm 2az + a^2 + b^2$ and there are no non-trivial, symplectic, B-invariant subspaces of $\mathbb{R}^4$.*

Proof. By the notation used in 5.0.5, $B^T = (JA)^T = A^T J^T = -AJ$ by multiplying J and J^{-1} we get:

$$J(B^T)J^{-1} = -JA = -B \tag{5.0.6}$$

which means B^T is similar to $-B$ so that they have same characteristic polynomials. Indeed, $P_B(\lambda) = P_{B^T}(\lambda) = P_{-B}(\lambda) = P_B(-\lambda)$ that implies if λ is eigenvalue, then $-\lambda$ is also an eigenvalue. In conclusion, the spectrum of eigenvalues are of the form $(\lambda, -\lambda)$. That is true also for the case complex eigenvalues.

For a system with one degree of freedom, either both λ and $-\lambda$ are real, or they are equal and complex; in other words, $\overline{\lambda} = -\lambda$. To show why these are the only cases that can occur, suppose that $\lambda \in \mathbb{C} - \mathbb{R}$, and λ, $-\lambda$, and $\overline{\lambda}$ are distinct eigenvalues. Then the polynomial $P_B(\lambda)$ would have a degree of at least 3, which cannot be the case.

For a system with two degrees of freedom, we have two possible scenarios:

- To begin with, $\mathbb{R}^4$ decomposes into a direct sum of two B-invariant subspaces which cannot be symplectic. Either the polynomial P_B possesses four distinct roots $\lambda_1, \overline{\lambda_1}, -\lambda_1, -\overline{\lambda_1}$, where $\lambda_1 = a + ib$ for $a, b \neq 0$ and $a > 0$. This results in the system having two copies of $\mathbb{C}$, one with $(\lambda_1, \overline{\lambda_1})$ for which corresponding sub-matrix has a positive determinant, i.e., $\lambda_1.\overline{\lambda_1} = a^2 + b^2$ and trace is $2a$. Since $a > 0$, then first copy is unstable spiral around the origin. Analogously, the second copy will be stable spiral.

- Or, the roots are $\lambda_1, -\lambda_1; \lambda_2, -\lambda_2$ where they are imaginary numbers and the corresponding equilibrium is center. In this case, we get two invariant copies of $\mathbb{C}$ crossing each other transversally and on each copy of $\mathbb{C}$ the Hamiltonian flow is rotation around the origin.

Example 5.0.8. Assume that (x_i, y_i) are canonical coordinates. Now look at the following cases:

- $H = \frac{a}{2}(x_1^2 + y_1^2) + \frac{b}{2}(x_2^2 - y_2^2)$ ends up with eigenvalues $\pm\lambda_1 = \pm ia, \pm\lambda_2 = \pm b$ which is a center-saddle.

The Hessian matrix is

$$\begin{pmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & -b \end{pmatrix} \tag{5.0.7}$$

and the corresponding Hamiltonian matrix (JA) is

$$\begin{pmatrix} 0 & 0 & a & 0 \\ 0 & 0 & 0 & -b \\ -a & 0 & 0 & 0 \\ 0 & -b & 0 & 0 \end{pmatrix}. \tag{5.0.8}$$

The eigenvalues and eigenvectors are

$$\lambda_{1,2} = \pm ai \Longleftrightarrow v_{1,2} = \begin{pmatrix} \mp i \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \lambda_{3,4} = \pm b \Longleftrightarrow v_{3,4} = \begin{pmatrix} 0 \\ \mp 1 \\ 0 \\ 1 \end{pmatrix}.$$

- $H = a(x_1y_1 - x_2y_2) + b(x_1x_2 + y_1y_2)$ ends up with eigenvalues $\lambda = \pm a \pm ib$ which is focus-saddle (complex saddle).

  The Hessian matrix is

$$\begin{pmatrix} 0 & b & a & 0 \\ b & 0 & 0 & -a \\ a & 0 & 0 & b \\ 0 & -a & b & 0 \end{pmatrix}. \tag{5.0.9}$$

The corresponding Hamiltonian matrix is

$$\begin{pmatrix} a & 0 & 0 & b \\ 0 & -a & b & 0 \\ 0 & -b & -a & 0 \\ -b & 0 & 0 & a \end{pmatrix}. \tag{5.0.10}$$

The eigenvalues and eigenvectors are

$$\lambda_{1,2} = a \pm bi \Longleftrightarrow v_{1,2} = \begin{pmatrix} \mp i \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \lambda_{3,4} = -a \pm bi \Longleftrightarrow v_{3,4} = \begin{pmatrix} 0 \\ \mp i \\ 1 \\ 0 \end{pmatrix}.$$

□

Remark 5.0.9. In the theorem 5.0.7, one can we infer that (in the first case) the dynamic around the equilibrium is sum of two copies of $\mathbb{C}$ which are Lagrangian subspace.

The above discussion establishes a firm building block for any further expansion of Hamiltonian around an equilibrium point as though the structure of the normal form is determined by the nature of the linear part of the vector field to a significant extend.

5.1 Normal form for vector fields

Consider the vector field

$$\dot{w} = S(w) \tag{5.1.1}$$

where $w \in \mathbb{R}^n$ and S is a C^r function for some choice of r.

5.1.1 Transformational Approach to Normal Form of Vector Fields

First off, we transform the fixed point w_0 to the origin by the translation

$$v = w - w_0 \qquad v \in \mathbb{R}^n \tag{5.1.2}$$

under which the vector field becomes:

$$\dot{v} = S(v + w_0) = G(v) \tag{5.1.3}$$

linearizing 5.1.3, we get:

$$\dot{v} = DG(0)v + \bar{G}(v), \tag{5.1.4}$$

where $\bar{G}(v) = G(v) - DG(0)v$. It should be clear that $\bar{G}(v) = O(|v|^2)$.

Suppose T transforms $DG(0)$ into (real) Jordan canonical form. Then, under the transformation $v = Tx$, equation 5.1.4 becomes

$$\dot{x} = T^{-1}DG(0)Tx + T^{-1}\bar{G}(Tx) \tag{5.1.5}$$

denoting the (real) Jordan canonical form of $DG(0)$ by J [1] we have,

$$J = T^{-1}DG(0)T. \tag{5.1.6}$$

Define also

$$F(x) = T^{-1}\bar{G}(Tx) \tag{5.1.7}$$

then we can rewrite 5.1.5 as follows,

[1]Please note that, exceptionally in this subsection, the notation J is employed to represent the Jordan canonical form, rather than its customary use as the symplectic matrix (5.0.5) .

$$\dot{x} = Jx + F(x), \qquad x \in \mathbb{R}^n. \tag{5.1.8}$$

Then by Taylor expansion of $F(x)$ we find

$$\dot{x} = Jx + F_2(x) + F_3(x) + \cdots + F_{r-1}(x) + O(|x|^r) \tag{5.1.9}$$

where $F_i(x)$ represents the i^{th} term in the Taylor expansion of F. To kill F_i we introduce a coordinate transformation

$$x = y + h_2(y), \tag{5.1.10}$$

where $h_2 = O(|y|^2), Dh_2 = O(|y|)$, so the assumption just that $|y|$ is sufficiently small. By substituting into 5.1.9 it gives

$$\begin{aligned} (I + Dh_2(y))\dot{y} = Jy + Jh_2(y) + F_2(y + h_2(y)) + F_3(y + h_2(y)) + \\ \cdots + F_{r-1}(y + h_2(y)) + O(|y|^r) \end{aligned} \tag{5.1.11}$$

note that, each term

$$F_k(y + h_2(y)), \qquad 2 \leqslant k \leqslant r - 1, \tag{5.1.12}$$

can be written as

$$F_k(y) + O(|y|^{k+1}) + \cdots + O(|y|^{2k}) \tag{5.1.13}$$

so that 5.1.11 becomes

$$(I + Dh_2(y))\dot{y} = Jy + Jh_2(y) + F_2(y) + F_3(y) + \cdots + F_{r-1}(y) + O(|y|^r). \tag{5.1.14}$$

Now for sufficient small value of $\|Dh_2(y)\|$, the following term exists

$$(I + Dh_2(y))^{-1} \tag{5.1.15}$$

and can be represented in a series expansion as follows

$$(I + Dh_2(y))^{-1} = I - Dh_2(y) + O(|y|^2) \tag{5.1.16}$$

substituting above into 5.1.14 we get:

$$\dot{y} = Jy + Jh_2(y) - Dh_2(y)Jy + F_2(y) + F_3(y) + \cdots + F_{r-1}(y) + O(|y|^r). \tag{5.1.17}$$

Now we are going to choose a specific form for h_2 in order to simplify the $O(|y|^2)$ terms of order two. Ideally, this would mean choosing h_2 such that

$$Dh_2(y)Jy - Jh_2(y) = F_2(y). \tag{5.1.18}$$

This would eliminate $F_2(y)$ from 5.1.17. We want to motivate that, when the equation 5.1.18 viewed in the correct way, it is in fact a linear equation acting on a linear vector space of homogeneous polynomials of order 2. This will be accomplished by 1) defining the appropriate linear vector space; 2) defining the linear operator on the vector space; and 3) describing the equation to be solved in this linear vector space.

Theorem 5.1.1. *[Wig03][Normal form theorem]*
Let H_k be the vector space of homogeneous polynomial vector fields on $\mathbb{R}^n$ of degree k. Let $A \in H_1$ and $\mathcal{L}_A : H_k \to H_k$ be the linear transformation induced by the Lie derivative,

$$\mathcal{L}_A(h)(y) = -d_y h \cdot A(y) + Ah(y) = [h, A](y), \quad y \in \mathbb{R}^n. \tag{5.1.19}$$

Let G_k be a subspace complementary to $\mathcal{L}_J(H_k)$. Then by a sequence of analytic coordinate changes, equation 5.1.9 can be transformed into

$$\dot{y} = Jy + F_2(y) + \ldots + F_{r-1}(y) + O(|y|^r) \tag{5.1.20}$$

where $F_k^r(y) \in G_k$, $2 \leqslant k \leqslant r-1$.

Remark 5.1.2.

- Equation 5.1.20 is said to be in normal form.
- The resonance terms are denoted by $F_k^r(y)$ for $2 \leqslant k \leqslant r-1$.
- As we simplify the terms at order k, no modifications are made to any lower-order terms. However, terms of order higher than k do get modified. This modification occurs at each step when the method is applied. If there's a need to calculate the coefficients of each term of the normal form in the context of the original vector field, it becomes necessary to keep track of how the coordinate transformations successively modify the higher-order terms.

While exploring the landscape of normal forms in vector fields, one might recognize patterns and structures that govern the behavior of these systems. When we shift our focus towards Hamiltonian functions, the Birkhoff Normal Form is considered form of normalized version of the Hamiltonian. The transition might seem superficially disconnected, but it's anchored in the underlying mathematics. For Hamiltonian vector fields, the approach to normal form theory, as discussed in subsection 5.1.1, undergoes some diffeomorphism transformations. Instead of dealing directly with vector fields, we can also analyze the Hamiltonian system's normal forms by a symplectic diffeomorphism transformation. This distinction arises because the Hamiltonian framework allows for a deeper exploration of energy conservation and symplectic geometry, offering a different perspective that connects the theory back to the broader understanding of normal forms. It also is a good exemplification of how mathematical concepts can take on various manifestations depending on the context and focus of the study. The next subsection unlike the previous explore into Normal Forms of Hamiltonian.

5.1.2 Linearized Hamiltonian and normalization by Williamson's theorem

Consider the quadratic Hamiltonian

$$H = \frac{1}{2}\langle Ax, x\rangle \tag{5.1.21}$$

where $x = (p_1, \ldots, p_n; q_1, \ldots, q_n)$ is a vector written in symplectic basis and A is a symmetric linear operator. Then canonical equations are as follows:

$$\dot{x} = JAx \tag{5.1.22}$$

where

$$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}. \tag{5.1.23}$$

Here the eigenvalues of the JA play the main role in normal form theory in terms of type and structure. We proceed by Williamson's theorem which theoretically ensures about the normal form.

Theorem 5.1.3 ([Arn89], page 382)**.** ***Williamson's theorem***
A real symplectic vector space with a given quadratic form H can be decomposed into a direct sum of pairwise skew orthogonal real symplectic subspaces so that the form H is represented as a sum of forms of different types on these subspaces.

For example for a pair of Jordan blocks of order k with eigenvalues $\pm a$ the Hamiltonian is

$$H = -a\sum_{j=1}^{k} p_j q_j + \sum_{j=1}^{k-1} p_j q_{j+1} \tag{5.1.24}$$

or for the case of $\pm a \pm bi$ the Hamiltonian is

$$H = -a \sum_{j=1}^{2k} p_j q_j + b \sum_{j=1}^{k} (p_{2j-1} q_{2j} - p_{2j} q_{2j-1}) + \sum_{j=1}^{2k-2} p_j q_{j+2} \tag{5.1.25}$$

or for the pair of Jordan blocks of order k with eigenvalue zero the Hamiltonian is

$$H = \sum_{j=1}^{k-1} p_j q_{j+1} \tag{5.1.26}$$

where for $j = 1$, $H = 0$.

Definition 5.1.4. Characteristic frequency, denoted by ω_j, are defined such that such that the quantities $\pm i\omega_j$ represent the eigenvalues of matrix B as described in equation 5.0.5.

5.2 Birkhoff Normal Form

We proceed by the following statement that will be explained and leveraged in upcoming sections. Suppose that in the linear approximation an equilibrium position of a Hamiltonian system with n degrees of freedom is stable, and that all n characteristic frequencies $\omega_1, \ldots, \omega_n$ are different. Then the quadratic part of the Hamiltonian can be reduced by a canonical linear transformation to the form:

$$H = \frac{1}{2}\Big(\omega_1(p_1^2 + q_1^2) + \cdots + \omega_n(p_n^2 + q_n^2)\Big) + O(3). \tag{5.2.1}$$

5.2.1 Frequency condition and statement of Birkhoff normal form

Definition 5.2.1. The characteristic frequencies $\omega_1, \ldots, \omega_n$ satisfy a **resonance relation of order** K if there exist integers k_l not all equal to zero such that

$$k_1\omega_1 + \cdots + k_n\omega_n = 0, \qquad |k_1| + \cdots + |k_n| = K. \tag{5.2.2}$$

The action-angle coordinate (I_l, θ_l) for the harmonic oscillator with $\omega = 1$ (example 4.2.14) are, by construction, action-angle coordinates for the Birkhoff normal form.

Definition 5.2.2. A **Birkhoff normal form** of degree s for a Hamiltonian is a polynomial of degree s in the canonical coordinates (P_l, Q_l) which is actually a polynomial of degree $[\frac{s}{2}]$ in the variables $I_l = (P_l^2 + Q_l^2)/2$.

For example, for a system with one degree of freedom the normal form of degree $2m$ or $2m+1$ is

$$H_{2m} = H_{2m+1} = a_1 I + a_2 I^2 + \cdots + a_m I^m \quad \textit{where} \quad I = \frac{1}{2}(P^2 + Q^2). \tag{5.2.3}$$

Birkhoff normal form gets a bit complicated with higher degrees of freedom. For instance, for a system with two degrees of freedom the Birkhoff normal form of degree 4 is

$$H_4 = a_1 I_1 + a_2 I_2 + a_{11} I_1^2 + a_{12} I_1 I_2 + a_{22} I_2^2 \tag{5.2.4}$$

the coefficients a_1 and a_2 in 5.2.4 are characteristic frequencies.

Theorem 5.2.3 ([Arn89], page 387)**.** *Assume that the characteristic frequencies ω_l do not satisfy any resonance relation of order s 5.2.2 or smaller. Then there is a canonical coordinate system in a neighborhood of the equilibrium point such that the Hamiltonian is reduced to a Birkhoff normal form of degree s up to terms of order $s+1$:*

$$H(p, q) = H_s(P, Q) + R \qquad \textit{where} \qquad R = O(|P| + |Q|)^{s+1}. \tag{5.2.5}$$

Birkhoff normal form of linearized Hamiltonian

As a matter of fact, if $\lambda_j = i\omega_j$, $\omega_j \in \mathbb{R} - \{0\}$ are n-distinct non-conjugate eigenvalues of B, then the Hamiltonian can be transformed to $H = \sum_j \frac{1}{2}\omega_j(P_j^2 + Q_j^2) + O(3)$ by a linear change of variables.

As a matter of simplicity, we will focus on finding the normal form of perturbed simple harmonic oscillator. It is notable to know the Birkhoff normal form of simple harmonic oscillator is as follows:

$$H_2(p,q) = \sum_{j=1}^{n} \omega_j \frac{p_j^2 + q_j^2}{2} = \sum_{j=1}^{n} \omega_j I_j \tag{5.2.6}$$

where $I_j := \frac{p_j^2+q_j^2}{2}$ is the j^{th} action variable.

Definition 5.2.4. P is in **normal form with respect to** H_0 if it Poisson commutes with H_0.

Theorem 5.2.5 ([BG22], page 12)**.** ***Birkhoff Normal Form***

Assume $H = H_0 + P$ where H_0 in 5.2.6 is the harmonic oscillator and P is C^∞ real value function with zero of order 3 at the origin and let $r \geqslant 3$ be fixed. Then there exist a analytic canonical change of coordinate $\tau : (p', q') \in U \longrightarrow (p, q) \in V$ from a neighborhood of origin to a neighborhood of origin which

$$H \circ \tau = H_0 + Z + R \tag{5.2.7}$$

with following properties:

- *Z is polynomial of order r and is in normal form with respect to H_0.*
- *$R \in C^\infty(M, \mathbb{R})$ and $R(p', q') = O(||(p', q')||)^{r+1}$.*
- *The map τ is close to identity in the sense that*

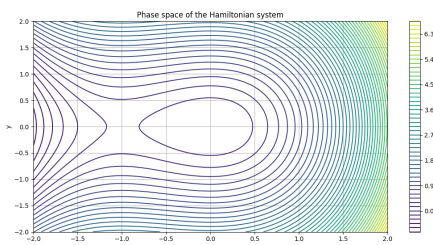
$$\tau(p', q') = (p', q') + O(||(p', q')||)^2. \tag{5.2.8}$$

- *If ω is non-resonant of order r then Z depends only on the new actions $Z = Z(I'_1, \ldots, I'_n)$ where $I'_j = (p'^2_j + q'^2_j)/2$.*

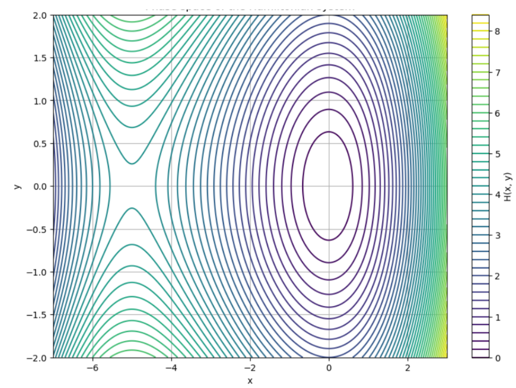
5.2.2 An application to an oscillator with a cubic potential

We now turn our attention to an example of cubic perturbation. Consider the harmonic oscillator H_0 perturbed by the term $\frac{x^3}{3}$, that is to say

$$H = \frac{1}{2}(x^2 + y^2) + \frac{1}{3}x^3. \tag{5.2.9}$$



(a) Harmonic oscillator perturbed by $\frac{1}{3}x^3$



(b) Harmonic oscillator perturbed by $\frac{1}{15}x^3$

Figure 5.1: Perturbation of harmonic oscillator with two different coefficients

We look for canonical transformation

$$\phi(x, y) = (X, Y). \tag{5.2.10}$$

5.2.3 Normalization to order 3

To find an expansion of ϕ, we use the method generating functions (section 3.3). Consider the generating function $V(x, Y)$ for which we have

$$ydx + XdY = dV(x, Y) \tag{5.2.11}$$

then we obtain y, X as follows:

$$\frac{\partial V}{\partial x} = y \quad , \quad \frac{\partial V}{\partial Y} = X \tag{5.2.12}$$

then by Birkhoff normal form theorem 5.2.5 we have

$$K = \frac{1}{2}(X^2 + Y^2) + O(4) \tag{5.2.13}$$

in other words we have

$$H \circ \phi = K(I) = I + \omega I^2 + O(6) \quad \text{where } I = \frac{1}{2}(X^2 + Y^2). \tag{5.2.14}$$

According to the theorem 5.2.5, the symplectic transformation is a near identity map and the generating function associated must look like V written as follows; We first off, take a homogeneous polynomial V of order 3:

$$V(x, Y) = xY + a_3x^3 + a_2x^2Y + a_1xY^2 + a_0Y^3 + O(4) \tag{5.2.15}$$

then by 5.2.12 we get

$$\begin{cases} y = Y + 3a_3x^2 + 2a_2xY + a_1Y^2 + O(3) \\ X = x + a_2x^2 + 2a_1xY + 3a_0Y^2 + O(3). \end{cases} \tag{5.2.16}$$

In this step, we need to determine the coefficients a_i. To this end, we substitute y into original Hamiltonian 5.2.9 and X into Birkhoff normal form 5.2.13. Then by equating the terms of order 3 we seek to find the a_i.

After substitution we will have

$$
\begin{aligned}
H(x,Y) &= \frac{1}{2}x^2 + \frac{1}{2}(Y + 3a_3x^2 + 2a_2xY + a_1Y^2)^2 + \frac{1}{3}x^3 + O(4) \\
&= \underbrace{\frac{1}{2}x^2 + \frac{1}{2}Y^2}_{H_2} + \underbrace{a_1Y^3 + 3a_3Yx^2 + 2a_2Y^2x + \frac{1}{3}x^3}_{H_3} \\
&+ \underbrace{2a_2^2x^2Y^2 + \frac{3}{2}a_3a_1x^2Y^2 + \frac{3}{2}a_3a_1Y^2x^2 + \frac{1}{2}a_1^2Y^4}_{\text{O(4)}} \\
&+ \underbrace{\frac{9}{2}a_3^2Y^4 + 6a_3a_2Yx^3 + 2a_2a_1Y^3x}_{\text{O(4)}}
\end{aligned} \tag{5.2.17}
$$

$$
\begin{aligned}
K(x,Y) &= \frac{1}{2}Y^2 + \frac{1}{2}(x + a_2x^2 + 2a_1xY + 3a_0Y^2)^2 + O(3) \\
&= \underbrace{\frac{1}{2}Y^2 + \frac{1}{2}x^2}_{H_2} + \underbrace{\frac{3}{2}a_0xY^2 + a_2x^3 + 2a_1x^2Y + \frac{3}{2}a_0Y^2x}_{H_3} \\
&= \underbrace{2a_2a_1x^3Y + \frac{3}{2}a_2a_0x^2Y^2 + 6a_1a_0xY^3 + \frac{3}{2}a_2a_0Y^2x^2}_{\text{O(4)}} \\
&+ \underbrace{\frac{1}{2}a_2^2x^4 + \frac{9}{2}a_0^2Y^4 + 2a_1^2x^2Y^2}_{O(4)}
\end{aligned} \tag{5.2.18}
$$

by equating the terms of order 3 we find the following table,

Terms	H	K
x^3	$\frac{1}{3}$	a_2
x^2Y	$3a_3$	$2a_1$
xY^2	$2a_2$	$3a_0$
Y^3	a_1	0

Table 5.1: Corresponding Monomials between H and K

then by equating them the coefficients turn out to be,

$$
a_1 = a_3 = 0, \qquad a_0 = \frac{2}{9}, \qquad a_2 = \frac{1}{3}. \tag{5.2.19}
$$

So the generating functions 5.2.15 becomes,

$$V(x,Y) = xY + \frac{1}{3}x^2Y + \frac{2}{9}Y^3 + O(4). \tag{5.2.20}$$

5.2.4 Normalization to order 4

To obtain the frequency of the system in Birkhoff normal we will do a completely analogous process but this time higher order terms are involved, i.e.,

$$H \circ \phi = K(I) = I + \omega I^2 + O(6). \tag{5.2.21}$$

Consider the generating function of order 4,

$$\begin{aligned} V(x,Y) &= xY + \frac{1}{3}x^2Y + \frac{2}{9}Y^3 \\ &\quad + a_4x^4 + a_3x^3Y + a_2x^2Y^2 + a_1xY^3 + a_0Y^4 + O(5) \end{aligned} \tag{5.2.22}$$

hence,

$$\begin{cases} y = \dfrac{\partial V}{\partial x} = Y + \dfrac{2}{3}xY \\ \qquad\quad + 4a_4x^3 + 3a_3x^2Y + 2a_2xY^2 + a_1Y^3 + O(4), \\ X = \dfrac{\partial V}{\partial Y} = x + \dfrac{1}{3}x^2 + \dfrac{2}{3}Y^2 \\ \qquad\quad + a_3x^3 + 2a_2x^2Y + 3a_1xY^2 + 4a_0Y^3 + O(4) \end{cases} \tag{5.2.23}$$

then by substituting y and X obtained above in original Hamiltonian 5.2.9 and the Birkhoff normal form 5.2.21; after expanding we get the followings:

$$
\begin{aligned}
H(x,Y) &= \frac{1}{2}x^2 + \frac{1}{3}x^3 + \frac{1}{2}[Y + \frac{2}{3}xY + 4a_4x^3 + 3a_3x^2Y + 2a_2xY^2 + a_1Y^3]^2 + O(5) \\
&= \underbrace{\frac{1}{2}x^2 + \frac{1}{2}Y^2}_{H_2} + \underbrace{\frac{1}{3}x^3 + \frac{2}{3}xY^2}_{H_3} + \underbrace{4a_4x^3Y + (3a_3 + \frac{2}{9})x^2Y^2 + 2a_2xY^3 + a_1Y^4}_{H_4} \\
&+ \underbrace{\frac{8}{3}a_4x^4Y + 2a_3x^3Y^2 + \frac{4}{3}a_2x^2Y^3 + \frac{2}{3}a_1xY^4 + \frac{1}{2}a_1^2Y^6 + 2a_1a_2Y^5x}_{\mathrm{O(5)}} \\
&+ \underbrace{\frac{1}{2}a_1^2Y^6 + 2a_1a_2Y^5x + (3a_1a_3 + 2a_2^2)Y^4x^2 + 4a_1a_4Y^3x^3 + 6a_2a_3Y^3x^3}_{\mathrm{O(5)}} \\
&+ \underbrace{8a_2a_4Y^2x^4 + \frac{9}{2}a_3^2Y^2x^4 + 12a_3a_4Yx^5 + 8a_4^2x^6 + 3a_1a_3Y^4x^2}_{\mathrm{O(5)}}
\end{aligned}
\tag{5.2.24}
$$

and,

$$
\begin{aligned}
K(x,Y) &= \frac{1}{2}\left[x + \frac{1}{3}x^2 + \frac{2}{3}Y^2 + a_3x^3 + 2a_2x^2Y + 3a_1xY^2 + 4a_0Y^3\right]^2 + \frac{1}{2}Y^2 \\
&+ \omega\left[\frac{1}{2}Y^2 + \frac{1}{2}\left[x + \frac{1}{3}x^2 + \frac{2}{3}Y^2 + a_3x^3 + 2a_2x^2Y + 3a_1xY^2 + 4a_0Y^3\right]^2\right]^2 + O(5) \\
&= \underbrace{\frac{1}{2}x^2 + \frac{1}{2}Y^2}_{H_2} + \underbrace{\frac{1}{3}x^3 + \frac{2}{3}xY^2}_{H_3} + \underbrace{(a_3 + \frac{\omega}{4} + \frac{1}{18})x^4 + 2a_2x^3Y + \dots}_{H_4} \\
&+ \underbrace{(\frac{2}{9} + 3a_1 + \frac{\omega}{2})x^2Y^2 + 4a_0xY^3 + (\frac{\omega}{4} + \frac{2}{9})Y^4}_{H_4} + \underbrace{\frac{8}{3}a_0Y^5 + \frac{2\omega}{9}Y^6 + \dots}_{O(5)} \\
&+ \underbrace{\frac{4\omega}{81}Y^8 + \frac{\omega}{3}x^5 + \frac{\omega}{6}x^6 + \frac{\omega}{27}x^7 + \frac{\omega}{324}x^8 + 8a_0^2Y^6 + \frac{2\omega}{3}Y^4x + a_3\omega x^7 + \dots}_{O(5)}
\end{aligned}
\tag{5.2.25}
$$

now by equating monomials of order 4 and taking them down in a table as below, then, by equating the coefficients of table above 5.2, we end up with a linear system

Terms	H	K
x^4	0	$a_3 + \frac{\omega}{4} + \frac{1}{18}$
x^3Y	$4a_4$	$2a_2$
x^2Y^2	$3a_3 + \frac{2}{9}$	$\frac{2}{9} + 3a_1 + \frac{\omega}{2}$
xY^3	$2a_2$	$4a_0$
Y^4	a_1	$\frac{\omega}{4} + \frac{2}{9}$

Table 5.2: Corresponding Monomials between H and K

of equations as follows:

$$\begin{aligned} -\tfrac{1}{18} &= && +a_3 && +\tfrac{1}{4}\omega \\ \tfrac{2}{9} &= +3a_1 && -3a_3 && +\tfrac{1}{2}\omega \\ -\tfrac{2}{9} &= -a_1 && && +\tfrac{1}{4}\omega \end{aligned}$$

and

$$\begin{aligned} 0 &= && -2a_2 && +4a_4 \\ 0 &= 4a_0 && -2a_2. \end{aligned}$$

That shows that the equations decouple into two subsystems with a_0, a_2, a_4 and a_1, a_3, ω.

$$a_0 = a_2 = a_4 = 0 \quad , \quad a_1 = \frac{17}{144} \quad , \quad a_3 = \frac{7}{144} \quad , \quad \omega = -\frac{5}{12}. \tag{5.2.27}$$

In conclusion, the Birkhoff normal form of the perturbed Hamiltonian is going to be:

$$H \circ \phi = K(I) = I - \frac{5}{12}I^2 + O(6) \quad \text{where } I = \frac{1}{2}(X^2 + Y^2). \tag{5.2.28}$$

The frequency of the system in action-angle coordinate is the partial derivative of H with respect to action variable

$$\frac{\partial H}{\partial I} = 1 - \frac{5}{6}I + O(5). \tag{5.2.29}$$

Since $\omega < 0$, the frequency is H_I is decreasing near the equilibrium $I = 0$. Indeed, due to the saddle critical point $x = -1, y = 0$, the frequency must tend to 0 as H goes to $\frac{3}{5}$. Also, it is worth mentioning that because H is integrable, the Birkhoff normalization process converges as $s \to \infty$ to a change of variables to action-angle variables for H.

5.2.5 Birkhoff normal form of order three for $H = \frac{1}{2}(x^2 + y^2) + \frac{1}{3}x^3$

To gain a deeper understanding about the behaviour of the system for longer time interval we aim to determine coefficients for higher order terms such as ω_1 in $K = I + \omega I^2 + \omega_1 I^3$. Up until now, we have derived the value of ω and we try to find ω_1 as well. Since the computations are similar to those in section 5.2.2, only the highlights are included.

To this end, we consider the following generating function

$$\begin{aligned} V(x,Y) &= xY + \frac{1}{3}x^2Y + \frac{2}{9}Y^3 + \frac{7}{144}x^3Y + \frac{17}{144}xY^3 \\ &\quad + a_5x^5 + a_4x^4Y + a_3x^3Y^2 + a_2x^2Y^3 + a_1xY^4 + a_0Y^5 + O(6) \end{aligned} \tag{5.2.30}$$

and the Birkhoff normal form is as follows

$$\begin{aligned} K(I) &= I - \frac{5}{12}I^2 + \omega_1 I^3 + O(8) \\ &= \frac{1}{2}(X^2 + Y^2) - \frac{5}{12}(\frac{1}{2}X^2 + \frac{1}{2}Y^2)^2 + \omega_1(\frac{1}{2}X^2 + \frac{1}{2}Y^2)^3 + O(8). \end{aligned} \tag{5.2.31}$$

Notice that the original Hamiltonian is

$$H(x,y) = \frac{1}{2}x^2 + \frac{1}{2}y^2 + \frac{1}{3}x^3. \tag{5.2.32}$$

The partial derivatives of the generating function V are:

$$\begin{aligned} y &= \frac{\partial V}{\partial x} \\ &= Y + \frac{2}{3}xY + \frac{21}{144}x^2Y + \frac{17}{144}Y^3 \\ &\quad + 5a_5x^4 + 4a_4x^3Y + 3a_3x^2Y^2 + 2a_2xY^3 + a_1Y^4 + O(5) \end{aligned} \tag{5.2.33}$$

$$\begin{aligned} X &= \frac{\partial V}{\partial Y} \\ &= x + \frac{1}{3}x^2 + \frac{2}{3}Y^2 + \frac{7}{144}x^3 + \frac{51}{144}xY^2 \\ &\quad + a_4x^4 + 2a_3x^3Y + 3a_2x^2Y^2 + 4a_1xY^3 + 5a_0Y^4 + O(5). \end{aligned} \tag{5.2.34}$$

We now substitute y and X into original Hamiltonian 5.2.32 and Birkhoff normal form 5.2.31 respectively. Equating terms of order 5 in each, as done in section 5.2.2, gives

$$a_0 = \frac{149}{1080}, \qquad a_1 = 0, \qquad a_2 = \frac{41}{144},$$
$$a_3 = 0, \qquad a_4 = \frac{53}{432}, \qquad a_5 = 0$$

to determine ω_1 we go for generating function of order 5,

$$\begin{aligned} V(x,Y) &= xY + \frac{1}{3}x^2Y + \frac{2}{9}Y^3 \\ &+ \frac{7}{144}x^3Y + \frac{17}{144}xY^3 \\ &+ \frac{53}{432}x^4Y + \frac{41}{144}x^2Y^3 + \frac{149}{1080}Y^5 \\ &+ a_6x^6 + a_5x^5Y + a_4x^4Y^2 + a_3x^3Y^3 + a_2x^2Y^4 + a_1xY^5 + a_0Y^6 + O(7) \end{aligned} \tag{5.2.35}$$

from above we derive y and X as follows:

$$\begin{aligned} y &= \frac{\partial V}{\partial x} \\ &= Y + \frac{2}{3}xY + \frac{21}{144}x^2Y + \frac{17}{144}Y^3 + \frac{212}{432}x^3Y + \frac{82}{144}xY^3 \\ &+ 6a_6x^5 + 5a_5x^4Y + 4a_4x^3Y^2 + 3a_3x^2Y^3 + 2a_2xY^4 + a_1Y^5 + O(6) \end{aligned} \tag{5.2.36}$$

$$\begin{aligned} X &= \frac{\partial V}{\partial Y} \\ &= x + \frac{1}{3}x^2 + \frac{2}{3}Y^2 + \frac{7}{144}x^3 + \frac{51}{144}xY^2 + \frac{53}{432}x^4 + \frac{123}{144}x^2Y^2 + \frac{149}{216}Y^4 \\ &+ a_5x^5 + 2a_4x^4Y + 3a_3x^3Y^2 + 4a_2x^2Y^3 + 5a_1xY^4 + 6a_0Y^5 + O(6) \end{aligned} \tag{5.2.37}$$

as we did in previous chapter, we substitute y and X into 5.2.32 and 5.2.31 respectively and then by equating terms of order 6 we determine the value of ω_1 through solving a system of linear equations. Since the process is completely analogous to the section 5.2.2 we avoid writing the details here.

$$\begin{aligned} &a_0 = a_2 = a_4 = a_6 = 0 \\ &a_1 = \frac{449}{1536}, \quad a_3 = \frac{24985}{62208}, \quad a_5 = \frac{4795}{41472} \\ &\omega_1 = -\frac{235}{432} \end{aligned} \tag{5.2.38}$$

then the Birkhoff normal form is

$$K(I) = I - \frac{5}{12}I^2 - \frac{32435}{41472}I^3 + O(8). \tag{5.2.39}$$

5.3 Perturbation of $H_0 = \frac{1}{2}(x^2+y^2)$ by general cubic homogeneous function

This subsection studies the Birkhoff normalization process where the cubic term $\frac{x^3}{3}$ of the previous section is replaced by a general, homogeneous cubic in

$$P(x,y) = b_3x^3 + b_2x^2y + b_1xy^2 + b_0y^3. \tag{5.3.1}$$

5.3.1 Normalization to order 3

The new harmonic oscillator is

$$H(x,y) = \frac{1}{2}x^2 + \frac{1}{2}y^2 + b_3x^3 + b_2x^2y + b_1xy^2 + b_0y^3 \tag{5.3.2}$$

also the first order Birkhoff normal form is

$$K(I) = I + O(4) \quad \text{where } I = \frac{1}{2}(X^2+Y^2). \tag{5.3.3}$$

The normalization algorithm we used so far in section 5.2.2 is going to be applied to H 5.3.3 in order to obtain the Birkhoff normal forms of low order.

As in equation 5.2.20, let the generating function V of the Birkhoff normalization transformation be

$$V(x,Y) = xY + a_3x^3 + a_2x^2Y + a_1xY^2 + a_0Y^3 + O(4) \tag{5.3.4}$$

$$\begin{cases} y = \frac{\partial V}{\partial x} = Y + 3a_3x^2 + 2a_2xY + a_1Y^2 + O(3), \\ X = \frac{\partial V}{\partial Y} = x + a_2x^2 + 2a_1xY + 3a_0Y^2 + O(3). \end{cases} \tag{5.3.5}$$

We then now substitute y into 5.3.2 and X into 5.3.3,

$$\begin{aligned} H(x,Y) = &\frac{1}{2}x^2 + \frac{1}{2}(Y + 3a_3x^2 + 2a_2xY + a_1Y^2)^2 + b_3x^3 \\ &+ b_2x^2(Y + 3a_3x^2 + 2a_2xY + a_1Y^2) + b_1x(Y + 3a_3x^2 + 2a_2xY + a_1Y^2)^2 \\ &+ b_0(Y + 3a_3x^2 + 2a_2xY + a_1Y^2)^3 + O(4). \end{aligned} \tag{5.3.6}$$

And regarding the Birkhoff normal form we obtain

$$K(x,Y) = \frac{1}{2}Y^2 + \frac{1}{2}(x + a_2x^2 + 2a_1xY + 3a_0Y^2)^2 + O(3). \tag{5.3.7}$$

By expanding above equations and underlying the terms of order 3 we find as follows,

$$H(x,Y) = H_0 + \underbrace{a_3x^3 + (3b_3 + a_2)x^2Y + (2b_2 + a_1)xY^2 + (b_1 + a_2)Y^3}_{\text{monomials of order 3}} + O(4) \tag{5.3.8}$$

and

$$K(x,Y) = H_0 + \underbrace{b_2x^3 + 2b_1x^2Y + 3b_0xY^2}_{\text{monomials of order 3}} + O(3). \tag{5.3.9}$$

By equating the underlined terms of 5.3.8 and 5.3.9 and perusing completely analogous process as in the section 5.2.2 and solving its linear system we get,

$$a_0 = \frac{1}{3}(2b_3 + b_1) \quad , \quad a_1 = -b_0 \quad , \quad a_2 = b_3 \quad , \quad a_3 = b_2. \tag{5.3.10}$$

5.3.2 Normalization to order 4

Seeking frequency of the perturbed system requires to calculate a Birkhoff normal form of higher order. Consider the generating function of order 4 and apply the same process as was applied in previous subsection 5.2.2,

$$V(x,Y) = xY + \frac{-1}{3}(2b_0 + b_2)x^3 + b_3x^2Y - b_0xY^2 + \frac{1}{3}(2b_3 + b_1)Y^3 \\ + a_4x^4 + a_3x^3Y + a_2x^2Y^2 + a_1xY^3 + a_0Y^4 + O(5) \tag{5.3.11}$$

from which we obtain the y and X as follows,

$$\begin{cases} y = Y + (-2b_0 - b_2)a_2x^2 + 2b_3xY - b_0Y^2 + 4a_4x^3 + 3a_3x^2Y \\ \qquad + 2a_2xY^2 + a_1Y^3 + O(4) \\ X = x + b_3x^2 - 2b_0xY + (2b_3 + b_1)Y^2 + a_3x^3 + 2a_2x^2Y \\ \qquad + 3a_1xY^2 + 4a_0Y^3 + O(4). \end{cases} \tag{5.3.12}$$

Substituting X and y into 5.3.2 and the Birkhoff normal form up to the order 2 ($K = I + \omega I^2 + O(6)$) and then by equating the terms of order 4, we try to evaluate b_i and ω,

$$\begin{aligned} H(x,Y) = &\frac{1}{2}\Big(Y + (-2b_0 - b_2)a_2x^2 + 2b_3xY - b_0Y^2 + 4a_4x^3 \\ &\quad + 3a_3x^2Y + 2a_2xY^2 + a_1Y^3\Big)^2 \\ &+ b_2x^2\Big(Y + (-2b_0 - b_2)a_2x^2 + 2b_3xY - b_0Y^2 + 4a_4x^3 \\ &\quad + 3a_3x^2Y + 2a_2xY^2 + a_1Y^3\Big) \\ &+ b_1x\Big(Y + (-2b_0 - b_2)a_2x^2 + 2b_3xY - b_0Y^2 + 4a_4x^3 \\ &\quad + 3a_3x^2Y + 2a_2xY^2 + a_1Y^3\Big)^2 \\ &+ b_0\Big(Y + (-2b_0 - b_2)a_2x^2 + 2b_3xY - b_0Y^2 + 4a_4x^3 \\ &\quad + 3a_3x^2Y + 2a_2xY^2 + a_1Y^3\Big)^3 \\ &+ \frac{1}{2}x^2 + b_3x^3 + O(6) \end{aligned} \tag{5.3.13}$$

and,

$$
\begin{aligned}
K(x,Y) = \frac{1}{2}Y^2 + \frac{1}{2}\Big(&x + b_3x^2 - 2b_0xY + (2b_3+b_1)Y^2 \\
&+ a_3x^3 + 2a_2x^2Y + 3a_1xY^2 + 4a_0Y^3\Big)^2 \\
+ \omega\Big(Y^2 + \frac{1}{2}(&x + b_3x^2 - 2b_0xY + (2b_3+b_1)Y^2 \\
&+ a_3x^3 + 2a_2x^2Y + 3a_1xY^2 + 4a_0Y^3)\Big)^2 + O(5).
\end{aligned}
\tag{5.3.14}
$$

A completely similar process to the section 5.2.2, corresponding the monomials of order 4 and equating the coefficients, leads to the following linear system of equations

$$
\begin{pmatrix}
0 & 0 & 0 & 1 & 0 & 1/4 \\
0 & 3 & 0 & -3 & 0 & 1/2 \\
0 & -1 & 0 & 0 & 0 & 1/4 \\
0 & 0 & 1 & 0 & -2 & 0 \\
2 & 0 & -1 & 0 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ \omega
\end{pmatrix}
=
\begin{pmatrix}
2b_0^2 - \frac{1}{2}b_2^2 - \frac{1}{2}b_3^2 \\
-6b_0^2 - 3b_0b_2 + 3b_3b_1 \\
-\frac{1}{2}b_1^2 - 2b_1b_3 - 2b_3^2 - \frac{5}{2}b_2 \\
-6b_0b_1 - b_0b_3 - 3b_1b_2 \\
-2b_0b_3
\end{pmatrix}.
\tag{5.3.15}
$$

As it can be seen, second row and the third row can be grouped and solved separately. Also, first, forth and fifth row can be grouped and solved independently.

$$
\begin{cases}
a_0 = 0 \\
a_1 = \frac{1}{8}\Big(\frac{25}{2}b_0^2 - 3b_2b_0 + \frac{5}{2}b_1^2 + 13b_3b_1 + \frac{17}{2}b_3^2 - \frac{3}{2}b_2^2\Big) \\
a_2 = -4b_3b_0 \\
a_3 = \frac{1}{8}\Big(\frac{47}{2}b_0^2 + 3b_0b_2 + \frac{3}{2}b_1^2 + 3b_1b_3 + \frac{7}{2}b_3^2 - \frac{5}{2}b_2^2\Big) \\
a_4 = -\frac{3}{2}b_0b_3 + b_0b_1 + \frac{1}{2}b_2b_1
\end{cases}
\tag{5.3.16}
$$

and eventually ω is,

$$
\omega = \frac{1}{2}\Big[-3b_0b_2 - \frac{15}{2}b_3^2 - \frac{15}{2}b_0^2 - \frac{3}{2}b_2^2 - \frac{3}{2}b_1^2 - 3b_3b_1\Big]. \tag{5.3.17}
$$

In our analysis of the Hamiltonian $H(x,y) = \frac{1}{2}x^2 + \frac{1}{2}y^2 + b_3x^3 + b_2x^2y + b_1xy^2 + b_0y^3$, a remarkable symmetry emerges in relation to its associated frequency ω. Specifically, when interchanging the coefficients b_0 and b_3, as well as b_1 and b_2 within the Hamiltonian, the expression for ω remains invariant. This symmetry, given by

$$\omega = \frac{1}{2}\Big[-3b_0b_2 - \frac{15}{2}b_3^2 - \frac{15}{2}b_0^2 - \frac{3}{2}b_2^2 - \frac{3}{2}b_1^2 - 3b_3b_1\Big], \tag{5.3.18}$$

elegantly underscores the intricate interplay between the terms in the Hamiltonian and their consequential impact on the dynamics of the system. Additionally, the presence of this symmetry between the Hamiltonian terms and the resulting frequency confirms the accuracy of our computations.

5.3.3 Kolmogorov–Arnol'd–Moser (KAM) theorem

It was originally introduced by Kolmogorov (1954)[Kol79] for the mathematics of planets orbiting around the sun for long time behaviour. He considered a perturbation setting applying to a fully integrable system:

$$H_0 + \epsilon H_1 \tag{5.3.19}$$

where H_0 is the fully integrable part, H_1 is the non-integrable part and ϵ is very small value.

5.3.4 Origins and the classical statement of the KAM theorem

Although the Hamiltonian under consideration doesn't strictly adhere to the assumptions of the KAM theorem as articulated in this thesis, it's imperative to underscore the historical significance of this theorem. Therefore, the inclusion of this material in the current discussion serves to highlight the profound historical and conceptual influence of the KAM theorem in the realm of dynamical systems.

Consider the perturbed Hamiltonian of the solar system induced by Newton's laws of motion:

$$H(X,Y) = \underbrace{\sum_{i=1}^{N}(\frac{|Y_i|^2}{2\mu_i} - G\frac{\mu_i m_0^i}{|X_i|})}_{H_0} + \underbrace{\epsilon\sum_{i<j}(\frac{Y_iY_j}{m_0} - G\frac{M_iM_j}{|X_i - X_j|})}_{H_1(perturabtion)}. \tag{5.3.20}$$

Notice that in above H_1 stands for non-integrable perturbation which is due to planet-planet interaction.

where

- m_0 : mass of the sun
- x_0 : position of the sun
- p_0 : momentum of the sun
- m_i, x_i and p_i for the other planets.
- $\epsilon = \frac{\sum_{i=1}^{N} m_i}{m_0}$
- $M_i = \frac{m_i}{\epsilon}$

$$m_0^i = m_0 + m_i \qquad , \qquad \mu_i = \frac{m_0 m_i}{\epsilon(m_0 + m_i)} \text{ reduced mass} \tag{5.3.21}$$

$$X_i = x_i - x_0 \qquad , \qquad Y_i = \frac{p_i}{\epsilon} \qquad and \qquad \sum_i p_i = 0. \tag{5.3.22}$$

It turns out the value of ϵ is 10^{-3} and the version of KAM-theory that Kolmogorov came up with was not accurate enough to be applied.

Consider the Hamiltonian $H(x, y) = H_0(y) + \epsilon H(x, y)$, where $H_0(y)$ is the integrable part and t $H(x, y)$ is the perturbation. The idea is that many trajectories of perturbed system can be constructed out of the trajectories of unperturbed system.

For the integrable part H_0, from Liouville-Arnol'd 4.2.1 theorem we know,

$$\begin{cases} y = \text{constant}, \\ \dot{x} = \underbrace{\nabla H_0(y)}_{\omega} \in \mathbb{T}^d. \end{cases} \tag{5.3.23}$$

Kolmogorov said for the non-resonant value of ω one can distort (twist) the trajectories into the trajectories of the perturbed system.

Definition 5.3.1. Let $\omega \in \mathbb{R}^d$. We say that a resonance occurs if there exists a non-trivial vector $k \in \mathbb{Z}^d$ such that the inner product $k \cdot \omega = 0$, indicating a linear dependence among the components of ω.

To prevent resonance and near resonances from destroying the convergence of a normalizing transformation, Kolmogrov came up with the Diophantine condition which states frequency (ω) of the unperturbed system is said to be Diophantine if there exist positive constants c and γ such that

$$|k \cdot \omega| \geqslant \frac{1}{c\|k\|^{\gamma}}. \tag{5.3.24}$$

Definition 5.3.2. The frequency map is called **non-degenerate** if

$$\det(\nabla^2 H_0(y)) \neq 0. \tag{5.3.25}$$

If the frequency vector $\nabla H(y)$ is Diophantine one can we construct solution for perturbed system out of unperturbed system via canonical change of coordinate. This means the majority (in the sense of Lebesgue measure) of invariant tori survive after perturbation. These tori are obtained by deformation of tori of the integrable part.

Originally the resonance issue is rooted to linearizing the whole system and applying a canonical transformation to make a new Hamiltonian independent of θ and it ends up the following equation,

$$\omega \cdot \nabla\phi = f \tag{5.3.26}$$

where f is given and ϕ is the transformation.

By applying Fourier transform we get:

$$2i\pi(k \cdot \omega)\tilde{\phi}(k) = \tilde{f}(k) \tag{5.3.27}$$

so in the quest for finding ϕ we face the divisor:

$$\tilde{\phi}(k) = \frac{\tilde{f}(k)}{2i\pi(k \cdot \omega)} \quad for \quad k \in \mathbb{Z}^d - \{0\} \tag{5.3.28}$$

that leads to Diophantine condition.

Consider the Hamiltonian system with real-analytic Hamiltonian-perturbation H_1 as given below,

$$H(\theta, I, \epsilon) = H_0(I) + \epsilon H_1(\theta, I) \tag{5.3.29}$$

where θ mod 2π belongs to n-dimensional torus $\mathbb{T}^n$ and I belongs to an open domain of $\mathbb{R}^n$.

Definition 5.3.3. The frequencies $\omega = \omega(I_0) = \frac{\partial H_0}{\partial I}(I_0)$ on the invariant torus

$$N_{I_0} = \{(\theta, I) : I = I_0\} \tag{5.3.30}$$

of the unperturbed system are said to be **Diophantine** if there exist positive constants c and γ such that

$$|k \cdot \omega| \geqslant \frac{1}{c\|k\|^{\gamma}}. \tag{5.3.31}$$

Let's denote by D_{γ} the set of frequencies satisfying for some $\gamma > 0$.

Remark 5.3.4. The meaning of conditions 5.3.31 is that the "small divisors" $k \cdot \omega$ are not too small. It not only guaranties $k.\omega$ is non-zero but also ensures it can not be approximated too well by rational numbers. The norm $\|.\|$ does not play an important role here. Usually it is taken as $\|v\| = max_j |v_j|$. The torus N_{I_0} with Diophantine frequencies $\omega(I_0)$ is called Diophantine.

Theorem 5.3.5 (Kolmogorov's version of KAM theorem)**.** *[TZ10] Suppose that the unperturbed system 5.3.29 is non-degenerate at the point I_0 and the torus N_{I_0} is Diophantine. Then N_{I_0} survives the perturbation. It is just slightly deformed and as before carries quasi-periodic motions with the frequencies $\omega = \omega(I_0) = \frac{\partial H_0}{\partial I}(I_0)$.*

Remark 5.3.6. Moser proved the theorem 5.3.5 on the preservation of quasi-periodic motions remains true also in the case of sufficiently smooth dependence of the Hamiltonian on phase variables. [Mos01]

5.3.5 Iso-energetic condition and existence of non-resonant invariant tori

Recall that the system with Hamiltonian $H_0(I)$ has n first integrals in involution (the n action variables). And every level set of all these integrals is an n-dimensional torus in $2n$-dimensional phase space. The motion of a phase point on the invariant torus $I = const$ is conditionally-periodic and the frequencies are

$$\omega_k = \frac{\partial H_0}{\partial I_k}. \tag{5.3.32}$$

Therefore, the phase curve densely fills a torus whose dimension is equal to the number of frequencies ω_k which are arithmetically independent. The case when the frequencies are functionally independent will be called the non-degenerate case.
A condition that guarantees the existence of many Diophantine tori is that the frequency map be a local diffeomorphism

$$\det(\frac{\partial \omega}{\partial I}) = \det(\frac{\partial^2 H_0}{\partial I^2}) \neq 0. \tag{5.3.33}$$

Thus, in the non-degenerate case, the unperturbed problem determines on the different invariant tori conditionally-periodic motions with different frequencies. In particular, the invariant tori on which the number of incommensurable frequencies is maximal, i.e., n form a dense set in phase space; such tori are referred to as non-resonant tori. It can be shown that the non-resonant tori form a set of full measure, i.e., the Lebesgue measure of the union of all invariant resonant tori of the unperturbed non-degenerate system is equal to zero.

On a non-resonant torus, the trajectory of a conditionally-periodic motion is dense. Thus, for almost all initial conditions, a phase curve of a non-degenerate unperturbed system densely fills an invariant torus whose dimension is equal to the number of degrees of freedom.

A second form of non-degeneracy, that is independent of the Kolmogorov non-degeneracy condition (eqn 5.3.33), is called iso-energetic non-degeneracy.

Regrading the system 5.3.29 the following condition is called iso-energetic non-degeneracy condition:

$$\det \begin{pmatrix} \frac{\partial^2 H_0}{\partial I^2} & \frac{\partial H_0}{\partial I} \\ \frac{\partial H_0}{\partial I} & 0 \end{pmatrix} \neq 0. \tag{5.3.34}$$

Geometrically, this condition is equivalent to the condition that the projectivized frequency map,

$$I \longrightarrow (\omega_1(I) : \cdots : \omega_n(I)) \in \mathbb{RP}^{n-1} \tag{5.3.35}$$

is a local diffeomorphism when constrained to an energy surface $H(I) = h$, is diffeomorphism of each unperturbed energy level hyper surface $H_0(I) = const$. For iso-energetically non-degenerate H_0, perturbed systems with the Hamilton functions H_1 admit many invariant tori on each energy level hyper surface $\{H(I, \phi) = const\}$.

Theorem 5.3.7 (KAM theorem)**.** *[TZ10][Arn63] Consider the system 5.3.29 and suppose that the invariant torus N_{I_0} of the unperturbed system lies on the energy level $\{H_0 = h\}$, the unperturbed system is iso-energetically non-degenerate at I_0:*

$$\det \begin{pmatrix} \frac{\partial^2 H_0}{\partial I^2}(I_0) & \frac{\partial H_0}{\partial I}(I_0) \\ (\frac{\partial H_0}{\partial I})^T(I_0) & 0 \end{pmatrix} \neq 0 \tag{5.3.36}$$

and the frequencies $\omega(I_0)$ are Diophantine. Then on the energy level $\{H_\epsilon = h\}$ of the perturbed system there is an invariant torus close to the original one. The frequencies on this torus are $\lambda\omega(I_0)$ where $\lambda = 1 + O(\epsilon)$.

6

On the Birkhoff Normal Form and KAM tori Persistence in Modified Jellinek-Berry Nosé Hamiltonian Thermostats for Ideal Gas

This chapter contains new and original results which are inspired by Nosé-thermostated mechanical systems on the n-torus [But18].

The Hamiltonian regarding to the mechanical system is $\Psi(p,q) = \frac{p^2}{2m} + \phi(q)$ where $q \in S^1$ and $p \in \mathbb{R}$. We are aiming to model the exchange of energy through heat between that system and the external heat bath at the constant temperature T.

To model the total energy of the heat bath and system, Jellinek and Berry introduced the Hamiltonian:

$$F(s, p_s, q, p) = \frac{p^2}{2mh^2(s)} + \phi(q) + \frac{p_s^2}{2Qu^2(s)} + kTv(s), \tag{6.0.1}$$

where

- p_s is the conjugate momentum of the thermostat respectively.
- ϕ is the potential of the physical system.
- $v(s), u(s)$ and $h(s)$ are a smooth non-vanishing function of the thermostat variable s.
- s is the thermostat variable.
- The parameter Q is the mass of the additional degree of freedom s.
- k is the Boltzmann constant.
- T is the desired constant temperature.

The net effect of the interaction with the heat bath is that it results in appearing $h(s)$ in the kinetic energy of the mechanical system. In other words, it is behaving as if the mass of the internal system were changing depending on the state of energy exchange.

Also, there is a complementary change in mass of the thermostat variable in 6.0.1, we can think of it as pseudo-particle which is carrying the heat energy of the heat bath.

Remark 6.0.1. By construction, the Hamiltonian F of the Jellinek-Berry thermostat (JB thermostat) is conserved by the Hamiltonian flow. That is, the total energy of the mechanical system and its thermal bath is constant in time.

Ergodic properties: The authors also examined the ergodic properties of the generalized Nosé method[JB89] and [LLM09]. Ergodicity is an extremely important property for a thermostat method, as it ensures that the time averages of physical quantities converge to the ensemble averages. In fact, without ergodicity the classical mechanics and statistical mechanics diverge. Jellinek and Berry discussed the ergodic properties of their generalized method and provided arguments to support the notion that the generalized Nosé method can be ergodic under certain conditions [Jel88].

Theorem 6.0.2. *Assume* v, u, h *are smooth functions of* s*;* Q, kT, m *are fixed positive parameters and* $F(s, p_s, q, p) = \frac{p^2}{2mh^2(s)} + \phi(q) + \frac{p_s^2}{2Qu^2(s)} + kTv(s)$*. Let*

$$\begin{aligned} A &= \frac{p^2}{2m}, & B &= kT, & s &= s_0 + x, \\ p_s &= y, & \alpha(s) &= \frac{1}{h^2(s)}, & \beta(s) &= \frac{1}{2Qu^2(s)}, \\ F_p &= H, & \gamma(x) &= A\alpha(x) + Bv(x). \end{aligned} \tag{6.0.2}$$

Then:

1. *If* $\gamma'' > 0$ *when* $\gamma' = 0$*, then there is a smooth function* $s = s_0(y; A, B)$ *defined for all* $y \neq 0$ *such that* $\frac{ds}{dt} = \frac{dp_s}{dt} = 0$ *when* $s = s_0$*.*

2. *The graph of* $s_0 \subset T^*(S^1 \times \mathbb{R}^+)$ *is fibred by periodic orbits that are normally elliptic with frequency* w *6.2.15.*

3. *There are action-angle variables* $(I_0, I_1, \theta_0, \theta_1)$ *defined in a deleted neighbourhood of the graph of* s_0 *such that the Hamiltonian* F *takes the normal form 6.2.30 where the coefficients are defined in equations 6.2.29.*

4. *The Hamiltonian* F *is iso-energetically non-degenerate for an open and dense set of parameter values.*

5. *For all parameter values in this set and all sufficiently small potentials* $\phi(q)$*,* $F + \phi$ *enjoys a positive measure set of invariant KAM tori in a neighbourhood of the graph of* s_0*. In particular, the thermostatted dynamical system is not ergodic.*

6.0.1 The JB-thermostated ideal gas

We consider idealized version of the Hamiltonian 6.0.1. Say that an ideal gas is a mechanical system with zero potential energy. In an idealized gas, atoms or molecules do not interact, so they only possess kinetic energy. Hence, in equation 6.0.1 we set ϕ zero.

From now on we consider the Jellinek-Berry ideal gas thermostat Hamiltonian

$$F(s, p_s, q, p) = \frac{p^2}{2mh^2(s)} + \frac{p_s^2}{2u^2(s)Q} + kTv(s). \tag{6.0.3}$$

in which the $h(s)$ and $u(s)$ are scaling functions which are smooth non vanishing functions. Regarding 6.0.3 Hamiltonian equations are,

$$\dot{q} = \frac{\partial F}{\partial p} = \frac{p}{mh^2(s)}, \qquad \dot{s} = \frac{\partial F}{\partial p_s} = \frac{p_s}{u^2(s)Q}, \tag{6.0.4}$$

$$\dot{p} = -\frac{\partial F}{\partial q} = 0, \qquad \dot{p_s} = -\frac{\partial F}{\partial s} = \frac{p^2 h'(s)}{mh^3(s)} + \frac{u'(s)p_s^2}{u^3(s)Q} - kTv'(s). \tag{6.0.5}$$

Since F is independent of q, the conjugate momentum p is conserved, as shown by (6.0.5). Subsequently, if $s(t)$ is known, $q(t)$ can be computed by integrating the equation for $\dot{q}$. This leads us to focus on the reduced system where q is ignored and p is treated as a parameter.

6.1 The reduced Hamiltonian thermostat in the (s, p_s)-plane

The reduced Hamiltonian 6.0.3 is:

$$F(s, p_s, q, p) = F_p(s, p_s) \tag{6.1.1}$$

in which p is considered to be parameter.

6.1.1 Linear analysis of the equilibrium point

The equilibrium point of the system is the zero of the equations $\dot{s}$ and $\dot{p}_s$ obtained from 6.0.4 and 6.0.5.

It is readily seen that the equilibrium point is $p_s = 0$ and s satisfies the following equation:

$$\psi(s) = -A\frac{h'(s)}{h^3(s)} + Bv'(s) = 0 \tag{6.1.2}$$

where $A = \frac{p^2}{m}$ and $B = kT$ are positive constants. If the graphs of $\frac{h'}{h^3}$ and v' cross transversely then the system has a unique solution by the implicit function theorem. The completely degenerate case takes place when $v'(s) = 0$ and $\frac{h'(s)}{h^3(s)} = 0$ at the same time.

In order to determine the sort of equilibrium point we find the Hessian matrix of the Hamiltonian around the equilibrium point $(s_0, 0)$

$$\begin{cases} \dot{s} = \dfrac{\partial F_p}{\partial p_s} = \dfrac{p_s}{Qu^2(s)} = f_1(s, p_s), \\ \dot{p_s} = -\dfrac{\partial F_p}{\partial s} = A\dfrac{h'(s)}{h^3(s)} + \dfrac{1}{Q}\dfrac{p_s^2 u'(s)}{u^3(s)} - Bv'(s) = f_2(s, p_s). \end{cases} \tag{6.1.3}$$

The Hamiltonian matrix, consequently around the equilibrium point $(s_0, 0)$ will be

$$\begin{pmatrix} \frac{\partial f_1}{\partial s} = -\frac{p_s u'(s)}{Qu^3(s)} & \frac{\partial f_1}{\partial p_s} = \frac{1}{Qu^2(s)} \\ \frac{\partial f_2}{\partial s} = -\psi'(s) & \frac{\partial f_2}{\partial p_s} = \frac{2}{Q}\frac{u'}{u^3}p_s \end{pmatrix}$$

$$= \begin{pmatrix} 0 & \frac{1}{Qu^2(s_0)} \\ -\psi'(s_0) & 0 \end{pmatrix}. \tag{6.1.4}$$

Now obviously the eigenvalues of 6.1.4 are going to be

$$\lambda_1 = \sqrt{\frac{-1}{Qu^2(s_0)}\psi'(s_0)} \qquad \lambda_2 = -\sqrt{\frac{-1}{Qu^2(s_0)}\psi'(s_0)}. \tag{6.1.5}$$

We want the equilibrium to be center. So, according to the 6.1.5 and theorem 5.0.5, if $\psi'(s_0) > 0$ then $T^2 - 4D < 0$, since $T = 0$, then the equilibrium is center.

6.1.2 Analysis to order four

We now go back to the reduced version of Hamiltonian 6.1.1 and for the sake of simplicity we rename the terms as follows,

$$A = \frac{p^2}{2m}, \qquad B = kT, \qquad s = s_0 + x, \tag{6.1.6}$$

$$p_s = y, \qquad \alpha(s) = \frac{1}{h^2(s)}, \qquad \beta(s) = \frac{1}{2Qu^2(s)}, \tag{6.1.7}$$

$$F_p = H. \tag{6.1.8}$$

then by choosing

$$\gamma(x) = A\alpha(x) + Bv(x) \tag{6.1.9}$$

the Hamiltonian 6.1.1 becomes

$$H(x, y) = \gamma(x) + \beta(x)y^2. \tag{6.1.10}$$

By assumption $x = 0$ is a critical point of γ (i.e. s_0 is a solution to equation 6.1.2). Hence a critical point of H is at $x = y = 0$. By invoking the Hamiltonian equations, one can find the equilibrium of 6.1.10

$$\begin{cases} H_x = 0 \Longrightarrow \gamma'(x) + \beta'(x)y^2 = 0 \Longrightarrow \gamma'(x_0) = 0, \\ H_y = 0 \Longrightarrow 2\beta(x)y = 0 \Longrightarrow y = 0 \end{cases} \tag{6.1.11}$$

where $(x_0, y = 0)$ is the equilibrium point. T he Maclaurin expansion of H to fourth order is:

$$
\begin{aligned}
H(x,y) =&\gamma(0) + \gamma'(0)x + \frac{\gamma''(0)}{2}x^2 + \frac{\gamma'''(0)}{6}x^3 + \frac{\gamma''''(0)}{24}x^4 + \\
&\beta(0)y^2 + \beta'(0)xy^2 + \frac{\beta''(0)}{2}x^2y^2 + O(5).
\end{aligned} \tag{6.1.12}
$$

Without loss of generality, it can be assumed that $\gamma(0) = 0$, and since $\gamma'(0) = 0$ (the equations of motion in Hamiltonian are partial derivative of Hamiltonian function), we get:

$$
H(x,y) = \frac{\gamma''(0)}{2}x^2 + \beta(0)y^2 + \frac{\gamma'''(0)}{6}x^3 + \beta'(0)xy^2 + \frac{\gamma''''(0)}{24}x^4 + \frac{\beta''(0)}{2}x^2y^2 + O(5). \tag{6.1.13}
$$

6.1.3 Birkhoff normal form

To normalize the system we apply the canonical transformation $T(x,y) = (\lambda x, \frac{1}{\lambda}y)$ to equate the coefficient of the quadratic parts.

$$
\frac{\gamma''(0)}{2}\lambda^2 = \frac{\beta(0)}{\lambda^2} \quad \Longrightarrow \quad \lambda^4 = \frac{2\beta(0)}{\gamma''(0)} \tag{6.1.14}
$$

Let's set $M = \sqrt{\frac{\gamma''(0)\beta(0)}{2}}$ applying the transformation T and dividing 6.1.13 by M we get

$$
\frac{1}{2M}H(x,y) = 1/2x^2 + 1/2y^2 + b_1x^3 + b_2xy^2 + b_3x^4 + b_4x^2y^2 + O(5) \tag{6.1.15}
$$

in which

$$
b_1 = \frac{\lambda^3\gamma'''(0)}{12M} \qquad b_2 = \frac{\beta'(0)}{2M\lambda} \tag{6.1.16}
$$

$$
b_3 = \frac{\lambda^4\gamma''''(0)}{48M} \qquad b_4 = \frac{\beta''(0)}{4M}. \tag{6.1.17}
$$

To put the Hamiltonian 6.1.15 into Birkhoff normal form, we use the algorithm demonstrated in section 5.3.

Suppose

$$V(x,Y) = xY + a_3x^3 + a_2x^2Y + a_1xY^2 + a_0Y^3 + O(4) \tag{6.1.18}$$

$$\begin{cases} y = \dfrac{\partial V}{\partial x} = Y + 3a_3x^2 + 2a_2xY + a_1Y^2 + O(3), \\ X = \dfrac{\partial V}{\partial Y} = x + a_2x^2 + 2a_1xY + 3a_0Y^2 + O(3). \end{cases} \tag{6.1.19}$$

By substituting y into the equation 6.1.15 and X into the Birkhoff normal form of order 2 ($K = \frac{1}{2}(X^2 + Y^2) + O(4)$) and equating the terms of order 3 and simplifying, we find

$$a_1 = a_3 = 0, \qquad a_2 = b_1, \qquad a_0 = \frac{1}{3}(2b_1 + b_2) \tag{6.1.20}$$

by which the generating function becomes

$$V(x,Y) = xY + b_1x^2Y + \frac{1}{3}(2b_1 + b_2)Y^3 + O(4). \tag{6.1.21}$$

Recall that the Birkhoff normal form of degree 4 with one degree of freedom is going to be

$$K(I) = I + \omega I^2 + O(6) = \frac{1}{2}X^2 + \frac{1}{2}Y^2 + \frac{\omega}{4}(X^2 + Y^2)^2 + O(6). \tag{6.1.22}$$

Then for the Birkhoff Normal Form of order 4 we consider the following generating function up to order four and follow analogous steps

$$\begin{aligned} V(x,Y) =& xY + b_1x^2Y + \frac{1}{3}(2b_1 + b_2)Y^3 + a_0x^4 + a_1x^3Y \\ &+ a_2x^2Y^2 + a_3xY^3 + a_4Y^4 + O(5). \end{aligned} \tag{6.1.23}$$

$$\begin{cases} y = Y + 2b_1xY + 4a_0x^3 + 3a_1x^2Y + 2a_2xY^2 + a_3Y^3 + O(4), \\ X = x + b_1x^2 + (2b_1 + b_2)Y^2 + a_1x^3 + 2a_2x^2Y + 3a_3xY^2 \\ \qquad + 4a_4Y^3 + O(4). \end{cases} \tag{6.1.24}$$

Similar to section 5.2.2, equating $H(X, y)$ (eqn 6.1.15) and $K(X, y)$ (eqn 6.1.22) modulo $O(5)$ yields five linear equations in six unknowns, $\omega, a_0, \ldots, a_4$:

$$\begin{aligned} a_0 = 0, \quad a_2 &= 0, \quad a_4 = 0, \\ a_1 &= \frac{1}{8}\left[\frac{5}{2}b_2^2 + 13b_1b_2 + b_4 + 3b_3 + \frac{17}{2}b_1^2\right], \\ a_3 &= \frac{1}{8}\left[\frac{3}{2}b_2^2 + 3b_1b_2 - b_4 + 5b_3 + \frac{7}{2}b_1^2\right], \\ \omega &= \frac{1}{2}\left[b_4 + 3b_3 - \frac{15}{2}b_1^2 - \frac{3}{2}b_2^2 - 3b_1b_2\right]. \end{aligned} \tag{6.1.25}$$

Then the Birkhoff normal form is as follows:

$$\frac{1}{2M}H(I) = I + \omega I^2 + O(6) \Longrightarrow H(I) = 2MI + 2M\omega I^2 + O(6). \tag{6.1.26}$$

and therefore the frequency of the system is:

$$\frac{\partial H}{\partial I} = 2M + 4M\omega I + O(4). \tag{6.1.27}$$

To fulfill the Kolmogorov non-degeneracy condition 5.3.25 of KAM theory we end this subsection with the following theorem.

Theorem 6.1.1. *If* $M\omega \neq 0$*, then* $H_I \neq 0$ *in a neighbourhood of {I=0}. In particular, the Kolmogorov non-degeneracy condition is satisfied if* $M\omega \neq 0$.

6.1.4 Non-vanishing of ω in terms of JB parameters

By definition, $M = 0$ if and only if $\gamma'' = 0$ since $\beta > 0$ everywhere. If the quantity $M\omega$ in equation 6.1.27 is zero then it implies that the parameters of the system lie in a real algebraic set with empty interior that in our case h, u and v are the functional non-constant parameters. In other words, we aim to prove the set of variables by which the conditions of KAM theorem hold, makes a dense real algebraic subset. The quantity ω, as seen in 6.1.25, is a smooth function of $b_1, \ldots, b_4$. Also, each b_i defined in 6.1.16. For ω to be non-zero, after breaking down in terms of their definition, implies the following relation:

$$\frac{\lambda^4\gamma^4(0)}{16M} - \frac{3\lambda^6(\gamma'''(0))^2}{144M^2} + \frac{\lambda^3\gamma'''(0)\beta'(0)}{8M^2} + \frac{\beta''(0)}{4M} - \frac{\lambda^3\gamma'''(0)}{2M} - \frac{3\beta'(0)}{2M\lambda} \neq 0 \quad (6.1.28)$$

which M and λ haven't been written in details yet. So, we break them down and write as an output of a smooth map that takes jets from space of jet bundle to a real number.

For the time being, let's call the following equation **KAM discriminant**

$$\Delta = \frac{96\gamma''\lambda\lambda'' + 288\gamma''(\lambda')^2 + 192\gamma'''\lambda\lambda' - 5\gamma''\lambda^4 - 12\gamma''\lambda^3 + 30\gamma''''\lambda^2 - 36\gamma''\lambda^2}{48\gamma''}. \quad (6.1.29)$$

As you can see, Δ is well-defined when $\gamma'' > 0$. Also, Δ is expressed as a rational function in the jets of γ, λ.

Example 6.1.2 (Nosé–Hoover thermostat)**.** The reduced Hamiltonian for the Nosé–Hoover thermostat is

$$H_{Nos\acute{e}} = F_p(s, p_s) = \frac{p^2}{2ms^2} + \frac{p_s^2}{2Q} + kT\ln(s). \quad (6.1.30)$$

The functional parameters are $h(s) = s, \quad u(s) = 1, \quad v(s) = \ln(s).$

Following the analogous process in 6.1.2 it turns out

$$\frac{1}{2M}H(x,y) = \frac{1}{2}x^2 + \frac{1}{2}y^2 + b_1x^3 + b_3x^4 + O(5). \tag{6.1.31}$$

Notice that $b_2 = b_4 = 0$ because β now is a constant function. According to 6.1.25

$$\omega = \frac{1}{4}(6b_3 - 15b_1^2) = \frac{6M\lambda^4\gamma''''(0) - 5\lambda^6(\gamma'''(0))^2}{192M^2}. \tag{6.1.32}$$

Now by calculating 6.1.32, Δ can be represented as a function of parameters

$$\Delta(p,m,Q) = \frac{-11\sqrt{2m}}{24p\sqrt{Q}}. \tag{6.1.33}$$

As seen above Δ has non-zero partial derivative with respect to m.

Reminding that the BNF of $F_p(s,p_s)$ is of the form $H(I) = 2MI + 2M\omega I^2 + O(I^3)$, by substituting we get:

$$H(I) = \frac{kT\sqrt{2m}}{p\sqrt{Q}}I - \frac{11mkT}{12Qp^2}I^2 + O(I^3). \tag{6.1.34}$$

Remark 6.1.3. The equilibrium condition is written as $\gamma'(0) = 0$, where we were thinking of the parameters m, T, Q—hence A, B, Q as fixed at some values, say m_0, T_0, Q_0—hence A_0, B_0, Q_0. However, we need to think of these as independent variables, so the equilibrium condition becomes $\gamma_x(x; A, B, Q) = 0$ and the centre condition is $\gamma_{xx}(x; A, B, Q) > 0$. Since $\gamma_{xx} = A\alpha'' + Bv'' > 0$ and $A, B > 0$, the equation $\gamma_x = 0$ implies that x is a smooth function $x = \chi(A, B)$ near $(x = 0, A = A_0, B = B_0, Q)$. The KAM discriminant is then $\Delta = \Delta(\chi(A, B); A, B, Q)$, which is a smooth function of the parameters in a neighbourhood of (A_0, B_0, Q).

6.2 Non-degeneracy condition for unreduced Hamiltonian

Going back to JB thermostat Hamiltonian on its phase space $T^*(\mathbb{T}^1 \times \mathbb{R}^+)$ is given as follows:

$$F(q, p, s, p_s) = \frac{p^2}{2m}\frac{1}{h^2(s)} + \frac{1}{2Q}\frac{p_s^2}{u^2(s)} + kTv(s) \tag{6.2.1}$$

where we assume that F is C^∞.

The Birkhoff normal form with two degrees of freedom up to the order 4 is:

$$G(I_0, I_1) = w_{10}I_0 + w_{01}I_1 + \frac{1}{2}(w_{20}I_0^2 + 2w_{11}I_0I_1 + w_{02}I_1^2) + O(6). \tag{6.2.2}$$

6.2.1 Birkhoff normal form with two degrees of freedom up to order four

First off, let's see what is the generating function for 2 degrees of freedom:

Let $(q_1, p_1, q_2, p_2) \equiv ((q_1, p_1), (q_2, p_2)) \in T^*\mathbb{R}^m \times T^*\mathbb{R}^n$ and $(Q_1, P_1, Q_2, P_2) = \phi(q_1, p_1, q_2, p_2)$ be a symplectic transformation. Then,

$$p_1 dq_1 - q_2 dp_2 + (Q_1 dP_1 - P_2 dQ_2) = d\nu \tag{6.2.3}$$

where $\nu = \nu(q_1, p_2, Q_2, P_1)$. Then,

$$p_1 = \frac{\partial \nu}{\partial q_1}, \qquad q_2 = -\frac{\partial \nu}{\partial p_2}, \tag{6.2.4}$$

$$Q_1 = \frac{\partial \nu}{\partial P_1}, \qquad P_2 = -\frac{\partial \nu}{\partial Q_2}. \tag{6.2.5}$$

Above will be used for following computations without further comments. However, notice that the order of p_i and q_i can be swapped since the symplectic form is invariant under this action.

For the sake of simplicity we rewrite the Hamiltonian 6.2.1 as follows:

$$F(x, y, s, p_s) = \frac{1}{2}y^2\Lambda(s) + \frac{1}{2}\Omega(s)p_s^2 + Tv(s) \tag{6.2.6}$$

where $\Lambda(s) = \frac{1}{mh^2(s)}$, $\Omega(s) = \frac{1}{Qu^2(s)}$, $(x, y, s, p_s) \in T^*S^1 \times T^*\mathbb{R}^+$. Here, for the sake of readability we can assume that our functional parameters in the Hamiltonian 6.2.6 are Λ, Ω and $v(s)$.

Since x is cyclic angle variable, y is a first integral. Assume that for some $s = s_0$, $y = y_0$ we have

$$\frac{1}{2}y^2\Lambda'(s) + Tv'(s) = 0 \tag{6.2.7}$$

then the Hamiltonian F has a relative equilibrium along $E_0 = S^1 \times \{(y_0, s_0, 0)\}$, i.e. a periodic orbit.

Lemma 9. The equation 6.2.7 implicitly defines a single-valued C^∞ function $s = s(y)$ defined on a neighborhood of $y = y_0$ such that $s_0 = s_0(y_0)$ if

$$\frac{1}{2}y^2\Lambda''(s) + Tv''(s) \neq 0. \tag{6.2.8}$$

Let $E = S^1 \times \{(y, s_0(y), 0) || y - y_0| < \delta\}$. This is an invariant set that is a union of periodic orbits of F and it contains E_0.

6.2.2 Adapting coordinates to E

Let's define a change of coordinates that puts E at the origin. Define:

$$s = s_0(y) + u, \qquad y = \hat{y}. \tag{6.2.9}$$

To find a symplectic change of variables $(x, y, s, p_s) \longrightarrow (\hat{x}, \hat{y}, u, p_u)$ we use a generating function (it might not be so obvious that how generating functions have been chosen in this section. However, we can find some close ones in [But16]). Since $s = -\frac{\partial \nu}{\partial p_s} = s_0(y) + u$ and $\hat{y} = \frac{\partial \nu}{\partial \hat{x}} = y$ we choose

$$\nu(y, p_s; \hat{x}, u) = \hat{x}y - p_s(s_0(y) + u) \tag{6.2.10}$$

so,

$$p_u = -\frac{\partial \nu}{\partial u} = p_s, \qquad x = \frac{\partial \nu}{\partial y} = \hat{x} - p_s s_0'(y). \tag{6.2.11}$$

Let's write a Taylor expansion of F 6.2.6 in a neighborhood of E using the symplectic coordinate $(\hat{x}, \hat{y}, u, p_u)$. According to 6.2.9, by expanding 6.2.6

$$\begin{aligned} F =& \frac{1}{2}\Lambda(s_0)y^2 + Tv(s_0) + \frac{1}{4}\Lambda''(s_0)y^2u^2 + \frac{1}{2}Tv''(s_0)u^2 \\ &+ \frac{1}{6}Tv'''(s_0)u^3 + \frac{1}{24}Tv''''(s_0)u^4 + \frac{1}{2}\Omega(s_0)p_u^2 \\ &+ \frac{1}{2}\Omega'(s_0)up_u^2 + \frac{1}{4}\Omega''(s_0)(up_u)^2 \\ &+ \underbrace{\frac{1}{12}\Lambda'''(s_0)\hat{y}^2u^3 + \frac{1}{120}Tv^{(5)}(s_0)u^5 + \frac{1}{12}\Omega'''(s_0)u^3p_u^2}_{\text{terms of order five}} + O(6). \end{aligned} \tag{6.2.12}$$

To simplify the presentation, we have employed a notation abuse: s_0 denotes the function $s_0(y)$.

Remark 6.2.1. In addressing the Birkhoff normal form of degree two with two degrees of freedom, it is pertinent to note that a Taylor expansion of F up to the fourth degree is adequately sufficient. This degree of expansion effectively captures the necessary detail for the analysis at hand. However, for enhanced precision, a remainder of order five is included. This will be considered from now on without further comment.

6.2.3 Normalizing the quadratic terms

A symplectic transformation $u = \hat{u}/\lambda$, $p_u = \lambda\hat{p_u}$ transforms the quadratic Hamiltonian to

$$G = \frac{1}{2}w(\hat{p_u}^2 + \hat{u}^2). \tag{6.2.13}$$

Considering the Hamiltonian 6.2.12 we seek to equate the coefficients in quadratic part

$$Tv''(s_0)\frac{1}{\lambda^2(y)} = \Omega(s_0)\lambda^2(y) \tag{6.2.14}$$

from there we find:

$$w(y) = \sqrt{\Omega(s_0)\Big(\frac{1}{2}\Lambda''(s_0)y^2 + Tv''(s_0)\Big)}. \tag{6.2.15}$$

The canonical transformation for the normalization of the quadratic part is a canonical map $(x, y, u, p_u) \longrightarrow (\hat{x}, \hat{y}, \hat{u}, \hat{p_u})$ which the equations are derived from the generating function $\nu(\hat{x}, y, u, \hat{p_u}) = \lambda(y)\hat{p_u}u - \hat{x}y$.

$$p_u = \frac{\partial \nu}{\partial u} = \lambda(y)\hat{p_u}, \qquad \hat{u} = \frac{\partial \nu}{\partial \hat{p_u}} = \lambda(y)u, \tag{6.2.16}$$

$$x = -\frac{\partial \nu}{\partial y} = \hat{x} - \lambda'(y)\hat{p_u}u, \qquad \hat{y} = -\frac{\partial \nu}{\partial \hat{x}} = y. \tag{6.2.17}$$

The Hamiltonian 6.2.12 in the new symplectic coordinates is:

$$\begin{aligned} F =& \frac{1}{2}\Lambda(s_0)y^2 + Tv(s_0) + \frac{1}{2}w(\hat{u}^2 + \hat{p_u}^2) \\ &+ \frac{1}{6}Tv'''(s_0)\frac{1}{\lambda^3(y)}\hat{u}^3 + \frac{1}{24}Tv''''(s_0)\frac{1}{\lambda^4(y)}\hat{u}^4 \\ &+ \frac{1}{2}\Omega'(s_0)\lambda(y)\hat{u}\hat{p_u}^2 + \frac{1}{4}\Omega''(s_0)(\hat{u}\hat{p_u})^2 \\ &+ \underbrace{\frac{1}{12}\frac{\Lambda'''(s_0)}{\lambda^3(y)}\hat{y}^2\hat{u}^3 + \frac{1}{120}\frac{1}{\lambda^5(y)}Tv^{(5)}(s_0)\hat{u}^5 + \frac{1}{12}\frac{1}{\lambda(y)}\hat{u}^3\hat{p_u}^2}_{\text{terms of order five}} + O(6). \end{aligned} \tag{6.2.18}$$

Since y Poisson commutes with F and is invariant under the canonical transformations that are used, F can be normalized by the frequency $w = w(y)$ of the thermostat oscillations. When this is done, we get:

$$\begin{aligned}\frac{1}{w}F =& \frac{1}{2w}\Lambda(s_0)y^2 + \frac{T}{w}v(s_0) + \frac{1}{2}(\hat{u}^2 + \hat{p_u}^2) \\ &+ b_1(y)\hat{u}^3 + b_2(y)\hat{u}\hat{p_u}^2 + b_3(y)\hat{u}^4 + b_4(y)(\hat{u}\hat{p_u})^2 \\ &+ \underbrace{D_1(y)\hat{y}^2u^3 + D_2(y)\hat{u}^5 + D_3(y)\hat{u}^3\hat{p_u}^2}_{\text{terms of order five}} + O(6)\end{aligned} \tag{6.2.19}$$

where,

$$b_1(y) = \frac{1}{6w}Tv'''(s_0)\frac{1}{\lambda^3(y)}, \qquad b_2(y) = \frac{1}{2w}\Omega'(s_0)\lambda(y), \tag{6.2.20}$$

$$b_3(y) = \frac{1}{24w}Tv''''(s_0)\frac{1}{\lambda^4(y)}, \qquad b_4(y) = \frac{1}{4w}\Omega''(s_0). \tag{6.2.21}$$

Once we have the Hamiltonian in this form 6.2.19, we perform the Birkhoff Normal-Form algorithm. As you can see from the generating functions above, the partial generating function of the transformation in the (u, p_u) plane is augmented by the generating function $\hat{x}y$. The resulting transformation normalizes F in the (u, p_u) variables with coefficients that depend on y and preserves y at each step. The angle variable x is changed during the normalization, but the new angle variable at each step preserves its canonical nature.

In equation 6.2.19 obviously first and second terms are normalized and the other terms will be normalized to the form:

$$\tilde{G} = I_0 + \omega I_0^2 + O(6) \tag{6.2.22}$$

through the following computation that we developed in section 5.2.2.

So the generating function is $\nu(\hat{x}, y, \hat{u}, \hat{p_u}) = \nu_1(\hat{u}, \hat{p_u}, y) + \hat{x}y$. Then, with respect to ν_1 which generates a canonical transformation $(\hat{u}, \hat{p_u}) \longrightarrow (\tilde{u}, \tilde{p_u})$ we are going to leverage the tool that we developed for reduced Hamiltonian which is given as follows,

$$\nu_1(\hat{u}, \tilde{p_u}, y) = \hat{u}\tilde{p_u} + a\hat{u}^2\tilde{p_u} + c\tilde{p_u}^3 + a_4\hat{u}^4 + a_3\hat{u}^3\tilde{p_u} + a_2\hat{u}^2\tilde{p_u}^2 + a_1\hat{u}\tilde{p_u}^3 + a_0\tilde{p_u}^4 + O(5) \tag{6.2.23}$$

by applying the same process of the chapter 5.2.2

$$a = b_1(y), \qquad c = \frac{1}{3}\Big(2b_1(y) + b_2(y)\Big), \tag{6.2.24}$$

$$a_4 = 0, \qquad a_3 = \frac{1}{8}\Big(\frac{3}{2}b_2^2 + 3b_1b_2 - b_4 + 5b_3 + \frac{7}{2}b_1^2\Big) \tag{6.2.25}$$

$$a_2 = 0, \qquad a_1 = \frac{1}{8}\Big(\frac{5}{2}b_2^2 + 13b_1b_2 + b_4 + 3b_3 + \frac{17}{2}b_1^2\Big) \tag{6.2.26}$$

$$a_0 = 0, \qquad \omega = \frac{1}{2}\Big(b_4 + 3b_3 - \frac{15}{2}b_1^2 - \frac{3}{2}b_2^2 - 3b_1b_2\Big). \tag{6.2.27}$$

By applying transformation obtained from generating function 6.2.23 and multiplying w to the both side of BNF we end up:

$$G = Tv(s_0) + wI_0 + \frac{1}{2}\Lambda(s_0)I_1^2 + w\omega I_0^2 + O(6) \tag{6.2.28}$$

where,

$$\begin{cases} w(y) = \sqrt{\Omega(s_0)\Big(\frac{1}{2}\Lambda''(s_0)y^2 + Tv''(s_0)\Big)}, \\ \omega(y) = \frac{1}{2}\Big[b_4 + 3b_3 - \frac{15}{2}b_1^2 - \frac{3}{2}b_2^2 - 3b_1b_2\Big]. \end{cases} \tag{6.2.29}$$

Recall that $I_1 = y, I_0 = \frac{1}{2}(\tilde{u}^2 + \tilde{p_u}^2)$. Given above let's rewrite the normalized Hamiltonian 6.2.28:

$$G = Tv(s_0(I_1)) + w(I_1)I_0 + \frac{1}{2}\Lambda(s_0(I_1))I_1^2 + (w\omega)(I_1)I_0^2 + O(6). \tag{6.2.30}$$

Example 6.2.2. The un-reduced Hamiltonian for the Nosé–Hoover thermostat is

$$F = H_{Nos\acute{e}}(q,p,s,p_s) = \frac{p^2}{2ms^2} + \frac{p_s^2}{2Q} + kTln(s). \qquad (6.2.31)$$

The functional parameters are $h(s) = s, \quad u(s) = 1, \quad v(s) = ln(s)$. We re-write Hamiltonian to

$$F(q,p,s,p_s) = \frac{1}{2}y^2\Lambda(s) + \frac{1}{2}\Omega(s)p_s^2 + Tv(s). \qquad (6.2.32)$$

Computing the equilibrium of F leads to the equation $s_0(y) = \frac{1}{\sqrt{mkT}}y$. Also the quantities λ, w, ω, b_1 and b_3 are as follow:

$$b_1(y) = \frac{m^{1/4}}{6Q^{1/4}y^{1/2}}, \qquad b_3(y) = \frac{m^{1/2}}{8Q^{1/2}y}, \qquad (6.2.33)$$

$$\lambda^4(y) = \frac{-mQ(kT)^2}{y^2}, \qquad w^2(y) = \frac{-4m(kT)^2}{Qy^2}, \qquad (6.2.34)$$

$$\omega(y) = \frac{-11m^{1/2}}{24iQ^{1/2}y}. \qquad (6.2.35)$$

Considering $I_1 = y$ and $I_0 = \frac{1}{2}(\tilde{u}^2 + \tilde{p_u}^2)$, now according to 6.2.30 we write the BNF of $H_{Nos\acute{e}}$ as follows:

$$G(I_0,I_1) = kT\ln(\frac{I_1}{\sqrt{mkT}}) + (\frac{(2m)^{1/2}kT}{Q^{1/2}})\frac{I_0}{I_1} + \frac{1}{2}\frac{kT}{I_1^2}I_1^2 - (\frac{11mkT}{12Q})\frac{I_0^2}{I_1^2} + O(6) \quad (6.2.36)$$

which simplifies to

$$G(I_0,I_1) = \frac{1}{2}kT - kT\ln(\sqrt{mkT}) + kT\ln(I_1) + (\frac{(2m)^{1/2}kT}{Q^{1/2}})\frac{I_0}{I_1} - (\frac{11mkT}{12Q})\frac{I_0^2}{I_1^2} + O(6). \qquad (6.2.37)$$

Now we can compute the iso-energetic non-degeneracy condition for un-reduced Nosé Hamiltonian.

6.2.4 Non degeneracy condition for unreduced Hamiltonian

By definition the frequency of the system $\nabla G : U \subset \mathbb{R}^2 \longrightarrow \mathbb{R}^2$

$$\nabla G(I_0, I_1) = \left(\frac{\partial G}{\partial I_0}, \frac{\partial G}{\partial I_1} \right). \tag{6.2.38}$$

To satisfy the iso-energetic condition for the unreduced Hamiltonian, the corresponding bordered Hessian matrix needs to be non-singular.

We now then compute the entries of the matrix:

$$\begin{aligned} \frac{\partial G}{\partial I_0} &= w(I_1) + 2I_0(w\omega)(I_1) + O(4), \\ \frac{\partial G}{\partial I_1} &= Ts_0'(I_1)v'(s_0(I_1)) + I_0 w'(I_1) \\ &\quad + \frac{1}{2}(\Lambda(s_0(I_1))I_1^2)' + (w\omega)'(I_1)I_0^2 + O(6) \end{aligned} \tag{6.2.39}$$

$$\begin{aligned} \frac{\partial^2 G}{\partial I_0^2} &= 2(w\omega)(I_1) + O(2), \\ \frac{\partial^2 G}{\partial I_1^2} &= (Ts_0'(I_1)v'(s_0(I_1)))' + I_0 w''(I_1) \\ &\quad + \frac{1}{2}(\Lambda(s_0(I_1))I_1^2)'' + (w\omega)''(I_1)I_0^2 + O(6), \\ \frac{\partial^2 G}{\partial I_1 \partial I_0} &= w'(I_1) + 2I_0(w\omega)'(I_1) + O(4). \end{aligned} \tag{6.2.40}$$

The resulting bordered Hessian is:

$$\nabla^2 G = \begin{pmatrix} \frac{\partial^2 G}{\partial I_0^2} & \frac{\partial^2 G}{\partial I_0 \partial I_1} & \frac{\partial G}{\partial I_0} \\ \frac{\partial^2 G}{\partial I_1 \partial I_0} & \frac{\partial^2 G}{\partial I_1^2} & \frac{\partial G}{\partial I_1} \\ \frac{\partial G}{\partial I_0} & \frac{\partial G}{\partial I_1} & 0 \end{pmatrix} = \tag{6.2.41}$$

$$
\begin{pmatrix}
2(w\omega)(I_1)+O(2) & w'(I_1)+2I_0(w\omega)'(I_1)+O(4) & w(I_1)+2I_0(w\omega)(I_1)+O(4) \\
 & & \\
* & \begin{array}{l}(Ts_0'(I_1)v'(s_0(I_1)))'+I_0w''(I_1)\\ +\frac{1}{2}(\Lambda(s_0(I_1))I_1^2)'' \\ +(w\omega)''(I_1)I_0^2+O(6)\end{array} & \begin{array}{l}Ts_0'(I_1)v'(s_0(I_1))+I_0w'(I_1)\\ +\frac{1}{2}(\Lambda(s_0(I_1))I_1^2)' \\ +(w\omega)'(I_1)I_0^2+O(6)\end{array} \\
 & & \\
* & * & 0
\end{pmatrix}
\tag{6.2.42}
$$

then, the condition $\det(H) \neq 0$ guarantees that the majority of KAM tori survive after the perturbation.

The domain of the "det" is a jet space, one k-jet for each functional parameter, in which the zero level set of "det" is an algebraic sub-variety of that jet space since it is a zero set of a polynomial function. This rises a helpful result since each connected component of the algebraic variety is stratified into a union of submanifolds and the top-dimensional stratum is dense in the component.

The condition $\det(H) \neq 0$ can be represented in terms of the jet bundle. In particular, $\det(H) = 0$ corresponds to the zero level set of the "det" map over the jet bundle $J^6(\mathbb{R}, \mathbb{R}^3)$ mapping to $\mathbb{R}$. Invoking Thom transversality theorem (see theorem 1.9.13), C is the set of three products of 6-jets for which $\det(H) = 0$ that is to say the algebraic variety according to the last paragraph. So any perturbation on f either is going to stay off the C, or crosses C transversally. Therefore by the Thom transversality theorem, C^r-mappings are dense in C^r-topology for $r > 6$. In particular the set of mappings $f : \mathbb{R} \to \mathbb{R}^3$ whose 6-jet bundle extension is transverse to C, is everywhere dense; resulting that the condition $\det(H) = 0$ does not typically occur which means we can perturb the functional parameters arbitrarily so that the determinant is non-zero.

Bibliography

[Arn63] V. I. Arnol d. “Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian”. In: *Uspehi Mat. Nauk* 18.5(113) (1963), pp. 13–40. ISSN: 0042-1316.

[Kol79] Academician A. N. Kolmogorov. “Preservation of conditionally periodic movements with small change in the Hamilton function”. In: *Stochastic Behavior in Classical and Quantum Hamiltonian Systems*. Ed. by Giulio Casati and Joseph Ford. Berlin, Heidelberg: Springer Berlin Heidelberg, 1979, pp. 51–56. ISBN: 978-3-540-35510-6.

[Arn88] V. I. Arnol d. *Geometrical methods in the theory of ordinary differential equations*. Second. Vol. 250. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Translated from the Russian by Joseph Szücs [József M. Szűcs]. Springer-Verlag, New York, 1988, pp. xiv+351. ISBN: 0-387-96649-8. DOI: 10.1007/978-1-4612-1037-5. URL: https://doi-org.uml.idm.oclc.org/10.1007/978-1-4612-1037-5.

[Jel88] Julius Jellinek. “Dynamics for nonconservative systems: ergodicity beyond the microcanonical ensemble”. In: *The Journal of Physical Chemistry* 92.11 (1988), pp. 3163–3173.

[JB88] Julius Jellinek and R. Stephen Berry. "Generalization of Nosé's isothermal molecular dynamics". In: *Phys. Rev. A* 38 (6 Sept. 1988), pp. 3069–3072. DOI: 10.1103/PhysRevA.38.3069. URL: https://link.aps.org/doi/10.1103/PhysRevA.38.3069.

[JB89] Julius Jellinek and R Stephen Berry. "Generalization of Nosé's isothermal molecular dynamics: necessary and sufficient conditions of dynamical simulations of statistical ensembles". In: *Physical Review A* 40.5 (1989), p. 2816.

[Mos01] J. Moser. "Remark on the paper: "On invariant curves of area-preserving mappings of an annulus" [Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II **1962**, 1–20; MR0147741 (26 #5255)]". In: *Regul. Chaotic Dyn.* 6.3 (2001), pp. 337–338. ISSN: 1560-3547,1468-4845. DOI: 10.1070/RD2001v006n03ABEH0001 URL: https://doi.org/10.1070/RD2001v006n03ABEH000181.

[Wig03] Stephen Wiggins. *Introduction to applied nonlinear dynamical systems and chaos*. Second. Vol. 2. Texts in Applied Mathematics. Springer-Verlag, New York, 2003, pp. xxx+843. ISBN: 0-387-00177-8.

[Gre06] Benoit Grebert. *Birkhoff Normal Form and Hamiltonian PDEs*. 2006. arXiv: math/0604132 [math.AP].

[LLM09] Frédéric Legoll, Mitchell Luskin, and Richard Moeckel. "Non-ergodicity of Nosé–Hoover dynamics*". In: *Nonlinearity* 22.7 (June 2009), p. 1673. DOI: 10.1088/0951-7715/22/7/011. URL: https://dx.doi.org/10.1088/0951-7715/22/7/011.

[TZ10] Dmitry Treschev and Oleg Zubelevich. *Introduction to the perturbation theory of Hamiltonian systems*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2010, pp. x+211. ISBN: 978-3-642-03027-7. DOI:

10.1007/978-3-642-03028-4. URL: https://doi.org/10.1007/978-3-642-03028-4.

[Lee13] John M. Lee. *Introduction to smooth manifolds*. Second. Vol. 218. Graduate Texts in Mathematics. Springer, New York, 2013, pp. xvi+708. ISBN: 978-1-4419-9981-8.

[But16] Leo T. Butler. "Invariant tori for the Nosé thermostat near the high-temperature limit". In: *Nonlinearity* 29.11 (2016), pp. 3454–3463. ISSN: 0951-7715. DOI: 10.1088/0951-7715/29/11/3454. URL: https://doi-org.uml.idm.oclc.org/10.1088/0951-7715/29/11/3454.

[But18] Leo T. Butler. "Nosé-thermostated mechanical systems on the n-torus". In: *Arch. Ration. Mech. Anal.* 227.2 (2018), pp. 855–867. ISSN: 0003-9527. DOI: 10.1007/s00205-017-1175-3. URL: https://doi-org.uml.idm.oclc.org/10.1007/s00205-017-1175-3.

[BG22] Joackim Bernier and Benoît Grébert. "Birkhoff normal forms for Hamiltonian PDEs in their energy space". In: *J. Éc. polytech. Math.* 9 (2022), pp. 681–745. ISSN: 2429-7100. DOI: 10.5802/jep.193. URL: https://doi-org.uml.idm.oclc.org/10.5802/jep.193.

[Arn89] V. I. Arnol d. *Mathematical methods of classical mechanics*. Vol. 60. Graduate Texts in Mathematics. Translated from the 1974 Russian original by K. Vogtmann and A. Weinstein, Corrected reprint of the second (1989) edition. Springer-Verlag, New York, [1989], pp. xvi+516. ISBN: 0-387-96890-3.