Lecture Notes on Conservative Dynamics and KAM Theory

Preliminary version February 1998

Introduction

These are notes, taken by Alexandre Baraviera, of a graduate course I taught at IMPA in January-February 1998. The goal was to give an introduction to Conservative Dynamics, including the proof of a KAM theorem. In the absence of a comprehensive text, I used several references that are listed in the bibliography. I also benefitted from a series of lectures given by M. Herman during his 1997 stay at IMPA.

Alexandre and I have been working on polishing the text, but the version you see here is still rather preliminary.

Marcelo Viana

Contents

1 Hamiltonian systems

Here we give a brief review of some basic notions and results in Classical Mechanics. Further information and proofs can be found in the book of Arnold [1].

1.1 The equations of Hamilton-Jacobi

In the Hamiltonian formulation of Classical Mechanics the laws of motion are described by the *equations of Hamilton-Jacobi:*

$$\frac{dq_i}{dt}(t) = \frac{\partial H}{\partial p_i}(q,p,t), \quad \frac{dp_i}{dt}(t) = -\frac{\partial H}{\partial q_i}(q,p,t), \qquad i = 1, ..., n. \tag{1}$$

where t denotes *time*, $q = (q_1, \dots, q_n)$ are the *configuration coordinates*, and $p = (p_1, \dots, p_n)$ are the *momenta*. For the time being we take q to vary in some open subset U of $\mathbb{R}^n$ and $p \in \mathbb{R}^n$, but in the next section we will introduce a broader setting. The function $H : M \to \mathbb{R}$, $M = U \times \mathbb{R}^n$ is the *Hamiltonian*, and can often be interpreted as the total energy of the mechanical system. We shall always consider the system to be autonomous, i.e., H does not depend explicitly on t. Moreover, we suppose that H is of class C^2.

Example 1.1. Take $n = 1$, $(q,p) \in \mathbb{R}^2$, and

$$H(q,p) = \frac{1}{2}p^2 - g\cos q$$

where g is a positive constant. This describes the pendulum with length and mass equal to 1, subject to a constant gravitational field. The coordinate q describes the angle relative to the position of (stable) equilibrium, and p corresponds to the linear momentum. The Hamiltonian can be interpreted as the mechanical energy of the system: $H =$ kinetic energy + potential energy, representing by g the gravitational acceleration. The equations of motion are

$$\begin{cases} \dot{p} = -g\sin q \\ \dot{q} = p \end{cases}$$

(dots represent derivative with respect to time). This first order system that can be written as $\ddot{q} = \dot{p} = -g\sin q$, corresponding to Newton's third law.

A differentiable function $I(q,p)$ is a *first-integral* of the system if it is constant along all the trajectories of the flow defined by (1): given any solution $(q(t),p(t))$ of the equations, then

$$\begin{aligned}\frac{d}{dt}I(q(t),p(t))=0 \Leftrightarrow \sum_{i=1}^{n}\frac{\partial I}{\partial q_i}(q(t),p(t))\frac{dq_i}{dt}(t)+\frac{\partial I}{\partial p_i}(q(t),p(t))\frac{dp_i}{dt}(t)=0\\ \Leftrightarrow \sum_{i=1}^{n}\left(\frac{\partial I}{\partial q_i}\frac{\partial H}{\partial p_i}-\frac{\partial I}{\partial p_i}\frac{\partial H}{\partial q_i}\right)(q(t),p(t))=0\end{aligned}$$

Let us define the *Poisson bracket* $\{F_1,F_2\}$ of two functions F_1 and F_2 by

$$\{F_1,F_2\}=\sum_{i=1}^{n}\frac{\partial F_1}{\partial p_i}\frac{\partial F_2}{\partial q_i}-\frac{\partial F_1}{\partial q_i}\frac{\partial F_2}{\partial p_i}.$$

It is immediate from the definition that $\{F_1,F_2\}$ is a bilinear and anti-symmetric function of F_1, F_2. Moreover, I is a first-integral of (1) if and only if $\{H,I\}=0$. In particular, the Hamiltonian H itself is always a first-integral (this is the well-known principle of conservation of energy in mechanical systems).

A system is said to be *integrable* (in the sense of Liouville) if it admits n first-integrals $I_1,\ldots,I_n$ which are

(a) *independent:* the vectors

$$\operatorname{grad} I_j=\left(\frac{\partial I_j}{\partial q_1},\frac{\partial I_j}{\partial p_1},\ldots,\frac{\partial I_j}{\partial q_n},\frac{\partial I_j}{\partial p_n}\right),\quad j=1,\ldots,n,$$

are linearly independent on an open and dense subset of $M=U\times\mathbb{R}^n$.

(b) *in involution:* the Poisson brackets $\{I_j,I_k\}$ are identically zero on M for all $1\le j\le n$ and $1\le k\le n$.

A classical theorem of Liouville says that *if the system is integrable then the equations of Hamilton-Jacobi can be solved by quadratures.*

Let us list a few important examples of integrable systems.

Example 1.2. Every (autonomous) Hamiltonian system with one degree of freedom, i.e. with $n=1$, is integrable: $I_1=H$ is always a first-integral , as noted before. In particular, this applies to Example 1.1.

Example 1.3. Given any $n \geq 1$, suppose that the Hamiltonian H depends only on the variables $p = (p_1, \ldots, p_n)$. Then the Hamilton-Jacobi equations reduce to

$$\dot{p}_i = -\frac{\partial H}{\partial q_i}(p) = 0 \quad \text{and} \quad \dot{q}_i = \frac{\partial H}{\partial p_i}(p).$$

The first equation means, precisely, that every p_i is a first-integral , and it is very easy to see that these first-integrals are independent and in involution. Observe also that, then the right hand side of the second equation is independent of time t, and so the solutions may be written

$$q_i(t) = q_i(0) + \frac{\partial H}{\partial p_i}(p(0))\, t.$$

Example 1.4. Lagrange's top: ***

Example 1.5. Motion in a central force: ***

Example 1.6. Toda's molecule: This is a system with 3 degrees of freedom and Hamiltonian function given by

$$H = \frac{1}{2}(p_1^2 + p_2^2 + p_3^2) + e^{q_1 - q_2} + e^{q_2 - q_3} + e^{q_3 - q_1}.$$

These are 3 first-integrals: the energy H, the momentum of the center of mass $P = p_1 + p_2 + p_3$, and

$$\begin{aligned} K = & \frac{1}{9}\,(p_1 + p_2 - 2p_3)(p_2 + p_3 - 2p_1)(p_3 + p_1 - 2p_2) - \\ & - (p_1 + p_2 - 2p_3)e^{q_1 - q_2} - (p_2 + p_3 - 2p_1)e^{q_2 - q_3} - (p_3 + p_1 - 2p_2)e^{q_3 - q_1} \end{aligned}$$

A few comments are in order on the role of condition (b) above. Let us define the Hamiltonian vector field of H:

$$X_H = \left(\frac{\partial H}{\partial p_1}, \ldots \frac{\partial H}{\partial p_n}, -\frac{\partial H}{\partial q_1}, \ldots - \frac{\partial H}{\partial q_n}\right).$$

Recall also that the *Lie bracket* of two vector fields $X = (X_1, \ldots, X_m)$ and $Y = (Y_1, \ldots, Y_m)$ in $\mathbb{R}^m$ is defined by $[X, Y] = ([X, Y]_1, \ldots, [X, Y]_n)$ with

$$[X, Y]_i = \sum_{j=1}^{m} Y_j \frac{\partial X_i}{\partial x_j} - X_j \frac{\partial Y_i}{\partial x_j}.$$

As noted before, a function I_j is a first-integral for the Hamiltonian system of H is and only if $\{H, I\}$. So, to say that I_j and I_k are in involution, is just to say that each one of them is constant along trajectories of the Hamiltonian flow of the other. Moreover, these flows commute:

$$X_{I_j}^s \circ X_{I_k}^t = X_{I_k}^t \circ X_{I_j}^s \quad \text{for every } s, t.$$

Indeed, this is the same as saying that

$$[X_{I_j}, X_{I_k}] = 0, \tag{2}$$

and this is a consequence of the following lemma (which, on its turn, can be proved by a direct calculation).

Lemma 1.7. *For any C^2 functions H_1, H_2 we have $X_{\{H_1,H_2\}} = [X_{H_1}, X_{H_2}]$.*

Let us push the consequences of condition (b) further on. We restrict ourselves to those points where the independence condition (a) holds, and so the Hamiltonian vector fields $X_{I_1}, \ldots, X_{I_n}$ are linearly independent.

By Frobenius theorem, the commutation property (2) implies that the distribution of 2-planes generated by $\{X_{I_j}, X_{I_k}\}$ is integrable: each point is contained in a unique surface that is everywhere tangent to the plane generated by the two vector fields. For the same reasons, the distribution of n-planes generated by $\{X_{I_1}, \ldots, X_{I_n}\}$ is also integrable: each point (q,p) belongs in an n-dimensional manifold $T = T_{(q,p)}$ formed by flow lines of these n vector fields $X_{I_1}, \ldots, X_{I_n}$. Moreover, the restriction of every I_j to this manifold T is constant. In fact, T must coincide with a level set $\{I = \text{const}\}$, at least locally, since this last set is also an n-dimensional submanifold (in view of the independence condition). We can always suppose $I_1 = H$, and then T is invariant by the Hamiltonian flow of H, and it is contained in an *energy surface* $\{H = \text{const}\}$.

Proposition 1.8. *If T is compact then T is diffeomorphic to the n-torus.*

Indeed, every compact n-dimensional manifold that supports n linearly independent fields and pairwise commuting vector fields is diffeomorphic to the n-torus. See [1, §49] for a proof.

In the proof one also constructs angular coordinates $\varphi = (\varphi_1, \ldots, \varphi_n)$ on the torus T such that the transformation

$$(q,p) \to (\varphi, I)$$

defines a change of coordinates that preserves the form of the Hamilton-Jacobi equations (a *canonical transformation*). Let us explain this in more detail. First of all, a differentiable map $\Psi : \mathbb{T}^n \times B_r^n \to M$, $(\varphi, I) \to \Psi(\varphi, I)$ is constructed which is a diffeomorphism from the product of the standard n-torus $\mathbb{T}^n$ by an r-ball $B_r^n \subset \mathbb{R}^n$ onto a tubular neighborhood of T. *Each set $\{I = cte\}$ is an invariant torus of the Hamiltonian flow,* which, in these coordinates, is still given by equations of the form (1):

$$\frac{d\varphi_i}{dt} = \frac{\partial H}{\partial I_i}(\varphi, I), \quad \frac{dI_i}{dt} = \frac{\partial H}{\partial \varphi_i}(\varphi, I)$$

(note that we continue to denote H the expression of the Hamiltonian in the new coordinates, an abuse of language will be recurrent throughout our text). Since the I_j are first-integrals of the system, the second equation gives

$$0 = \frac{dI_i}{dt} = \frac{\partial H}{\partial \varphi_i}(\varphi, I),$$

which means that H does not depend on the variables φ (H is constant on each torus). Compare Example 1.3, the right hand side of the first equation is constant in time. As a consequence, *the Hamiltonian flow on each invariant torus is linear*, in φ coordinates

$$\varphi_j(t) = \varphi_j(0) + \omega_j(I)t, \quad \omega_j(I) = \frac{\partial H}{\partial I_i}(I).$$

Such I, φ are usually called *action-angle variables*. Observe that what we have been saying implies that typical solutions of integrable systems can be written in the form

$$t \to \Psi(\varphi(0) + \omega(I)t, I) = \psi_{\varphi(0),I}(\omega(I)t) \tag{3}$$

where the function $\psi_{\varphi(0),I} : \mathbb{R}^n \to M$ is $\mathbb{Z}^n$-periodic.

1.2 Non-integrability. The N-body problem

Having obtained such a complete description of the behaviour of integrable flows, it is natural to ask how typical is the property of integrability among Hamiltonian systems. Let us discuss this problem in the setting of the *N-body problem*, for which it was first formulated, and where it is directly related to the problem of the *stability of the Solar system*.

One considers N massive bodies of neglectable dimensions moving in d-dimensional space (usually, $d = 2$ or $d = 3$) and interacting through Newtonian gravitation. That is, denoting m_i, $1 \leq i \leq N$, the masses and

$Q_i = (q_{i,1}, \dots, q_{i,d})$, $1 \leq i \leq N$, the positions, the jth body is subjet to a force

$$F_j = \sum_{i \neq j} \frac{G m_i m_j}{\|Q_i - Q_j\|^2} \frac{Q_i - Q_j}{\|Q_i - Q_j\|}$$

where G is the universal gravitational constant. It is not difficult to see that this force field derives from a potential:

$$F_j = \partial U / \partial Q_j \quad \text{where} \quad U(Q_1, \dots, Q_N) = -\sum_{i \neq j} G \frac{m_i m_j}{\|Q_i - Q_j\|}.$$

The system has $n = Nd$ degrees of freedom, with $q_{i,l}$ as configuration coordinates and $p_{i,l} = m_i \dot{q}_{i,l}$ as momenta. The Hamiltonian is just the total energy

$$H = \sum_{i=1}^{N} \frac{1}{2} p_i^2 + U(Q_1, \dots, Q_N).$$

A few first-integrals can be found from physical laws of conservation (the energy, linear momenta, angular momenta), from which one can show that the system is integrable when $N = 2$.

However, it was soon realized that this is a rather special, and that "solving" the equations of motion for systems with more than 2 bodies, including our Solar system, posed formidable difficulties. A source of inspiration came from the fact that the mass of the Sun, denoted m_1, is much larger than the mass of any other object in the Solar system. Thus, as a first approximation, one may try to solve the equations of motion neglecting the interaction between these other objects: the potential U is replaced by

$$U_0(Q_1, \dots, Q_N) = -\sum_{i \neq 1} G \frac{m_i m_1}{\|Q_i - Q_1\|}.$$

This approximation corresponds, simply, to $N - 1$ uncoupled 2-body problems, and so it is completely integrable.

The problem is to understand to what extent the mutual attraction between planets, comets, and asteroids, modifies the overall evolution of the system. In particular, *is the Solar system stable, that is, will it keep the present form forever ?* Or can these secondary gravitational effects cause trajectories to change so much over long periods of time that *some of the planets will eventually leave the system, or collide with each other ?* With this question in mind, people like Laplace, Lagrange, Leverrier, and others, devoted a great deal of effort to calculating with increasing accuracy the corrections (*secular terms*) introduced in the solutions of the equations of motion by the interactions between other bodies (specially the massive planets Jupiter and Saturn).

Writing the potential U as a perturbation of the integrable one U_0,

$$U = U_0 + \varepsilon U,$$

where $\varepsilon > 0$ corresponds to the largest quotient of masses m_i/m_1, $i \neq 1$, one may find successive approximations to the actual motion of the form which are polinomial in the parameter ε, with quasi-periodic coefficients. Cf. (3). In fact, the solutions of the original, unsimplified, system may be formally expressed as

$$\phi_0(\omega t) + \sum_{j=1}^{\infty} \varepsilon^j \phi_j(\omega^{(j)} t), \tag{4}$$

where $\omega, \omega^{(j)} \in \mathbb{R}^n$ and each ϕ_j is a $\mathbb{Z}^n$-periodic function. More precisely, such a formal series expansion of the solutions can be found for all frequency vectors $\omega = (\omega_1, \ldots, \omega_n)$ which are *non-resonant:*

$$k \cdot \omega = \sum_{i=1}^{n} k_i \omega_i \neq 0 \quad \text{for all } k = (k_1, \ldots, k_n) \in \mathbb{Z}^n - \{0\} \tag{5}$$

We illustrate this, as well as the role played by ressonances, through a simplified problem. Let us consider the system of complex differential equations

$$\begin{cases} \dot{z}_1 = i\omega_1 z_1 \\ \dot{z}_2 = i\omega_2 z_2 + \varepsilon z_1^{k_1} z_2^{k_2} \end{cases}$$

where $\varepsilon \geq 0$, $\omega_1, \omega_2 \in \mathbb{R}$, and k_1, k_2 are integers. For $\varepsilon = 0$ the solutions are quasi-periodic functions

$$z_1(t) = z_1(0) e^{i\omega_1 t} \quad \text{and} \quad z_2(t) = z_2(0) e^{i\omega_2 t}.$$

Then, one may try to find the solution of the systems for $\varepsilon > 0$ by adding a perturbation term, analytic in ε,

$$z_2(t) = z_2(0) e^{i\omega_2 t} + \sum_{j=1}^{\infty} \epsilon^j w_j(t). \tag{6}$$

Replacing in the equation of z_2,

$$\sum_{j=1}^{\infty} \varepsilon^j \dot{w}_j(t) = i\omega_2 \sum_{j=1}^{\infty} \varepsilon^j w_j(t) + \varepsilon \left(z_1(0)e^{i\omega_1 t}\right)^{k_1} \left(z_2(0)e^{i\omega_2 t} + \sum_{j=1}^{\infty} \epsilon^j w_j(t)\right)^{k_2}.$$

Comparing the terms of order 1, one gets a linear non-homogeneous equation

$$\dot{w}_1(t) = i\omega_2 \sum_{j=1}^{\infty} w_1(t) + Ae^{i(k_1\omega_1 + k_2\omega_2)t}, \quad A = z_1(0)^{k_1} z_2(0)^{k_2},$$

whose solutions are given by

$$w_1(t) = e^{i\omega_2 t} \int Ae^{i(k_1\omega_1 + (k_2-1)\omega_2)s}\, ds.$$

If $k_1\omega_1 + (k_2 - 1)\omega_2 \neq 0$ then

$$w_1(t) = \frac{A}{k_1\omega_1 + (k_2-1)\omega_2} e^{i(k_1\omega_1 + k_2\omega_2)t} + Be^{i\omega_2 t} \qquad \text{(for some } B\text{)},$$

is quasi-periodic and, thus, bounded. On the other hand, the "perturbation term" $\epsilon w_1(t)$ *may take large values if the divisor* $k_1\omega_1 + (k_2 - 1)\omega_2$ *happens to be small*. If $k_1\omega_1 + (k_2 - 1)\omega_2 = 0$ then

$$w_1(t) = e^{i\omega_2 t}(At + B),$$

is not even bounded. Similar conclusions apply to all the $\varepsilon^j w_j(t)$, $j \geq 1$: they are all well-defined and quasi-periodic, if one supposes $\omega = (\omega_1, \omega_2)$ to be non-resonant; yet, even under this assumption, it is not clear whether the series in (6) should be convergent, due to the presence of these small divisors.

By the end of last century, king Oscar II of Sweden decided to sponsor an international mathematical competition. Weierstrass, invited to join the jury, suggested the formulation of the Prize Problem: *to prove that the formal series (4) do converge, and thus express the general solution of the* N*-body problem.* Poincaré was to be the winner of the competition but, to Weierstrass surprise, he concluded just the opposite!

More specifically, Poincaré was considering the planar restricted 3-body problem: one of the bodies has mass $m_3 = 0$ (more precisely, one considers the limit system when $m_3 \to 0$) and so it does not affect the motion of the other two m_1, m_2; consequently, m_1, m_2 move according to solutions of a 2-body problem, that is, along conic curves; one supposes that these conics are

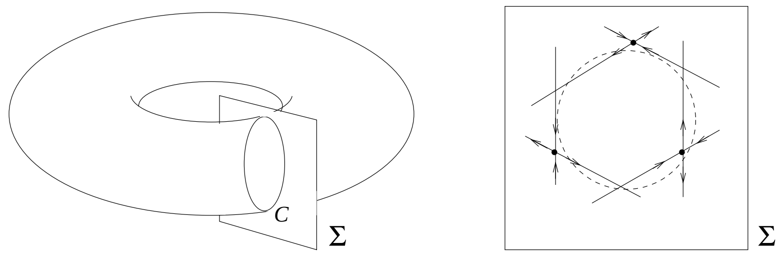

Figure 1:

circles, and that m_3 moves in the same plane as m_1, m_2. The problem, to describe this motion of m_3 has $n = 2$ degrees of freedom (the coordinates of m_3 in the plane). The special case $m_2 = 0$ is integrable, in fact it corresponds to two uncoupled 2-body systems. However, as shown by Poincaré, no first-integrals can be expressed as *convergent* power series $\sum_{k=0}^{\infty}(m_2)^k \phi_k$ for small positive m_2.

Going on to try and understand this phenomenon from a geometric point of view, Poincaré realized that *resonant invariant tori of an integrable system tend to be destroyed when the system is slightly perturbed*. Part of what is going on can be described by considering a cross-section Σ to the flow, and the first-return map P to Σ (precise statements wil appear later). Invariant tori of the unperturbed (integrable) flow intersect Σ in circles $\mathcal{C}$ which are invariant under the first-return map. If the frequency vector $\omega = (\omega_1, \omega_2)$ is resonant, in other words, if ω_2/ω_1 is rational then P is conjugate to the rigid rotation of angle $p/q = \omega_2/\omega_1$. In particular, all the points in C are periodic for P, with a same period q.

Such a continuum of periodic points is easily destroyed, by arbitrarily small perturbation of the Hamiltonian. On the other hand, some periodic orbits do persist in the perturbed system, including at least one hyperbolic periodic orbit of period q. Moreover, the corresponding stable and unstable manifolds intersect each other transversely. See Figure 1. Poincaré realized that in the presence of such *homoclinic intersections*, the behaviour of nearby trajectories must be extremely complex, and so the system can not be integrable. See [**?**, vol 3, p. 389].

In this respect, Poincaré's intuition went beyond what he could obtain rigorously at the time. He did not actually prove the existence of homoclinic orbits in the restricted 3-body problem (which was done only recently). More important, his methods were not conclusive to prove divergence in the non-resonant case:

"... les séries ne pourraient-elles pas, par exemple, converger quand x_1^0 et x_2^0 ont été choisis de telle sorte que le raport n_1/n_2 soit incommensurable, et que son carré soit au contraire commensurable (ou quand le raport n_1/n_2 est assujetti à une autre condition analogue à celle que je viens d'énoncer un peu au hasard) ? Les raisonnements de ce Chapitre ne me permettent pas d'affirmer que ce fait ne se présentera pas. Tout ce qu'il m'est permis de dire, c'est qu'il est fort invraisemblable. [**?**, vol 2, pp. 104-105].

The problems of the solvability of the general N-body problem, and of the stability of the Solar system, were to remain wide open for yet a half a century.

1.3 The theorem of Kolmogorov-Arnold-Moser

An answer to these problems was finally given by the theorem of Kolmogorov, Arnold, Moser (KAM): *most invariant tori of persist when a (nondegenerate) integrable system is slightly perturbed*. More precisely, a positive fraction (in volume) of phase space is occupied by invariant tori of the perturbed system, and this fraction goes to 1 when the size of the perturbation decreases to zero. A formal statement is given below.

It turns out that persistence or not of an invariant torus is directly related to the corresponding frequency vector ω, more specifically, to its arithmetic properties. As already mentioned, torus with resonant frequencies

$$k \cdot \omega \quad \text{for some } k \in \mathbb{Z}^n \setminus \{0\}$$

can be easily destroyed (precise statements will appear later in the text). So, one may expect that robust tori such as granted by the KAM theorem should correspond to vectors ω which are "far from being resonant", and this is indeed so.

Given $c > 0$ and $\tau > 0$, we say that $\omega \in \mathbb{R}^n$ is (c, τ)-*Diophantine* if

$$|k \cdot \omega| \geq \frac{c}{\|k\|^\tau} \quad \text{for all } k \in \mathbb{Z}^n,$$

where $\|k\| = |k_1| + \cdots |k_n|$. It is a classical fact that this condition is satisfied by some $\omega \in \mathbb{R}^n$ if and only if $\tau \geq n-1$. On the other hand, if τ is fixed strictly larger than $n-1$, then the set of all ω that are (c,τ)-Diophantine for some $c > 0$ has full measure in $\mathbb{R}^n$.

Let H_0 be an integrable C^∞ Hamiltonian, with $n \geq 1$ degrees of freedom. Let (φ, I) be action-angle variables for H_0, as in Section 1.1: H_0 depends only on $I = (I_1, \ldots, I_n)$, each I_j is a first-integral for H_0, and each φ_j evolves linearly with time t:

$$\varphi_j(t) = \varphi(0) + t\,\omega_j(I), \quad \omega_j(I) = \frac{\partial H_0}{\partial I_j}(I).$$

Recall that φ lives in the n-torus $\mathbb{T}^n$, whereas I may be taken in an open ball $D_r^n \subset \mathbb{R}^n$.

Let us say that H_0 is *non-degenerate* (or that it *satisfies a twist condition*) if the map $I \mapsto \omega(I) = (\omega_1(I), \ldots, \omega_n(I))$ is a local diffeomorphism, that is, if

$$\left(\frac{\partial \omega_j}{\partial I_i}\right)_{i,j}(I) = \left(\frac{\partial^2 H_0}{\partial I_i \partial I_j}\right)_{i,j}(I) \tag{7}$$

is a linear isomorphism everywhere on D_r^n.

Theorem 1.9. *Let H_0 be a non-degenerate Hamiltonian as before. Fix $c > 0$ and $\tau \geq n-1$ and a compact subset B of the image $\omega(D_r^n)$ of $I \mapsto \omega(I)$. Then there exists a neighbourhood $\mathcal{V}$ of H_0 in $C^\infty(\mathbb{T}^n \times D_r^n, \mathbb{R})$ such that, given any $H \in \mathcal{V}$ and any (c,τ)-Diophantine vector $\omega \in B$, there exists an embedded n-torus $T_{H,\omega} \subset \mathbb{T}^n \times D_r^n$ such that*

a) $T_{H,\omega}$ is invariant under the Hamiltonian flow of H, and

b) the restriction of this Hamiltonian flow to $T_{H,\omega}$ is C^∞ conjugate to the linear flow $t \mapsto z(t) = z(0) + t\omega$ on $\mathbb{T}^n$.

Moreover, the union of these tori T_ω has positive volume in $\mathbb{T}^n \times D_r^n$.

More detailed information can be given, specially on the way these tori depend with the Hamiltonian. This will be done later, in Section 2.3, where we also state a version of this theorem for discrete-time systems (symplectic diffeomorphisms).

2 Symplectic formalism

In general, the configuration space of a given system can not be described by an open subset of a Euclidean space. For this reason, one needs to extend the notion of Hamiltonian flow introduced above to the manifolds context. A good reference is [1].

2.1 Symplectic manifolds

Let M be a C^∞ m-dimensional manifold, $m \geq 1$. A differential k- form is a mapping that associate to each point $z \in M$ a k-linear alternate transformation

$$\omega_z : T_zM \times \ldots \times T_zM \to \mathbb{R}$$

Given a local system of coordinates $z = (x_1, \ldots, x_n)$ we can write a k-linear alternate transformation ω_z in a unique form

$$\omega_z = \sum_{1 \leq i_1 < \cdots < i_k \leq n} \alpha_{i_1 \cdots i_k}(z) dx_{i_1} \wedge \ldots dx_{i_k}$$

where $dx_i : T_zM \to \mathbb{R}$ is the differential at z of the coordinate function $\pi_i(x_1, \ldots, x_m) = x_i$ and

$$dx_{i_1} \wedge \ldots dx_{i_k}(v_1, \ldots, v_k) = \sum_\sigma dx_{i_1}(v_{\sigma(1)}) \cdots dx_{i_k}(v_{\sigma(k)})(-1)^{\epsilon(\sigma)}$$

where the sum is over all permutations σ of the set $\{1, \ldots, k\}$ and $\epsilon(\sigma)$ denotes the parity of the respective permutation. The differential form ω is said to be C^k if the functions $\alpha_{i_1 \cdots i_k}(z)$ are C^k for every choice of coordinates. In this text, all the forms will be supposed C^∞.

The *exterior derivative* of the k-form ω is the $(k+1)$-form $d\omega$ given, in coordinates, by

$$d\omega_z = \sum_{1 \leq j \leq m} \sum_{1 \leq i_1 < \cdots < i_k \leq m} \frac{\partial \alpha_{i_1, \ldots, i_k}}{\partial x_j} dx_j \wedge dx_{i_1} \wedge \ldots dx_{i_k}$$

Note that a 0-form is just a C^∞ function $f : M \to \mathbb{R}$. The exterior derivative is the 1-form df defined by $df_z = Df(z)$. In coordinates

$$df_z = \sum_{1 \leq i \leq m} \frac{\partial f}{\partial x_i}(z) dx_i$$

It is also easy to see that any k-form with $k > m$ must be identically zero.

Example 2.1. Let $M = \mathbb{R}^{2n}$, with coordinates $z = (q_1, \dots, q_n, p_1, \dots, p_n)$. Then

$$\alpha_z = \sum_{1 \leq i \leq n} p_i dq_i \quad \text{and} \quad \omega_z = \sum_{1 \leq i \leq n} dp_i \wedge dq_i$$

define, respectively, a 1-form α and a 2-form ω in M. Furthermore $\omega = d\alpha$.

A k-form ω is *closed* if $d\omega = 0$, and *exact* if there is a $(k-1)$-form α such that $\omega = d\alpha$. An exact form is necessarily closed, since $d(d\alpha) = 0$ for every differential form α. Every m-form is closed and, by convention, every 0-form is exact.

Let ω be a differential 2-form. We say that ω is *non-degenerate* if for all $z \in M$ and for all $v_1 \in T_zM$, there exists $v_2 \in T_zM$ such that $\omega_z(v_1, v_2) \neq 0$. In other words, for all $z \in M$ the 2-form ω_z induces an isomorphism

$$T_zM \to (T_zM)^*$$

$$v_1 \mapsto \omega_z(v_1, .) : T_zM \to \mathbb{R}$$

that maps V_2 to $\omega_z(v_1, v_2)$. It's easy to verify that if such a form exists then the dimension of the vector space T_zM is necessarily even, $m = 2n$, and $\omega_z^{\wedge n} = \omega_z \wedge \dots \wedge \omega_z$ is a non-null form.

Now we may define the main notions introduced in this section. A 2-form ω on a manifold M is a *sympletic form* if it is closed and non-degenerate. A *symplectic manifold* is a pair (M, ω) where ω is a symplectic form on the smooth manifold M.

If (M, ω) is a symplectic manifold then $dimM$ is even and $\omega^{\wedge n}$ is a volume form on M, i.e., $\omega_z^{\wedge n} \neq 0$ in every point $z \in M$. In particular, M must be orientable.

Example 2.2. The 2-form ω in Example 2.1

$$\omega_z = \sum_{1 \leq i \leq n} dp_i \wedge dq_i$$

is a sympletic form on $M = \mathbb{R}^{2n}$. Note that

$$\omega_z^{\wedge n} = (dp_1 \wedge dq_1) \wedge \dots \wedge (dp_n \wedge dq_n)$$

is the usual volume form on $\mathbb{R}^{2n}$ (up to a sign).

The main examples are provided by the following construction, that generalizes Example 2.2

Example 2.3. Let N be an n-dimensional manifold and $M = T^*N$ be its cotangent bundle. Given $q \in N$ and $p \in (T_qN)^*$ there is a canonical identification

$$T_{(q,p)}(T^*N) = T_qN \times T_p(T_q^*N) = T_qN \times T_q^*N.$$

Then

$$\alpha_{(p,q)} : T_{(q,p)}(T^*N) \to \mathbb{R}, \quad (u,v) \mapsto p(u)$$

defines a canonical 1-form α on $M = T^*N$. We claim that $\omega = d\alpha$ is a symplectic form on M. The fact that ω is non-degenerate can be read from its expression in appropriate coordinates (q,p). Taking the q_i arbitrary local coordinates on N, and choosing the p_i conjugate coordinates on T_q^*N, i.e., $p_i = dq_i$ for $i = 1, \dots, n$, one gets

$$\alpha_z = \sum_{1 \le i \le n} p_i dq_i \quad \text{and} \quad \omega_z = \sum_{1 \le i \le n} dp_i \wedge dq_i$$

Theorem 2.4. (Darboux) *Every symplectic manifold (M,ω) admit an atlas of local coordinates (q,p) in which the symplectic form ω is given by*

$$\omega_z = \sum_{1 \le i \le n} dp_i \wedge dq_i$$

We shall refer to such (q,p) as *canonical coordinates*.

2.2 Hamiltonian flows and symplectic maps

Let (M,ω) be a symplectic manifold. Given a C^k function $H : M \to \mathbb{R}$, $k \ge 2$, the *Hamiltonian vector field* X_H associated to H is the C^{k-1} vector field on M defined by

$$\omega_z(X_H(z), \cdot) = dH_z\,.$$

It is clear that

$$dH_z X_H(z) = \omega_z(X_H(z), X_H(z)) = 0$$

meaning that the field $X_H(z)$ is tangent to the level hypersurface of H at z for every point in M, and so the flow leaves H invariant.

Example 2.5. $M = \mathbb{R}^{2n}$, $\omega = \sum_{1 \leq i \leq n} dp_i \wedge dq_i$. Given $H : M \to \mathbb{R}$ let us calculate the corresponding Hamiltonian vector field. One can write

$$X_H = \sum_{i=1}^{n} A_i \frac{\partial}{\partial p_i} + B_i \frac{\partial}{\partial q_i}$$

and try to find the coefficients A_i, B_i through the relation

$$\omega_z(X_H(z), \sum_{i=1}^{n} \alpha_i \frac{\partial}{\partial p_i} + \beta_i \frac{\partial}{\partial q_i}) = dH_z(\sum_{i=1}^{n} \alpha_i \frac{\partial}{\partial p_i} + \beta_i \frac{\partial}{\partial q_i})$$

valid for all vectors $\sum_{i=1}^{n} \alpha_i \frac{\partial}{\partial p_i} + \beta_i \frac{\partial}{\partial q_i} \in T_z M$. The result is $A_i = \frac{\partial H}{\partial p_i}$ and $B_i = -\frac{\partial H}{\partial q_i}$, and so we have the expression of the field:

$$X_H = \left(\frac{\partial H}{\partial p_1}, \dots, \frac{\partial H}{\partial p_n}, -\frac{\partial H}{\partial q_1}, \dots, -\frac{\partial H}{\partial q_n} \right)$$

Proposition 2.6. *The flow $\phi^t = X_H^t$ preserves the symplectic form ω, i.e.,*

$$(\phi_*^t \omega)_z(v_1, \dots, v_n) = \omega_{\phi^t(z)}(d\phi_z^t v_1, \dots, d\phi_z^t v_n) = \omega_z(v_1, \dots, v_n)$$

for all t. In particular, ϕ^t preserves the volume form $\omega^{\wedge n}$.

Proof. Using the theorem of Darboux, we may write $\omega_z = \sum_{i=1}^{n} dp_i \wedge dq_i$, in appropriate coordinates $(q_1, \dots, q_n, p_1, \dots, p_n)$. In other words,

$$\omega_z((q,p),(Q,P)) =< (q,p), (-P,Q) >=< (q,p), J(Q,P) >$$

where

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

Note that $J^t = -J = J^{-1}$. Then we have for all v

$$\omega_z(X_H, v) = dH_z(v) \Longleftrightarrow$$

$$< X_H(z), Jv >=< gradH_z, v > \Longleftrightarrow$$

$$< J^t X_H, v >=< gradH_z, v >\Longleftrightarrow$$

$$J^t X_H(z) = gradH_z \Longleftrightarrow X_H = JgradH_z$$

A preserving ω_z is equivalent to say that for all v, w

$$\omega_z(Av, Aw) = \omega_z(v, w) \Longleftrightarrow$$

$$< Av, JAw >=< v, Jw >\Longleftrightarrow$$

$$< v, A^t JAw >=< v, Jw >\Longleftrightarrow$$

$$A^t JA = J$$

Finally

$$\frac{d\phi^t(x)}{dt} = X_H(\phi^t(x)) \Rightarrow$$

$$\frac{d\Phi(t,x)}{dt} = dX_H(\phi^t(x))\Phi(t,x)$$

$$dX_H(\phi^t(x)) = d(JgradH)(\phi^t(x)) =$$

$$JhessH(\phi^t(x))$$

Hence

$$\frac{d(\Phi^t J\Phi)}{dt} = (dX_H\Phi)^t J\Phi + \Phi^t J(dX_H\Phi) =$$

$$(JhessH\Phi)^t J\Phi + \Phi^t JJhessH\Phi$$

$$\Phi^t hessH(-J)J\Phi + \Phi^t JJhessH\Phi = 0$$

and then $\Phi^t J\Phi = J$, showing that Φ preserves ω_z. □

Example 2.7. (Geodesic flow): Let N be a riemannian manifold, $M_1 = T^*N$ and $M_2 = TN$; $\omega = d\alpha$ is the canonical structure intoduced above. By means of the riemannian metric we can identify

$$M_1 = T^*N = TN = M_2$$

through the map $(q, v) \in M_2 \mapsto (q, p =< v, . >) \in M_1$, where $<, >$ is the inner product. With this identification we can consider the form α as acting on M_2:

$$\alpha_{(q,v)} : T_{(q,v)}(TN) = T_qN \times T_qN \to \mathbb{R}$$

mapping (u, w) to $< v, w >$. Then $\omega = d\alpha$ defines a symplectic structure in M_2. Consider $H : M_2 \to \mathbb{R}$, $H(q, v) = \frac{1}{2}|v|^2$. The hamiltonan flow of H is known as the geodesic flow of N, since the projection of its trajectories are the geodesics of the manifold.

Definition 2.8. Given a symplectic manifold (M, ω), a transformation

$$f : M \to M$$

is *symplectic* if preserves ω, i.e., $f_*\omega = \omega$.

There are two important classes of examples: flows transformations and Poincaré transformations associated with Hamiltonian vector fields. The first class corresponds to the proposition above. The second one is object of more considerations in what follows.

Let X_H be a Hamiltonian field with n degrees of freedom. As was seen, the trajectories lies on a level hypersurface $H = cte$. Let $z \in M$ be such that $dH_z \neq 0$ (and then $X_H \neq 0$) and let N be the submanifold $H = cte$ containing z. Locally N is a codimension 1 submanifold, i.e., $dimN = 2n-1$. Let z' be a point in the same trajectory of z and close to it, and consider the sections Σ and Σ', transversal to the flow and containing, resp., z and z'. $dim\Sigma = dim\Sigma' = 2n - 2$. Then, $\omega|\Sigma$ is a symplectic form on Σ (and $\omega|\Sigma'$ is a symplectic form on Σ'):

a) $\omega|\Sigma$ is closed: $d(\omega|\Sigma) = 0$

b) $\omega|\Sigma$ is nondegenerate. Prove it!!

Proposition 2.9. *The Poincaré map associated to the flow*

$$P : \Sigma \to \Sigma'$$

is symplectic.

Proof. Take two surfaces $S \subset \Sigma$ and $S' = P(S) \subset \Sigma'$. Let Ω be the region of the space bounded by S,S' and R. The Stoke's theorem states that

$$\int_S \omega + \int_R \omega + \int_{S'} \omega = \int_\Omega d\omega = 0$$

If $\xi \in R$ and $\{X_H(\xi), v\}$ is a basis for $T_\xi R$ then $v \in T_x N$ and hence,

$$\omega(X_H(\xi), v) = dH_\xi(v) = 0$$

i.e., $\omega|_R = 0$. Therefore

$$\int_S \omega - \int_{S'} \omega = 0$$

for a suitable choice of orientation. *** □

Given H_1, H_2 we define the *Poisson bracket*:

$$\{H_1, H_2\} = \omega(X_{H_1}, X_{H_2}) = dH_1(X_{H_2})$$

The following are important properties of the Poisson bracket

a) $\{,\}$ is bilinear and skew symmetric;

b) $\{H, KL\} = \{H, K\}L + K\{H, L\}$

c) $[X_H, X_K] = X_{\{H,K\}}$

Now we can define, in this more general context, the notions of *first-integral* , *integrals in involution*, and *integrable system*, just in the same way as in the last section.

Definition 2.10. A symplectic transformation $f : \mathbb{T}^{n-1} \times \mathbb{R}^{n-1} \to \mathbb{T}^{n-1} \times \mathbb{R}^{n-1}$ is integrable if it has the form

$$f(\theta, r) = (\theta + \omega(r), r)$$

Proposition 2.11. *If $f : (N, \omega) \to (N, \omega)$ is symplectically isotopic to identity, then there exists a Hamiltonian flow in $N \times S^1 \times \mathbb{R}$ that admits Poincaré transformations conjugated to f.*

Proposition 2.12. *If f is a sympletic transformation with fixed point z then:*

a) if λ is eigenvalue of df_z then λ^{-1} is also eigenvalue of df_z;

b) if 1 or -1 are eigenvalue of df_z then their multiplicity is even.

Proof. First, note that if λ is eigenvalue of A then λ is also eigenvalue of A^t. If λ is eigenvalue of B, B symplectic, then there exists v such that $Bv = \lambda v$.

$$B^t J B v = Jv \Rightarrow B^t J(Jv) = Jv \Rightarrow \lambda B^t(Jv) = Jv$$

$$\Rightarrow B^t(Jv) = \lambda^{-1}(Jv)$$

showing that λ^{-1} is eigenvalue of B^t and, by the remark above, also of B. ☐

2.3 Generic symplectic systems

3 Birkhoff normal form

Let $f : (M, \omega) \to (M, \omega)$ be a symplectic transformation with fixed point z. We suppose that all the eigenvalues of df_z are distinct.

Definition 3.1. ω is said to be non-resonant of order $s \geq 1$ if $k \cdot \omega) \neq 0$ for all $k \in \mathbb{Z}^n$ such that $0 \leq \|k\| \leq s$, where $\|(k_1, \dots, k_n)\| = \sum_{i=1}^n |k_i|$

Theorem 3.2. *If ω is non-resonant of order $s \geq 2$ then there exists canonical coordinates $(q_1, \dots, q_n, p_1, \dots, p_n)$ so that*

$$H(q,p) = \sum_{i=1}^n \omega_i(p^2 + q^2) + H_3(r_1, \dots, r_n) + \cdots + H_m(r_1, \dots, r_n) + O(|(q,p)|^{s+1})$$

where $r_i = \frac{1}{2}(p_i^2 + q_i^2)$, H_j is a homogeneous polynomial of degree j and $m = [s/2]$.

Example 3.3. For systems with $n = 1$ degrees of freedom, the normal form of

order $s = 2$ is $H(q, p) = 2\omega r + O(|(q, p)|^3)$;

order $s = 3$ is $H(q, p) = 2\omega r + O(|(q, p)|^4)$;

order $s = 2$ is $H(q, p) = 2\omega r + \omega_2 r^2 + O(|(q, p)|^5)$.

For systems with $n = 2$ degrees of freedom, the normal form of order $s = 4$ is $H(q, p) = (2\omega_1 r_1 + 2\omega_2 r_2) + \alpha_{11} r_1^2 + \alpha_{12} r_1 r_2 + \alpha_{22} r_2^2 + O(|(q, p)|^5)$

Proposition 3.4. *Given H and $s \geq 2$ the normal form of order s*

$$H^{(s)}(q, p) = H_1(r) + \cdots + H_m(r)$$

defines an integrable system.

Proof. Consider the *symplectic polar coordinates* (θ_i, r_i) given by

$$q_i = 2r_i \cos\theta_i \quad \text{and} \quad p_i = 2r_i \sin\theta_i$$

In these new coordinates the Hamiltonian H depends only of r. So, the r_i are first-integrals. $\square$

Definition 3.5. ω is (c, τ)-diophantine if

$$|k \cdot \omega| \geq c\|k\|^{-\tau}$$

for all $k \in \mathbb{Z}^n - 0$.

Theorem 3.6. (KAM) *Let $H_0(I)$ be an integrable on $\mathbb{T}^n \times \mathbb{R}^n$, nondegenerate; let $r < R$, $c > 0$ and $\tau > 0$ be fixed. Then there exists a neighborhood W of H_0 in $C^\infty(\mathbb{T}^n \times \mathbb{R}^n)$ such that for all $H \in W$ and ω (c, τ)-diophantine there is a n-torus $T_\omega \subset \mathbb{T}^n \times \mathbb{R}^n$ such that:*

a) T_ω is a graphic on $\mathbb{T}^n$, lagrangean and invariant.

b) The flow of $X_H|_{T_\omega}$ is conjugate to the linear flow $t \mapsto Z_0 + \omega t$ on the torus.

Let $f : (M, \omega) \to (M, \omega)$ be symplectic, $z = f(z)$ be an elliptic point with eigenvalues $\lambda_1, \ldots, \lambda_n, \lambda_1^{-1}, \ldots, \lambda_n^{-1}$, $|\lambda_i| = 1$. Then we have the following result:

Theorem 3.7. *If $\lambda = (\lambda_1, \dots, \lambda_n)$ is non-resonant of order s then there exist coordinates (θ, r) so that*

$$f(\theta, r) = (\theta + S(r), r) + O(r^{(\frac{s+1}{2})})$$

where S is a polynomial of degree $m = [s/2]$ in r.

Example 3.8. For $n = 1$ the normal form is

$$f(\theta, r) = (\theta + a_0 + a_1 r + \cdots + a_m r^m, r) + O(r^{(\frac{s+1}{2})})$$

Proof. Here we give only the main steps in the proof, that is an indutive procedure.

First: We have

$$H(q,p) = \sum_{i=1}^{n} 2\omega_i r_i + O((q,p)^3)$$

Second: Let's suppose that we have

$$H(q,p) = \sum_{i=1}^{n} 2\omega_i r_i + H_2(r) + \cdots + H_m + \mathcal{P}_s(q,p) + O((q,p)^{s+1})$$

(note that $\mathcal{P}_s(q,p) + O((q,p)^{s+1}) = O((q,p)^s$. We want to find symplectic coordinates (Q, P) such that:

$$H(Q,P) = \sum_{i=1}^{n} 2\omega_i R_i + H_2(R) + \cdots + H_m(R) + O((Q,P)^{s+1})$$

for $s = 2m + 1$, and

$$H(Q,P) = \sum_{i=1}^{n} 2\omega_i R_i + H_2(R) + \cdots + H_m(R) + H_{m+1}(R) + O((Q,P)^{s+1})$$

if $s = 2m + 2$. We consider the change of coordinates:

$$(q,p) \to (z,w) \to (Z,W) \to (Q,P)$$

given by

$$z_j = p_j + iq_j, w_j = p_j - iq_j$$

and

$$Z_j = P_j + iQ_j, W_j = P_j - iQ_j$$

In the new coordinates (z, w) the variable r_i assumes the form $r_i = \frac{1}{2}(p_i^2 + q_i^2) = \frac{1}{2}z_i w_i$ and the is

$$\begin{aligned} H(z,w) = \sum_{j=1}^{n} \omega_j(z_j w_j) + H_2(r) + \cdots + H_m(r) \\ + \sum_{(k,l)=s} \alpha(k,l) z^k w^l + O((z,w)^{s+1}) \end{aligned}$$

where z^k means $z_1^{k_1} \dots z_n^{k_n}$. Now, in order to define the map $(z, w) \mapsto (Z, W)$ we introduce the generating function

$$S(z, W) = \sum_{j=1}^{n} z_j W_j + \sum_{(k,l)=s} \beta_{(k,l)} z^k W^l$$

Let us consider

$$\begin{cases} w_j = \frac{\partial S}{\partial z_j} = W_j + \sum_{(k,l)=s} \beta(k,l) k_j z^{k(j)} w^l \\ Z_j = \frac{\partial S}{\partial w_j} = z_j + \sum_{(k,l)=s} \beta(k,l) l_j z^k w^{l(j)} \end{cases}$$

where $k(j) = k - (0, \dots, 1_j, \dots, 0)$ and $l(j) = l - (0, \dots, 1_j, \dots, 0)$. The expressions above define implicitly

$$z_j = Z_j - \sum_{(k,l)=s} \beta(k,l) l_j z^k w^{l(j)} + O((z,w)^s)$$

and, due to its definition by means of generating functions this change of coordinates is in fact symplectic. Using this variables in the expression of H we obtain

$$H(Z,W)=\sum_{j=1}^{n}\omega_j Z_j W_j+\sum_{j=1}^{n}\left(\sum_{(k,l)=s}[\beta(k,l)k_j-\beta(k,l)l_j]Z^kW^l\right)$$

$$+O((Z,W)^{s+1})+H_2(R)+\cdots+H_m(R)+\sum_{(k,l)=s}\alpha Z^kW^l$$

The terms of order s can be collected:

$$\sum_{(k,l)=s}\left[\alpha(k,l)+\beta(k,l)[\sum_{j=1}^{n}\omega_j(k_j-l_j)]\right]Z^kW^l$$

Now, if $\sum_{j=1}^{n}\omega_j(k_j-l_j)\neq 0$ we can take

$$\beta(k,l)=-\frac{\alpha(k,l)}{\sum_{j=1}^{n}\omega_j(k_j-l_j)}$$

vanishing the corresponding term, that is of order s. As there is no resonance of order j,

$$\sum_{j=1}^{n}\omega_j(k_j-l_j)=0 \iff l_1=k_1,\ldots,l_n=k_n$$

So we can eliminate all terms, except the ones of the form

$$Z_1^{k_1}\ldots Z_n^{k_n}W_1^{k_1}\ldots W_n^{k_n}.$$

But those terms exist only when s is even, since $s=k_1+\ldots+k_n+k_1+\ldots+k_n$. Moreover, $Z_jW_j=R_j$, and so these remaining terms are precisely the $R_1^{k_1}\ldots R_n^{k_n}$, with $k_1+\ldots k_n=s/2$. □

4 Twist maps

Here we consider the transformation $f:A\to A$, C^k (for $k\geq 1$), homotopic to identity, where A is one of the following manifolds: $S^1\times\mathbb{R}$, $S^1\times[0,\infty)$ or $S^1\times[0,1]$.

Given a point $z\in A$ and an angle c we can introduce the cones $C^+(c,z)$ and $C^-(c,z)$ in the space T_zA, that we identify with $\mathbb{R}^2$.

Definition 4.1.

$$C^+(c,z) = \{v \in \mathbb{R}^2 : v_x \geq (\tan c)|v_y|\}$$

$$C^-(c,z) = \{v \in \mathbb{R}^2 : v_x \leq -(\tan c)|v_y|\}$$

Definition 4.2. f is said to be a twist map if there exists $c > 0$ such that the following conditions hold:

$$df_z \frac{\partial}{\partial r} \in C^+(c,z)$$

$$df_z^{-1} \frac{\partial}{\partial r} \in C^-(c, f^{-1}(z))$$

Example 4.3. Symplectic transfomations at the neighborhood of an elliptic fixed point z. If $df_z = \lambda$, λ non-resonant of order s, then we can find symplectic coordinates (θ, r) such that f assumes the form

$$f(\theta, r) = (\theta + a_0 + a_1 r + \cdots + a_m r^m, r) + O(r^{(s-1)/2})$$

with $m = [(s-1)/2]$.

Example 4.4. Convex billiards:

$$B : S^1 \times (0,\pi) \to S^1 \times (0,\pi)$$

$$(s_0, \alpha_0) \mapsto (s_1, \alpha_1)$$

It's a well known fact that this transformation preserves the form $\omega(s,\alpha) = \sin \alpha d\alpha \wedge ds$

4.1 Birkhoff theory

Theorem 4.5. *Let $f : A \to A$ be a C^k twist map, homotopic to identity and C be an invariant curve that is a graph on S^1 (i.e., $\exists \psi : S^1 \to \mathbb{R}$, C^0, such that $C = graph(\psi)$). Then ψ is Lipschitz with constant depending only of f.*

Proof. Consider the lift $\tilde{f} : \tilde{A} \to \tilde{A}$ on the universal covering of A. Let g be a function such that

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

Given $\theta \in \mathbb{R}$ $\exists \epsilon > 0$ such that if $\theta \leq \tau \leq \theta + \epsilon$ then

$$\tilde{\psi}(\tau) \leq \tilde{\psi}(\theta) + \frac{1}{c}(\tau - \theta)$$

$$\tilde{\psi}(\tau) \geq \tilde{\psi}(\theta) - \frac{1}{c}(\tau - \theta)$$

As S^1 is compact we can assume $\epsilon > 0$ independent from θ and then, given θ_1, θ_2 in S^1 we have

$$|\psi(\theta_1) - \psi(\theta_2)| \leq \frac{1}{c}|\theta_1 - \theta_2|$$

□

Definition 4.6. The rotation number $\rho(f)$ is the real number given by

$$\rho(f) = \lim_{n\to\infty} \frac{f^n(x) - x}{n}$$

The limit is, in fact, independent of the chosen point x.

Theorem 4.7. *Let f be Lebesgue preserving. If $C = graph(\psi)$ is invariant and $\rho(C) \in \mathbb{R} - \mathbb{Q}$ then given another invariant curve C' (also graphic of some function), $C \cap C' = \emptyset$ and $\rho(C') \neq \rho(C)$*

Proof. The Theorem of Denjoy states that either g has dense orbit (and, in fact, all the orbits are dense) or g is a Denjoy counter-example.

First case: Given another invariant curve $C' \neq C$ let us suppose that $C \cap C' \neq \emptyset$. Then $C \cap C'$ is f-invariant and, due to the irrational rotation number, is dense in C. Thus, $C \cap C' = C$, a contradiction.

Second case: $C \cap C'$ is compact. If I is a conex component of $C - [C \cap C']$ then the images of I by f are disjoint (otherwise they coincide implying the existence of a fixed point, what is a contradiction since we assume that ρ is irrational). Let R be the conex region bounded by I and an arc of C'. The images of R are also disjoint and then

$$Leb(\bigcup_{j\geq 0} f^j(R)) = \sum_{j\geq 0} Leb(f^j(R)) = \sum_{j\geq 0} Leb(R) = \infty$$

a contradiction.

Now let us prove that $C \cap C' = \emptyset \Rightarrow \rho(C) \neq \rho(C')$. $C = graph(\psi)$ and $C' = graph(\psi')$, allowing us to define a positive number $\delta = \min |\psi - \psi'|$. Then there exists ϵ such that $g'(\theta) \geq g(\theta) + \epsilon$ for all θ. In the lift:

$$\tilde{g'}(\theta) \geq \tilde{g}(\theta) + \epsilon$$

implying, since ρ is irrational, that $\rho(\tilde{g'}) > \tilde{g}$. □

Theorem 4.8. (Birkhoff) *Let $U \subset S^1 \times [0, \infty)$ such that*

a) U is diffeomorphic to $S^1 \times \mathbb{R}_+$ and contain some subset $S^1 \times [0, \delta]$.

b) U is f-invariant.

c) $closure(U)$ is compact and $U = interior(closure(U))$

Then there exists $\psi : S^1 \to \mathbb{R}$ continuous such that $\partial U = graph(\psi)$

See [5, Chp I] for a proof.

Corollary 4.9. *Let C_1 and C_2 (both graphs of, resp., ψ_1 and ψ_2) be invariant curves and $\not\exists \psi_1 < \psi < \psi_2$ that has graphic f-invariant. Then $\forall V_1$ neighborhood of C_1 and $\forall V_2$ neighborhood of C_2 there exists $z_1 \in V_1$, $z_2 \in V_2$, $n_1, n_2 \geq 1$ such that $f^{n_1}(z_1) \in V_2$ and $f^{n_2}(z_2) \in V_1$.*

Proof. Let $W = \bigcup_{n\geq 0} f^n(V_1)$ and take

$$U = W \bigcup (\text{domain bounded by } C \text{ and } r = 0)$$

Then $f(W) = W \Rightarrow f(U) = U$ and then U fullfil the conditions of the Theorem above, implying that $\partial U = graph(\psi)$, $\psi > \psi_1$. But we assumed that $\not\exists \psi_1 < \psi < \psi_2 \Rightarrow \exists \theta_0 \in S^1$ such that $\psi(\theta_0) = \psi(\theta_0)$. Now the existence of the desired point z_1 follows easily. For z_2 the procedure is similar and is left to the reader. □

4.2 The Poincaré-Birkhoff theorem

Theorem 4.10. (Poincare-Birkhoff) *Let $f : S^1 \times [0,1] \mapsto S^1 \times [0,1]$ be an homeomorphism, Lebesgue preserving, that admits a lift $\tilde{f} : (\theta, r) \mapsto (\Theta, R)$ such that $\Theta(\theta, 0) > \theta$ and $\Theta(\theta, 1) < \theta$ for all $\theta \in \mathbb{R}$. Then f has at least 2 fixed points.*

Proof. For all θ there exists a unique $r(\theta)$ such that

$$f(\theta, r(\theta)) \in \text{vertical line defined by } \theta$$

and this $r(\theta)$ is differentiable. Define the curve

$$C = \{(\theta, r(\theta) : \theta \in S^1\}$$

This is a differentiable circle. As f is Lebesgue preserving, $f(C) \cap C$ hat at least 2 points. If z is one of them, its image is at the same vertical line showing that $f(z) = z$. □

Corollary: Let $f : S^1 \times [0,1] \mapsto S^1 \times [0,1]$ be a C^1 twist map and C_1, C_2 two invariant curves, both graphic of some function, with $\rho(C_1) < \rho(C_2)$. Then, for all $p/q \in (\rho(C_1), \rho(C_2))$ there exists 2 periodic orbits of period q with $\rho = p/q$.

5 The Aubry-Mather theorem

Definition 5.1. a) $\omega \in \mathbb{R}^n$ is (c, τ)-diophantine if

$$|k.\omega| \geq \frac{c}{|k|^\tau}$$

for all $k \in \mathbb{Z}^n - \{0\}$.

b) $\omega \in \mathbb{R}^{n-1}$ is (c, τ)-diophantine in the strong sense if $(\omega_1, \omega_2, \ldots, 1)$ is (c, τ)-diophantine.

Obs: If $\omega \in \mathbb{R}$ is (c, τ)-diophantine then

$$|\omega - \frac{p}{q}| \geq \frac{c}{|q|^{\tau+1}}$$

for all rationals p/q. For a fixed τ,

$$\bigcup_c D(c,\tau)$$

is a full-measure set.

Definition 5.2. $E \subset A$ is an Aubry-Mather set if:

a) E is f-invariant.

b) There exists $K \subset S^1$ (compact) and $\psi : K \to [0,1]$ such that $E = graph(\psi)$.

Example 5.3. A periodic orbit given by the proposition above for some $p/q \in [\rho_0, \rho_1]$ is an Aubry-Mather set with $\rho = p/q$.

Example 5.4. Invariant curves that are graphic.

Theorem 5.5. *Take $A = S^1 \times [0,1]$. Let $f : A \to A$ be a C^1 twist map, homotopic to identity and Lebesgue preserving. Defines $\rho_0 = \rho(f|r=0)$ and $\rho_1 = \rho(f|r=1)$. Then, for all $\rho \in [\rho_0, \rho_1]$ there exists an Aubry-Mather set $E = E_\rho$ with rotation number ρ.*

For the proof we will need some results.

Proposition 5.6. *For all $p/q \in [\rho_0, \rho_1]$ there exists an Aubry-Mather set E such that $\rho(E) = p/q$.*

Proof: Let us define

$$\xi(n+q) = \xi(q) + 1$$

$$\tilde{f}(\xi(n), r(n)) = (\xi(n+p), r(n+p))$$

Now we introduce the space $\Phi_{p,q}$ whose elements are the sequences $\xi(n) \in \mathbb{R}$ satisfying the following properties:

p1) $\xi(n)$ is non-decreasing.

p2) $\xi(n+q) = \xi(n) + 1$ for all $n \geq 1$.

p3) $g_0(\xi(n)) \leq \xi(n+p) \leq g_1(\xi(n))$

and identifying two sequences if one is only a shift of the other. $\Phi_{p,q}$ can be embedded on $\mathbb{R}^q/\mathbb{Z}$ (this equivalence identify $(x_0, \dots, x_{q-1})$ and $(x_0 + k, \dots, x_{q-1} + k)$) through

$$\xi = \{\xi(n)\}_n \mapsto (\xi(0), \dots, \xi(q-1))$$

This procedure endows $\Phi_{p,q}$ with a topology that makes it compact.

Definition 5.7.

$$L_{p,q} : \Phi_{p,q} \to \mathbb{R}$$

$$\xi \mapsto L_{p,q}(\xi)$$

is the continuous map given by

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

where A is a f-invariant measure, positive on open sets.

Lemma 5.8. *If $L_{p,q}$ reaches its minimum at ξ then $h^+(n) = h^-(n)$ for all n and so $(\xi(n), h^{\pm}(n))$ corresponds to a trajectory of $\tilde{f}$:*

$$\tilde{f}(\xi(n), h^{\pm}(n)) = (\xi(n+p), h^{\pm}(n+p))$$

Proof. Let us suppose $h^+(n) > h^-(n)$ for some n and $\xi(n-1) < \xi(n) < \xi(n+1)$. Define $\tilde{\xi}(j) = \xi(j)$ for $j \neq n$ and $\tilde{\xi}(n) = \xi(n) - \delta$. Then,

$$L_{p.q}(\tilde{\xi}) = L_{p,q} - \alpha - \beta + \gamma \Rightarrow$$

$$L_{p,q}(\tilde{\xi}) = L_{p,q} - \beta$$

and then ξ is not a minimum, a contradiction. □

Proposition 5.9. *If E is an Aubry-Mather set, graphic of some $\psi : K \to [0,1]$, then ψ is C-Lipschitz, C depending only of f.*

Proposition 5.10. *The set $AM = \{E : E$ is an Aubry-Mather set$\}$ is closed and invariant in the Hausdorff topology, and $\rho(.)$ is a continuous function of E.*

Proof. Let $z \in closure(AM) \Rightarrow \exists E_n \in AM$ such that $z_n \to z$, $z_n \in E_n$. By compacity, we can find a convergent subsequence $E_{n_k} \to E$, $E \in AM$. Then $z \in E$, showing that $z \in AM$.

□

Corollary 1: The union of all Aubry-Mather sets is closed and invariant.

Corollary 2: For all $\rho \in [\rho_0, \rho_1]$ there exists an Aubry-Mather set with rotation number ρ.

Proof. ρ can be approximated by a sequence of rationals p_n/q_n. For each one of this numbers there exists $E_n = \{\text{set of periodic orbits}\}$, that is an Aubry-Mather set with $\rho(E_n) = p_n/q_n$. We can now take a subsequence $E_{n_k} \to E$, E an Aubry-Mather set, and from the continuity of the rotation numbers follows that

$$\frac{p_{n_k}}{q_{n_k}} \to \rho = \rho(E)$$

that is exactly the contents of the Theorem. □

6 KAM theory

6.1 Local linearization theorems

Given $\omega = (\omega_1, \dots, \omega_n) \in \mathbb{R}^n$ take the constant vector field $X_\omega = \omega$ on $\mathbb{T}^n$.

Theorem 6.1. *If ω is diophantine then there exists a neighborhood V of X_ω and for each $X \in V$ $\exists \lambda \in \mathbb{R}^n$ and $h \in Diff^\infty(\mathbb{T}^n)$ such that $X = h_* X_\omega + \lambda$.*

Define $T_\omega : \mathbb{T}^n \to \mathbb{T}^n$ by $T_\omega(\theta) = \theta + \omega$.

Theorem 6.2. *If ω is diophantine in the strong sense then there exists neighborhood V of T_ω in $Diff^\infty(\mathbb{T}^n)$ such that for all $f \in V$ $\exists \lambda \in \mathbb{R}^n$ and $h \in Diff^\infty(\mathbb{T}^n)$ so that $f = T_\lambda(h^{-1} T_\omega h)$.*

Obs:
a) If $\lambda = 0$ then f is C^∞-conjugate to T_ω.
b) If $\rho(f) = \omega$ then $\lambda = 0$.
Let us now prove the Theorems above.
Proof of Theorem 6.1: Consider

$$\Phi = \Phi_\omega : \mathbb{R}^n \times Diff^\infty(\mathbb{T}^n) \to \mathcal{X}^\infty(\mathbb{T}^n)$$

$$(\lambda, h) \mapsto \lambda + h_* X_\omega$$

Clearly, $\Phi(0, id) = X_\omega$. Is Φ a bijection between some neighborhoods of $(0, id)$ and X_ω?

Given $h \in Diff^\infty(\mathbb{T}^n, 0)$ we can take one of its lifts $\tilde{h} : \mathbb{R}^n \to \mathbb{R}^n$,

$$\tilde{h}(\theta + k) = \tilde{h}(\theta) + k \quad \forall k \in \mathbb{Z}^n$$

Let $\tilde{u} = \tilde{h} - id : \mathbb{R}^n \to \mathbb{R}^n$. This function induces $u : \mathbb{T}^n \to \mathbb{R}^n$ that can be chosen in order to satisfy $u(0) = 0$. This give us an embedding from $Diff^\infty(\mathbb{T}^n, 0)$ into $C^\infty(\mathbb{T}^n, \mathbb{R}^n, 0)$, that is a Fréchet topological vector space. This means that the topology is defined by an increasing sequence of norms, being, in this case:

$$||u||_n = ||u||_{C^n}$$

The convergence is defined as follows:

$$(u_j)_j \to u \iff \forall n \geq 0 \ \ ||u_j - u||_n \longrightarrow 0$$

Definition 6.3. Given E_1, E_2 topological vector spaces and $\psi : U_1 \subset E_1 \to E_2$ (U_1 open set), ψ is C^1-Gâteaux if there is $D\psi : U_1 \times E_1 \to E_2$ continuous and linear in the second variable such that

$$D\psi(x, v) = \lim_{t \to 0} \frac{1}{t}[\psi(x + tv) - \psi(x)]$$

for all $x \in U_1$ and $v \in E_1$.

Definition 6.4.

$$L_\omega : C^\infty(\mathbb{T}^n, \mathbb{R}^n, 0) \to C^\infty(\mathbb{T}^n, \mathbb{R}^n)$$

$$u(x) \mapsto Du(x).\omega$$

Lemma 6.5. *a)* $L_\omega(C^\infty(\mathbb{T}^n, \mathbb{R}^n, 0)) = C_0^\infty(\mathbb{T}^n, \mathbb{R}^n) = \{v \in C^\infty(\mathbb{T}^n, \mathbb{R}^n) : \int_{\mathbb{T}^n} v = 0\}$

b) ω *nonresonant* $\Rightarrow L_\omega$ *is injective.*

c) ω *diophantine* $\Rightarrow L_\omega$ *is a bijection on* $C_0^\infty(\mathbb{T}^n, \mathbb{R}^n)$.

d) If ω is diophantine then there exists $r > 0$ and $A_j > 0$ such that $||L_\omega^{-1} v||_j \leq A_j ||v||_{j+r}$ with $r = n + \tau + 1$.

Proof. a)

$$\int_{\mathbb{T}^n} (L_\omega u) d\theta = \int_{\mathbb{T}^n} (Du_i).\omega d\theta = 0$$

since $Du_i\omega$ is an exact 1-form.

b) $u : \mathbb{T}^n \to real^n$ has the Fourier expansion

$$u(\theta) = \sum_{k \in \mathbb{Z}^n} \hat{u}(k) \exp(2\pi i k.\theta)$$

where

$$\hat{u}(k) = \int_{\mathbb{T}^n} \exp(-2\pi i k.\theta) u(\theta) d\theta$$

Now we have the equivalence:

$$u \to (\hat{u}(k))_{k \in \mathbb{Z}^n} \iff L_\omega u \to (2\pi i k.\omega \hat{u}(k))_{k \in \mathbb{Z}^n}$$

Thus, $L_\omega u = 0 \Rightarrow 2\pi i k.\omega \hat{u}(k) = 0$. But ω is nonresonant and then $\hat{u}(k) = 0$ for all $k \neq 0$, that implies $u = cte$. As $u(0) = 0$, $u \equiv 0$ and L_ω is injective.

c) $\exists C_j(n) > 0$ $j \in N_+$ such that $\forall \in C^\infty(\mathbb{T}^n, \mathbb{R}^n)$:

$$\frac{1}{C_j} \sup_{k \in \mathbb{Z}^n} \left[(1 + |k|)^j |\hat{f}(k)| \right] \leq |f|_j \leq C_j \sup_{k \in \mathbb{Z}^n} \left[(1 + |k|)^{j+n+1} |\hat{f}(k)| \right]$$

Let us suppose ω (c, τ)-Diophantine and $v \in C_0^\infty(\mathbb{T}^n, \mathbb{R}^n)$. Take

$$\hat{u}(k) = \frac{\hat{v}(k)}{2\pi i k.\omega} \quad k \neq 0$$

($\hat{u}(0)$ will be determined later). Then the function

$$u(\theta) = \sum_{k \in \mathbb{Z}^n} \hat{u}(k) \exp(2\pi i k.\theta)$$

is C^∞:

$$||u_j|| \leq C_j \sup_{k\in\mathbb{Z}^n} \left((1+|k|)^{n+j+1} \frac{|\hat{v}(k)|}{2\pi k|k.\omega|} \right)$$

But $|k.\omega| \geq c|k|^{-\tau}$. Thus

$$||u_j|| \leq \frac{C_j}{2\pi c} \sup_{k\in\mathbb{Z}^n} \left((1+|k|)^{n+j+1}|k|^\tau |\hat{v}(k)| \right) \leq$$

$$\frac{C_j}{2\pi c} \sup_{k\in\mathbb{Z}^n} \left((1+|k|)^{n+j+1+\tau}|k|^2 |\hat{v}(k)| \right) \leq$$

$$\frac{C_j}{2\pi c}[C_{j+n+1+\tau}|v|_{n+j+1+\tau}] < \infty$$

This calculation also shows that $|u|_j \leq A_j|v|_{j+n+1+\tau}$ and $|L_\omega^{-1}v|_j \leq A_j|v|_{j+n+1+\tau}$, proving d). □

6.2 The Nash-Moser inverse function theorem

Let E and F be Fréchet spaces, i.e., topological vector spaces with topology defined by an increasing sequence of norms $|\ |_j$ that we assume fixed. In this situation the space is said to be graded.

Example 6.6. $C^\infty(\mathbb{T}^n, \mathbb{R}^n)$ with the sequence $|\ |_j = |\ |_{C^j}$.

Definition 6.7. Let $U \subset E$ be an open set. $f : U \to F$ is tame if it's continuous and for all $p \in U$ $\exists r \geq 0$ and $C_j > 0$ such that $|f(x)|_j \leq C_j(1+|x|_{j+r+1})$ $\forall x \in V(p)$ and all $j \geq 0$.

Example 6.8. The differential operator $L : C^\infty(\mathbb{R},\mathbb{R}) \to C^\infty(\mathbb{R},\mathbb{R})$, $Lu = D^r u$ is tame.

Definition 6.9. a) $f : U \to F$ is C^1-Gâteaux if it is continuous and there exists $Df : U \times E \to F$ continuous, linear in the second coordinate, such that

$$Df(x,v) = \lim_{t\to 0} \frac{1}{t}[f(x+tv) - f(x)]$$

b) f is C^k-Gâteaux if Df is C^{k-1}-Gâteaux.

c) f is C^∞-Gâteaux if it is C^k-Gâteaux for all k.

Lemma 6.10. *Let E, F, G be Fréchet spaces, $U \subset E$, $V \subset F$ open sets, $f : U \to V$, $g : V \to G$ C^k-Gâteaux. Then $g \circ f : U \to G$ is also C^k-Gâteaux and $D(g \circ f) = Dg(f(x), Df(x, v))$. Furthermore, if f and g are C^k-tame, then $g \circ f$ is C^k-tame.*

Example 6.11. $L_\omega : C^\infty(\mathbb{T}^n, \mathbb{R}^n, 0) \to C_0^\infty(\mathbb{T}^n, \mathbb{R}^n, 0)$ with ω Diophantine. L_ω^{-1} is tame.

Example 6.12. Differential operator $D^r : C^\infty(M, \mathbb{R}^k) \to C^\infty(M, \mathbb{R}^k)$.

Definition 6.13. E is a tame space if it is a graded Fréchet space that admits a family $(S(t))_{t\in(1,\infty)}$ of smoothing operators $S(t) : E \to E$ such that for all $1 \le k \le n$ there exists $C(k, n) > 0$

a) $|S(t)f|_k \le C(k, n)t^{n-k}|f|_n$

b) $|f - S(t)f|_k \le C(k, n)t^{k-n}|f|_n$ $\forall f \in E$.

Example 6.14. Take $E = C^\infty(\mathbb{T}^n, \mathbb{R}^p)$. Fix some $\varphi : \mathbb{R}^n \to \mathbb{R}$, $\varphi \in C^\infty$ with compact support. Define $\varphi_t : \mathbb{R}^n \mapsto \mathbb{R}$ by $\varphi_t(z) = t^n\varphi(tz)$. Then

$$S(t)f(x) = \varphi_t * f = \int_{\mathbb{R}^n} \varphi_t(x - y)f(y)dy$$

The proof is a long calculation, and here we give just the first steps.
$k = 0, n = 0$:

$$|S(t)f|_0 \le |\int \varphi(\xi)f(x - \frac{1}{t}\xi)d\xi| \le |f|_0 \int \varphi(\xi)d\xi$$

$$\le C(0, 0)|f|_0$$

$$|f - S(t)f|_0 \le |f|_0 + |S(t)f|_0 \le (C(0, 0) + 1)\,|f|_0$$

$k = 0, n = 1$:

$$|D(S(t)f)|_0 = \sup_x \left|\int D\varphi_t(x - y)f(y)dy\right| \le |f|_0 \sup_x \left|\int D\varphi_t(x - y)dy\right| =$$

$$|f|_0 \sup_x \left| \int t^{n+1} D\varphi_t(t(x-y))dy \right| = t|f|_0 \sup_x \left| \int D\varphi(y)dy \right| \leq Ct|f|_0$$

$$|f - S(t)f|_0 = \sup_x \left| f(x) - \int \varphi(\xi) f(x - \frac{1}{t}\xi) d\xi \right|$$

$$= \sup_x \left| \int \varphi(\xi)[f(x) - f(x - \frac{1}{t}\xi)] d\xi \right|$$

$$\leq |Df|_0 \frac{1}{t} R \leq Ct^{-1}|Df|_0 \tag{8}$$

Theorem 6.15. (Hadamard inequality) *If exists a family of smoothing operators then for all* $1 \leq k \leq n$ $\exists C(k,n) > 0$ *such that*

$$|f|_l \leq \hat{C}(k,n)|f|_k^{\frac{n-l}{n-k}} |f|_n^{\frac{l-k}{n-k}}$$

for all $k \leq l \leq n$.

Proof. Take $t = |f|_k^{-\frac{1}{n-k}} |f|_n^{\frac{1}{n-k}}$. Then

$$|f|_l \leq |S(t)f|_l + |f - S(t)f|_l \leq$$

$$C(k,l) \left(|f|_k^{-\frac{1}{n-k}} |f|_n^{\frac{1}{n-k}} \right)^{l-k} |f|_k + C(l,n) \left(|f|_k^{-\frac{1}{n-k}} |f|_n^{\frac{1}{n-k}} \right)^{l-n} |f|_n \leq$$

$$C(k,l)|f|_k^{1-\frac{l-k}{n-k}} |f|_n^{\frac{l-k}{n-k}} + C(l,n)|f|_k^{\frac{n-l}{n-k}} |f|_n^{1-\frac{n-l}{n-k}}$$

Now define

$$\hat{C}(k,n) = \max_{k \leq l \leq n} [C(k,l) + C(l,n)]$$

□

Theorem 6.16. (Inverse function) *Let* E, F *be tame Fréchet spaces,* U *an open set of* E*,* $f : U \to F$ C^k*-tame* ($k \geq 2$)*. Let* $x_0 \in U$*,* $y_0 = f(x_0)$*. If* $Df(x)$ *is a bijection for all* $x \in W$ *(*W *neighborhood of* x_0*) and*

$$W \times F \to E$$

$$(x, w) \mapsto Df^{-1}(x)w$$

is tame then there exist neighborhoods V_1 and V_2, containing respectively x_0 and y_0, and $g : V_2 \to V_1$ C^k-tame such that $f \circ g = g \circ f = id$. Furthermore, $Dg(y) = Df^{-1}(g(y))$.

this result follows the next theorem.

Theorem 6.17. (Implicit function) *Let E, F, G be tamed spaces , U an open set of $E \times F$, $f : U \to G$ C^k-tame, $(x_0, y_0) \in U$, $z_0 = f(x_0, y_0)$. If there is W neighborhood of (x_0, y_0) such that $D_2 f(x, y)$ is a bijection for all $(x, y) \in W$ and $D_2 f(x, y)^{-1}$ is continuous and linear in y then we can find a neighborhood V of x_0 and $g : V \to F$ C^k-tame such that $(x, g(x)) \in U$ $\forall x \in V$ and $f(x, g(x)) = z_0$ for all $x \in V$.*

Proof. The proof is an application of the preceding theorem to

$$h : U \to E \times G$$

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

□

The proof of the inverse function theorem consists of a modification of the Newton method to the context of Fréchet spaces. In this case the alghoritm takes the following form:

$$g_0(y) = x_0$$

$$g_{n+1}(y) = g_n(y) - S(t_n) Df^{-1}(g_n) \left(f(g_n(y)) - y \right)$$

where t_n tends to infinity in a fast way, for exmple, $t_n = \exp\left(\frac{3}{2}n\right)$.

6.3 Proof of the KAM theorem

Theorem 6.18. (KAM) *Let $H_0 \in C^\infty(\mathbb{T}^n \times D^n_r, \mathbb{R})$ be nondegenerate (i.e., $\frac{\partial \omega_0}{\partial I}$ is an isomorphism at all points $I \in D^n_r$),*

$$H_0 : \mathbb{T}^n \times D^n_r \to \mathbb{R}$$

$$(\theta, I) \mapsto H_0(I)$$

Let $\alpha = \omega_0(I_0)$ with $I_0 \in int(D^n_r)$ and α is Diophantine. Then there exists a neighborhood W of H_0 in $C^\infty(\mathbb{T}^n \times D^n_r, \mathbb{R})$ such that for all $H \in W$ there is an embedded n-torus T_α satisfying:

a) T_α is a graphic on $\mathbb{T}^n$, lagrangean, invariant under the flow of X_H.

b) The flow of X_H resticted to T_α is C^∞ conjugate to the flow of X_α. Furthermore, $T_{\alpha,H}$ depends continuously on H.

Lemma 6.19. *Let $u \in C^\infty(\mathbb{T}^n, \mathbb{R}^n)$ and $T = graph(u)$. Then the following propositions are equivalent:*

a) T is lagrangean.

b)

$$\left(\frac{\partial u_i}{\partial \theta_j}\right)_{i,j}$$

is symmetric.

c) The 1-form $\sum_{i=1}^n u_i d\theta_i$ is closed.

d) $(\theta, I) \mapsto (\theta, I + u(\theta))$ is symplectic.

Proof. a) $\iff$ b): T is lagrangean $\iff$

$$\left(\sum_{i=1}^n dI_i \wedge d\theta_i\right)\left(\left(\frac{\partial}{\partial \theta_j}, \frac{\partial u}{\partial \theta_j}\right), \left(\frac{\partial}{\partial \theta_k}, \frac{\partial u}{\partial \theta_k}\right)\right) = 0 \iff$$

$$\frac{\partial u_k}{\partial \theta_j} - \frac{\partial u_j}{\partial \theta_k} = 0 \iff$$

$$\left(\frac{\partial u}{\partial \theta}\right) \text{is symmetric}$$

b) $\iff$ c): $\sum_{i=1}^n u_i d\theta_i$ is closed $\iff$

$$\sum_{i=1}^{n}\left(\frac{\partial u_i}{\partial \theta_j}\right)\wedge d\theta_i = 0 \iff \sum_{i,j}\frac{\partial u_i}{\partial \theta_j}d\theta_j \wedge d\theta_i = 0 \iff$$

$$\left(\frac{\partial u}{\partial \theta}\right) \text{ is symmetric}$$

c) $\iff$ d): $(\theta, I) \mapsto (\theta, I+u(\theta))$ is symplectic $\iff$

$$\sum_{i=1}^{n} d(I+u(\theta))_i \wedge d\theta_i = \sum_{i=1}^{n} dI_i \wedge d\theta_i \iff$$

$$\sum_{i=1}^{n} du_i \wedge d\theta_i = 0 \iff$$

$$\left(\frac{\partial u}{\partial \theta}\right) \text{ is symmetric}$$

□

Lemma 6.20. *Let $T = graph(u)$ be a lagrangean torus. Then T is invariant $\iff$ $H|_T = cte$.*

Theorem 6.21. *Given H_0, I_0, α as in the KAM theorem there exists a neihgborhood W of H_0 in $C^\infty(\mathbb{T}^n \times D_r^n, \mathbb{R})$ and C^∞-tame applications*

$$W \to \mathbb{R}^n \times C^\infty(\mathbb{T}^n, \mathbb{R}, 0) \times Diff^\infty(\mathbb{T}^n, 0)$$

$$H \mapsto (t_H, f_H, g_H)$$

such that $t_{H_0} = I_0$, $F_{H_0} = 0$, $g_{H_0} = id$ and if we write $u_H = t_H + \nabla f_H$ then:
a) $graph(u_H) \subset \mathbb{T}^n \times int(D_r^n)$ and $graph(u_H)$ is a lagrangean torus.
b) $H_0(id, u_H) = cte$. In other words, this torus is invariant.
c)

$$\frac{\partial H}{\partial I}(id, u_H) = (g_H)_* X_\alpha$$

The flow of X_H restricted to $graph(u_H)$ is conjugate to the flow of X_α.

Lemma 6.22. *Let*

$$\mathcal{U} = \{(u,t) \in C^\infty(\mathbb{T}^n, \mathbb{R}^n, 0) \times \mathbb{R}^n : graph(t+u) \subset int(\mathbb{T}^n \times D_r^n)\}$$

and

$$\Phi : C^\infty(\mathbb{T}^n \times \mathbb{R}, \mathbb{R}) \times \mathcal{U} \times Diff^\infty(\mathbb{T}^n, 0) \to \mathcal{X}(\mathbb{T}^n)$$

$$(H,u,t,g) \mapsto \frac{\partial H}{\partial I}(id, u+t) - g_* \alpha$$

Then:

a) Φ *is a* C^∞*-tame application with* $\Phi(H_0, I_0, 0, id) = 0$

b) $D_{3,4}\Phi(H,u,t,g) : R^n \times C^\infty(\mathbb{T}^n, \mathbb{R}^n) \to \mathcal{X}^\infty(\mathbb{T}^n)$ *is invertible for all* (H,u,t,g) *close to* $(H_0, 0, I_0, id)$*; furthermore,*

$$(H,u,t,g,\Delta t, \Delta g) \mapsto (D\Phi_{3,4}(H,u,t,g))^{-1}(\Delta t, \Delta g)$$

is tame.

Corollary 6.23. *There exists a neighborhood* V *of* (H_0, I_0) *and a* C^∞*-tame applications*

$$T : V \to \mathbb{R}^n$$

$$(H,u) \mapsto T(H,u)$$

and

$$G : V \to Diff^\infty(\mathbb{T}^n, 0)$$

$$(H,u) \mapsto G(H,u)$$

such that:

i)

$$\frac{\partial H}{\partial I}(id, u + T(H,u)) = G(H,u)_* \alpha$$

ii) $(u, T(H,u)) \in \mathcal{U}$*, i.e.,* $graph(u+T(H,u)) \subset int(\mathbb{T}^n \times D_r^n) \forall (H,u) \in V$

iii) $T(H_0, I_0) = 0$ *and* $G(H_0, I_0) = id$*.*

Proof of the lemma (ideas):

$$(D_{3,4}\Phi)(H,u,t,g)(\Delta t,\Delta g) = \frac{\partial^2 H}{\partial I^2}(id,u+t)\Delta t - (Dgg^{-1})(D\dot{g}g^{-1})\alpha$$

Hence,

$$(D_{3,4}\Phi)(H,u,t,g)(\Delta t,\Delta g) = \Delta\eta$$

if snf only if

$$\frac{\partial^2 H}{\partial I^2}(id,u+t)\circ g\Delta t - DgD\dot{g}\alpha = \Delta\eta\circ g$$

if and only if

$$(Dg)^{-1}\frac{\partial^2 H}{\partial I^2}(id,u+t)\circ g\Delta t - D\dot{g}\alpha = (Dg)^{-1}\Delta\eta\circ g.$$

But $D\dot{g}\alpha = L_\alpha\dot{g}$ and L_α is an isomorphism when α is Diophantine. So we can choose Δt such that

$$\left(\int_{\mathbb{T}^n}(Dg)^{-1}\frac{\partial^2 H}{\partial I^2}(id,u+t)\circ g\right)\Delta t = \int_{\mathbb{T}^n}(Dg)^{-1}(\Delta\eta\circ g)$$

Then we take

$$\dot{g} = L_\alpha^{-1}\left[(Dg)^{-1}\frac{\partial^2 H}{\partial I^2}(id,u+t)\circ g\Delta t - (Dg)^{-1}(\Delta\eta\circ g)\right]$$

and

$$\Delta g = Dg.\dot{g}$$

Now let us state some results about tame functions that will be used soon.

Proposition 6.24. *Let E,F,G be Fréchet spaces, U an open set of E, V open set of F, $f:U\to V$ and $g:V\to G$ C^k-tame. Then $g\circ f:U\to G$ is C^k-tame and*

$$D(g\circ f)(x,v) = Dg(f(x),Df(x,v))$$

Proposition 6.25.

$$\phi : Diff^\infty(\mathbb{T}^n, 0) \to Diff^\infty(\mathbb{T}^n, 0)$$

$$h \mapsto h^{-1}$$

is C^k-tame and

$$D\phi(h)\Delta h = -D(h^{-1})(\Delta h \circ h^{-1}) = -((Dh)^{-1}\Delta h) \circ h^{-1}$$

Proposition 6.26. *Let $\pi_1 : V_1 \to M$, $\pi_2 : V_2 \to M$ be fiber bundles of finite dimension, C^∞ on the compact manifold M, U an open set of V_1 and $\varphi : U \to V_2$ a C^∞ application that is fiber preserving. Then*

$$\Phi : \Gamma^\infty(M, U) \to \Gamma^\infty(M, V_2)$$

$$s \mapsto \varphi \circ s$$

is C^∞-tame. ($\Gamma^\infty(M, U) = \{s : M \to U,\ s \in C^\infty : \pi_1 \circ s = id\}$)

Proposition 6.27. *Define J^rV as the set of jets of order r ($r \in \mathbb{N}$) in V. Then*

$$j^r : \Gamma^\infty(M, V) \to J^rV_1$$

$$s \mapsto j^r(s)$$

is C^∞-tame.

Lemma 6.28. (inversion) *Let E, F, G be Fréchet spaces, U an open set of E,*

$$A : U \times F \to G$$

linear in the second variable and

$$u : U \times F \to G$$

such that

a) $A(x) : F \to G$ *is invertible for all* $x \in U$ *and*

$$U \times G \to F$$

$$(x, z) \mapsto A^{-1}(x)z$$

is C^1*-tame.*

b) There exists $x_0 \in U$ *such that* $u(x_0) = 0$ *and* u *satisfies the following property:*

($\mathcal{P}$*)* $\exists l, s \geq 1$ *and* $A_j > 0$ *such that*

$$|u(x)y|_j \leq A_j(1 + |x|_{j+s})|y|_l$$

$$|u(x_1)y - u(x_2)y|_j \leq A_j\left(|x_1 - x_2|_{j+s} + |x_1|_{j+s}|x_1 - x_2|_s\right)|y|_l$$

Then $L(x) = A(x) + u(x)$ *is invertible for all* x *in some neighborhood* W *of* x_0 *and*

$$W \times G \to F$$

$$(x, z) \mapsto L^{-1}(x)z$$

is tame.

Lemma 6.29. *Let* E, F, G *be Fréchet spaces,* U *an open set of* E *and*

$$L : U \times F \to G$$

$$(x, y) \mapsto L(x)y$$

tame and linear in the second variable. Then, given $x_0 \in U$ *there exists* $V(x_0)$*,* $r \geq 1$*,* $k \geq 1$ *and* $A_j > 0$ *such that*

$$|L(x)y|_j \leq A_j\left(|x|_{j+r}|y|_k + |y|_{j+r}\right)$$

Proof. Take y such that $|y|_k = 1 \Rightarrow |\lambda y|_k = |\lambda|$. Then

$$|L(x)(\lambda y)|_j = |\lambda||L(x)y|_j \le |\lambda| B_j \left(1 + |x|_{j+r} + |y|_{j+r}\right) =$$

$$B_j \left(|\lambda y|_k + |\lambda y|_k |x|_{j+r} + |\lambda y|_{j+r}\right) \le B_j \left(|\lambda y|_k |x|_{j+r} + 2|\lambda y|\right) \le$$

$$2B_j \left(|\lambda y|_k |x|_{j+r} + |\lambda y|_{j+r}\right)$$

for $r \ge k - j$. □

Lemma 6.30. *Let*

$$\mathcal{V} = \{(H, f) \in C^\infty(\mathbb{T}^n \times D_r^n, \mathbb{R}) \times C^\infty(\mathbb{T}^n, \mathbb{R}, 0) : (H, u_f = df + I_0) \in V\}$$

where V is the neighborhood given by lemma 6.22 and

$$\Psi : \mathcal{V} \times R \to C^\infty(\mathbb{T}^n, \mathbb{R})$$

$$(H, f, E) \mapsto H_0(id, u_f + T(H, u_f)) - E$$

Then

a) Ψ is C^∞-tame and $\Psi(H_0, 0, H_0(I_0)) = 0$

b) the derivative $D_{2,3}\Psi(H, f, E) : C^\infty(\mathbb{T}^n, \mathbb{R}, 0) \to C^\infty(\mathbb{T}^n, \mathbb{R})$ is invertible for all (H, f, E) close to $(H_0, 0, H_0(I_0))$ and

$$((H, f, E), (\Delta f, \Delta E)) \mapsto (D_{2,3}\Psi(H, f, E))^{-1} (\Delta f, \Delta E)$$

is tame.

Corollary 6.31. *There exists a neighborhood W of H_0 in $C^\infty(\mathbb{T}^n \times D_r^n, \mathbb{R})$ and a C^∞-tame application*

$$W \to C^\infty(\mathbb{T}^n, \mathbb{R}, 0) \times \mathbb{R}$$

$$H \mapsto (F(H), E(H))$$

such that

i) $F(H_0) = 0,\ E(H_0) = H_0(I_0)$

ii) $(H, d(F(H)) + I_0) \in \mathcal{V}$

iii) $H_0(id, d(F(H)) + I_0 + T(H, df + I_0)) - E(H) = 0$

Proof of the Theorem: Take $f_H = F(H)$, $t_H = I_0 + T(H, df + I_0)$, $u_H = df_H + t_H$ and $g_H = G(H, u_H)$.

Proof of Lemma (?):

$$D_{2,3}\Psi(H,f,E)(\Delta f, \Delta E) = \frac{\partial H}{\partial I} \circ (id, df + I_0 + T(H, df + I_0))d(\Delta f) - \Delta E$$

$$+\frac{\partial H}{\partial I} \circ (id, df + I_0 + T(H, df + I_0))D_2T(H, df + I_0)d(\Delta f)$$

Let us call

$$L(H,f,E)(\Delta f, \Delta E) = \frac{\partial H}{\partial I} \circ (id, df + I_0 + T(H, df + I_0))d(\Delta f) - \Delta E$$

and $l(H,f,E)(\Delta f, \Delta E)$ is equal to

$$\frac{\partial H}{\partial I} \circ (id, df + I_0 + T(H, df + I_0))D_2T(H, df + I_0)d(\Delta f)$$

Claim 1: $L(H,f,E)$ is invertible for all (H,f,E) close to $(H_0, 0, H_0(I_0))$ and

$$((H,f,E),(\Delta f, \Delta E)) \mapsto L(H,f,E)^{-1}(\Delta f, \Delta E)$$

is C^1-tame.

Claim 2:

$$l(H_0, 0, H_0(I_0)) = 0$$

and l satisfies $\mathcal{P}$.

Lemma 6.32. *Claim 1 + Claim 2 $\Rightarrow$ $(L+l)(x)$ is invertible for all x close to x_0 and*

$$(x,z) \mapsto [(L+l)(x)]^{-1}z$$

is tame.

Proof.

$$L : U \times F \to G \quad \text{and} \quad l : U \times F \to G$$

First case: Let us suppose $L(x) = id$ and $F = G$.

$$U \times F \to F$$

$$(x, y) \mapsto (id + l(x))y$$

Let us consider $l(x) : (F, |.|_l) \to (F, |.|_l)$. $l(x)$ is a bounded linear operator on $(F, |.|_l)$, since

$$|l(x)y|_l \leq A_l(1 + |x|_{l+s})|y|_l$$

Furthermore,

$$|l(x)y - l(x_0)|_l \leq A_l(|x - x_0|_{l+s} + |x|_{l+s}|x - x_0|_s)|y|_l$$

Hence, $l(x)$ is close to $l(x_0) = 0$ if x is close to x_0. Choosing a neighborhood $V(x_0)$ such that $|l(x)|_l < 1$ for all $x \in V(x_0)$ then $\forall x \in V$

$$(id + l(x)) : (F, |.|_l) \to (F, |.|_l)$$

is invertible and

$$|(id + l(x))^{-1}|_l \leq (1 - |l(x)|_l)^{-1}$$

From $(id + l(x))^{-1} = id - l(x)(id + l(x))^{-1}$ we conclude that

$$\begin{aligned}
|(id{+}l(x))^{-1}y|_j \leq& \\
&\leq |y|_j + |l(x)(id + l(x))^{-1}y|_j \\
&\leq |y|_j + A_j(1 + |x|_{j+s})|(id + l(x))^{-1}y|_l \\
&\leq |y|_j + A_j(1 + |x|_{j+s})(1 - |l(x)|_l)^{-1}|y|_l \\
&\leq |y|_j + A_j(1 + |x|_{j+s})[1 - (A_l(|x - x_0|_{l+s} + |x|_{l+s}|x - x_0|_s)]^{-1}|y|_l \\
&\leq B_j(1 + |x|_{j+s+l} + |y|_{j+s+l})
\end{aligned}$$

General case: write $(L + l)(x) = L(x)(id + L^{-1}(x)l(x)) = L(x)(id + \hat{l}(x))$. Then we are the first case of the lemma, since $\hat{l}(x)$ satisfies $\mathcal{P}$. □

Proof of the Claim 1:

$$L(H,f,E)(\Delta f,\Delta E) = (g_*\alpha)\circ d(\Delta f) - \Delta E = d(\Delta f)(dg\circ g^{-1})\alpha - \Delta E =$$

$$[d(\Delta f\circ g)\alpha - \Delta E]\circ g^{-1}$$

Then

$$L(H,f,E)(\Delta f,\Delta E) = \Delta\eta \iff (d(\Delta f\circ g)\alpha - \Delta E)\circ g^{-1} = \Delta\eta \iff$$

$$L_\alpha(\Delta f\circ g) = \Delta E + (\Delta\eta\circ g)$$

This equation is true if we choose

$$\Delta E = -\int_{\mathbb{T}^n}(\Delta\eta\circ g)$$

$$\Delta f = L_\alpha^{-1}(\Delta E + (\Delta g))\circ g^{-1}$$

Proof of the Claim 2:

$$l(H_0,0,E)(\Delta f,\Delta E) = \frac{\partial H_0}{\partial I}(id, I_0 + T(H_0,I_0)).D_2T(H_0,I_0)d(\Delta f) =$$

$$\frac{\partial H_0}{\partial I}(I_0)D_2T(H_0,I_0)d(\Delta f) = 0$$

is true if $D_2T(H_0,I_0)d(\Delta f) = 0$. But

$$\frac{\partial H}{\partial I}\circ(id, u + T(H,u)) = G(H,u)_*\alpha \Rightarrow$$

$$\frac{\partial^2 H}{\partial I^2}\circ(id, u+T(H,u))(\Delta u + D_2T(H,u)\Delta u) = (D\dot{g}\circ g^{-1})\alpha$$

For $H = H_0$ and $u = I_0$ we have

$$\frac{\partial^2 H}{\partial I^2}(I_0)(\Delta u + D_2T(H_0, I_0)\Delta u) = D(D_2G(H_0, I_0)\Delta u)\alpha$$

Integrating over $\mathbb{T}^n$:

$$\frac{\partial^2 H}{\partial I^2}(I_0)\int_{\mathbb{T}^n}(\Delta u + D_2T(H_0, I_0)\Delta u) = \int_{\mathbb{T}^n} D(D_2G(H_0, I_0)\Delta u)\alpha = 0 \Rightarrow$$

$$\int_{\mathbb{T}^n} \Delta u + D_2T\Delta u = 0 \Rightarrow D_2T\Delta u = -\int_{\mathbb{T}^n} \Delta u$$

So,

$$D_2T(d(\Delta f)) = -\int_{\mathbb{T}^n} d(\Delta f) = 0$$

proving the claim.

7 The closing lemma and the ergodic hypothesis

In the first part of this section we state and prove a theorem of Herman, implying that the C^∞-closing lemma is false for flows and symplectic maps in dimension greater than 4.

In the second part, we discuss the relation between the ergodic hypothesis of Boltzmann and the KAM theorem. We also state another result of Herman, showing that even the much weaker version of the ergodic hypothesis known as the quasi-ergodic hypothesis is false.

Theorem 7.1. *Given n even there exists a constant symplectic form ω on $M = \mathbb{T}^{n+2}$ and given $k > 2n$ there exists an open set $\mathcal{U} \subset C^{k+1}(M, \mathbb{R})$ and an open interval $A \subset \mathbb{R}$ such that for all $H \in \mathcal{U}$ and $c \in A$, $H^{-1}(c)$ has no periodic orbit.*

Proof. We give the proof for $n = 2$, but the general case is analogous.

Fix $\alpha = (1, \alpha_2, \alpha_3)$ Diophantine (in particular, $\alpha_3 \neq 0$). Consider

$$B = \begin{bmatrix} 0 & 1 & 0 & -1 \\ -1 & 0 & 0 & \alpha_2 \\ 0 & 0 & 0 & \alpha_3 \\ -1 & -\alpha_2 & -\alpha_3 & 0 \end{bmatrix}$$

that is antisymmetric, i.e., $B^t = -B$, and $\det B = \alpha_3^2 \neq 0$. Take the form

$$\omega(v_1, v_2) =< v_1, Bv_2 >$$

defined on $M = \mathbb{T}^4$. Consider

$$H_0(\theta_1, \theta_2, \theta_3, \theta_4) = \sin(2\pi\theta_4)$$

$\mathcal{U}$ a neighborhood of H_0 and $A = (-1, +1)$. Then

$$H_0^{-1}(c) = \{(\theta_1, \dots, \theta_4) : \sin(2\pi\theta_4) = c\} = T_1(H_0) \cup T_2(H_0)$$

Then for $H \in \mathcal{U}$ and $c \in A$, $H^{-1}(c)$ is the union of two 3-torus close to T_1, T_2. Hence, they are graphic on $(\theta_1, \dots, \theta_3) \in \mathbb{T}^3$. Let $Z = P_*(X_H | T_j(H))$, $j = 1$ or $j = 2$, where P denotes the projection on $(\theta_1, \dots, \theta_3)$. If H is close to H_0 then Z is close to $Z_0 = (2\pi \cos(2\pi\theta_4)).\alpha$. From the Local Linearization Theorem follows that there exists $\lambda \in \mathbb{R}^3$ and $h \in Diff^\infty(\mathbb{T}^3, 0)$ such that $Z = \lambda + h_*(Z_0)$.

Claim: $\lambda = 0$

Proof: The rotation number of a field on $\mathbb{T}^k$ is defined as

$$\rho(Y) = \int_{\mathbb{T}^k} Y$$

with respect to the Lebesgue measure. To show the Claim is sufficient to prove that $\rho(Z) = \rho(h_*(Z_0))$.

$$X_H = B\nabla H = \begin{bmatrix} 0 & 1 & 0 & -1 \\ -1 & 0 & 0 & \alpha_2 \\ 0 & 0 & 0 & \alpha_3 \\ -1 & -\alpha_2 & -\alpha_3 & 0 \end{bmatrix} = \begin{bmatrix} \frac{\partial H}{\partial \theta_1} \\ \frac{\partial H}{\partial \theta_2} \\ \frac{\partial H}{\partial \theta_3} \\ \frac{\partial H}{\partial \theta_4} \end{bmatrix} = \begin{bmatrix} -\frac{\partial H}{\partial \theta_2} + \frac{\partial H}{\partial \theta_4} \\ \frac{\partial H}{\partial \theta_1} + \alpha_2 \frac{\partial H}{\partial \theta_4} \\ \alpha_3 \frac{\partial H}{\partial \theta_4} \\ \dots \end{bmatrix}$$

By the other hand, $T_j(H) = graph(\psi)$ and $H(\theta_1, \theta_2, \theta_3, \psi(\theta_1, \theta_2, \theta_3)) = cte$

$$\Rightarrow \frac{\partial H}{\partial \theta_1} + \frac{\partial H}{\partial \theta_4}\frac{\partial H}{\partial \theta_1} = 0$$

$$\Rightarrow Z = \frac{\partial H}{\partial \theta_4}(\frac{\partial \psi}{\partial \theta_2} + 1, -\frac{\partial \psi}{\partial \theta_1} + \alpha_2, \alpha_3) = \frac{\partial H}{\partial \theta_4}\hat{Z}$$

If H is close to H_0 then $\hat{Z}$ is close to $\hat{Z}_0 = (1, \alpha_2, \alpha_3) = X_\alpha$. Hence, there exists $\lambda \in \mathbb{R}^3$ and $H \in Diff^\infty(\mathbb{T}^3)$ such that $\hat{Z} = \lambda + h_*(\hat{Z}_0)$ and $\lambda = 0$.

$$\rho(\hat{Z}_0) = \int_{\mathbb{T}^3} \hat{Z} = \int_{\mathbb{T}^3} (\frac{\partial \psi}{\partial \theta_2} + 1, -\frac{\partial \psi}{\partial \theta_1} + \alpha_2, \alpha_3) = (1, \alpha_2, \alpha_3)$$

Now we only need to prove that $\rho(h_*(\hat{Z}_0)) = \alpha$.

Claim: $\hat{Z}$ preserves the Lebesgue measure and then $h_*\hat{Z}_0$ also preserves it.

Proof:

$$div\hat{Z} = \frac{\partial^2 \psi}{\partial \theta_1 \partial \theta_2} - \frac{\partial^2 \psi}{\partial \theta_2 \partial \theta_1} = 0 \Rightarrow$$

$$div(h_*\hat{Z}_0) = 0 \Rightarrow$$

both preserve the Lebesgue measure on $\mathbb{T}^3$.

But $\hat{Z}_0 = \alpha$ and $h_*\hat{Z}_0$ are uniquely ergodic, showing that h is Lebesgue preserving. So,

$$\rho(h_*\hat{Z}_0) = \int_{\mathbb{T}^3} h_*\hat{Z}_0 = \int_{\mathbb{T}^3} \hat{Z}_0 = \int_{\mathbb{T}^3} \alpha = \alpha$$

finishing the construction. □

References

[1] V. I. Arnold. *Mathematical methods of classical mechanics.* Springer Verlag, 1978.

[2] G. D. Birkhoff. On the periodic motions of dynamical systems. *Acta. Math.*, 50:359–379, 1927.

[3] J.-B. Bost. Tores invariants des systèmes hamiltoniens. *Astérisque*, 133–134:113–157, 1986.

[4] R. S. Hamilton. The inverse function theorem of Nash and Moser. *Bull. Am. Math. Soc.*, 7:65–222, 1982.

[5] M. Herman. Sur les courbes invariantes par les diffeomorphismes de l'anneau. *Astérisque*, 103–104, 1983.

[6] A. Katok. Some remarks on Birkhoff and Mather twist map theorems. *Ergod. Th. & Dynam. Sys.*, 2:185–194, 1982.

[7] A. N. Kolmogorov. General theory of the dynamical systems and classical mechanics. *Procs. Intern. Cong. Math. Amsterdam*, pages 315–333, 1954. (Russian).

[8] R.S. MacKay and J.D. Meiss, editors. *Hamiltonian dynamical systems.* IOP Publ., 1987.

[9] J. Moser and E. Zehnder. Lecture notes on dynamical systems. Courant Institut, NYU, 1980.

[10] J. Pöschel. Lecture on the classical KAM theorem. ICTP, Trieste, 1992.

[11] J.-C. Yoccoz. Travaux de Herman sur les tores invariants. *Astérisque*, 206:311–344, 1992.