Università degli Studi di Roma

" La Sapienza"

Dottorato di Ricerca in Matematica **XIII ciclo**

Estimates on Hamiltonian splittings

Tree techniques in the theory of homoclinic splitting and Arnold diffusion for a-priori stable systems

Michela Procesi

May 7, 2003

Abstract

We consider the problem of the splitting of invariant hyperbolic manifolds for close to integrable, Hamiltonian systems and consequently " Arnold diffusion".
Following Chierchia- Gallavotti:*Drift and diffusion in phase space* and Gallavotti: *Twistless KAM tori, quasi flat homoclinic intersections...* we work on a Hamiltonian which is a model for small analytic perturbations of stable, integrable Hamiltonian system near a simple resonance. We will call the small perturbation parameter ε.
Roughly speaking the model Hamiltonian represents a set of $n \geq 2$ rotators and clocks, weakly (ε^P with $P > 2$) coupled to a generalized pendulum with Lyapunov exponent $\sqrt{\varepsilon}$.
Namely if $I, \psi \in \mathbb{R}^n \times \mathbb{T}^n$, $p, q \in \mathbb{R} \times \mathbb{T}$ are pairs of conjugate action-angle variables, a set of rotators and clocks is given by a quadratic Hamiltonian of the type: $IA(\varepsilon)I + b(\varepsilon)I$ where $A(\varepsilon)$ is semi-positive definite and $\lim_{\varepsilon\to 0} A(\varepsilon) = A$.
Finally a generalized pendulum is a two dimensional Hamiltonian system $H(p,q) = \frac{1}{2}p^2 + \varepsilon F(q)$, with $F(q)$ analytic on $\mathbb{T}$, having $p = q = 0$ as the only unstable fixed point on the energy level $E = 0$.

The initial data and the matrices $A(\varepsilon), b(\varepsilon)$ are suitably chosen so that there are are at least three relevant time scales for the uncoupled system: namely there will be $m \neq 0$ order one (fast) frequencies, $n-m$ slow frequencies of order $\varepsilon^{\frac{1}{2}+a}$ (with $a \leq \frac{1}{2}$) and finally the Lyapunov exponent of the pendulum $\sqrt{\varepsilon}$.

KAM-like results show that the presence of the small (ε^P with $P > 2$) coupling term preserves a set of n dimensional unstable tori together with their $n+1$ dimensional local stable and unstable manifolds. In general such manifolds intersect in a curve; proving such intersection and evaluating the transversality of the manifolds is the so called problem of homoclinic splitting which is the basis for proving Arnold instability.

The thesis is mostly dedicated to the study of upper and lower bounds for the determinant of the splitting matrix, which is a measure of the "angles " of the homoclinic splitting.

We use perturbative theory and in particular, following Gallavotti:*Twistless KAM tori, quasi flat homoclinic intersections...*, and Gallavotti, Gentile, Mastropietro: *Separatrix splitting for systems with three time scales* we construct a suitable tree representation to evidence the cancellations in the perturbative expansion of the splitting determinant.

The main results are:
1) We prove that the splitting determinant is exponentially small in ε, for systems interacting through an analytic function depending only on the angle variables.
We present two alternative methods of proving the assertion, one is direct, using the cancellations; while the second (following the strategy of Berti, Bolle: *A functional analysis approach to Arnold diffusion*) constructs perturbatively a suitable set of coordinates, where the generating function of the splitting has a simpler form which implies that the the splitting determinant (which is the Hessian of the generating function at the intersection point) is smaller that any power of ε.

2) We give lower bounds for systems with one fast variable ($m = 1$) and satisfying a set of conditions which are sufficient to prove that the first order of perturbation in the splitting determinant (the Melnikov term) dominates, thus providing a lower bound.
3) We find lower and upper bounds on the splitting determinant for "D'Alembert like" Hamiltonians similar to those proposed in Gallavotti, Gentile, Mastropietro: *Hamilton-Jacobi equation and existence of heteroclinic chains in three time scales systems.* Such Hamiltonians carry a "large" (i.e. order ε) unimodal perturbation.

For completeness the last chapter is dedicated to showing the construction of the transition chains for systems where the Melnikov term dominates.

Contents

1 Preliminaries **1**

1.1 Whiskered KAM tori for anisochronous Hamiltonian systems 1

1.1.1 The KAM construction, definitions of splitting vector and splitting matrix. 3

1.1.2 Perturbative construction of the homoclinic trajectories 6

1.1.3 Whisker calculus 8

1.1.4 Analytic expansions for the whiskers 12

1.2 Trees 16

1.2.1 Trees, symmetry groups and admissible trees 16

1.2.2 Functions on admissible trees 23

2 Tree expansion for the homoclinic trajectory **31**

2.1 Holomorphic properties of tree representations 31

2.1.1 Linear operators on trees, ... 31

2.1.2 Holomorphic and non holomorphic contributions (trees with and without fruits) 35

2.2 Equivalent trees and cancellations 37

3 Basic estimates on tree expansions **43**

3.1 Upper bounds on the values of trees 44

3.2 Identities for asymptotic power series. 56

4 Upper bounds on homoclinic splittings I **59**

4.1 Cancellations and splitting determinants 59

4.1.1 The generating function of the splitting 60

4.1.2 Cancellations due to energy conservation 62

4.1.3 Relation between trees with and without fruits 63

4.1.4 Formal power series relations involving the splitting matrix . . . 64

4.2 Extension to partially isochronous systems 69

4.2.1 Tree representation 71

4.2.2 Formal power series relations involving the splitting matrix . . . 72

5 Upper bounds on homoclinic splittings II **75**
5.1 Moving Poincaré sections . . . 76
5.2 Tree representations of the auxiliary dynamics . . . 80
5.3 Connection between the auxiliary dynamics and the splitting . . . 86

6 Lower bounds on the splitting **93**
6.1 Basic lower estimates . . . 93
6.1.1 Lower bounds on the Melnikov integral . . . 95
6.1.2 Systems with one fast frequency . . . 97
6.2 Examples of Melnikov dominance . . . 100
6.2.1 Systems with prefixed Lyapunov exponent . . . 100
6.2.2 Improved bounds for three dimensional systems . . . 106
6.3 D'Alembert-like problems . . . 116
6.3.1 Big uni-modal perturbations . . . 117

7 Systems with more general unperturbed separatrices **123**
7.1 Acceptable functions $F(q)$ . . . 124
7.2 Computation of the Wronskian matrix . . . 130
7.3 Homoclinic splitting for the generalized pendulum . . . 133

8 Arnold diffusion **135**
8.1 Heteroclinic chains . . . 136
8.1.1 Heteroclinic intersection for systems with one fast frequency . . . 136
8.2 Transition chains . . . 141

A Appendices **143**
A.1 Examples of functions in $B(b,d)$ . . . 143
A.2 Evaluation of the coefficients $T(k), N(k)$ . . . 146
A.3 Notions on lattices in $\mathbb{Z}^n$ . . . 147
A.4 Normal form theorem . . . 150
A.5 Fast averaging Theorem for uni-modal perturbations . . . 155
A.6 Proof of Theorem 4.20 . . . 155
A.7 Cancellations due to integration by parts . . . 157
A.8 Properties of the matrix $\mathcal{M}$ due to cancellations . . . 160

Introduction

Generalities and a class of models

The problem of the stability under perturbations of dynamical systems is a "fundamental" problem of classical mechanics (as formulated by Poincaré in [P]).

For integrable Hamiltonian systems with n degrees of freedom, it was long believed (up to the 50'ies) that maximal (i.e. n dimensional) invariant tori were usually destroyed by most perturbations[1]. This was disproved for non-degenerate Hamiltonians[2] in the theorem by Kolmogorov, proved in full detail by Arnol'd for real-analytic flows and for smooth maps by Moser. The Kolmogorov, Arnold, Moser (KAM for short) theorem states that those invariant tori with sufficiently incommensurate (diofantine[3]) frequencies $\omega(I)$, persist for sufficiently small perturbations of a non-degenerate integrable system. Such tori form a set of positive measure in the phase space, and as the system approaches to integrable the measure of the complementary set approaches zero[4].

One expects that the, dense but zero measure, set of maximal tori of the unperturbed system with commensurate frequencies is not preserved, in general, under perturbations no matter how small.

Such sets of tori with commensurate frequencies are called "resonant". In particular

---

[1]**Some standard definitions:** An n degrees of freedom Hamiltonian system is characterized by a Hamiltonian function h defined on a $2n-$dimensional manifold $\mathcal{M}$ endowed with a symplectic structure i.e. a closed alternate and non-degenerate two-form w. In local coordinates $(p,q) \in \mathbb{R}^{2n}$ such that the two form is $dp \wedge dq$ we call the flow $\phi_h^t(q_0,p_0)) \in \mathbb{R}^{2n}$ the solution of:

$$\dot{q} = \partial_p h(p,q)\,, \quad \dot{p} = -\partial_q h(p,q).$$

The change of coordinates which preserve w are called symplectic.
A Hamiltonian system is called integrable if there exists a symplectic change of coordinates: $p,q \to I,\psi \in \mathbb{R}^n \times \mathbb{T}^n$, $\mathbb{T}^n$ being the n dimensional torus, where the Hamiltonian $h(p(I,\psi),q(I,\psi) = H(I)$. In such case the flow is confined on n dimensional tori:

$$I(t) = I(0)\,, \qquad \psi(t) = \psi(0) + (\nabla_I H)|_{I(0)} t.$$

The vector $\omega(I) = (\nabla_I H)|_I$ is called the frequency of the torus with initial datum I.

[2]we say that $H(I)$ is non degenerate in a domain $D \subset \mathbb{R}^n$ if $\det \partial^2 H(I) \neq 0$ for all $I \in D$.

[3]A vector $\omega \in \mathbb{R}^n$ is diofantine, with constants C,τ, if it satisfies a relation of the form $|m \cdot \omega| \geq C/|m|^\tau$ for all integer vectors $m \neq 0$.

[4]This means that the constant C can be taken to be small with the perturbation parameter.

if there there exists a $k \times n$ entire matrix N, of rank k, such that

$$N\omega(I) = 0 \ \text{ with } \ N \in \ \mathrm{Mat}_{k\times n}(\mathbb{Z}) \,, \quad \mathrm{Rank}\ (N) = k \tag{0.1}$$

we will call I an order k resonance.[5].
For iso-energetically non-degenerate[6] systems with two degrees of freedom the existence of a positive measure set of two-dimensional persistent tori forces the behavior of the whole system to be stable for purely topological reasons as the two-dimensional tori separate the three-dimensional energy surface. The, possibly chaotic, behavior near the resonances is thus confined in the layers between persistent tori.

On the other hand there is no a-priori objection to the possibility of action-unstable motions for higher dimensional systems, as the complementary set of the preserved tori is connected.
Arnold, for the first time in the appendix of [A1], formulates the problem and states the following conjecture(see [Dy]):

"*...A typical case in many-dimensional problems of perturbation theory is topological instability: through an arbitrarily small neighborhood of any point there pass phase trajectories along which the action variables drift away from the original value by a quantity of order one...*"

Such topological instability is known as Arnold Diffusion.

In this thesis we shall consider the $n+1$ degrees of freedom Hamiltonian:

$$H(I,p,\psi,q,\varepsilon,\mu) = \frac{1}{2}p^2 + \frac{1}{2}I \cdot AI + b \cdot I + \varepsilon F(q) + \mu f(\psi,q), \tag{*}$$

where

$$((I,p)\,,\,(\psi,q)) \in (U \times (-1,1)) \times (\mathbb{T}^n \times \mathbb{T})\,, \quad U \subset \mathbb{R}^n$$

are a set of conjugate action-angle variables ($\mathbb{T}^n$ being the standard torus $\mathbb{R}^n/2\pi\mathbb{Z}^n$), ε, μ are small parameters and the matrix A is semi positive definite (A and b can depend on ε in a fashion we will specify in the following). The functions $F(q)$ and $f(\psi,q)$ are real analytic and even. Moreover we choose the function $F(q)$ so that $p^2 = -2\varepsilon F(q)$ is the graph of a separatrix having $p = q = 0$ as the only (unstable) fixed point.
The basic problem addressed here is the study of homo/heteroclinic transversal intersections and finding upper and lower bounds on suitable measures for the transversality.
It should be clear that such problems are much simpler if one considers $\varepsilon > 0$ and $\mu \ll \varepsilon$ an independent parameter; in such case Hamiltonian (*) is called *a-priori unstable* .

[5]This means that the unperturbed motion with initial datum $I(0) = I$ is on a $n-k$ dimensional torus.

[6]An Hamiltonian $H(I)$ is called iso-energetically non degenerate if

$$\det \begin{pmatrix} \partial_I^2 H & \partial_I H \\ \partial_I H & 0 \end{pmatrix} \neq 0$$

on the energy surface.

In this thesis we will mainly consider the *a-priori stable* case, which means setting $\mu = \varepsilon^P$ for some $P > 1$ (in some special case we will consider also $P = 1$).

To motivate the choice of Hamiltonian (*) we briefly review the properties of an iso-energetically non degenerate Hamiltonian near an order k resonance. We will argue that the a-priori stable Hamiltonian (*) is a "natural" model for iso-energetically non degenerate Hamiltonians near a simple, i.e. order one, resonance ($k = 1$ in relation (0.1)).

Resonant Hamiltonians

We consider a close to integrable, analytic system

$$H(I, \psi, \varepsilon) = H_0(I) + \varepsilon F(I, \psi),$$

in action-angle variables $I \in U \subset \mathbb{R}^d$ $\psi \in \mathbb{T}^d$. .

Classical averaging theory (see for instance [Dy]) shows that near an order k resonance described by the matrix N as in relation (0.1), H is modeled, in appropriate local action-angle coordinates, by a Hamiltonian:

$$\bar{H}(I', \varepsilon) + \varepsilon g_N(I', \psi') + \mu f(I', \psi'), \tag{0.2}$$

with

$$g_N(I', \psi') = \sum_{k \in \Lambda_N} g_k(I') e^{ik\cdot\psi'},$$

Λ_N being the lattice generated by the rows of N.
The functions $g_N(I', \psi')$, $f(I', \psi')$ are analytic in some $U' \times \mathbb{T}^d$ (U' close to U) and $\mu = \varepsilon^P$ with $P > 1$.

As remarked for Hamiltonian (*), it is simpler to study Hamiltonian (0.2) considering μ and ε as independent parameters.
For simple resonances, it is easily seen that the Hamiltonian (0.2) is still "analytically soluble" for $\varepsilon > 0$ and $\mu = 0$. Up to a linear symplectic change of coordinates, $I', \psi' \to J, \varphi$, one can assume that g_N depends only on one angle, say φ_d:

$$\bar{H}(J, \varepsilon) + \varepsilon g_N(J, \varphi_d) + \mu f(J, \varphi). \tag{0.3}$$

Let us study Hamiltonian (0.3) for $\mu = 0$. The actions $J_1, \ldots, J_{d-1}$ are still constants of motion; the time evolution of J_d, φ_d does not depend on the φ_i with $i < d$ and so is soluble (by integrations and inversions).

Notice that Hamiltonian (0.3) with $\mu = 0$ is not integrable, in the classical sense (i.e. in the sense of footnote (1)), as the resonant variable φ_d can have unstable fixed points and one cannot define action-angle variables near the hyperbolic trajectories.
A model for Hamiltonian (0.3) with $\mu = 0$ is:

$$H(J, \varphi_1, \varepsilon) = \bar{H}(J_1, \ldots, J_n) + \frac{1}{2}p^2 + \varepsilon(\cos q - 1), \tag{0.4}$$

with $n = d - 1$ and $J_d, \psi_d = p, q$. Notice that Hamiltonian (*) with $F(q) = \cos q - 1$ is of the type (0.4).
On the other hand, if we consider higher order resonances, the μ independent Hamiltonian

$$\bar{H}(I', \varepsilon) + \varepsilon g_N(I', \psi'),$$

is generally not analytically soluble so that, in connection to the problem of Arnold diffusion, most authors consider only simple resonances (see [LMS] for an approach to general resonances).

The dynamics of Hamiltonian (0.4)

The trajectories of Hamiltonian (0.4) are the direct product of an integrable motion on $n = d - 1$ dimensional tori and of the motion of the pendulum.
In our notation the pendulum has a stable fixed point in $q = \pi$, $p = 0$ and an unstable one in $q = p = 0$. The stable and unstable manifolds of such fixed point coincide and are represented in phase space by a curve, called the separatrix $p^2 = 2\varepsilon(\cos q - 1)$.
We have n dimensional unstable tori $\mathcal{T}(J)$ (direct product of the motion of the $\varphi_1, \dots, \varphi_n$ and $q = p = 0$) and their ($n + 1$ dimensional) stable/unstable manifolds $\mathcal{W}^{\pm}(\mathcal{T}(J))$ which are the direct product of the motion of the $\varphi_1, \dots, \varphi_n$ with the motion on the separatrix.
If $H(J, p, q, \varepsilon)$ is iso-energetically non degenerate for $\varepsilon \neq 0$ a set of such unstable n dimensional tori $\mathcal{T}(J, \mu)$, survive (on fixed energy levels) the onset of the μ dependent perturbation, together with their $n + 1$ dimensional stable-unstable manifolds $\mathcal{W}^{\pm}(\mathcal{T}(J, \mu))$. Such manifolds however may intersect transversally in a curve, this is the so called "homoclinic splitting" and is known to be related to "chaotic" behavior.

General techniques for proving Arnol'd diffusion

The existence of Arnold diffusion is usually proved by following the mechanism proposed by Arnold in [A2], where the author considers a model of an a-priori unstable almost integrable system near a simple resonance. Interest on the subject was renewed in [CG], followed by several papers; see for instance [GGM3], [BB1] and references therein.

To illustrate the mechanism used for proving Arnold diffusion, let us state some definitions taken from [C], where "Arnold diffusion" is named "Arnold instability".

Definition 0.1 (Heteroclinic chains). *A heteroclinic chain is a set of* $N \geq 1$ *trajectories* $z^1(t), \dots, z^N(t)$ *together with* $N + 1$ *different minimal sets*[7] $T_0, \dots, T_N$ *such that for all* $1 \leq i \leq N$

$$\lim_{t \to -\infty} dist\ (z^i(t), T_{i-1}) = 0 = \lim_{t \to \infty} dist\ (z^i(t), T_i).$$

[7] A closed subset of the phase space is called minimal (with respect to a Hamiltonian flow ϕ_h^t) if it is non-empty, invariant for Φ_h^t and contains a dense orbit. In our case the minimal sets will be unstable tori $\mathcal{T}(I)$ with $\omega(I)$ diofantine.

Definition 0.2 (Transition chains). *A heteroclinic chain is called a transition chain if for any $r > 0$ there exists a trajectory $z(t)$ and a time $T > 0$ such that*

$$dist\ (z(0), T_0) \le r\,, \quad dist\ (z(T), T_N) \le r\,, \quad \sup_{0\le t\le T} dist\ (z(t), Z) < r$$

where Z is the closure of the union over i of the $\{z^i(t) : t \in \mathbb{R}\}$. The sets T_0 and T_N are said to be connected by as transition chain.

Definition 0.3 (Arnold instability). *Given $E \in \mathbb{R}$ consider an Hamiltonian h_ε (with Hamiltonian flow $\phi^t_{h_\varepsilon}$) such that h_0 represents an integrable system.*
The system $(\phi^t_h, h_\varepsilon^{-1}(E))$ is called Arnold unstable if there exist two positive numbers ε_0 and d_0 such that for all $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$ there exist (closed) invariant sets $T(\varepsilon)$, $T'(\varepsilon) \subset h_\varepsilon^{-1}(E)$ satisfying the following conditions:

(i) $T(\varepsilon)$, $T'(\varepsilon)$ are continuous, at $\varepsilon = 0$, in the Hausdorff metric and if Π_I denotes the natural projection over the action variables then

$$\Pi_I T(0) = \{I\}\,, \quad \Pi_I T'(0) = \{I'\}\,, \quad \textit{with} \quad |I' - I| > d_0\,;$$

(ii) for each $0 < |\varepsilon| < \varepsilon_0$ $T((\varepsilon)$, $T'(\varepsilon)$ are connected by a transition chain.
Finally the system is said uniformly Arnold unstable in a region $V \in \mathbb{R}^n$ for any $E \in [E_1, E_2]$ if the invariant sets $T(\varepsilon)$, $T'(\varepsilon)$ have the property:

$$\Pi_I T(0), \Pi_I T'(0) \in V$$

and the constants ε_0, d_0 depend only on V and on E_1, E_2.

Thus, to prove Arnold instability for system(*) one typically proceeds in three steps:

1. Homoclinic intersection:

   - Prove that the systems (*) admit a set of unstable n dimensional tori together with their $n+1$ dimensional stable-unstable (Lagrangian) manifolds $W^\pm$; for system (*) such manifolds are graphs on the angles:

   $$W^\pm = \{I^\pm(\psi, q), p^\pm(\psi, q), \psi \in \mathbb{T}^n, q \in (-a, a) \ \text{ with } \ a \in (0, \pi)\}.$$

   - Prove that such manifolds intersect transversally in a curve (as expected). In the case of Hamiltonian (*), the assumed parity conditions imply that $\psi = 0$ $q = \pi$ is a Homoclinic point, i.e. lies on the intersection curve.

   - Provide estimates on the measure of the transversality in appropriate (order one) regions in the action variables.

2. Prove the existence of *heteroclinic chains* of n-dimensional tori by showing that the persistent tori are "close enough" with respect to the transversality measure in the prescribed regions.

3. Prove that such heteroclinic chains are *transition chains* for which the action variables undergo an $O_\varepsilon(1)$ variation in a finite time[8].

A natural question that arises in this scheme for proving Arnold instability is what is a good measure of transversality.
For system (*) (in the coordinates (I, ψ)) one may consider (as in [A2]) the splitting determinant, i.e. the determinant of:

$$\Delta = \partial_{\psi_i}(I_j^+(\psi, \pi) - I_j^-(\psi, \pi)),$$

whose eigenvalues estimate (in local coordinates) the angles of the intersection of $\mathbb{W}^+$ and $\mathbb{W}^-$ at the Poincaré section $q = \pi$.

Then, if the gaps on the persistent tori are smaller than $|\det \Delta|^a$, one can use the Implicit Function Theorem to prove heteroclinic intersections for persistent n dimensional tori at distances of order $|\det \Delta|^b$ (for suitable $a, b > 0$).

Remark 0.4. *This is a local point of view. However, since the symplectic group acts transitively on the couples of transverse Lagrangian manifolds, estimates on the "Euclidean " angles of the intersection are expected to be coordinate dependent*[9].

Analytical proofs of Arnold instability rely strongly on the choice of an appropriate region of the initial data in the action variables (and on the characteristic frequencies); to illustrate this let us return to Hamiltonian (*) which we describe in full detail.

$$H(I, p, \psi, q, \varepsilon, \mu) = \frac{1}{2}\big(p^2 + I \cdot A(\varepsilon)I\big) + b(\varepsilon) \cdot I + \varepsilon F(q) + \mu f(\psi, q), \qquad (*)$$

where, as we said before, $A(\varepsilon)$ is an $n \times n$ semi positive definite matrix.
The integrable part of Hamiltonian (*) (with $\mu = 0$) can model both completely anisochronous systems of rotators (i.e. $A(\varepsilon)$ is positive definite) and isochronous systems of harmonic oscillators ($A(\varepsilon) \equiv 0$); moreover by varying the $\varepsilon-$dependence of $A(\varepsilon)$, $b(\varepsilon)$ one can model both non-degenerate and degenerate Hamiltonians.
To prove Arnol'd instability (following the scheme proposed in page ix) in a region of the action space, one needs conditions on the order of magnitude of the frequencies $\omega(I) = AI + b$ in such region. In particular we require that the components $\omega_j(I)$ are "not too slow" by setting $|\omega_j(I)| \geq C\varepsilon$ for some order one[10] C. Moreover it is

---

[8] Having performed these three steps one can rise the question of finding (good) estimates on the instability time

[9] [LMS] provides an intrinsic definition of the transversality measure (which coincides with Δ in local coordinates at the Poincaré section $q = \pi$) and its variation laws through symplectic change of coordinates and different choices of Poincaré sections.

[10] One could prove Arnold instability under less restrictive conditions $|\omega_j(I)| \geq C\varepsilon^b$ for some $b > 1$, we set $b = 1$ only for simplicity.

useful to distinguish between **fast** (i.e. order one in ε) and **slow** (going to zero with ε) components of the frequency vector $\omega(I)$; we will call $V_m(E) \subset \mathbb{R}^n$ a region (of action space) having "energy" E (i.e. $IAI+bI=E$) and m fast components for the frequency vectors (often referred to as m fast frequencies or variables) and consequently $n-m$ slow frequencies.
The regions where there are *at least two* different orders of magnitude for the frequencies are particularly relevant in proving Arnol'd instability (in this thesis we will mainly consider such regions).

Quantitatively we set the following (non minimal but already quite cumbersome) conditions on $A(\varepsilon)$, $b(\varepsilon)$ and $V_m(E)$:

Condition 0.5. • *The functions $F(q)$ and $f(\psi,q)$ are real analytic and even. Moreover we choose the function $F(q)$ so that $p^2=-2\varepsilon F(q)$ is the graph of a separatrix having $p=q=0$ as the only (unstable) fixed point.*

• *$A(\varepsilon)$ is diagonal.*

• *The eigenvalues*[11] *of $A(\varepsilon)$, a_i (where $i=1,\dots,n$)are either identically zero or:*

$$a_i = C\varepsilon^{\alpha_i} \quad \textit{with} \quad 0 \le \alpha_i \le 1\,, \quad \textit{and } C \textit{ non zero and } \varepsilon\textit{-independent.}$$

• *Without loss of generality we will suppose that $a_i \neq 0$ for all $i \le h$ for some $0 \le h \le n$. The remaining $n-h$ eigenvalues a_i are zero.*

• *$b(\varepsilon)$ is an n-dimensional* **diofantine** *vector $b=(b_1,\dots,b_n)$ such that $b_1=\cdots=b_h=0$ and the remaining b_i have the form:*

$$b_i = C\varepsilon^{\beta_i} \quad \textit{with} \quad 0 \le \beta_i \le 1\,, \quad \textit{and} \quad C = O_\varepsilon(1).$$

• *We consider the system near a simple resonance for the variable p: $p \in B_{\sqrt{\varepsilon}}(0)$.*

• *Let $\omega(I)=AI+b$,we assume that the I variables are in a domain*

$$\mathbb{R}^n \supset V_m(E) := \{I : O_\varepsilon(1) \ge |I_j| \ge O_\varepsilon(\varepsilon)\,,\ j=1,\dots,n\,, \quad AI\cdot I + bI = E \\ \textit{there exist } i_1,\dots,i_m \textit{ such that } |\omega_{i_j}(I)| = O_\varepsilon(1)\},$$

with $m \le n$ and $E=O_\varepsilon(1)$; we will call V_m a domain with m fast frequencies $\omega_{i_1},\dots,\omega_{i_m}$ (or fast variables) as the orbits of Hamiltonian (*) with $\mu=0$ are tori run with frequency $\omega(I)=A(\varepsilon)I+b(\varepsilon)$.

In the domains V_0 where there are no fast variables ($m=0$), the Hamiltonian (*) can be written (via an appropriate change of variables) as a-priori unstable and then solved via classical perturbation theory (see [CG] and [C]). In fact, for small enough

[11] Now and in the following we will say that $C(\varepsilon)=O_\varepsilon(f(\varepsilon))$ if $\lim_{\varepsilon\to 0} C(\varepsilon)/f(\varepsilon) = l \neq 0$.

ε the matrix Δ is well approximated by its first order perturbation in μ, the so-called Melnikov integral:

$$M_{i,j} = \int_{-\infty}^{\infty} (\partial_{\psi_i}\partial_{\psi_j} f)(\psi_0(t), q_0(t)). \tag{0.5}$$

Where $\psi_0(t), q_0(t)$ is the motion on the separatrix for $\mu = 0$.
The presence of $m \neq 0$ fast variables makes matters much more difficult as *the determinant of M is exponentially small in ε while (if $m \neq n$) the higher order truncations of Δ have generally only polynomially small entries* so that one should consider μ exponentially small w.r.t ε in order to have

$$\Delta \sim M.$$

As we have said the natural value of μ is ε^P for some $P > 0$. In such case M is not a good approximation of Δ and it is not a trivial matter to show that $\det\Delta$ is exponentially small in ε for $m \neq 0$.
The first step in estimating $\det\Delta$ is finding exponentially small upper bounds for systems with fast frequencies.
Then one would like to prove that $\det M$ is large enough to dominate on the higher order terms in the μ expansion of $\det\Delta$ and consequently prove lower bounds on $\det\Delta$.
Upper bounds, with $\mu \leq \varepsilon^P$, are derived in [G1] for $m = n$ in [GGM1] for systems with three degrees of freedom and three time scales and in [BB1] for isochronous Hamiltonians and generic n, m.
All the cited articles set $F(q) = \cos q - 1$ and require that the perturbating function $f(\psi, q)$ is a trigonometric polynomial in q while we shall allow more general functions $F(q)$ and analytic assumptions on $f(q, \psi)$ (see Condition 0.5). The cited articles provide as well lower bounds on $\det\Delta$ (see as well [GGM1]- [GGM4]) for systems with one fast time scale.
The problem of upper bounds is considered as well in [LMS] for quite general ($n + k$ dimensional) systems, in the presence of an order k resonance in a region characterized by two time scales ($m = n$). The results of [LMS], applied to Hamiltonian (*) (so to a simple resonance) lead to the results of [G1], however the proof contained in [LMS] is coordinate independent so it would be interesting to see if it applies to three time scale systems.
A system with $m = n = 2$ is considered in [DGJS] providing upper and lower bounds on the distance between stable and unstable manifolds, it is not however clear if this estimates can be used to prove the existence of heteroclinic chains (see the discussion in [GGM2]).

In this thesis we generally follow the strategy proposed in [CG], [G1] and [GGM1]. These articles use perturbation theory to construct the "homoclinic trajectories" (i.e. the trajectories which are bi-asymptotic to an n dimensional torus run with prescribed frequency). This approach by series expansion in the parameter μ (with fixed $\varepsilon > 0$) is quite old; it is a generalization to the partially hyperbolic setting of Hamiltonian (*), of the *Lindstedt series* proposed by Poincaré, Lindstedt et al. in the 19'th century.

Proving the convergence of such series is quite complicated and was indeed an open problem, even in the non hyperbolic setting, up to the '80-ies when it was solved by Eliasson [E], see as well [G1] and [CF] (moreover see [GGM4] for a proof of the convergence in the hyperbolic setting). The main point is to find sufficient *compensations* between "big" terms of the Lindstedt series in order to ensure the convergence.

While one can use KAM theory to prove the local existence of the manifolds $\mathbb{W}^{\pm}$ (and then extend them via the Hamiltonian flow), one cannot use the KAM algorithm to estimate the splitting determinant as the computations involved are unmanageable.

The problem of convergence of the perturbation series is avoided, in [CG] [G1], by combining Lindstedt series and KAM theory. Namely one considers suitable truncations of the Lindstedt series whose remainder is bounded via a KAM theorem (which ensures, under appropriate conditions, that the homoclinic trajectories exist and are analytic in $\mu \le \mu_0$). To study a large but finite number of terms in the perturbation series it is natural to use a "graph theoretical" (tree) representation (see [GJ] for applications of tree representations to Taylor series). The tree representation, which contains information on the symmetries of the Taylor series, is well suited to show the cancellations which are necessary to prove the exponential smallness of the splitting matrix.

Roughly speaking the exponentially small terms in the splitting matrix appear via the following "shift of contour" formula (c_1 and c_2 are positive ε independent parameters):

$$|\int_{-\infty}^{\infty} e^{-\frac{c_1}{\varepsilon}t} g(t)| \le O_{\varepsilon}(e^{-\frac{c_2}{\varepsilon}}) \,,$$

for all the analytic $g(t) \in L_1$.

This formula proves for instance that the Melnikov term (defined in (0.5))is exponentially small.

The main problem is that the terms of order higher than one in the expansion of the homoclinic trajectories are in general not analytic in t (for $t = 0$ as they come from the time evolution of $\mathbb{W}^{+}$ for $t > 0$ and $\mathbb{W}^{-}$ for $t < 0$). so that, even if all the frequencies are fast ($m = n$), the splitting matrix apparently contains "big" (i.e. polynomially small in ε) terms, arising from integrals of non analytic functions. In [G1] the author shows that such "big" terms *cancel* so that a suitable (say order $K(\varepsilon)$) truncation of the splitting matrix is exponentially small in ε. Bounds on the splitting matrix are then derived by showing that one can choose $k \le K(\varepsilon)$ so that the remainder (estimated via KAM theory) at order k is small with respect to the bounds on the order k truncation.

A different approach is to prove directly the convergence of the Lindstedt series by proving via the tree representation both cancellations and compensations (see [GGM4]).

Brief description of the main results and of the techniques used

In this thesis we consider mainly the items 1) and 2) at page ix and we simply give a brief review of the construction of Arnold unstable orbits (taken from [CV]). We will not attempt any estimate on the diffusion time. For such estimates see for instance [BB1], [Be], [BB2], [BCV] and references therein.

• We prove exponentially small upper bounds for $\det\Delta$ for Hamiltonians in the class (*) in regions V_m with $m \neq 0$ fast variables.

Theorem 0.6 (Upper bounds). *Assume Conditions 0.5. The Hamiltonian (*) , considered in the domains $V_m(E)$ with $E = O_\varepsilon(1)$ $m \neq 0$ has an homoclinic point at $q = \pi, \psi = 0$. The determinant of the splitting matrix in such point is*

$$\det\Delta \leq O_\varepsilon(e^{-c/\varepsilon^b}).$$

where c and b depend on the domain V_m and on the analytic properties of the perturbing function $f(\psi, q)$.

This Theorem generalizes [GGM1] and [BB1] which consider respectively a partially isochronous and partially degenerate Hamiltonian (*) with three degrees of freedom, and a completely isochronous Hamiltonian (*) with n degrees of freedom. Both references set $F(q) = \cos q - 1$ and $f(\psi, q)$ a trigonometric polynomial (at least in the q variables).

• For systems with $m = 1$ fast variables (say $\psi_{\bar{\imath}}$) we prove lower bounds for the splitting determinant for the Hamiltonians (*) satisfying the following conditions:

Condition 0.7. *a) The function $f(\psi, q)$ is a trigonometric polynomial in the ψ_i*

$$f(\psi, q) = \sum_{|\nu|\leq N} f_\nu(q) e^{i\nu\cdot\psi},$$

and all the functions $f_\nu(q(t))$, where $\psi(t), q(t)$ is the solution of Hamiltonian (*) for $\mu = 0$, are rational function of $e^{\lambda\sqrt{\varepsilon}t}$. $\sqrt{\varepsilon}\lambda > 0$ is the Lyapunov exponent of the generalized pendulum (see next item).

b) The Hamiltonian $\frac{1}{2}p^2 + \varepsilon F(q)$ has the following trajectories:

1. *$q = \dot{q} = 0$ is an hyperbolic fixed point and the separatrix*

$$\frac{\dot{q}^2}{2} + \varepsilon F(q) = 0$$

contains only this fixed point.

2. *On the separatrix, we can chose a sign for $\dot{q}$ and the equation of motion on the separatrix is:*

$$\dot{q} = \pm\sqrt{2}\sqrt{F(q)} = \pm G(q)$$

where $G(q) \geq 0$ and $G(q) = 0$ if and only if $q = 0, 2\pi$. Notice that $\dot{q}(t)$ is even and $q(t)$ is odd provided that we set $q(0) = \pi$.

3. *The time evolution on the separatrix $q(t)$ (on a prefixed branch), satisfies*

$$e^{iq(t)} = R(e^{-\lambda\sqrt{\varepsilon}t}) \quad \textit{where } R(y) \textit{ is a rational function .} \tag{**}$$

c) The function $f(\psi(t), q(t))$ satisfies appropriate "non-degeneracy conditions", which we describe later, here let us state simple sufficient conditions:

1. *The functions $f_\nu(q(t))$ all have the same poles $t_1, \dots, t_M$. The function $q(t)$ has poles $\tau_1, \dots, \tau_N$ and:*

$$D = \min_{i=1,\dots,M} |Im\ t_i| \leq \min_{j=1,\dots,N} |Im\ \tau_j|.$$

2. *The Melnikov matrix defined in (0.5) is non degenerate and the $f_{e_i}(q)$ $i = 1, \dots, n$ are all different from zero.*

A simple example of functions $F(q)$ satisfying Condition 0.7 b) are the following:

$$F(q) = -\frac{1}{2}\big(\sin^2 q + a(\cos q - 1)^2\big),$$

with $a \in [1, \infty)$ ($a = 1$ is the standard pendulum).
Under this conditions we prove the existence of heteroclinic intersections provided that $\mu \leq \varepsilon^P$ where P depends on the poles of $q(t)$ and of the functions $f_\nu(q(t))$.

Theorem 0.8 (Lower bounds). *Consider Hamiltonian (*) under conditions 0.7. Given $0 \leq \alpha \leq \frac{1}{2}$ consider the domains $W(E, \bar\imath, \alpha) =$*

$$\{I \in V_1(E)\,,\ \ |\omega_{\bar\imath}(I)| = O_\varepsilon(1)\,,\ \ O_\varepsilon(\varepsilon) \leq |\omega_j(I)| \leq O_\varepsilon(\varepsilon^\alpha)\,,\ j = 1, \dots, n\ j \neq \bar\imath\}.$$

The determinant of the splitting matrix at the homoclinic point, $q = \pi, \psi = 0$, is bounded from below by a quantity of the order of the Melnikov integral:

$$|\det \Delta| \geq C\varepsilon^{-Q} e^{-D/\sqrt{\varepsilon}},$$

provided that

$$\mu \leq \varepsilon^P,$$

where $P = \max(p+5, 4\tau+4)$ τ being the diofantine exponent of the frequency vector ω. The parameters p, Q depend on the degree of the poles of the $f_\nu(q(t))$ (p is the degree of the pole closest to the real axis) and D is defined in Condition 0.7 c).

After proving this Theorem we provide a Normal Form Theorem for Hamiltonian (*). Such theorem, restricted to systems with one fast frequency implies the existence of heteroclinic chains.

Theorem 0.9 (Arnold instability). *Given $E \in [E_1, E_2]$ with $E_1, E_2 = O_\varepsilon(1)$, the Hamiltonians (*), satisfying Conditions 0.7 and having at least one degenerate variable (namely one or more of the a_i are of order ε), are uniformly Arnold unstable in each of the domains $D(E, \bar{\imath})$, for all values of μ such that:*

$$\mu \leq \varepsilon^P,$$

where P depends on the constants p and Q of the preceding theorem.

The bounds on μ proposed it this Theorems *are not optimal*, in particular one can obtain better bounds by using the techniques proposed in [Ge]. We illustrate this on examples of three degrees of freedom systems[12] where we prove Arnold instability for

$$\mu \leq \varepsilon^{p+5/2}.$$

• Finally we consider some special systems with three degrees of freedom and three time scales which we call "D'Alembert-like" Hamiltonians as they are quite similar to the Hamiltonian proposed in [CG] (see as well [GGM3]) as a model for the D'Alembert problem. Such problem, of interest in celestial mechanics, is characterized by the presence of three relevant time scales and of a big (i.e. order ε) uni-modal (i.e. the lattice generated by the frequencies of $f(\psi, q)$ is one dimensional) perturbation. To be explicit let us write down the simplified D'Alembert Hamiltonian proposed in [GGM3]:

$$\frac{1}{2}(\varepsilon J^2 + p^2) + I\omega_1 + \varepsilon[(\cos q - 1) + \alpha A(\phi + \psi)B(q))] + \mu f(\phi, \psi, q), \qquad (0.6)$$

where the functions $A(x)$, $B(x)$ are trigonometric polynomials of degree N and α is a free (order one in ε) parameter. The technically difficult question is to prove lower bounds on the splitting determinant (Melnikov dominance) when α is of order one in ε, and so clearly does not satisfy the conditions of Theorem 0.8.
The article [GGM3] proves a semi-hyperbolic KAM theorem and consequently upper bounds on the splitting determinant for Hamiltonian (0.6). The problem of lower bounds is left open as it requires proving appropriate cancellations in the series representation of the splitting determinant. We prove such cancellations (and so lower bounds and Arnold instability) provided that $f(\phi, \psi, q)$ is NOT a trigonometric polynomial and respects the following:

Condition 0.10. *the function f is a trigonometric polynomial in ψ, ϕ and rational in e^{iq} with at least one pole for finite values of Imq and $Req \neq 0$.*

Theorem 0.11. *The Hamiltonian (0.6), respecting Condition 0.10, is uniformly Arnold unstable in the domain:*

$$W(E) := \{H(I, J, \psi, \phi) = E\,, \quad O_\varepsilon(1) = b \leq |I|, |J| \leq a = O_\varepsilon(1)\}$$

for $E \in [E_1, E_2]$ with $E_1, E_2 = O_\varepsilon(1)$, provided that $\mu \leq \varepsilon^{p+5/2}$ and $\alpha \ll 1$ but still $O_\varepsilon(1)$.

[12] This restriction is only to give explicit examples, we show that one can apply the same procedure to systems with n degrees of freedom.

Brief review of the techniques

Following [CG], [G1] and [C] we use perturbation theory to construct the "homoclinic trajectories" (i.e. the trajectories which are bi-asymptotic to an n dimensional torus run with prescribed frequency). This leads to recursive equations for the coefficients of the Taylor expansion of the homoclinic trajectories (in the parameter μ with $\varepsilon > 0$ a fixed parameter).
Then, still following [G1], we introduce a suitable "graph theoretical" representation of the homoclinic trajectories useful to identify *cancellations*. We use quite standard notions on trees: mainly labeled rooted trees and their automorphisms groups. With respect to [G1], we use a different grouping algorithm for the tree representation; in particular we use the isomorphism groups of trees which, we believe, make computations on trees more explicit and we hope simpler. We assign quite a few labels to the trees to represent directly on the trees the relevant structure of the homoclinic trajectory. As mentioned before, the terms of order higher than one in the expansion of the homoclinic trajectories are in general not analytic in t. Following [G1] we represent this by introducing specially labeled nodes (called fruits); such nodes are responsible for the appearance of the non analytic terms. An accurate study of the tree representation (and some notions on asymptotic power series) enable us to prove Theorem 0.6. One of the main tools is a formal linear equation for the splitting matrix (which generalizes the one proposed in [GGM1]). This formal linear equation directly implies exponential smallness and, we think, simplifies significantly the procedure of [GGM1] (as wee as extending the results of [GGM1] to Hamiltonian*).

We provide as well an alternative proof of Theorem 0.6, following the strategy of [BB1] adapted to perturbative series and tree representation (so we generalize the results of [BB1] to anisochronous Hamiltonians although our bounds are less sharp that those obtained in the cited article).

The proofs of Theorems 0.8 and 0.9 follow the general strategies proposed in [GGM1] and [GGM3] which we refine and develop so to apply them to our more general Hamiltonian (*)).
To prove Theorem 0.8 we provide "accurate enough" bounds on the coefficients of the series representation of the homoclinic trajectories. We explicitly compute the first order term and use Cauchy estimates to find upper bounds on the terms of order higher than one. The fact that $f(\psi, q)$ is not taken to be a trigonometric polynomial creates various technical problems. For instance one cannot Fourier expand $f(\psi, q)$

$$f(\psi, q) = \sum_{n\in\mathbb{Z}, m\in\mathbb{Z}^n} f_{n,m} e^{iqn} e^{i\psi\cdot m},$$

and bound it (and its derivatives) on annular domains $\mathbb{T}^{n+1} \times i(-r, r)$; instead one has to choose suitable (in general non annular) domains on which to perform Cauchy estimates.
To prove Theorem 0.9 we provide a Normal Form Theorem (which generalizes the corresponding Theorem proposed in [GGM3]).

Finally, in the proof of Theorem 0.11 we use the "improved" tree representation introduced in [Ge]. The main idea is to apply the improved bounds coming from this tree representation to the "analytic" terms (related to "fruitless" trees as said above), this is quite delicate and requires, for instance, an attentive use of the formal linear equation used for the proof of Theorem 0.6.

The thesis is organized as follows:

In **Chapter 1** we provide some basic notions.
In **Section 1.1** We consider an *anisochronous* Hamiltonian of type (*), namely with $A(\varepsilon)$ positive definite for $\varepsilon \neq 0$ (consequently $b(\varepsilon) = 0$). Moreover we set $F(q) = \cos q - 1$.
For such system we state a KAM theorem; following [CG], we define the homoclinic trajectories

$$z(\varphi, \omega, t) := (I(\varphi, \omega, t)\,,\ \ \psi(\varphi, \omega, t)\,,\ \ q(\varphi, \omega, t)),$$

running for positive (resp. negative) times t on the to unstable (resp. stable) manifolds of the persistent torus of diofantine frequency $\omega \in \mathbb{R}^n$.The initial data are $\psi(\varphi, \omega, 0) = \varphi \in \mathbb{T}^n$, $q(\varphi, \omega, 0) = \pi$. The value of $I(\varphi, \omega, 0^+)$ (resp. $I(\varphi, \omega, 0^-)$) is fixed by requiring that the homoclinic trajectory is on the unstable manifold for positive times so that the homoclinic trajectory is possibly discontinuous for $t = 0$ and analytic in $\mathbb{R}^\pm$.
We finally define the splitting vector:

$$\Delta I_j(\varphi, \omega) = I(\varphi, \omega, 0^-) - I(\varphi, \omega, 0^+)$$

and the splitting matrix which is the Jacobian of the splitting vector at the intersection point $\varphi = 0$.
The KAM theorem ensures that the S/U manifolds are analytic in μ for small enough μ. Then we find a recursive algorithm for computing the Taylor expansion of the manifolds in μ:

$$z_\mu(\varphi, \omega, t) = \sum_{k=0}^{\infty} z^k(\varphi, \omega, t).$$

To do so, again following [CG], we introduce a suitable generalization of the improper integration we call it the operator $\Im^t$. This definitions are essentially taken from [G1] and only slightly modified in order to deal with non trigonometric perturbations.

In **Section 1.2** we give some definitions of trees, labeled trees and their symmetry groups. We then define admissible trees, which are a set of labeled trees whose labels satisfy suitable conditions. Finally we define the order of an admissible tree $k > 0$.
Such trees carry quite a few labels (sometimes referred to as "decorations"); they will be used in Chapters 2 and 4 to prove cancellations in the perturbation series of the S/U manifolds. The decorations are necessary to infer the cancellations directly from the trees.
In particular, following [G1] we consider special end-nodes, called fruits, which carry a different set of labels from the ordinary nodes, called free nodes. Such distinction is

useful to evidence the holomorphic parts of the homoclinic trajectory. We call the set of admissible trees $\mathcal{T}$ *trees with fruits* and call the subset of $\mathcal{T}$ of trees without fruits $\mathcal{A}$.

In **Chapter 2** we define linear operators on fruitless trees $\mathcal{A}$, called the tree *values*, which set the homoclinic trajectories, splitting vector and splitting matrix, in correspondence with particular linear combinations of trees.
Consequently the tree values are appropriate (generally non analytic) functions of time and of the initial data (φ, ω). We then define suitable linear combinations of trees of order k whose values are in correspondence with the order k term in the expansion of the homoclinic trajectory or of the splitting vector...

We repeat the same scheme on the trees with fruits $\mathcal{T}$, defining "holomorphic tree values"; again such values set the homoclinic trajectories, splitting vector and splitting matrix, in correspondence with particular linear combinations of trees with fruits.

The "holomorphic tree values" are called so as the value of all fruitless trees $\mathcal{A}$ is a real analytic function in t.
The presence of the fruits generates the possibly non analytic terms which are responsible for the complexity of the problem of evaluating the splitting determinant.

We are mainly interested in cancellations for the splitting vector and for the splitting matrix. We view such cancellations on the trees by setting two trees to be equivalent if they have the same value.

In **Chapter 3** we define trees with prefixed total frequency $\nu \in \mathbb{Z}^n$, $A(\nu)$ where $A \in \mathcal{A}$ and their values.

Setting appropriate (non minimal) hypothesis on the function $f(\psi, q)$, we provide bounds for the contribution to the splitting matrix of a tree $A(\nu)$ of order k.

Given $a \geq 0$ and $d < \pi/2$, consider the domain:

$$C(a,d) = \{t \in \mathbb{C} : |\mathrm{Re}t| \leq a\,,\ |\mathrm{Im}t| < d\} \cup \{t \in \mathbb{C} : |\mathrm{Re}t| > a\,,\ |\mathrm{Im}t| < 2\pi\};$$

we consider perturbing functions $f(\psi, q)$ such that:
1) $f(\psi_0(t), q_0(t))$ is analytic inside a domain $C(a, D)$ and has poles on the border.
2) There exists $p \geq 0$ such that:

$$\max_{t \in C(2a, D-\sqrt{\varepsilon})} |f(\psi_0(t), q_0(t))| \leq C\sqrt{\varepsilon}^{-p}. \tag{0.7}$$

For such systems we prove that, for $\mu \leq \varepsilon^P$, the contribution to the splitting matrix of a tree $A(\nu)$ of order k is bounded from above by

$$(k!)^{c_1}(C\frac{\mu}{\varepsilon^{c_2}})^k[e^{-D|\omega\cdot\nu|/\sqrt{\varepsilon}}],$$

with C, c_1 and c_2 are appropriate constants not depending on the tree. P depends on the meromorphic properties of $f(\psi_0(t), q_0(t))$.

In **Chapter 4** we use the bounds of Chapter 3 and the formalism of Chapter 2 to prove exponentially small bounds on the splitting determinant. We follow the techniques proposed in [GGM2] which we have generalized and, we hope, simplified.
In **Section 4.1** we consider the completely anisochronous systems treated in the previous Chapters. In Subsection 4.1.2 we prove that the splitting vector is a Lagrangian manifold generated by a function $S(\varphi)$ called the *generating function of the splitting*. Subsection 4.1.3 contains some technical identities on $\mathcal{A}_1$. Finally, in Subsection 4.1.4, we prove that the splitting matrix satisfies two formal linear non homogeneous equations which ensure the exponential smallness of the splitting determinant.
In **Section4.2** we consider Hamiltonian (*) with $F(q) = \cos q - 1$ and show that we can repeat the procedure proposed in the preceding section and prove the same exponentially small upper bounds.
Finally we discuss (non optimal) exponentially small upper bounds for the splitting determinant of Hamiltonian (*) and prove Theorem 0.6.

In **Chapter 5** we give an alternative method for computing the upper bounds on the splitting determinant for the completely anisochronous case. Following [BB1], we construct recursively a transformation $\vartheta : \mathbb{T}^n_s \ni \varphi \to \alpha \in \mathbb{T}^n_s$ such that in the induced symplectic coordinates the generating function of the splitting (which we prove is $S \circ \vartheta$) is the integral $\Im$ of an analytic function $F(\alpha, t)$ plus a remainder of order η^K with $K = O(\varepsilon^{-b})$ for an appropriate b depending on the number of fast variables. This implies that the splitting determinant, i.e. the determinant of the Hessian of S, is $O_\varepsilon(\infty)$. So this Chapter provides a possibly simpler proof of the upper bounds on the splitting determinant. Moreover the existence of ϑ implies a stronger condition, which is useful to prove fast diffusion (see [BB2]). For each $\alpha \in \mathbb{T}^n_s$ the Hessian matrix of $S \circ \vartheta$ has the following block structure:

$$M(\alpha) = \left| \begin{array}{c|c} M_F & N_F \\ \hline N_F^t & M_S \end{array} \right|$$

where M_F is an $m \times m$ matrix whose entries are $O_\varepsilon(\varepsilon^\infty)$, N_F is a $n - m \times m$ matrix whose entries are $O_\varepsilon(\varepsilon^\infty)$ and M_S contains terms which are polynomial in $\varepsilon, \varepsilon^{-1}$.

In **Chapter 6** we find lower bounds on the splitting determinant and on the eigenvalues of the splitting matrix, for systems with one fast frequency. This can be done independently by using the results of Chapter 4 or of Chapter 5.

First we compute the Melnikov integral for perturbations $f(\psi, q)$ satisfying the Condition 0.7 with $F(q) = \cos q - 1$; then we use the upper bounds proved in Chapter 3, restricted to systems with one fast frequency, to infer that the Melnikov integral dominates on the higher order remainder if $\mu \leq \varepsilon^P$. We obtain Theorem 0.9 for the pendulum (i.e. for $F(q) = \cos q - 1$).

In **Section 6.2** we consider systems with three degrees of freedom and adapt the

techniques of [Ge] and [GGM4] to prove better bounds on P (which depends on the poles of the function $f(\psi_0(t), q_0(t))$.
Finally in **Section 6.3** we apply our results to D'Alembert- like Hamiltonians thus obtaining Theorem 0.11.

In **Chapter 7** we generalize the dependence of the q variable of the (μ)-unperturbed pendulum. We can consider the full Hamiltonian (*) with conditions 0.7.

We find non perturbative conditions on F such that one can "shadow" the procedure used in the preceding chapters and prove lower and upper bounds on the splitting determinant (we show the procedure explicitly on an example). The conditions on F will be quite technical but the fact that they require no closeness conditions with the pendulum is, possibly, interesting.

In **Chapter 8** we prove the existence of heteroclinic chains and we sketch the procedure for proving that such chains are transition chains.

The **Appendices** contain particularly technical proofs and some notions and definitions which are useful in the thesis.

In **Appendix A.1** we give examples of functions with essential singularities which satisfy the bounds (0.7). Moreover we prove that the only entire functions $f(q)$ satisfying (0.7) are trigonometric polynomials.

In**Appendix A.2** we provide some computations on trees, useful in Chapter 3.

In **Appendix A.3** we provide some basic notions on lattices in $\mathbb{Z}^n$.

In **Appendix A.4** we prove the Normal Form Theorem needed to solve the "gap bridging problem".

In **Appendix A.5** we report a proof (taken from [GGM3]) of the convergence of a KAM theorem for the D'Alambert-like Hamiltonian of Chapter 6.

In **Appendix A.6** we give the complete proof of Theorem 0.6 extending the proof of Chapter 4 to general analytic functions $f(\psi, q)$.

In **Appendices A.7- A.8** we review some cancellations on trees, which are not strictly needed in the thesis but which we find nonetheless interesting.

Acknowledgments

I wish to thank prof. Luigi Chierchia who introduced me to the subject of Arnold diffusion and advised me in the writing of the thesis with helpful discussions and suggestions.
I thank as well my fellow PHD students, in particular Domenico Finco and Riccardo Adami; finally I thank prof. Claudio Procesi who gave me useful insights on compact Riemann surfaces.

Chapter 1

Preliminaries

1.1 Whiskered KAM tori for anisochronous Hamiltonian systems

We discuss a completely anisochronous version of Hamiltonian (*) and present a brief review of known results on the problem of homoclinic splitting.

In Subsection 1.1.1 we will first state a classical KAM Theorem for partially hyperbolic systems (see [CG]) which ensures the existence of unstable tori and of their local S/U manifolds and then prove the existence of functions

$$I^{+}_{\mu}(\psi, q, \omega)\ ,\ \ I^{-}_{\mu}(\psi, q, \omega) \tag{1.1}$$

that parameterize respectively the unstable and stable manifolds for all $\psi \in \mathbb{T}^n$ *and* $q \in (-\pi+\delta, \pi-\delta)$*. These are well known results which can be found in most of the references so we will give no proofs of the KAM theorem.*

In Subsection 1.1.2 we discuss the perturbative construction of the manifolds (1.1), by studying the trajectories that are asymptotically quasi-periodic for $t \to \pm\infty$*. These are known results as well, we will briefly report the proofs as they will be useful in the following sections.*

Consider the model Hamiltonian:

$$\frac{(I, A(\varepsilon)I)}{2} + \frac{p^2}{2} + \varepsilon(\cos(q) - 1) + \mu f(\psi, q) \tag{1.2}$$

the pairs $I \in \mathbb{R}^n, \psi \in \mathbb{T}^n$ and $p \in \mathbb{R}, q \in \mathbb{T}$ are conjugate action-angle coordinates, ε, μ are small parameters. For the moment we will consider this parameters as independent and finally prove that we can take $|\mu| < \varepsilon^P$ for some appropriate positive P.

As said in the introduction A is a diagonal matrix, whose eigenvalues $a_i \leq O_\varepsilon(1)$. For $\varepsilon \neq 0$ the matrix A is positive definite, and for $\varepsilon = 0$ it can have some zero eigenvalues. We have in mind a matrix with eigenvalues of the type $a_j(\varepsilon) = \varepsilon^{\alpha_j}$ with $0 \leq \alpha_j \leq 1$

for $j=1,\dots,n$. Some of the α_j will be zero; in particular we set $\alpha_1,\dots,\alpha_m=0$ for $0\le m\le n$.
We will consider the system at energy E of order $E=O_\varepsilon(1)$, $\varepsilon\neq 0$ is a fixed parameter, and we will construct a perturbation theory in μ.
The system (1.2) is integrable for $\varepsilon\neq 0$, $\mu=0$. It represents a list of n uncoupled rotators and a pendulum. We will denote the frequency of the rotators (which determines the initial data $I(0)$) by ω so that:

$$I(t)=I(0)=A^{-1}\omega\,,\quad \psi(t)=\psi(0)+\omega t.$$

The initial data are chosen in an appropriate domain D_m so that there are at least two characteristic orders of magnitude for the frequencies of the unperturbed system. Given $0\le\alpha\le\frac{1}{2}$ and $\delta\in\mathbb{R}^n$ such that:

$$\alpha\ge\max_{j=1,\dots,n}(\alpha_j-\frac{1}{2})\,,\quad \delta_1,\dots,\delta_m=0\,,\quad \frac{1}{2}+\alpha-\alpha_j\le\delta_j\le 1-\alpha_j\,,\ \text{for}\ \ j=m+1,\dots,n$$

and there exists $i\in\{m+1,\dots,n\}$ such that $\delta_i=1+\alpha-\alpha_i$, we consider the domain:

$$D_m(\alpha,\delta):=\begin{array}{l}\{I\ :\ \ I\cdot AI=2E\,,\quad r\varepsilon^{\delta_j}<|I_j|<R\varepsilon^{\delta_j}\,,\\ \ \text{for all}\ \ i=0,\dots,n\ \ \text{and for some}\ R,r=O_\varepsilon(1)\}.\end{array}$$

This implies that the corresponding frequencies are in a domain

$$\Omega\equiv\{\omega:\textstyle\sum_{i=1}^n\omega_i^2/a_i=2E\,,\ \ \omega=(\omega_1,\varepsilon^{\frac{1}{2}+\alpha}\omega_2)\ \text{with}\ |\omega_i|\ge\varepsilon\,,\ \ \omega_1\in\mathbb{R}^m\ ;$$

$$r<|\omega_{1,i}|<R\ \ \text{and}\ \ r<|\omega_2|<R\ \text{for some}\ R,r=O_\varepsilon(1)\}.$$

Notice that, for $n-m\ge 2$ not all the components of ω_2 are necessarily of order one in ε.
There are at least three characteristic time scales $O_\varepsilon(1)$, $O_\varepsilon(\varepsilon^{\frac{1}{2}+\alpha})$ and $\sqrt{\varepsilon}$ which is the Lyapunov exponent of the unperturbed pendulum.
We will call $\psi_1,\cdots,\psi_m$ the fast variables and we will sometimes denote them as $\psi_F\in\mathbb{T}^m$. Conversely we will call $\psi_{m+1},\cdots,\psi_n$ slow variables $\psi_S\in\mathbb{T}^{n-m}$.
Notice that we can consider indifferently systems that are degenerate or non-degenerate for $\varepsilon=0$. The only (obvious) restriction is that if the system is degenerate in some of its action variables, for $\varepsilon=0$, then these are necessarily slow variables with characteristic frequency $\omega_{2,j}\le a_{m+j}$.

The perturbating function $f(\psi,q)$ is a trigonometric polynomial of degree N in the rotators ψ, it is analytic in q in a domain $\mathbb{T}\times i(-R,R)$, for simplicity we take it even and with zero mean value; this means that:

$$f(\psi,q)=\sum_{n,\nu\in\mathbb{Z}^{n+2},|\nu|\le N}f_{n,\nu}e^{i(nq+\nu\cdot\psi)}$$

where $f_{0,0}=0$, $f_{n,\nu}=f_{-n,-\nu}$ and $|f_{n,\nu}|\le C_\nu e^{-R|n|}\le Ce^{-R|n|}$.

These conditions are sufficient to ensure the convergence of the local KAM theorem and to provide exponentially small upper bounds on m eigenvalues of the splitting matrix. In the case of one fast frequency ($m = 1$)we will restrict our attention to perturbating functions $f(q,\psi)$ that are rational in e^{iq} (with no singularity on the unit circle). In this case we will give lower bounds on the eigenvalues of the splitting matrix and finally consider the problem of heteroclinic intersections.

For each $\omega \in \mathbb{R}^n$ the unperturbed system has an unstable fixed torus :

$$p(t) = q(t) = 0\,, \quad I(t) = I(0) = A^{-1}\omega, \qquad \psi(t) = \psi(0) + \omega t.$$

The stable and unstable manifolds of such tori coincide and can be expressed as graphs on the angles:

$$p = \pm\sqrt{2\varepsilon}\sqrt{1-\cos q}\,, \qquad q(t) = 4\text{arctan}\, e^{\pm\sqrt{\varepsilon}t}\,;$$

$$I = A^{-1}\omega\,, \qquad\qquad \psi(t) = \psi(0) + \omega t.$$

It is known that for diofantine values of the frequencies the unstable tori, with their S/U manifolds, survive the onset of a small perturbation (and so does the property of being graphs over the angles) but generally the two manifolds will no longer coincide and one should expect a transversal intersection; evaluating the "intersection angle" will be the purpose of the following sections.

1.1.1 The KAM construction, definitions of splitting vector and splitting matrix.

Definition 1.1. *given any $\gamma \in \mathbb{R}$, $0 < \gamma \le O(\varepsilon^{\frac{1}{2}+\alpha})$ and a fixed $\tau > n-1$, we define the set*

$$\Omega_\gamma \equiv \{\omega \in \Omega : |\omega\cdot n| > \frac{\gamma}{|n|^\tau} \quad \forall n \in \mathbb{Z}^n/\{0\}\}$$

of γ,τ diofantine vectors in Ω. Now we consider

$$\Omega^*_\gamma \equiv \Omega_\gamma \times (-\frac{1}{2},\frac{1}{2})$$

and for all $(\omega,\rho) \in \Omega^*_\gamma$ we set $\omega_\rho = (1+\rho)\omega$.
For all $(\omega,\rho) \in \Omega^*_\gamma$ an for all $n \in \mathbb{Z}^n/\{0\}$ $|\omega_\rho \cdot n| > \frac{\gamma}{2|n|^\tau}$.
$\omega \in \Omega_\gamma$ implies that ω_1 and ω_2 are diofantine as well; we will call τ_F and τ_S their exponents.

Theorem 1.2. *There exists*[1] *$\mu_0(\varepsilon,\gamma)$ such that if $|\mu| \le \mu_0$ and if $(\omega,\rho) \in \Omega^*_\gamma$, there*

[1]in the Appendix A.4, we will specify $\mu_0(\varepsilon,\gamma)$; generally speaking, if we consider only those $\omega \in \Omega_\varepsilon(\gamma)$ which are as well in $B_s : \{\omega : |\omega\cdot n| > CE_2 \quad \forall |n| < s\}$, one obtains, by combining classical perturbation theory and KAM techniques, that $\mu_0(\varepsilon,\varepsilon^m) = \varepsilon^L$ with $L = \max(2, \frac{2(m+1)}{s})$. This estimates can be much refined by using the existence of separate time scales, see for instance Theorem 1.4 of [GGM4]. In that article the authors consider a system with three degrees of freedom (and three time scales $1, \sqrt{\varepsilon}, \varepsilon^{\frac{1}{2}+\alpha}$); they obtain $\mu_0(\varepsilon,\gamma) = \varepsilon^3$ for all $\gamma < e^{-\frac{1}{\varepsilon^{\frac{1}{2}+\alpha}}}$.

exists one and only one n-dimensional H_μ-invariant torus $T_\mu(\omega,\rho)$ whose Hamiltonian flow is analytically conjugated to the flow $\mathbb{T}^n \ni \vartheta \to \vartheta + \omega_\rho t$.
The torus $T(\omega,\rho)$ admits local stable/unstable manifolds $W^{\pm}_{\mu,loc}(\omega,\rho)$, described by a function:

$$\mathbb{T}^n \times B^2_{2r} \times B^1_{\mu_0} \times \Omega_\gamma \ni (\vartheta,(x^+,x^-),\mu,\omega_\rho) \to \xi_\mu(\vartheta,x^+,x^-,\omega_\rho) \tag{1.3}$$

C^3 in all its arguments. For fixed (ω,ρ) the function is analytic on $\mathbb{T}^n_k \times \hat{B}^2_{2r} \times \hat{B}^1_{\mu_0}$;where k is some ε-independent constant and $r = O_\varepsilon(\varepsilon^{\frac{1}{4}})$. In terms of the function (1.3) one has:

$$\begin{array}{l} T_\mu(\omega,\rho) \equiv \{\xi_\mu(\vartheta,0,0,\omega_\rho)\vartheta \in \mathbb{T}^n\} \\ W^+_{\mu,loc}(\omega,\rho) \equiv \{\xi_\mu(\vartheta,x^+,0,\omega_\rho)\vartheta \in \mathbb{T}^n,\ |x^+| < 2r\} \\ W^-_{\mu,loc}(\omega,\rho) \equiv \{\xi_\mu(\vartheta,0,x^-,\omega_\rho)\vartheta \in \mathbb{T}^n,\ |x^-| < 2r\} \end{array} \tag{1.4}$$

on the local stable/unstable manifolds the flow is:

$$\Phi^{\mp t}_\mu \xi^{\pm}_\mu(\vartheta,x^+,x^-,\omega_\rho) = \xi_\mu(\vartheta + \omega_\rho t, x^+ e^{-\Lambda t}, x^- e^{\Lambda t}, \omega_\rho)$$

where the Lyapunov exponent $\Lambda \equiv \Lambda_\mu(\vartheta,x^+,x^-,\omega_\rho)$ has the same regularity as ξ.

The proof of this theorem can be found, for example in [CG].

We have introduced the variable ρ in order to fix the energy of the perturbed system equal to[2] E (for all $\omega \in \Omega_\gamma$).

Proposition 1.3. *There exists a function $\rho = \rho(\mu,\omega)$, analytic in μ, such that for $E \in [E_1,E_2]$:*

$$H_\mu(\xi_\mu(0,0,0,\omega_{\rho(\mu,\omega)}),\varepsilon,\mu) = E.$$

Proof. As

$$H_0(\xi_\mu(0,0,0,\omega_0),\varepsilon,0) = \frac{1}{2}\omega \cdot A^{-1}\omega = E$$

and

$$\partial_\rho H_0(\xi_\mu(0,0,0,\omega_\rho),\varepsilon,0)|_{\rho=0} = \left[\partial_\rho \frac{1}{2}\omega \cdot A^{-1}\omega(1+\rho)^2\right]_{\rho=0} = 2E > 0$$

we can apply the implicit function theorem and obtain $\omega_\rho(\mu,\omega)$. □

Notice that Theorem 1.1 is local in the hyperbolic variables x^+, x^- (it holds in a domain $|x^\pm| \le 2r = O(\varepsilon^{\frac{1}{4}})$),to find extended stable/unstable manifolds we "follow the flow" i.e. we apply the Hamiltonian flow $\Phi^{\mp T}_\mu$ to the stable/unstable local manifolds, where T is sufficiently large (positive for the unstable manifold and negative for the stable one).

[2] The final goal is to find heteroclinic intersections on the fixed energy surface, and so "Arnold diffusion", but in the following sections we will discuss only homoclinic intersections and so we will drop the parameter ρ

The time $T = \varepsilon^{-\frac{1}{2}} \log \varepsilon^{-1}$ is such that given a point $\bar{z} \in \mathbb{T} \times \mathbb{R}$ inside the local unstable manifold of the pendulum, $\Phi^T_{\mu=0}(\bar{z}) = (2\sqrt{\varepsilon}, \pi)$,

Now the extended stable unstable manifolds are:

$$\xi^{\pm}_{\mu}(\vartheta, x^+, x^-, \omega) = \Phi^{\mp T}_{\mu} \xi^{\pm}_{\mu}(\vartheta, x^+, x^-, \omega).$$

And by the choice of T:

$$\pi_q\{\xi_\mu(\vartheta, x^+, 0, \omega_\rho) \ : \ |x^+| < 2r\} \supset [-\pi, 0) \ \forall \vartheta, \tag{1.5}$$

$$\pi_q\{\xi_\mu(\vartheta, 0, x^-, \omega_\rho) \ : \ |x^-| < 2r\} \supset (0, \pi] \ \forall \vartheta. \tag{1.6}$$

Proposition 1.4. *The branches of the stable/unstable manifolds can be represented as graphs on the rotator angles, for instance for $p < 0$*[3]*:*

$$\xi^{+}_{\mu}(\vartheta, x^+, 0, \omega) = \psi, I^{+}_{\mu}(\psi, q, \omega), q, p^{+}_{\mu}(\psi, q, \omega), \tag{1.7}$$

$$\xi^{-}_{\mu}(\vartheta, 0, x^-, \omega) = \psi, I^{-}_{\mu}(\psi, q, \omega), \ q, p^{-}_{\mu}(\psi, q, \omega). \tag{1.8}$$

A proof of this Proposition can be found, for example, in [C].

Definition 1.5. *We will study the difference between the stable and unstable manifolds on an hyper-plane transverse to the flow (a Poincaré section). In the following Sections we will use $\psi = \varphi \in \mathbb{T}^n, q = \pi$ and call $I^{\pm}_{\mu}(\varphi, \omega)$ the graphs of the S/U manifolds at the Poincaré section. We will call*

$$\Delta I(\varphi, \omega) = I^{-}(\varphi, \omega) - I^{+}(\varphi, \omega)$$

the splitting vector. We will prove that $\Delta I(\varphi = 0, \omega) = 0$. We will call

$$M : \partial_{\varphi}(I^{+}_{\mu}(\varphi, \omega) - I^{-}_{\mu}(\varphi, \omega)|$$

the splitting matrix and detM the splitting determinant.

It is convenient to re-scale the time and action variables so that the Lyapunov exponent of the unperturbed pendulum is equal to one. Namely we will consider the following Hamiltonian:

$$\frac{(\tilde{I}, A(\varepsilon)\tilde{I})}{2} + \frac{\tilde{p}^2}{2} + (\cos(\tilde{q}) - 1) + \eta f(\tilde{\phi}, \tilde{q}), \tag{1.9}$$

which generates the same Hamilton equations as (1.2), provided that:

$$\begin{aligned} \tilde{I}(t) &= \frac{I(\frac{t}{\sqrt{\varepsilon}})}{\sqrt{\varepsilon}}, \quad \tilde{\psi}(t) = \psi(\frac{t}{\sqrt{\varepsilon}}), \quad \eta = \frac{\mu}{\varepsilon} \\ \tilde{p}(t) &= \frac{p(\frac{t}{\sqrt{\varepsilon}})}{\sqrt{\varepsilon}}, \quad \tilde{q}(t) = q(\frac{t}{\sqrt{\varepsilon}}). \end{aligned} \tag{1.10}$$

We re-scale the domains D_m and Ω consequently so obtaining a rescaled frequency $\tilde{\omega} = (\frac{\omega_1}{\sqrt{\varepsilon}}, \varepsilon^a \omega_2)$. In the following sections we will consider the system after this change of variables, but we will omit the tilde (except in ω).

To retrieve the true size of I we must only remember to multiply by $\sqrt{\varepsilon}$, the inverse for the variable t, to have the correct estimates on the diffusion times.

[3] notice that $p^{+}_{\mu}(\psi, q, \omega)$ is obtained via the energy conservation one we have fixed the sign of p.

1.1.2 Perturbative construction of the homoclinic trajectories

In this Subsection we will use perturbation theory to find the (analytic for $\mu \leq \mu_0$) trajectories on the S/U manifolds of Hamiltonian[4] (1.9)

$$(z_\eta^\pm(\pm|t|,\varphi,\tilde\omega) \equiv \Phi_\eta^{\pm|t|}(I_\eta^\pm(\varphi,\tilde\omega),\varphi,\pi) = \sum_k (\eta)^k z^{k\pm}(t,\varphi,\tilde\omega).$$

The basic ideas, which go back to Poincaré, consist mainly in determining the trajectories on the S/U manifolds by requiring boundedness as $t \to \pm\infty$.
Namely given the Hamilton equations of system (1.9):

$$\begin{aligned} &\dot I_j = -(\eta) f_{\psi_j}(\psi,q)\,, \quad && \dot p = \sin(q) - (\eta) f_q(\psi,q)\,, \\ &\dot\psi_j = a_j I_j, && \dot q = p, \end{aligned} \tag{1.11}$$

an initial datum $\varphi, I_\eta^\pm(\varphi,\tilde\omega), \pi, p_\eta^\pm(\varphi,\tilde\omega)$ is on the stable (unstable) manifold if and only if its flow approaches the invariant torus of frequency $\tilde\omega$ for [5] $t \to \pm\infty$. This requirement is sufficient to determine the initial datum as a power series in η.

Definition 1.6. *To avoid the $\pm$ apex we will set*[6]*:*

$$z_j(t) = \begin{cases} z^+(t) & \text{if } t > 0 \\ z^-(t) & \text{if } t < 0 \end{cases} .$$

Moreover as we will now consider $\tilde\omega$ as fixed we will omit $\tilde\omega$ in the expansion coefficients.

Inserting in the Hamilton equations the convergent power series representation:

$$I(t,\varphi,\eta) = \textstyle\sum_{k=0}^\infty (\eta)^k I^k(t,\varphi)\,, \quad \psi(t,\varphi,\eta) = \sum_{k=0}^\infty (\eta)^k \psi^k(t,\varphi)\,,$$

$$p(t,\varphi,\eta) = \textstyle\sum_{k=0}^\infty (\eta)^k p^k(t,\varphi)\,, \quad q(t,\varphi,\eta) = q^0(t) + \sum_{k=1}^\infty (\eta)^k \psi_0^k(t,\varphi)$$

we obtain, for $k > 0$, the hierarchy of linear non-homogeneous equations[7]:

$$\begin{aligned} &\dot I_j^k = F_j^k(\{\psi_i^h\}_{\substack{i=0,\ldots,n\\ h<k}})\,, \quad \dot\psi_j^k = a_j I_j^k\,, \quad \text{for } j = 1,\ldots,n\,; \\ &\dot p^k = \cos q^0 \psi_0^k + F_0^k(\{\psi_i^h\}_{\substack{i=0,\ldots,n\\ h<k}})\,, \quad \dot\psi_0^k = p^k\,, \end{aligned} \tag{1.12}$$

[4]Notice that the apex k on the functions I,ψ represents the order in the expansion in η NOT an exponent. To avoid confusio, when we need to exponentiate we always set the argument in parentheses.

[5]and so tends, as $t \to \pm\infty$, to a quasi-periodic function with frequency $\tilde\omega$ at an exponential rate given by the Lyapunov exponent

[6]note that the functions so defined are possibly non continuous in $t = 0$ as each boundedness condition ($t \to \pm\infty$ determines uniquely the value in $t = 0$

[7]when it is not strictly necessary we will omit the prefixed initial data of the angles $\varphi = \psi_1(0),\cdots,\psi_n(0)$; $\psi_0(0) = \pi$

where the functions F_i^k are defined as follows. Set:

$$[\cdot]_k = \frac{1}{k!}\frac{d^k}{d\eta^k}(\,\cdot\,)|_{\eta=0},$$

we have for $j = 0,\dots,n$

$$F_j^k(t) = -[f_j(\sum_{h=1}^{k-1}(\eta)^h\psi^h(t))]_{k-1} + \delta_{j0}[\sin(\sum_{h=1}^{k-1}(\eta)^h\psi_0^h(t))]_k,$$

where δ_{ji} denotes the Kronecker delta and $\psi^h(t)$ is the vector $\psi_0^h(t),\dots,\psi_n^h(t)$. For $k=0$ we obtain the unperturbed homoclinic trajectory:

$$z^0(t) = (\varphi + \frac{\omega}{\sqrt{\varepsilon}}t\,, A^{-1}\frac{\omega}{\sqrt{\varepsilon}}\,, q^0(t)\,, p^0(t)),$$

$(q^0(t), p^0(t))$ is the lower branch of the pendulum separatrix starting at $q=\pi$:

$$\begin{aligned} q^0(t) &= 4\arctan e^{-t}\,, \quad p^0(t) = -\frac{2}{\cosh t}\,, \\ \cos q^0 &= 1 - \frac{2}{(\cosh t)^2}\,, \quad \sin q^0 = \frac{2\sinh t}{(\cosh t)^2}. \end{aligned} \tag{1.13}$$

For $k > 0$ we have a linear non-homogeneous ODE that we solve by variation of constants.
The fundamental solution of the linearized pendulum equation is given by:

$$W(t) = \begin{pmatrix} (1 - \frac{\omega^0(t)}{4}\frac{\sinh t}{(\cosh t)^2} & -\frac{\sinh t}{(\cosh t)^2} \\ \frac{\omega^0(t)}{4} & \frac{1}{\cosh t} \end{pmatrix}$$

$$\omega^0(t) = 2\frac{t + \sinh 2t}{\cosh t},$$

so that integrating equations (1.12) we have:

$$\begin{aligned} p^k(t) &= w_{11}(t)p^k(0^{\pm}) + w_{11}(t)\int_0^t w_{22}(\tau)F_0^k(\tau)d\tau - w_{12}(t)\int_0^t w_{21}(\tau)F_0^k(\tau)d\tau, \\ \psi_0^k(t) &= w_{21}(t)p^k(0^{\pm}) + w_{21}(t)\int_0^t w_{22}(\tau)F_0^k(\tau)d\tau - w_{22}(t)\int_0^t w_{21}(\tau)F_0^k(\tau)d\tau \\ I_j^k(t) &= I^k(0^{\pm}) + \int_0^t F_j^k(\tau)d\tau \\ \psi_j^k(t) &= a_j(I^k(0^{\pm})t + \int_0^t (t-\tau)F_j^k(\tau)d\tau), \end{aligned} \tag{1.14}$$

the functions w_{ij} are the entries of $W(t)$ and we have used the fact that, for $k>0$, $\psi_i^k=0$ for all $i=0,n$.

Remark 1.7. *This procedure can be repeated for any generalized pendulum; see Section 7.2 for the construction of the Wronskian matrix. One obtains a matrix $W'(t)$ having the same qualitative properties a W.*

To give meaning to the $t\to\pm\infty$ limit ,following [CG], in the following Subsection we shall introduce a suitable generalization of the standard improper integration.

1.1.3 Whisker calculus

Let $\mathcal{D}$ be the class of functions f smooth for $t\neq 0$, such that for any $k\geq 0$, there exist $a>0>b$ for which, given $t\in\mathbb{R}$ the function:

$$u\to F_k(u,t)\equiv\int_{\sigma(t)\infty}^{t} e^{-u|t|}f^{(k)}(\tau)d\tau \quad \text{where } \sigma(t)=\text{sign}(t) \tag{1.15}$$

is analytic on the complex domain $\{u\in\mathbb{C}:\Re u>a\}$ and admits an analytic continuation which is meromorphic in $\{u\in\mathbb{C}:\Re u>b\}$ and analytic in a neighborhood of $u=0$. If $f\in\mathcal{D}$ we set $\Im^t=F_0(0,t)$.

Notice that if $\limsup_{t\to\sigma(t)\infty}e^{r|t|}|f^{(k)}(t)|<\infty$ for some $r>0$, then $f\in\mathcal{D}$ and $\Im^t(f)=\int_{\sigma(t)\infty}^t f$.
It is easy to check that $f\equiv t^je^{ct}\in\mathcal{D}$ for any j and any non-zero complex number c. Polynomials are clearly not contained in $\mathcal{D}$. Nevertheless, we extend the operator $\Im$ on

$$\tilde{\mathcal{D}}\oplus \text{ ring of Polynomials in t}$$

by defining $\Im^t\tau^j=\frac{t^{j+1}}{j+1}$.

Now set $\tilde{H}$ to be the largest subset of $\tilde{\mathcal{D}}$ which is closed under product, derivative and integration $\Im$ and

$$\tilde{M}\equiv\{f\in\tilde{H}:\ \pi_P f=0\}$$

where π_P is the natural projection onto polynomials.

On $\tilde{H}$ one can set

$$\Im^t f=\oint\frac{du}{2i\pi u}\int_{\sigma(t)\infty}^{t} e^{-u|\tau|}f(\tau)d\tau, \tag{1.16}$$

where the integration in the u variable is performed on a suitably small $|u|\leq\delta$ circle around $u=0$ It is easily seen that expression (1.22) works as well on polynomials in t the only difference being that

$$\int_{\sigma(t)\infty}^{t} e^{-u|\tau|}f(\tau)d\tau,$$

will no longer be analytic in $u = 0$.
For all $f \in \tilde{H}$, $\Im^t f$ is a primitive of f as:

$$\Im^t f - \Im^s f = \int_s^t f(\tau) d\tau, \tag{1.17}$$

for any $f \in \tilde{H}$ and any s, t such that $\sigma(t) = \sigma(s)$.
Let us consider some interesting subspaces of $\tilde{H}$.

Definition 1.8. *(i) H is the vector space (on $\mathbb{C}$) generated by monomials of the form:*

$$m = \sigma(t)^a \frac{|t|^j}{j!} x^h e^{i(\varphi+\omega t)\cdot\nu} \quad \textit{where } h \in \mathbb{Z}, \ \nu \in \mathbb{Z}^n, \ j \in \mathbb{N},$$
$$x = e^{-|t|}, \ a = 0, 1, \ \sigma(t) = \ \textit{sign}(t). \tag{1.18}$$

(ii) Given two positive constants b and d, $H(b, d)$ is the subset of (couples of) function(s) $f(t)$ that admit a (unique) representation:

$$f(t) = \sum_{j=0}^{k} \frac{|t|^j}{j!} M_j^{\sigma(t)}(x, \varphi + \omega t), \tag{1.19}$$

with $M_j(x, \varphi)$ trigonometric polynomials in φ.
The Fourier coefficients $M_{j\,\nu}(x)$ are all holomorphic in the x-plane in an annulus $0 < |x| < e^{-b}$ and satisfy the following properties.

1) The $M_{j\,\nu}(x)$ have possible singularities outside the disk $|x| < e^{-b}$ and outside the cone $|\arg x| < d$.
2) The $M_{j\,\nu}(x)$ have possible polar singularities at $x = 0$.
If $M_k^{\sigma(t)} \neq 0$ then k is called the t degree of f. In Figure 1.1 we have represented a possible "candy"shaped domain of analyticity for the $M_{\nu\,j}$

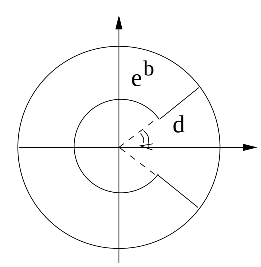


Figure 1.1:

Notice that H is contained in all the spaces $H(b, d)$; moreover if $|t| > b$, $f(t)$ can be represented as an absolutely convergent series of monomials of the type m.

One can easily check that the integration $\Im$ acts on monomials m of the form (1.18) as:

$$\Im^t(m) = \begin{cases} -\sigma^{a+1} x^h e^{i(\psi+\omega t)\cdot\nu} \displaystyle\sum_{p=0}^{j} \frac{|t|^{j-p}}{(j-p)!(h - i\sigma\omega\cdot\nu)^{p+1}} & \text{if } |h| + |\nu| \neq 0 \\ -\dfrac{\sigma^{a+1}|t|^{j+1}}{(j+1)!} & \text{if } |h| + |\nu| = 0 \end{cases} \tag{1.20}$$

This and equation (1.17) show that $\Im^t$ acts on $H(b,d)$ as (1.20) if $|t| > b$ and if $|t| \le b$ as

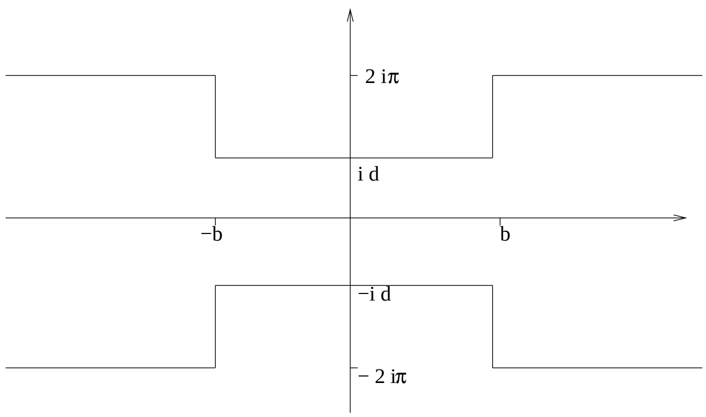
$$\Im^{2\sigma(t)b} + \int_{2\sigma(t)b}^{t} , \tag{1.21}$$

obviously the choice of $2b$ is arbitrary.

On $H(b,d)$ we can extend $\Im^t$ to complex values of t such that $t \in C(b,d)$ where:

$$C(b,d) := \{t \in \mathbb{C} : |\operatorname{Im} t| \le d\,,\ |\operatorname{Re} t| \le b\} \cup \{t \in \mathbb{C} : |\operatorname{Im} t| \le 2\pi\,,\ |\operatorname{Re} t| > b\},$$

is the domain in Figure 1.1 in the t variables.



To extend $\Im^t$ simply consider the definition 1.16, for $t \in C(b,d)$ so that if $t = t_1 + is$, with $t_1, s \in \mathbb{R}$, the integral is performed on the line $\operatorname{Im}\tau = s$.

$$\Im^t f = \oint \frac{du}{2i\pi u} \int_{\sigma(t)\infty + is}^{t} e^{-\sigma(\tau)u\tau} f(\tau)d\tau, \tag{1.22}$$

where $\sigma(t) = \operatorname{sign}(\operatorname{Re} t)$.
This definition does not modify the expressions (1.18) (one simply sets $t = t_1 + is$, $x = x_1 e^{i\sigma(t_1)s}$). The following property holds:

Lemma 1.9. *$H(b,d)$ is closed under the application of $\Im^t$.*

Proof. Let us expand f as in (1.19) and consider the single term

$$t^j e^{i\omega\cdot\nu t} M_{\nu j}^{\sigma(t)}(x), \tag{1.23}$$

moreover, if $|t| \le b$, we divide $\Im^t$ as in (1.21). For $|t| > b$ we can expand $M_{\nu j}^{\sigma(t)}(x)$ in convergent power series of x and apply (1.18). The radius of convergence is the same and the degree of the pole in zero is the same. Moreover

$$\int_{2\sigma(t)b+is}^{t} \tau^j e^{i\omega\cdot\nu\tau} M_{\nu j}^{\sigma(t)}(x')$$

is well defined and finite provided that $s < d$ and $|\,\mathrm{Re}\, t| < b$.
Finally as f is a finite combination of terms like (1.23) so $\Im^t f$ is still in $H(b,d)$. □

Definition 1.10. *We define as $\tilde{H}_0$ the subspace of $\tilde{H}$ of functions that can be extended to an analytic function in some strip around the real axis.*
$H_0(b,d)$ is the subspace of $H(b,d)$ of functions that can be extended to analytic functions in $C(b,d)$.

Notice that f is in $H_0(b,d)$ if it is in $H(b,d)$ and $f^+(t)$ and f^- join analytically at $t = 0$.

Remark 1.11. *Notice that if $f \in H_0(b,d)$ then generally $\Im f \notin H_0(b,d)$ and has a discontinuity in $t = 0$. For instance if $f \in \mathcal{L}_1$ is even, then:*

$$\Im(f) := \Im^{0^-} - \Im^{0^+} f = \int_{-\infty}^{\infty} f \neq 0.$$

We can construct operators which preserve $H_0(b,d)$; let $\Im = \Im^{0^-} - \Im^{0^+}$ and

$$\Im_+^t = \begin{cases} \Im^t & \text{if } t \geq 0 \\ \Im^t - \Im & \text{if } t < 0\,, \end{cases}$$

$$\Im_-^t = \begin{cases} \Im^t & \text{if } t \leq 0 \\ \Im^t + \Im & \text{if } t > 0\,. \end{cases}$$

The operator

$$\frac{1}{2} \sum_{\rho = \pm 1} \Im_\rho^t = \Im^t - \frac{1}{2}\sigma(t)\Im \tag{1.24}$$

preserves the analyticity.

Now let us cite two important properties of $H_0(b,d)$, whose proofs are taken from [G1].

Lemma 1.12. *In $H_0(b,d)$ we have the following shift of contour formulas:*
$\forall f \in H_0(b,d)$ and for all $d > s \in \mathbb{R}$

$$(i)\ \ \Im f(\tau) = \Im f(\tau + is)\,,$$

$$(ii)\ \ \sum_{\rho=\pm 1} \Im_\rho^{t+is} f(\tau) = \oint \frac{dR}{2i\pi R} \sum_{\rho = \pm 1} \int_{\rho\infty}^{t} e^{-R\sigma(\tau)(\tau+is)} f(\tau + is) d\tau\,.$$

The integrals in the right hand side have to be considered to be the analytic continuation on R from R positive and large.

Proof. (i) If f is a polynomial one can check by direct calculation that the relation is $0=0$.

For R large and positive

$$\int_{-\infty}^{0} e^{R\tau} f(\tau)d\tau + \int_{\infty}^{0} e^{-R\tau} f(\tau)d\tau \tag{1.25}$$

is well defined and can be shifted by is for all $s<d$. It is equal to

$$\int_{-\infty}^{0} e^{R(\tau+is)} f(\tau+is)d\tau + \int_{\infty}^{0} e^{-R\tau-is} f(\tau+is)d\tau - i\int_{0}^{s}(e^{-iR\ tau} - e^{iR\tau})f(i\tau)d\tau.$$

This differs from

$$\int_{-\infty}^{0} e^{R\tau} f(\tau+is)d\tau + \int_{\infty}^{0} e^{-R\tau} f(\tau+is)d\tau \tag{1.26}$$

precisely by:

$$(e^{iRs}-1)\int_{-\infty}^{0} e^{R(\tau)} f(\tau+is)d\tau + (e^{iRs}-1)\int_{\infty}^{0} e^{-R\tau} f(\tau+is)d\tau - \tag{1.27}$$

$$i\int_{0}^{s}(e^{-iR\tau} - e^{iR\tau})f(i\tau)d\tau.$$

This implies (i) by taking the residues at $R=0$.

We consider only f with no polynomial component, so the t-integrals are all analytic in R for $R=0$. This implies that the residue of (1.27) is zero.

(ii) The two sides differ by the residue at $R=0$ of

$$-i\int_{0}^{s}(e^{-iR\tau} - e^{iR\tau})f(i\tau)d\tau$$

which vanishes. □

1.1.4 Analytic expansions for the whiskers

Let us consider some (probably non minimal) conditions on the perturbing function $f(q,\psi)$. Namely we will consider only those functions $f(q,\psi)$ which are trigonometric polynomials in ψ and such that $f(q(t),\psi(t)) \in H_0(b,d)$ for some b,d.

Remember that $q(t)$ and $\psi(t)$ are the motions on the unperturbed separatrix after the change of variables (1.10).
The trajectory $q(t)$ can be analytically extended to $t \in \mathbb{R}\times(-\pi/2, \pi/2)$, in Figure 1.2 we show some $q(t+id)$, $t, d \in \mathbb{R}$ for various values of $|d| \leq \pi/2$.

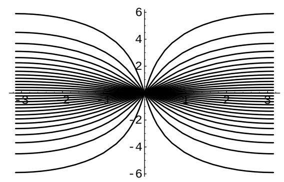


Figure 1.2:

We have in mind functions f such that

$$f(q,\psi) = \sum_{\nu<N} e^{i\nu\cdot\psi} F_\nu(e^{iq})$$

and there exist $\alpha < 1 < \beta$ and c such that the $F_\nu(y)$ are all analytic in the domain

$$\alpha < |y| < \beta, \quad |y-1| < c.$$

Given f we define

$$\begin{array}{ll} a = & \inf_{b\in\mathbb{R}^+} \{b : F_\nu(e^{iq(t)}) \in H_0(b,d) \text{ for some } d \text{ and } \forall\nu\} \\ D = & \sup_{d\in\mathbb{R}^+} \{d : F_\nu(e^{iq(t)}) \in H(a,d) \ \ \forall\nu\}. \end{array} \tag{1.28}$$

In Figure1.3 we have represented in light-blue the region $\frac{1}{2} < |y| < 2$ and in dark blue the image through $e^{iq(t)}$ of the region $C(3, \pi/16)$. if the $F_\nu(y)$ have no poles inside this region and have a pole on the border both of the circle around $y = 1$ and on the "annulus" around S^1 then $a = 3$ and $D = \pi/16$.

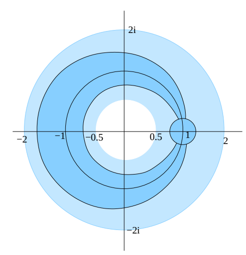


Figure 1.3:

We easily see that $z^k(t)$ and hence F^k belong to $\tilde{H}$ for all $k > 0$ so that:

$$I^k(t) - \Im^t F^k = I^k(0^{\sigma(t)}) - \Im^{0^{\sigma(t)}} F^k. \tag{1.29}$$

Remark 1.13. *(i) The quasi-periodic average,*

$$\lim_{T\to\pm\infty} \frac{1}{|T|}\int_0^T f^\pm = (< f^- >, < f^+ >) \equiv < f >,$$

of an asymptotically quasi-periodic (couple of) function(s) $f \equiv \psi + g$, where g is exponentially decreasing, coincides with the quasi-periodic average of ψ;
(ii) if f is asymptotically quasi-periodic with $< f >= 0$ then both f and $\Im^t f$ belong to $\tilde{H}$ and $< \Im f >= 0$ as well.

Thus taking the quasi-periodic average in the first line of (1.11), one sees that both F^k and I^k , which are asymptotically quasi-periodic, have vanishing quasi-periodic average. Therefore taking the quasi periodic average in (1.29) we obtain $I^k_j(0) = \Im^0 F^k_j$ and so:

$$I^k_j(t) = \Im^t F^k_j \quad \psi^k_j = a_j(\Im^t[\Im^\tau F^k_j] - \Im^0[\Im^\tau F^k_j]).$$

With similar arguments (and keeping in mind the asymptotics of $W(t)$, we find that

$$p^k(0^\pm) = \int_{\pm\infty}^{0} w_{22}(\tau) F^k_0(\tau) d\tau.$$

Finally we summarize the equations for the stable/unstable manifolds as:

$$I^k_i(t) = \Im^t(F^k_i) \qquad \psi^k_j(t) = a_j O^t_j(F^k_j) \tag{1.30}$$

where $i = 1, n$ and $j = 0, n$ and $a_0 = 1$.
The operators O^t_j are defined in terms of $\Im^t$:

$$O^t_j = Q^t_j + R^{0\,t}_j + R^{1\,t}_j$$

$$Q^t_j(g) = \tfrac{1}{2} \sum_{\rho=\pm} \Im^t_\rho(w_j(t,\tau) g(\tau))$$

$$R^{i\,t}_j(g) = -\tfrac{1}{2} x^{[i]}_j(t) \Im(x^i_j(\tau) g(\tau)) \quad [i] = |i-1|$$

$$w_j(t,\tau) = \sigma(t) x^1_j(t) x^0_j(\tau) - \sigma(\tau) x^0_j(t) x^1_j(\tau))$$

$$x^1_j = \begin{cases} |t| & j \neq 0 \\ \dfrac{|t|x}{x^2+1} - \dfrac{1}{4}(x - x^{-1}) & j = 0 \end{cases}, \qquad x^0_j = \begin{cases} 1 & j \neq 0 \\ \dfrac{2x}{x^2+1} & j = 0\,. \end{cases} \tag{1.31}$$

Notice that x^0_j belongs to $H_0(0, \pi/2)$ x^1_j belongs to $H(0, \pi/2)$ and that $w(t,\tau)$ is in $H_0(0, \pi/2) \times H_0(0, \pi/2)$. By our assumptions F^1_j $j = 0, n$ belongs to $H_0(a, D)$. Thus Lemma 1.11 guarantees that $H(a, D)$ is closed under the application of $\Im^t$ and O^t_j.

Remark 1.14. *If $f(\psi, q)$ is a trigonometric polynomial then F^1_j $j = 0, n$ belongs to $H(0, \pi/2)$ which is closed under the action of $\Im^t$ and O^t_j.*
In the following Section we will work symbolically on I, ψ, so we will not note weather we are working in $H(b, d)$ or in $\tilde{H}$. Then in Chapter 3 , where we estimate the integrals, we will need to keep track of the action of Q_j on $H(b, d)$.

Remark 1.15. *We have expressed the operators O_j in terms of Q_j and R^i_j to keep track of the occurrence of terms not in H_0; actually we start with $f(\varphi + \tilde{\omega} t, q(t))$ and $\cos(q(t))$ which are in H_0, but the operators $R^{0\,t}_j$ produce x^1_j which is clearly not in H_0.*

The following proposition contains some important properties of the operators Q_j all proved in [G1].

Proposition 1.16 (Chierchia). *(i)The operators Q_j and O_j are "symmetric" on $\tilde{H}$:*

$$\Im(f\,Q_j g) = \Im(g\,Q_j f)\,, \quad \Im(f\,O_j g) = \Im(g\,O_j)f).$$

(ii) $H_0(a, D)$ is closed under the application of Q_j^t.
(iii) The operators Q_j preserve parities and if $f \in \tilde{H}$ is odd then $\Im f = 0$
(iv)If $F, G \in H$ are such that $\pi_P F \cdot G$ has no constant component, then:

$$\Im^{0^\sigma} G(\tau) d_\tau F(\tau) = F(0^\sigma)G(0^\sigma) - \Im^{0^\sigma} F(\tau) d_\tau G(\tau)$$

Proof. (i) Consider the bilinear forms:

$$(Fe^{-R_1|t|}, Q_j G e^{-R_2|t|}) \equiv \int_{-\infty}^{\infty} e^{-R_1|t|} dt \sum_{\rho=\pm} \int_{\rho\infty}^{t} w_j(t,\tau) e^{-R_2|\tau|} G(\tau) d\tau.$$

For sufficiently large values of R_1, R_2 the integrals are proper and the bilinear form is symmetric (as $w_j(t,\tau)$ is odd). So taking the residues at $R_1, R_2 = 0$ we obtain the symmetry of the operators Q_j on $\tilde{H}$.
(ii) We are simply restating Lemma1.11 and remarking that the operators Q_j preserve the analyticity in $t = 0$.
(iii) The operator $\Im$ changes the time parity (it is the inverse of a derivative); moreover we remember that

$$w_j(t,\tau) = x_j^0(t)\sigma(\tau)x_j^1(\tau) - \sigma(t)x_j^1(t)x_j^0(\tau),,$$

where both the x_j^i are even.
(iv) We want to compute:

$$\oint \frac{du}{2\pi i u} \int_{\sigma(t)\infty}^{t} e^{-u|t'|} F(t') d_{t'} G(t') = F(t)G(t) -$$

$$\oint \frac{du}{2\pi i u} \int_{\sigma(t)\infty}^{t} e^{-u|t'|} G(t') d_{t'} F(t') + \oint \frac{du}{2\pi i} \int_{\sigma(t)\infty}^{t} e^{-u|t'|} F(t') G(t')$$

the third summand is clearly zero if $\Pi_P FG = 0$ as in that case

$$\int_{\sigma(t)\infty}^{t} e^{-u|t'|} F(t') G(t')$$

is analytic in $u = 0$. If FG is a polynomial a direct computation on t^k shows that

$$\oint \frac{du}{2\pi i} \int_{\sigma(t)\infty}^{t} e^{-u|t'|} t^k = \delta(k, 0).$$

□

Corollary 1.17. *For any even $f, g \in \tilde{H}$:*

$$\Im f \Im^t w_j(t,\tau) g = \Im x_j^0 f \Im x_j^1 g - \Im x_j^1 f \Im x_j^0 g + \Im g \Im^t w_j(t,\tau) f.$$

Proof. We simply substitute (1.24) in Proposition 1.16(i) and then use 1.16(iii) to set he integrals of odd functions to zero. □

1.2 Trees

We supply the necessary definitions of trees, labeled trees, rooted trees and introduce sets of trees (which we will call admissible) having labels and grammatical rules adapted to our dynamics. We construct a vector space $\mathbb{V}$ on $\mathbb{Q}$ generated by the sets of admissible trees and define on $\mathbb{V}$ linear and multi-linear functions. The definitions are adapted to the problem of describing the homoclinic trajectories with the aid of trees; therefore many definitions could be given in more general terms and maybe appear then more natural (for a general presentation see for instance [GR]).

We hope however that the notation will become more clear when we define the connection with the dynamics in Chapter 2.
All the definitions of trees are standard, notice however that we are using a different notation from that of [G1] and the subsequent papers, which use numbered trees. This minor modification enables us to follow the combinatorics more explicitly.

1.2.1 Trees, symmetry groups and admissible trees

The definitions contained in this Subsections are all adapted from [GR].

Definition 1.18. *A* **graph** *G consists of two sets $V(G)$ (vertices), $\mathcal{E}(G)$ (edges) such that $\mathcal{E}(G)$ is a subset of the unordered pairs of distinct elements of G. We will always consider finite graphs, i.e. graphs such that $N(G) = |V(G)|$ is finite. Two vertices $i, j \in V(G)$ are said to be adjacent if $(i,j) \in \mathcal{E}(G)$. It is customary to write $n \in G$ in place of $n \in V(G)$ and $(i,j) \in G$ in place of $(i,j) \in \mathcal{E}(G)$.*
Two graphs G_1, G_2 are equal if and only if they have the same vertex set and the same edge set.

Definition 1.19. *A* **path** *joining the vertices $i, j \in G$ is a subset $\mathcal{P}_{ij}$ of $\mathcal{E}(G)$ of the form*

$$\mathcal{P}_{ij} := \{(i, v_1), (v_1, v_2), \cdots, (v_k, j)\}.$$

A graph G is **connected and without loops** *if for all $i, j \in G$ there exists one and only one path that connects them. Such graphs are called* **trees**. *Their vertices are called nodes and their edges are called branches.*

A tree T such that the set $V(T) = \{1, 2, \ldots, N(T)\}$ is called a numbered tree.

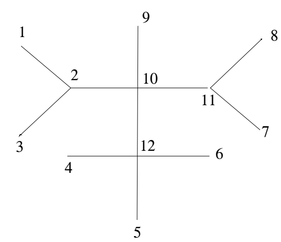


Figure 1.4: A numbered tree

Definition 1.20. *A* **labeled** *tree is a tree A plus a label $\mathcal{L}_A(v) \geq 0$ which is generally a set of functions $f^i_A(v)$ defined on the nodes.*

When possible we will omit the subscript A in the functions f^i.

Definition 1.21. *Two labeled trees X, Y are isomorphic if there is a bijection, h say, from $V(X)$ to $V(Y)$ such that for all $a \in V(X)$, $\mathcal{L}_X(a) \equiv \mathcal{L}_Y(h(a))$, moreover $(a, b) \in \mathcal{E}(X)$ if and only if $(h(a), h(b)) \in \mathcal{E}(Y)$.*

We say that h is an isomorphism from X to Y. Notice that since h is a bijection h^{-1} is well defined and is an isomorphism from Y to X.
We will call **symmetries** *or automorphisms of X, the isomorphisms from X to X.*

It is often convenient and more compact to represent a tree by a diagram, with points for the nodes and lines for the branches, as in Figure 1.5. In this diagrams the positions of the points and lines do not matter - the only information it conveys is which pairs of nodes are joined by a branch. This means that the two diagrams in Figure 1.5are equal by definition.

Strictly speaking these diagrams do not define graphs, since the set V is not specified. However, if the diagram has N points, we may assign distinct natural numbers $1, 2, \ldots, N$ to the points (which we still call nodes), so obtaining a labeled numbered tree.
Then it is easily seen that the two trees in Figure 1.5 are isomorphic.

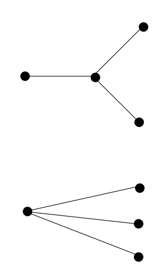


Figure 1.5:

Definition 1.22. *Formally we can define such diagrams as the equivalence classes of labeled trees via the relation $A \cong B$ if and only if A and B are isomorphic.*
An obvious consequence of this definition is that, $\mathcal{L}_A(v)$ and $N(A)$ are well defined on the equivalence classes.

We can choose a representative A' of the equivalence class A by giving a numbering $1, 2, \cdots, N(A)$ to the nodes of A.

Remark 1.23. *Given an equivalence class of labeled trees A let A' be a numbering, and let $\mathcal{S}(A')$ be the group of automorphisms of A'.*
This means that $\mathcal{S}(A')$ is the subgroup of the permutations $\sigma \in S_{N(A)}$ which fix both $\mathcal{E}(A)$ and the labels j_A, δ_A. Namely $\sigma \in \mathcal{S}(A) \to \sigma\mathcal{L} = \mathcal{L}$ and $j_A(v) = j_A(\sigma(v)), \delta_A(v) = \delta_A(\sigma(v))$.
Given two isomorphic trees A', A'' of A, let h be the bijection such that $\mathcal{E}(A') = \sigma\mathcal{E}(A'')$. The groups $\mathcal{S}(A')$ and $\mathcal{S}(A'') = h^{-1}\mathcal{S}(A')h$ are isomorphic. We will improperly call the equivalence classes via this relation the **symmetry** *group $\mathcal{S}(A)$ of the diagram A.*

Using standard notation (see for instance [L]) we denote by $a := (i_1, i_2, \ldots, i_m)$ with $\mathbb{N} \ni i_j \leq N(A)$ the permutation such that $a(i_h) = i_{h+1}$, $a(i_m) = i_1$, and $a(n) = n$ for all $\mathbb{N} \ni n \leq N(A)$ such that $n \notin \{i_1, i_2, \ldots, i_m\}$. Moreover $(i, j, k)(l, m)$ is the composition of $a = (i, j, k)$ and $b = (l, m)$.

As an example in Figure 1.6 consider the numbered tree A ($N(A) = 6$), its symmetries are the identity and: $a := (1, 4)$; $b := (2, 3)$; $c \equiv a \circ b$; $d \equiv (5, 6)(1, 2)(4, 3)$, $e := (5, 6)(1, 3)(2, 4)$; $f := (5, 6)(1, 2, 4, 3)$; $g := f \circ a$. Clearly any other numbering on A, would give an isomorphic symmetry group.

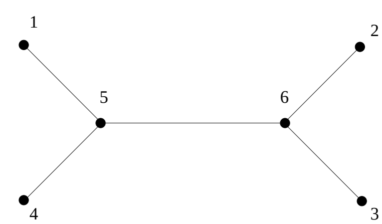


Figure 1.6:

Given a node $v \in A$,we define its **orbit**:

$$[v] := \{w \in A : w = g(v) \text{ for some } g \in \mathcal{S}(A),$$

i.e. the list of nodes obtained by applying the whole group $\mathcal{S}(A)$ to v, notice that this is an equivalence relation (a proof of this statement is in [GR]). In the example of Figure1.6 there are two orbits, which in the chosen numbering are:

$$[1] \equiv \{1, 2, 3, 4\} \text{ and } [5] \equiv \{5, 6\}.$$

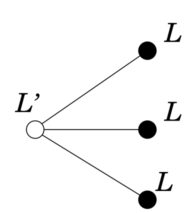


Figure 1.7:

Remark 1.24. *The orbits are well defined on the equivalence classes of labeled trees, it should be clear , for instance, that the nodes signed in black in the diagram of Figure 1.7 are an orbit*

Definition 1.25. *A* **rooted** *labeled tree is a labeled tree A plus one of its nodes called the first node (v_A or v_0); this gives a partial ordering to the tree, namely we say that $i > j$ if $\mathcal{P}_{v_0 j} \subset \mathcal{P}_{v_0 i}$ (see Figure1.8). Moreover choosing a first node induces a natural ordering on the couples of nodes representing the branches namely $(a, b) \in \mathcal{E}(A)$ implies that $a < b$.*
We recall some definitions on rooted trees:
a) the level of v $l(v)$ is the cardinality of $\mathcal{P}_{v_0 v}$;

b) the nodes subsequent to v, $s(v)$, are the nodes adjacent to v and of higher level; the node preceding v is the only node adjacent to v and of lower level;

c) given v node of A, we call $A^{\geq v}$ the rooted tree (with first node v) of the nodes $w \geq v$; we call $A^{\setminus v}$ the remaining part of the tree A.

An isomorphism between rooted trees (A, v_A), (B, v_B) is an isomorphism between A and B which sends v_A in v_B.
The symmetries of a rooted labeled tree (A, v_A) ,which we denote again by $\mathcal{S}(A, v_A)$ are the subgroup of the symmetries of the corresponding unrooted tree, that fix the first node v_A. As done for trees, we can represent the equivalence classes of rooted trees with diagrams, representing by convention the first node on the left and all the nodes of the same level aligned vertically (it should be obvious that the definitions $v > w$, $A^{\setminus v}$ and $A^{\geq v}$ are well posed on the equivalence classes).

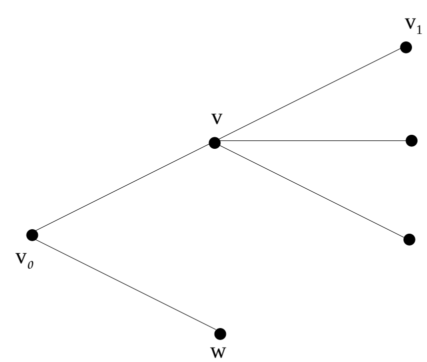


Figure 1.8: A rooted tree, $l(v) = 1$, $l(v_1) = 2$, in this example the nodes subsequent to v, $s(v)$ are the orbit of v_1.$|\mathcal{S}_{v_0}(A)| = 6$, the tree $A^{\setminus v}$ is (v_0, w)

Remark 1.26. *By the Lagrange theorem (see [L] or [GR]) we have that:*

$$|\mathcal{S}(A, v_A)| = \frac{|\mathcal{S}(A)|}{|[v_A]|},$$

where $[v_A]$ is the orbit of v_A considered as a node in the unrooted tree A.

$\mathcal{S}(A, v_A)$ is a group, so we can define its orbits on the nodes v which we call again $[v]$ (see Figure1.8). Notice that now $[v_A] \equiv \{v_A\}$. Moreover if $v_1 \in [v]$:

$$l(v) = l(v_1) \quad \text{and } A^{\geq v_1} = A^{\geq v} \equiv A^{\geq [v]};$$

We call $[v]_l$ the cosets of level l and $m[v] = |[v]|$.

Lemma 1.27. *The order of the symmetry group $|\mathcal{S}(A, v_A)|$ is:*

$$|\mathcal{S}(A, v_A)| = \prod_{[v]_1} m[v]!|\mathcal{S}(A^{\geq [v]}, v)|^{m[v]}$$

Proof. We apply the Lagrange theorem repeatedly: first we choose a node v of level one, and prove that the order of its stabilizer (in $\mathcal{S}(A, v_A)$) is the product of $|\mathcal{S}(A^{\geq v}, v)|$ and $|\mathcal{S}(A^{\setminus v}, v_A)|$;then in $A^{\setminus v}$ choose a node $w \in [v]$ and so on until all the nodes in $[v]$ are canceled; one gets

$$m[v]!|\mathcal{S}_v(A^{\geq [v]})|^{m[v]}||\mathcal{S}(A^{\setminus [v]}, v_A)|,$$

where $(A^{\setminus [v]}, v_A)$ is the rooted tree A deprived of all the subtrees $A^{\geq w}$ with $w \in [v]$. So in $A^{\setminus [v]}$ we consider another coset $[v'] \neq [v]$ and repeat the procedure. A more detailed proof is in [GR]. □

Now we will fix the label functions and restrict our attention to trees respecting some rules (a grammar) which reflect the properties of our perturbative expansion of the homoclinic trajectory.

Definition 1.28. *We consider rooted labeled trees such that some nodes are distinguished by having a different set of labels* [8]*. An* **admissible** *tree is a symbol:*

$$A, \{v_A\}, \{v_1, \ldots, v_m\}, \{w_1, \ldots, w_h\}$$

such that A is a tree, all the v_i, w_j and v_A are nodes of A, the v_i are all end-nodes,

$$\{v_i\}_{i=1}^m \cap \{w_j\}_{j=1}^h = \varnothing$$

and the v_i are all different.

We call $\{v_i\}_{i=1}^m \equiv \mathcal{F}(A)$ the fruits of A, $\{w_j\}_{j=1}^h \equiv \mathcal{M}(A)$ the marked[9] *nodes of A and the set*

$$\overset{0}{A} : \{v \notin \mathcal{F}(A)\}$$

the free nodes of A.

The labels are distributed in the following way:

a) For each node $v \neq v_A$ one angle label $j_v \in \{0, \ldots, n\}$ (remember that we are considering a system with $n+1$ degrees of freedom).

[8] The dynamical meaning of the labels will be clear when we will define the “value” of a tree

[9] a node v can appear many times in $\mathcal{M}(A)$ we will say it carries more than one marking.

b) For each node v one order label $\delta_v = 0, 1$ if $v \in \overset{0}{A}$ and $\delta_v \in \mathbb{N}$ otherwise.

c) For each node $v \in \mathcal{M}(A)$ one angle-marking $J = 0, \ldots, n$ and one function-marking $h(t) \in H$.

d) For each node $v \in \mathcal{F}(A)$ one type label $i = 0, 1$.

We set a grammar on the so defined labeled rooted trees, namely:

$$\delta_v = 0 \rightarrow \{j_v = J_v = 0,\ |s(v)| \geq 2\ ,\ j_{v'} = 0\ \forall v' \in s(v)\}.$$

To draw the diagrams without writing down the labels we give a color to each $j = 1, n$ (which forces $\delta = 1$) and two different colors for the couples of labels $j = 0, \delta = 1$ and $j = 0, \delta = 0$.
In all the pictures we will set $n = 2$ and choose the colors blue, green, black and white, see Figure1.9. The fruits $\mathcal{F}(A)$ will be represented as "bigger" end-nodes colored with the color corresponding to their component label and with their order and type written on a side. The marked nodes will be distinguished by a box of the color corresponding to their angle-marking and with their function-marking written on a side.If the function marking is $h(t) = 1$ we will omit the function marking.

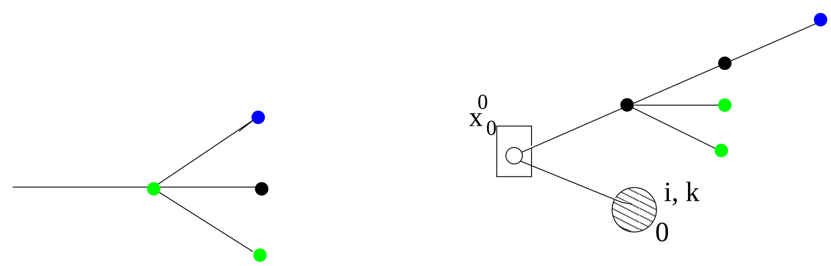


Figure 1.9: Examples of trees in $\mathcal{A}^5$ and in $\overset{0}{\mathcal{T}}{}^5$

Definition 1.29. *1) We will call fruitless trees the (labeled rooted trees) A such that $\mathcal{F}(A)$ is empty. We will say that a fruit v stems from w if $v \in s(w)$.*

2)We will call $\mathcal{T}$ the set of equivalence classes (as in definition 1.22) of admissible trees, $\overset{0}{\mathcal{T}}$ the subset of $\mathcal{T}$ of trees with at least a free node and $\mathcal{A}$ the subset of $\overset{0}{\mathcal{T}}$ of "fruitless" trees.
Finally we will call $\overset{m}{\mathcal{A}}$ the subset of $\mathcal{A}$ of fruitless trees with no marking.

3) We will call $\mathcal{F}_j^{ik}$ the "tree" composed of one fruit of order k angle j and type i; clearly

$$\mathcal{T} \equiv \overset{0}{\mathcal{T}} \bigcup_{\substack{i=0,1 \\ j=0,\ldots,n \\ k>0}} \mathcal{F}_j^{ik}.$$

Notational Convention 1. *Using standard notation we represent the equivalence classes by $[A]$ where A is an admissible tree.*
Moreover givea a tree A we will write $A \in \mathcal{T}$ if it is a representative of an equivalence class in $\mathcal{T}$.

Definition 1.30. *The order of an element $[A] \in \mathcal{T}$ is:*

$$o(A) = \sum_{v \in A} \delta_v .$$

The order of a node v of A is $o(v) = o(A^{\geq v})$.

Given a tree $A \in \overset{0}{\mathcal{T}}$ and one of its nodes v we call $A^{\geq v}$ the tree composed of the nodes greater or equal to v; if $A^{\geq v}$ is not a fruit then it is not admissible as it carries a label j in the first node. In such case, we conventionally set $A^{\geq v} \in \mathcal{T}$ by setting a mark $J(v) = j_v$, $h(v,t) = 1$ on v and subsequently "forgetting" the label j_v.

It is easily seen that $o(A) > 0$ for all $A \in \mathcal{T}$ and that

$$\mathcal{T}^k \equiv \{A \in \mathcal{T} \text{ t.c. } o(A) = k\}$$

is a finite set (see also Proposition 1.37); clearly the same is true in $\overset{0}{\mathcal{T}}$ and in $\mathcal{A}$

Notational Convention 2. *in all our sets an apex k means we consider the subset of trees of order k.*

We list here all the subsets of $\mathcal{T}$, $\overset{0}{\mathcal{T}}$ and $\mathcal{A}$ that we will need in the following sections.

Definition 1.31. *a) $\mathcal{T}_*$ (resp $\overset{0}{\mathcal{T}}_*$ and $\mathcal{A}_*$) is the subset of $\mathcal{T}$ (resp $\overset{0}{\mathcal{T}}$, $\mathcal{A}$) such that v_A appears exactly once in $\mathcal{M}(A)$ and $h(v_A,t) = 1$ or $v_A \equiv \mathcal{F}(A)$.*

b) $\overset{0}{\mathcal{T}}_j$ ($\mathcal{A}_j$) is the subset of $\overset{0}{\mathcal{T}}_*$ ($\mathcal{A}_*$) such that $J(v_A) = j$ and $\mathcal{M}(A) \equiv \{v_A\}$;

$$\mathcal{T}_j = \overset{0}{\mathcal{T}}_j \cup_{k \in \mathbb{N}, i=0,1} \mathcal{F}_j^{ik} .$$

c) $\mathcal{A}_{(j,f(t))}$is the subset of $\mathcal{A}$ such that $\mathcal{M}(A) \equiv \{v_A\}$ and $J(v_A) = j$, $h(v_A,t) = f(t)$.

d) $\mathcal{A}_{(i,h(t)),(j,f(t))}$ is the subset of $\mathcal{A}$ such that $\mathcal{M}(A) \equiv \{v_A, v\}$ for some $v \in A$ moreover $J(v_A) = i$, $h(v_A,t) = h(t)$, $J(v) = j$, $h(v,t) = f(t)$.

For each of these sets we will consider a vector space on $\mathbb{Q}$ generated by the set; if S is the set we represent it by $\mathbb{V}(S)$.

Definition 1.32. $\mathbb{V}(S)$ *is the vector space of linear combinations of elements of S with rational coefficients.*

$$[A] \in S \to [A] \in \mathbb{V}(S)\ , \qquad [A],[B] \in \mathbb{V}(S) \to q_1[A] + q_2[B] \in \mathbb{V}(S)\ , \quad \forall q_1, q_2 \in \mathbb{Q}.$$

$\mathbb{V}(S)$ is an infinite dimensional vector space and can be expressed as direct sum of finite dimensional spaces generated by the sets S^k (we call these spaces $\mathbb{V}^k(S)$). For example (remember that . $n = 2$) $\mathcal{A}_1^3$ is the set in Figure1.10. The values of the labels j_i of the nodes $1, \ldots, 4$ are free: they can be $0, 1, 2$ while δ_i is fixed to one because $s(1) = 1$ and the nodes $2, 3, 4$ are end-nodes ; the dimension of $\mathbb{V}_1^3 = \mathbb{V}^3(A_1)$ is $|\mathcal{A}_1^3| = 19$.

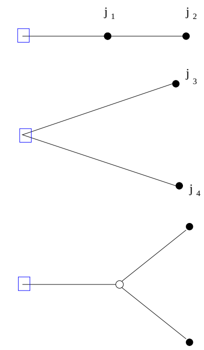


Figure 1.10:

1.2.2 Functions on admissible trees

We define some functions on the subspaces of $\mathcal{T}$ which will be useful in the following sections. The definitions are very much "ad hoc" so they will necessarily seem quite unnatural.

Such functions will then be extended to linear functions on the corresponding $\mathbb{V}$.

Definition 1.33. *Consider a rooted labeled marked tree A, with first node v_A angle-marked $J(v_A)$ and function-marked $h(t) = 1$ (A is not necessarily in $\overset{0}{\mathcal{T}}_*$). We define $\tilde{A}$ as the tree obtained from A by setting $j(v_{\tilde{A}}) = J(v_A)$ and subsequently forgetting the marking $J, h = 1$ of v_A so that the first node does not have a different labeling from the other free nodes.*
Given a tree $B \in \overset{0}{\mathcal{T}}$ plus one of its nodes $v \neq v_B$ let w be the node preceding v, we define:

$$\bar{g}_A(B, v) = \mathcal{E}(\tilde{A}) \cup \{\mathcal{E}(B) \setminus (w, v)\} \cup (w, v_A) \cup (v_A, v),$$

and

$$g_A(B, v) = \begin{cases} \bar{g}_A(B, v) & \text{if } \bar{g}_A(B, v) \in \overset{0}{\mathcal{T}} \\ 0 & \text{otherwise.} \end{cases}$$

Finally we can define $g_A(B) = \sum_{v \in B} g_A(B, v)$, this is a function $g_A : \overset{0}{\mathcal{T}} \to \mathbb{V}(\overset{0}{\mathcal{T}})$ so we can extend it linearly on $\mathbb{V}(\overset{0}{\mathcal{T}})$.

Definition 1.34. *For all $k \in \mathbb{N}$ we define functions on unordered k-ples of trees in $\mathcal{T}_*$. Let A be a labeled rooted marked tree with at least one free node, and $\{B_i\}_{i=1}^k$ be an unordered set of trees $B_i \in \mathcal{T}_*$.*

We call as usual v_A the first node of A and v_{B_i} the first nodes of the B_i.
If $B \in \{B_i\}$ is not a fruit and $J(v_B)$ is the marking of v_B we call $\tilde{B}$ the tree obtained from B by setting $j_{v_{\tilde{B}}} = J(v_B)$ and forgetting the marking. Then we define

$$f_A(B_1, \ldots, B_k) = \begin{cases} \cup_i (v_{\tilde{A}}, v_{\tilde{B}_i}) \cup \mathcal{E}(\tilde{A}) \cup_i \mathcal{E}(\tilde{B}_i) & \text{if it is in } \overset{0}{\mathcal{T}} \\ 0 & \text{otherwise.} \end{cases}$$

This functions as well can be extended to $\mathbb{V}(\mathcal{T}_*)$ by linearity.

It should be clear that the definition is invariant by permutations of the B_i. Notice moreover that it is not necessary that $A \in \mathcal{T}_*$ to obtain a (linear or multi-linear) function f_A: $\mathcal{T}_* \to \mathcal{T}_*$ (or g_A). Consider for instance the trees in Figure 1.11 respectively for linear functions $g_A(B)$ and for multi-linear functions ($k \geq 2$).

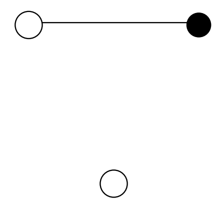


Figure 1.11:

Definition 1.35. *We will use functions whose arguments are in some specified subspaces*[10] $\mathbb{V}^k(\mathcal{T}_j)$ *(this means that the image is in some fixed* $\mathbb{V}^h(\mathcal{T}_i)$ *as well):*

$$f^A_{\{p_i^h\}_k} : \otimes^{n\,,k}_{\substack{i=0\\h=1}} \underbrace{\mathcal{T}_i^h \times \cdots \times \mathcal{T}_i^h}_{p_i^h} \to \mathcal{T}_j^{k+o(A)},$$

namely there is an ordering of the set $\{B_i\}$ such that $B_1, \ldots, B_{p_0^1} \in \mathcal{T}_0^1$ then $B_{p_0^1+1}, \ldots,$ $B_{p_0^1+p_1^1} \in \mathcal{T}_1^1$ and so on, see Figure1.13.

All this this functions are well defined on the equivalence classes, namely if

$$A \cong A'\,, \;\; B \cong B' \;\to g_A(B) \cong g_{A'}(B') \ldots;$$

this implies that the functions can be represented graphically on the diagrams. Functions $g_A(B)$ use the marking of the first node of A as angle label and substitute the branch w, v with A by joining the first node of A to v and w (we set the result to zero if we obtain a tree not in $\mathcal{T}$).

Functions $f_A(\{B_i\})$ use the marking of the first node of the B_i as an angle label and join with a branch the first nodes of A and of the B_i. They have v_A as first node. As an example if A_1 and A_2 are the two trees in Figure6.3.1,then $f_{A_1}(A_2)$ is the tree in Figure1.12(a) while $f_{A_2}(A_1) = 0$ as all nodes with $\delta(v) = 0$ must be followed by nodes with $j = 0$.
Let us define the multi-linear functions:

$$\Gamma^\delta_{\{p_i^h\}_k} = f^{\alpha(\delta)}_{\{p_i^h\}_k} \;\text{ where the tree } \alpha(\delta) \in \overset{0}{\mathcal{T}} \text{ is } \overset{\delta}{\circ};$$

these functions are used to construct recursively the sets $\overset{m}{\mathcal{A}}{}^k$.

[10]remember that $\{p_i^h\}_k$ is a weighted partition of k: a list of numbers in $\mathbb{N}_0$ such that

$$\sum_{i,h} h p_i^h = k.$$

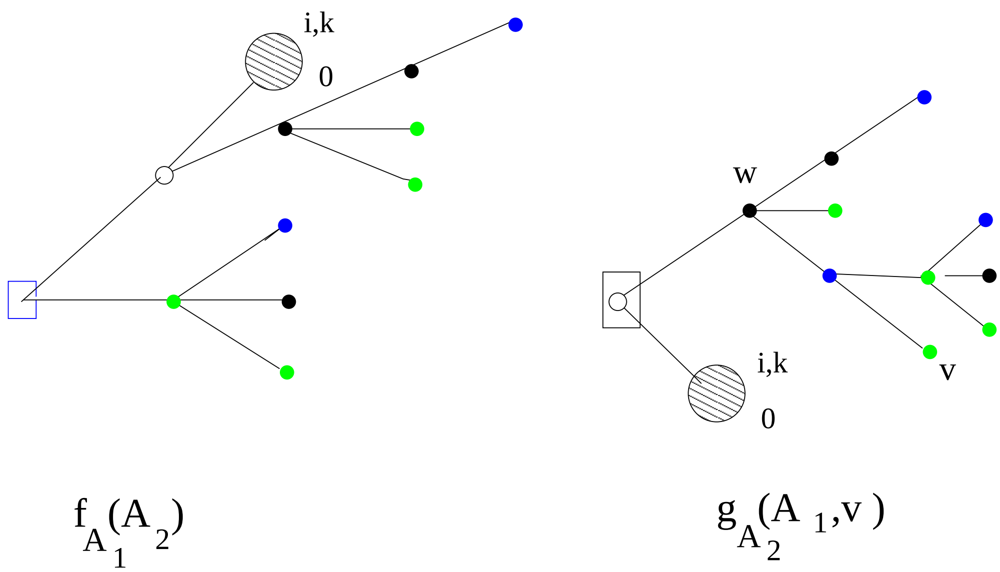


Figure 1.12: Linear functions on $\mathbb{V}$; the diagrams A_1 and A_2 are those of Figure 1.9.

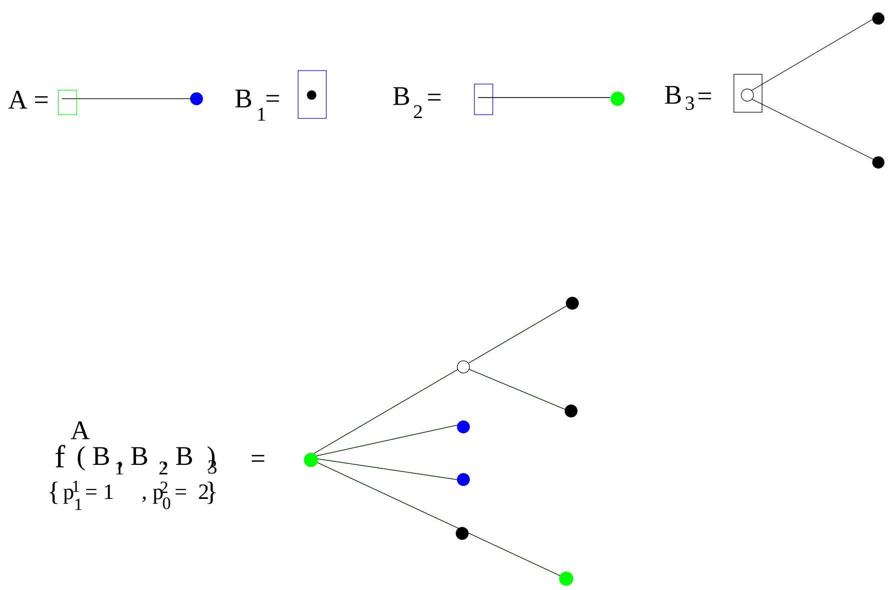


Figure 1.13: A multi-linear function $f : \mathcal{A}_1^1 \times \mathcal{A}_0^2 \times \mathcal{A}_0^2 \to \mathcal{A}_1^7$

For each $A \in \overset{m}{\mathcal{A}}^k$, let v_A be its first node and $v_1, \dots, v_m$ the nodes of level one. Now for each $l = 0, \dots n$ and $h = 1, \dots, k-1$ let $p_l^h(A)$ be the number of elements w of the list $s(v)$ such that $A^{\geq w} \in \mathcal{A}_l^h$; notice that

$$\sum_{\substack{l=0\\h=1}}^{n\,,\,k} h\, p_l^h(A) = k - \delta_{v_A}$$

.

Remark 1.36. *Given $A \in \overset{m}{\mathcal{A}}^k$:*

$$A = \Gamma^{\delta_{v_A}}_{\{p_i^h(A)\}}(A^{\geq v_1}, \dots, A^{\geq v_m}).$$

Conversely the set $\delta \in (0,1)$, $\{A_l\}_{l=1}^K \in \cup_{j=0}^n \mathcal{A}_j$ with $K \geq 1$, represents one and only one (non zero) tree: namely for any $0 \leq i \leq n$, $h \geq 1$, set $\{p_i^h\}_{\{A_l\}}$ to be the number of trees in the list $\{A_l\}$ belonging to $\mathcal{A}_i^h$, and consider the tree

$$A = \Gamma^{\delta}_{\{p_i^h\}_{\{A_l\}}}(\{A_l\}). \tag{1.32}$$

Clearly there are many lists $\delta \in (0,1)$, $\{A_l\}_{l=1}^K$ such that expression (1.32) gives zero .

This simple Remark leads to a constructive algorithm for constructing the sets $\overset{m}{\mathcal{A}}^k$ from the sets $\mathcal{A}_j^h$ with $h < k$.

Proposition 1.37 (Recursive construction of $\overset{m}{\mathcal{A}}^k$). *For all $k \in \mathbb{N}$:*

$$\overset{m}{\mathcal{A}}^k = \bigcup_{\substack{\delta=0,1\,;\,\{t_h^i\}_{k-\delta}\\ A_i^h(\alpha)\in\mathcal{A}_i^h}} \Gamma^{\delta}_{\{t_i^h\}}\big(A_0^1(1), \cdots, A_0^1(t_0^1), A_1^1(1), \cdots, A_n^{k-1}(t_n^{k-1})\big) \tag{1.33}$$

Proof. This follows directly from Remark1.36 as expression (1.33) generates all the lists $\delta \in (0,1)$, $j \in (0, \dots, n)$, $\{A_l\}_{l=1}^k$. ☐

Now to generate $\mathcal{A}$ (and in particular the sets $\mathcal{A}_j^h$) we consider linear functions which add extra markings to a tree; given $A \in \overset{0}{\mathcal{T}}$ the symbol:

$$h(v,t)\partial_l^v A$$

represents the application of an angle-marking $J(v) = l$ and a function-marking $h(v,t)$ in the node v; formally

$$A, \{v_A\}, \{v_i\}_{i=1}^m, \{w_j\}_{j=1}^h \to A, \{v_A\}, \{v_i\}_{i=1}^m, \{\{w_j\}_{j=1}^h \cup \{v\}\}.$$

We can define the linear function:

$$D_j(h(t))[A] := \sum_{v \in \overset{0}{A}} h(v,t)\partial_j^v A. \tag{1.34}$$

Lemma 1.38. *The set $\mathcal{A}$ is obtained from $\overset{m}{\mathcal{A}}$ by successive applications of the mark-adding functions. In particular $\mathcal{A}_j$ is generated by*

$$\partial_j^{v_0}\Gamma_{\{p_i^h\}}.$$

To generate $\overset{0}{\mathcal{T}}$ we can consider functions which add fruits to a tree: given $A \in \overset{0}{\mathcal{T}}$ The function $d_j^{i,k}(v)$ adds a fruit F_j^{ik} to the node v by adding a node y labeled (i,k,j) to the list $\mathcal{F}(A)$ and setting $y \in s(v)$. Then naturally we can define the linear function:

$$D_j^{(i,k)}[A] := \sum_{v \in \overset{0}{A}} d_j^{i,k}(v)[A].$$

This is not the only possible way of adding fruits namely if $\alpha^{i,k}$ is the tree in Figure 1.14 then we consider the linear function:

$$B^{(i,k)}[A] := \sum_{\substack{v \in A \\ j_v = 0}} g_{\alpha^{i,k}}(A, v).$$

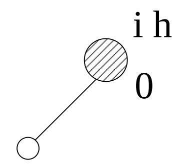


Figure 1.14:

Finally to generate all the possible trees with one fruit we consider the function:

$$F^{i,k}[A] = f_{\alpha^{i,k}}(A),$$

defined on trees $A \in \mathcal{A}_0$

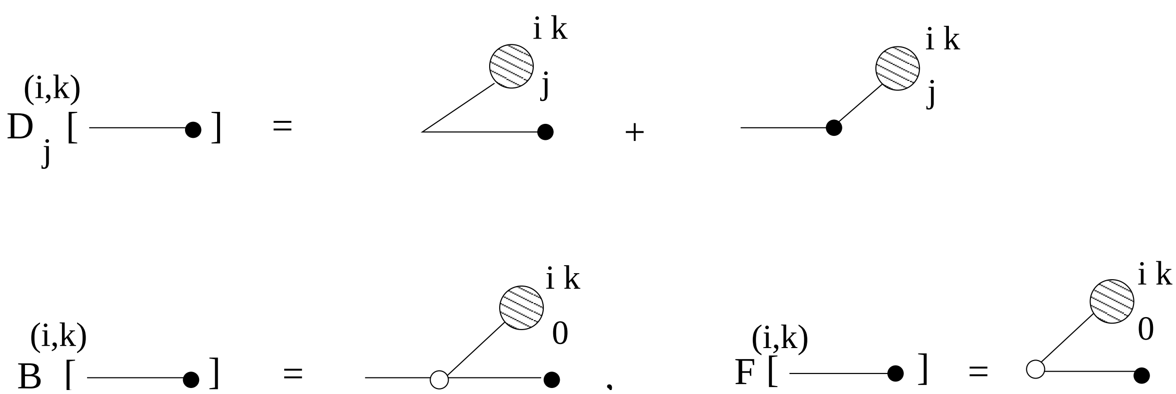


Figure 1.15: The adding fruits functions

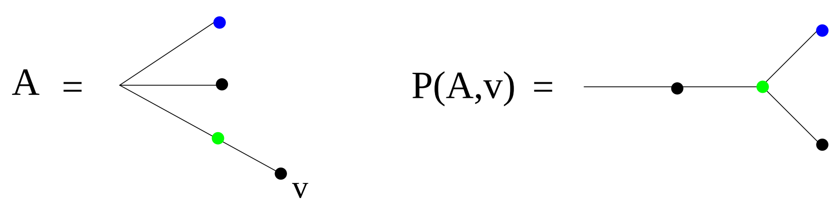
Lemma 1.39. *The set* $\overset{0}{\mathcal{T}}$ *is obtained from* $\overset{m}{\mathcal{A}}$ *by successive applications of the fruit-adding and mark-adding functions; in particular:*

$$\bigcup_{\substack{i=0,1 \\ j=0,\ldots,n \\ k \in \mathbb{N}}} D_j^{(i,k)}[\overset{m}{\mathcal{A}}] \bigcup L^{(i,k)}[\overset{m}{\mathcal{A}}] \bigcup F^{(i,k)}[\overset{m}{\mathcal{A}}] \equiv \overset{m}{\mathcal{A}}(1F),$$

where $\overset{m}{\mathcal{A}}(1F)$ *are the trees without markings and with only one fruit.*

Another way of manipulating trees is to change the first node (which is distinguishable only as it does not have the label j). Generally one can obtain various trees in $\overset{0}{\mathcal{T}}$ by simply changing the uncolored node (for example one can shift the angle labels down along a path joining any node v to the uncolored one v_A). However not all the trees obtained in such a way are in $\mathcal{T}$

Definition 1.40. *Given a tree $A \in \overset{0}{\mathcal{T}}$ let v_A be the first node and v a free node; the change of first node $P(A,v) : \overset{0}{\mathcal{T}} \to \overset{0}{\mathcal{T}}$ is so defined:*
let $v_A = v_0, v_1, \ldots, v_m = v$ be the nodes of the path $\mathcal{P}_{v_A,v}$. $P(A,v)$ is obtained from $A, \{v_A\}, \{v_i\}_{i=1}^{m}, \{w_j\}_{j=1}^{h}$ by shifting only the j labels of the nodes of $\mathcal{P}_{v_A,v}$ in the direction of v_A. This automatically implies that v is left $j-$uncolored and is the first node of $P(A,v)$. If we obtain a tree not in $\mathcal{T}$ we set $P(A,v) = 0$.
$P : \mathbb{V}(\mathcal{T}) \to \mathbb{V}(\mathcal{T})$ is the linear function such that $\forall A \in \mathcal{T}$, $P(A) = \sum_{v \in \overset{0}{A}} P(A,v)$.



Lemma 1.41. *$P(A,v) = 0$ if and only if $\delta_{v_A} = 0$, $|s(v_A)| = 2$. This means that the possibility of applying the change of first node does not depend on the chosen $v \neq v_A$.*

Proof. Consider the trees A and $P(A,v)$ and the nodes $v_A = v_0, v_1, \ldots, v_m = v$ of the path $\mathcal{P}_{v_A,v}$. For each $i = 0, m-1$ v_i precedes v_{i+1} in A and follows it in $P(A,v)$. So for each node $w \neq v_A, v$ the number of following nodes $s(w)$ is the same in A and $P(A,v)$; $s(v_A)$ decreases by one and $s(v)$ consequently increases by one. This implies that all trees A with $\delta_{v_A} = 0$ and $|s(v_A)| = 2$ have $P(A,v) = 0$ for all v. Moreover if v_i has $\delta = 0$ then it has $j = 0$ as well as all the nodes (including v_{i+1}) following it. This means that in $P(A,v)$ it will still have $\delta = j = 0$, the same $s(v_i) \geq 2$; moreover v_{i-1} that follows v_i in $P(A,v)$ has $j = 0$.

□

Notational Convention 3. *We will call $\overset{r}{\mathcal{T}}$ the subspace of $\overset{0}{\mathcal{T}}$ of trees whose first node can be changed. In general an apex r on a tree set S means that we consider only trees in S whose first node can be changed.*

Definition 1.42 (change of nodes in $\mathcal{T}_{(i,h),(j,f)}$). *Given a tree $A \in \mathcal{T}_{(i,h),(j,f)}$ let v_A and v be respectively the first node and the other marked node. We define $P_1 \equiv P(A,v) : \mathcal{T}_{(i,h),(j,f)} \to \mathcal{T}_{(j,f),(i,h)}$; see Figure1.16.*

Remark 1.43. *Notice that given a tree $A \in \overset{r}{\mathcal{T}}$ and one of its nodes v there exists a unique B such that $P(A,v) = B$. This means that for all i,j and for all the functions $h, f \in H$:*

$$\overset{r}{\mathcal{T}}_{(i,h),(j,f)} \leftrightarrow \overset{r}{\mathcal{T}}_{(j,f),(i,h)}.$$

If $i, j \neq 0$ then $\overset{r}{\mathcal{T}}_{(i,h),(j,f)} \equiv \mathcal{T}_{(i,h),(j,f)}$ and so $\mathcal{T}_{(i,h),(j,f)} \leftrightarrow \mathcal{T}_{(j,f),(i,h)}$.

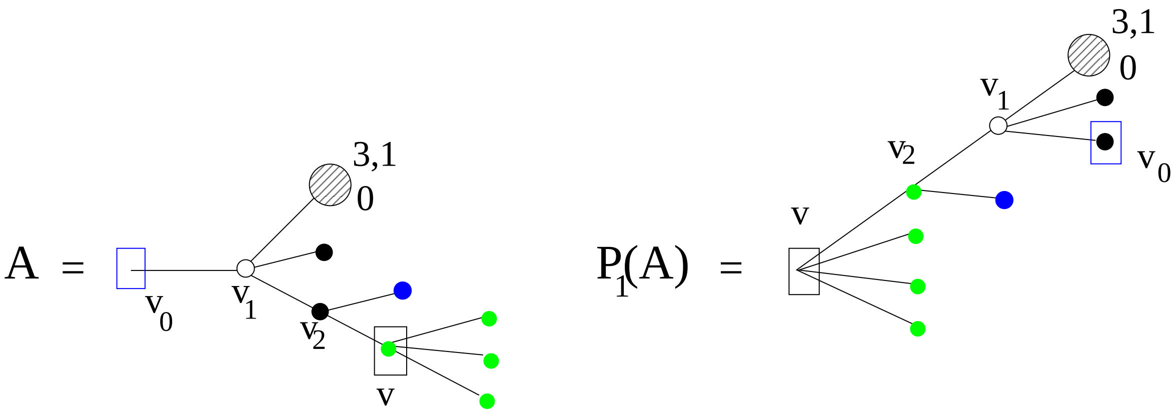
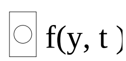


Figure 1.16: Example of $P_1(A)$; we have evidenced the path joining the two marked nodes.

In $\mathcal{T}_{(0,f),(i,h)}$ $(i \neq 0)$ we have trees not in $\overset{r}{\mathcal{T}}$; i.e. trees with $\delta_{v_0} = 0$ and $|s(v_0)| = 2$. Call $\mathcal{T}^{(0)}_{(0,f),(i,h)}$ such subset.

Lemma 1.44. *$\mathcal{T}^{(0)}_{(0,f),(i,h)}$ is the image of $\mathcal{T}_{(i,h)}$ by a suitable linear function, (similar to $g_A(B)$).*

Proof. We choose

$\alpha_f =$  where $f \in H$

(this is a marked rooted tree with one node y) and consider the application $\bar{g}_{\alpha_f}(A, v)$. Then we apply the change of first node in y. We have a “linear function”:

$$L_f(A) = \sum_{\substack{v \in A \\ j_v = 0}} P(\bar{g}_{\alpha_f}(A, v), y)$$

Consider a tree $A \in \mathcal{T}_{(i,h)}$, as α_f has degree zero in k the degree of $\bar{g}_{\alpha_f}(A, v)$ is the same as that of A. However the trees $\bar{g}_{\alpha_f}(A, v)$ are never in $\mathcal{T}$. We then apply the change of first node and obtain the linear function $\mathcal{T} \to \mathcal{T}$ $L_f(A)$ whose first node is y marked zero and $s(y) = 2$; the node v (that follows y) is labeled $j = 0$ by the definition of g, while the node that precedes v in A (that now follows y) gets the label $j = 0$ from y by the shift of labels; the trees we obtain are in $\mathcal{T}$. Notice that L_f is an injective linear function $\mathcal{T}^k_{(i,h)} \to \mathbb{V}(\mathcal{T}^{k\,0}_{(0,f),(i,h)})$ and that each tree $B \in \mathcal{T}^{(0)k}_{(0,f),(i,h)}$ uniquely identifies the couple A, v where $A \in \mathcal{T}^k_{(i,h)}$ and v is one of its nodes. □

Corollary 1.45. *Consider the set*

$$\hat{\mathcal{A}} = \overset{m}{\mathcal{A}} \cap \overset{r}{\mathcal{A}},$$

$\hat{\mathcal{A}}$ generates $\overset{m}{\mathcal{A}}$ (and consequently $\mathcal{A}$ and $\overset{0}{\mathcal{T}}$).

Proof. We simply consider the linear function $L(A) := P(\bar{g}_{\alpha}(A, v), v)$ where $\alpha = \circ^{\delta=0}$, and proceed as in Lemma 1.44. $\square$

Chapter 2

Tree expansion for the homoclinic trajectory

In the preceding Chapter we have defined all the necessary spaces of trees; now we finally set the trees in correspondence with the dynamics. In particular we will define two applications $\mathcal{V}$ and $\mathcal{W}$ defined on $\mathcal{A}$ and two applications $\mathcal{V}^1$ and $\mathcal{W}^1$ defined on $\mathcal{T}$. Correspondingly we will define two vectors

$$\mho_j^k \in \mathbb{V}(\mathcal{A}_j^k)\,, \quad and \quad \Lambda_j^k \in \mathbb{V}(\overset{0}{\mathcal{T}}{}_j^k)$$

this vectors will have the property:

$$\mathcal{V}(\mho_j^k) = \mathcal{V}^1(\Lambda_j^k) = \psi_j^k\,, \quad \Im\mathcal{W}(\mho_j^k) = \Im\mathcal{W}^1(\Lambda_j^k) = \Delta I_j^k .$$

Moreover $\mathcal{V}^1(A), \mathcal{W}^1(A) \in H_0$ for all $A \in \mathcal{A}$, while the presence of fruits introduces non analytic terms.

2.1 Holomorphic properties of tree representations

2.1.1 Linear operators on trees,

To establish a correspondence between each function $\psi_j^k(t)$ and a vector of $\mathbb{V}(\mathcal{A}_j)$, let us first write the functions F_j^k explicitly (using well known formulas on the derivatives of composite functions):

$$F_j^k = -\sum_{\vec{m}\in\mathbb{N}_0^n} (\nabla^{\vec{m}+e_j} f(t)) \sum_{\{p_j^h\}_{\vec{m},k-1}} \prod_{\substack{j=0\\h=1}}^{n,k-1} \frac{1}{p_j^h!} (\psi_j^h)^{p_j^h} -$$

$$\delta_{j0} \sum_{n\geq 2} (d^n \sin q(t)) \sum_{\{p_h\}_{n,k}} \prod_{h=1}^{k-1} \frac{1}{t_h!} (\psi_0^h)^{p^h}$$

where $\{p_j^h\}_{\vec{m},k}$ is a list of numbers in $\mathbb{N}_0 \equiv \mathbb{N} \cup \{0\}$ and respect the relation

$$\sum_h p_j^h = m_j\,, \quad \sum_{j,h} h p_j^h = k;$$

similarly $\{p_h\}_{n,k}$ is a list of numbers in $\mathbb{N}_0$ such that $\sum_h p_h = n$, $\sum_h h p_h = k$, finally

$$\nabla^{\vec{m}} f(t) = [\prod_{j=0}^{n} \partial_{\psi_j}^{m_j} f(\psi)]_{\substack{\psi_i=\varphi+\omega_i t\\ \psi_0=q_0(t)}}\,, \quad d^n g(q(t)) = \frac{d^n}{d\psi_0} g(\psi_0)|_{\psi_0=q_0(t)}.$$

We have that:

$$\psi_j^k(t) = a_j O_j^t(F_j^k) = a_j \left[Q_j(F_j^k) + \frac{1}{2}\sum_{i=0,1}(\Im(x_j^i F_j^k)\right]\,, \qquad \Delta I_j^k(t=0) = \Im(F_j^k).$$

Given a list $\{p_j^h\}$ we set:

$$P\{p_j^h\} = \prod p_j^h!\,, \quad m_i = \sum_h p_i^h\,, \quad f^1 = f\,, \quad f^0 = cos(q)\,, \quad M = \sum_{i,h} p_i^h\,;$$

and define the multi-linear functions on $x_i \in H$:

$$F^{\delta,j}_{\{p_i^h\}}(x_1, x_M) = (-1)^\delta \nabla^{\vec{m}+e_j} f^\delta \prod x_i.$$

Notice that $F^{\delta,j}_{\{p_i^h\}} = 0$ if $\delta = 0$ and $j \neq 0$ as in that case $\partial_j f^\delta = 0$. We can write:

$$\psi_j^k = a_j O_j[\sum_{\delta=0,1} \sum_{\{p_i^h\}_{k-\delta}} \frac{1}{P(\{p_i^h\})} F^{\delta,j}_{\{p_i^h\}}(\bar{x}_1, \cdots, \bar{x}_M)]$$

where $\bar{x}_1 = \cdots = \bar{x}_{t_0^1} = \psi_0^1(t)$,$\bar{x}_{t_0^1+1} = \cdots = \bar{x}_{t_0^1+t_1^1} = \psi_1^1(t) \cdots$ and the ordering is arbitrary.

We now construct the linear functions

$$\mathcal{V}_\varphi : \mathbb{V}(\cup_{j=0}^n \mathcal{A}_j) \to H\,, \quad \mathcal{W}_\varphi : \mathbb{V}(\mathcal{A}) \to H$$

such that for each $j \in \{0, 1, \ldots, n\}$ and for each k there is a unique $\mho_j^k \in \mathbb{V}_j^k$ such that $\psi_j^k(t,\varphi) = \mathcal{V}_\varphi(\mho_j^k)$ and $\Delta I_J^k = \Im \mathcal{W}_\varphi(\mho_j^k)$.

The function $\mathcal{W}$ is defined recursively on the the finite sets $\overset{m}{\mathcal{A}}{}^h$ and then extended to $\mathcal{A}$ via the mark adding functions and to $\mathbb{V}(\mathcal{A})$ by linearity. $\mathcal{V}$ is directly defined on the sets $\mathcal{A}_j$.

First we define the functions on trees of order one $\mathcal{W}(\overset{m}{\mathcal{A}}{}^1)$ and $\mathcal{V}(\mathcal{A}_j^1)$.

Remember that $\overset{m}{\mathcal{A}}{}^1$ is the tree: $\mho^1 = \overset{\delta=1}{\circ}$ and $\mathcal{A}_j^1$ the tree $\mho_j^1 = \partial_j^{v_0} \mho^1$

$$\mathcal{W}_\varphi(\mho^1) = -(\eta)(f^1(q(t), \varphi + \tilde{\omega} t))\,, \;\; \mathcal{V}_\varphi(\mho_j^1) = -(\eta) a_j Q_j(\nabla^{e_j} f^1(q(t), \varphi + \tilde{\omega} t)) \quad (2.1)$$

Then, using Remark1.36, we see that by setting:

$$\mathcal{V}_\varphi\big[\partial_j^{v_0}\Gamma^\delta_{\{p_i^h\}}(\{A_l\})\big) = a_j O_j(-\eta)^\delta(\nabla^{\vec{m}+e_j} f^\delta \prod_l \mathcal{V}(A_l)\big],$$

for each list $\delta = 0,1$ $j = 0,n$ and $\{A_l\} \in \cup_j \mathcal{A}_j$, we can define $\mathcal{V}$ recursively on all the $\mathcal{A}_j^k$. In the following we will omit the initial data φ whenever it is possible.
Namely the value $\mathcal{V}$ of a tree $A \in \mathcal{A}_j$ is found recursively from the value of its level one subtrees[1] $A^{\geq v}$.

We have seen in the preceding section that we can obtain $\mathcal{A}$ from $\overset{m}{\mathcal{A}}$ by successively adding marks, so given a tree with no marks on the first node we add the marks $j_1,\dots,j_l$ $h_1(v_0,t),\dots,h_l(v_0,t)$ and set:

$$\mathcal{W}(\prod_{i=1}^{l} h_i(v_0,t)\partial_{j_i}^{v_0}\mho^1) = (-\eta)\prod_i h_i(v_A,t)\nabla^{\sum_i e_{j_i}} f^1,$$

$$\mathcal{W}(\prod_{i=1}^{l}\partial_{j_i}^{v_0} h_i(v_0,t)A) \equiv (-\eta)^{\delta_{v_0}}\prod_i h_i(v_0,t)(\nabla^{\vec{m}(v_0)+\sum_i e_{j_i}} f^\delta)\prod_{v\in s(v_0)} a_{j_v} O_{j_v}[\mathcal{W}(A^{\geq v})],$$

where $\vec{m}_i(v)$ is the number of nodes $v' \in s(v)$ having $j_{v'} = i$. This extends $\mathcal{W}$ to $\mathcal{A}$.

Definition 2.1. *We define recursively the vectors $\mho_j^k$ that we will prove to be in correspondence with ψ_j^k:*

$$\mho^k = \sum_{\delta=0,1}\sum_{\{p_i^h\}_{k-\delta}}\frac{1}{P\{p_i^h\}}\Gamma^\delta_{\{p_i^h\}}(\underbrace{\mho_0^1,\dots,\mho_0^1}_{p_0^1},\mho_0^2,\dots,\mho_n^{k-1});\quad \mho_j^k = \partial_j^{v_0}\mho^k.$$

notice that each $\mho_j^h$ appears p_j^h times.

The definition immediately implies

$$\mathcal{V}(\mho_j^k) = (\eta)^k\psi_j^k.$$

Proposition 2.2 (Determination of $\mho_j^k$). *For each j,k*

$$\mho_j^k = \sum_{A\in\mathcal{A}_j^k}\frac{1}{|\mathcal{S}(A)|}A \equiv \sum_{A\in\mathcal{A}_j^k} c(A)A, \tag{2.2}$$

where the sum $A \in A_j^k$ means choosing one representative from each equivalence class of $\mathcal{A}_j^k$ (clearly this is well defined as $|\mathcal{S}(A)|$ does not depend on the chosen representative).

[1] remember that, if v is a node of level one, we consider $A^{\geq v} \in \mathcal{A}_*$ and the angle label j_v becomes an angle-marking with function-marking $h(v,t) = 1$.

Proof. We proceed by induction. The assertion is trivially true for $\mho_j^1$ so we suppose it true for all i and $\forall h < k$. By Definition 2.1 and Proposition 1.37 $\mho_j^k$ is the sum of all the trees in $\mathcal{A}_j^k$ and we only have to prove that the coefficients are those of expression (2.2).

Given a tree $A \in \mathcal{A}_j^k$ let v_1 v_m be its level one nodes and $A^1, \cdots A^m$ its level one subtrees; by the definition of $\mho$ we have to prove that:

$$\frac{1}{|\mathcal{S}(A)|} = \frac{N(A^1, \cdots, A^m)}{P\{p_i^h(A)\}} \prod_{i=1}^{m} \frac{1}{|\mathcal{S}(A^i)|}$$

where $\{p_i^h(A)\}$ is the number of trees $\{A^j\}$ in $\mathcal{A}_i^h$ and $N(A^1, \cdots, A^m)$ is the number of ways in which one can choose one summand form each $\mho_0^1, \ldots, \mho_n^{k-1}$ and obtain the unordered list $(A^1, \cdots, A^m)$.

Now if $m[v_i]$ is the cardinality of the orbit of v_i (so there are $m[v_1]$ subtrees equal to A^1 ...),

$$N(A^1, \cdots, A^m) = \frac{P\{p_i^h(A)\}}{\prod_{[v]_1} m[v]!} \quad \text{and} \quad \prod_{i=1}^{m} \frac{1}{|\mathcal{S}(A^i)|} = \prod_{[v]_1} \frac{1}{|\mathcal{S}(A^{\geq [v]})^{m[v]}|}.$$

□

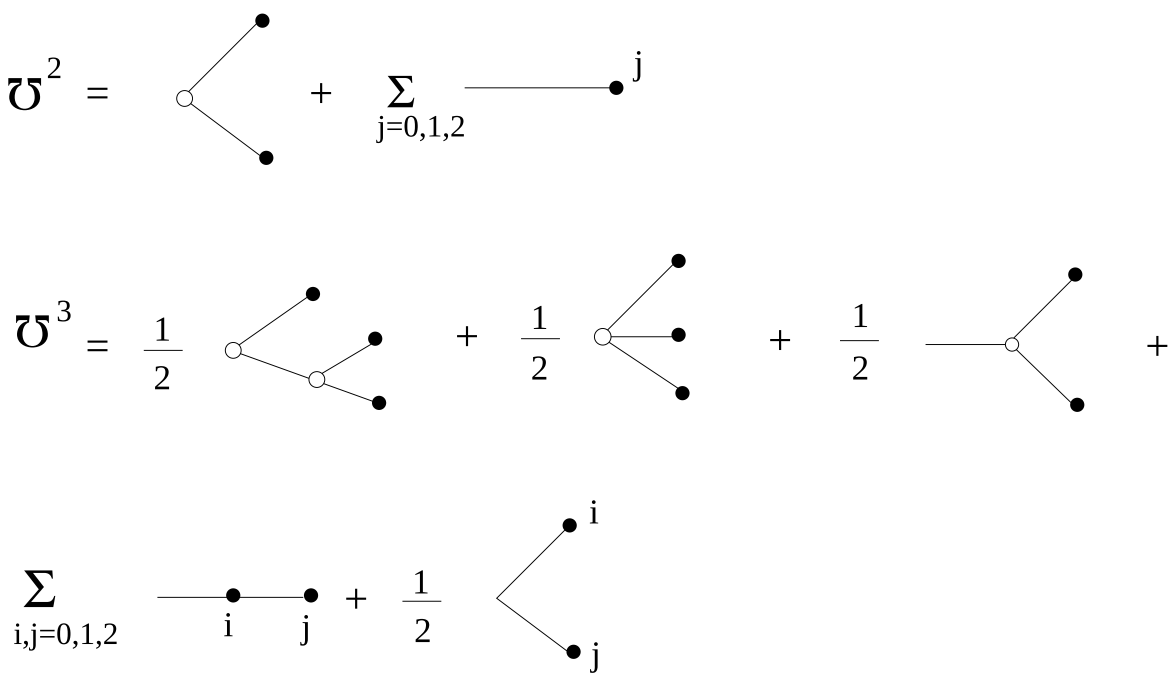


Figure 2.1: the vectors $\mho_0^2$ and $\mho_0^3$, as in Figure1.10 some of the labels j are left free, they can be equal to $(0, 1, 2)$, in any case $\delta = 1$

Notice that

$$\mathcal{V}_\varphi(A) = \mathcal{V}_\varphi(B) \text{ if and only if } A \cong B,$$

so we will always consider isomorphic trees as equal and make no difference between the tree and its diagram.

To compute the expansion of the action variables we use Expression (1.30). It is easily seen that the splitting vector of order k $(\eta)^k \Delta I_j^k$ can be expressed as the value $\Im \mathcal{W}_\varphi$ of $\mho_j^k$.

2.1.2 Holomorphic and non holomorphic contributions (trees with and without fruits)

We have mentioned in Remark 1.15 that it can be useful to divide the series expansion of $\psi(\eta, t, \varphi)$ in an analytic part, due to successive applications of Q_j and a part not in H_0 due to the appearance of the operators R_j^i. To represent this choice we use the full space $\mathcal{T}$. In particular the fruits will represent the choice of one of the R_j^i.
We set:

$$\psi_j^k = a_j O_j(F_j^k) = a_j Q_j F_j^k + \sum_{i=0,1} x_j^{[i]} G_j^{ik}, \quad \text{where } G_j^{ik} = \frac{1}{2} a_j \Im x_j^i F_j^k. \tag{2.3}$$

Then as in the preceding subsection we define a vector in $\mathbb{V}(\mathcal{T}^k)$ which we want to set in correspondence with the angles ψ_j^k.

Definition 2.3. *We define recursively* $\overset{0}{\Lambda}{}_j^1 = \mho_j^1$,

$$\overset{0}{\Lambda}{}_k^j = \sum_{\delta=1,0} \sum_{\{p_j^h\}_{\vec{m},k-\delta}} \frac{1}{P\{p_j^h\}} \partial_j^{v_0} \Gamma^\delta_{\{p_j^h\}}(\Lambda_0^1, \ldots, \Lambda_n^{k-1})$$

and finally

$$\Lambda_j^k = \sum_{i=0,1} \mathcal{F}_j^{ik} + \overset{0}{\Lambda}{}_j^k.$$

Proposition 2.4. *As in Proposition 2.2:*

$$\overset{0}{\Lambda}{}_j^k = \sum_{A \in \overset{0}{\mathcal{T}}{}_j^k} \frac{1}{|\mathcal{S}(A)|} A$$

Proof. We proceed by induction:

$$\Lambda_1^j = \mho_j^1 + \sum_{i=0,1} \mathcal{F}_j^{i,1}$$

verifies the Proposition, so we suppose it to be true $\forall\, j = 0, n$ and $\forall\, h < k$. By the inductive hypothesis:

$$\overset{0}{\Lambda}{}^{j}_{k} = \sum_{\delta=1,0} \sum_{\{p^h_j\}_{\vec{m},k-\delta}} \frac{1}{P\{p^h_j\}} \partial^{v_0}_j \Gamma^{\delta}_{\{p^h_j\}}(\Lambda^1_0, \ldots, \Lambda^{k-1}_n)$$

where the Λ^h_i $h \le k-1$ are in $\mathbb{V}(\mathcal{T}_i{}^h)$, and we proceed as in Proposition2.2 . $\square$

We now give a value $\mathcal{W}^1$ to trees in $\mathcal{T}$ and the define a function $\mathcal{V}^1$ so that $\mathcal{V}^1(\Lambda^k_j) = \psi^k_j$. As usual we proceed recursively on trees of increasing order and decorations: given a tree in $\overset{0}{\mathcal{T}}$ with no marks on the first node we add the marks $j_1, \ldots, j_l$, $h_1(t)$, $\ldots, h_l(t)$ and set:

$$\begin{aligned}
\mathcal{W}^1(\prod_{i=1}^{l} h_i(v_A,t)\partial^{v_A}_{j_i}\mho^1) &= (-\eta)\prod_i h_i(v_A,t)\nabla^{\sum_i e_{j_i}} f^1, \\
\mathcal{W}^1(\prod_{i=1}^{l} \partial^{v_A}_{j_i} h_i(v_A,t)A) &= \\
(-\eta)^{\delta_{v_A}} \prod_i h_i(v_A,t)(\nabla^{\vec{m}+\sum_i e_{j_i}} f^{\delta}) &\prod_{v\in s_0(v_A)} a_{j_v} Q_{j_v}[\mathcal{W}^1(A^{\ge v})] \prod_{v\in\mathcal{F}(v_A)} x^{[i_v]}_{j_v}(t) G^{i_v,k_v}_{j_v}, \\
G^{i,k}_j &= \frac{1}{2} a_j \Im x^i_j \mathcal{W}^1(\overset{0}{\Lambda}{}^k_j),
\end{aligned}$$

where $\vec{m}_i$ is the number of nodes in $s(v_0)$ having $j_v = i$ and $s_0(v)$ is the number of free nodes following v. As seen in the previous Section this defines $\mathcal{W}^1$ on $\overset{0}{\mathcal{T}}$.

Then we define the values $\mathcal{V}^1$, recursively on $\cup_j \overset{0}{\mathcal{T}}_j$ by setting:

$$\mathcal{V}^1(A) \equiv a_j Q_j[\mathcal{W}^1(A)].$$

for all $A \in \overset{0}{\mathcal{T}}_j$. Finally we extend the definition to fruits by setting:

$$\mathcal{V}^1(\mathcal{F}^{i,k}_j) = x^{[i]}_j G^{i,k}_j.$$

By Definition 2.3 the "value" of a fruit of order h label j and type i is as well :

$$-\frac{1}{2} a_j x^{[i]}_j \Im\{x^i_j \mathcal{W}[\mho^h_j]\}.$$

Notice that the fruits bring non analytic terms namely

$$\mathcal{V}^1(\mathcal{F}^{0,k}_j) \notin H_0\,, \quad \text{and } G^{1,k}_j = \Im x^1_j \mathcal{W}^1(\overset{0}{\Lambda}{}^k_j),$$

with $x^1_j \mathcal{W}^1(\overset{0}{\Lambda}{}^k_j) \notin H_0$. In general, for trees in $\overset{0}{\mathcal{T}}$, it is useful to consider the following function

$$\Psi_{\varphi}(A) = \prod_{v\in \overset{0}{A}} (-\frac{1}{2}\eta)^{\delta_v} \nabla^{\sum_{j=0}^n m_v(j)e_j} f^{\delta_v} \prod_{\alpha\in\mathcal{F}(v)} x^{[i_\alpha]}_{j_\alpha} \prod_{\beta\in\mathcal{M}(v)} h_\beta(v,\tau_v) w(\tau_w,\tau_v) \prod_{\alpha\in\mathcal{F}(A)} G^{o(\alpha),i(\alpha)}_{j(\alpha)},$$

$\mathcal{F}(v)$ are the fruits stemming from v, $\mathcal{M}(v)$ is the list of markings of the node v and finally $m_v(j)$ is the number of elements in $\{v, s_0(v), \mathcal{F}(v), \mathcal{M}(v)\}$ having angle label (or angle marking) equal to j. We write $s_0(v), \mathcal{F}(v)$ instead of $s(v)$ to remark that the fruits are not considered proper nodes. Notice that $\Psi^1_\varphi(A)$ contains the kernels of the integral operators Q_j so that $\mathcal{W}$ is obtained by "integrating" on the times τ_v $v > v_0$.

$$\mathcal{W}^1_\varphi(A) = 2\prod_{v>v_0} \Im^{\tau_w}_+ + \Im^{\tau_w}_- \Psi_\varphi(A) \equiv O \circ \Psi_\varphi(A).$$

Remark 2.5. *For the splitting vector we have $j = 1, n$ and:*

$$\Delta I^k_j \equiv 2a_j^{-1} G^{0k}_j = \Im \mathcal{W}^1[\Lambda^k_j]. \tag{2.4}$$

The angles ψ^k_j for $j = 0, n$ are:

$$\psi^k_j \equiv \mathcal{V}^1(\Lambda^k_j)) = a_j(\Im^t_+ + \Im^t_-) w_j(t, \tau_{v_0}) \mathcal{W}^1(\Lambda^k_j) \tag{2.5}$$

2.2 Equivalent trees and cancellations

In this section we use the tree expansion to prove properties of the homoclinic trajectory and of the homoclinic splitting matrix.
First let us define some particular vectors in $\mathbb{V}(\overset{0}{\mathcal{T}})$.

Definition 2.6. *For $i, j = 0, n$, and given $f(t), h(t) \in H$, we define for any $k \in \mathbb{N}$:*

$$\mathbb{V}(\mathcal{T}^k_{(i,f)}) \ni \Lambda^k_{(i,f)} = \sum_{A \in \overset{0}{\mathcal{T}}{}^k} c(A) f(v_0, t) \partial_i^{v_A} A = \sum_{A \in \mathcal{T}^k_{(i,f)}} c(A) A,$$

$$\mathbb{V}(\mathcal{A}^k_{(i,f)}) \ni \mho^k_{(i,f)} = \sum_{A \in \mathcal{A}} c(A) f(v_0, t) \partial_i^{v_A} A = \sum_{A \in \mathcal{A}_{(i,f)}} c(A) A,$$

$$\mathbb{V}(\mathcal{T}^k_{(i,f),(j,h)}) \ni \Lambda^k_{(i,f),(j,h)} = \sum_{A \in \overset{0}{\mathcal{T}}{}^k} \sum_{[v] \in A} c(A, v) f(v_A, t) \partial_i^{v_A} h(v, t) \partial_j^v A = \sum_{A \in \mathcal{T}^k_{(i,f),(j,h)}} c(A) A,$$

$$\mathbb{V}(\mathcal{A}^k_{(i,f),(j,h)}) \ni \mho^k_{(i,f),(j,h)} = \sum_{A \in \mathcal{A}^k} \sum_{[v] \in A} c(A, v) f(v_A, t) \partial_i^{v_A} h(v, t) \partial_j^v A = \sum_{A \in \mathcal{A}^k_{(i,f),(j,h)}} c(A) A,$$

where

$$c(A) = \frac{1}{\mathcal{S}(A)}, \quad c(A, v) = \frac{m[v]}{\mathcal{S}(A)}$$

for all labeled trees A. By convention we will omit the marking function if it is equal to one.

Notice that

$$\Im\mathcal{W}_\varphi(\mho_j^k) = \Im\mathcal{W}^1_\varphi(\Lambda_j^k) = \Delta I_j^k(\varphi)\,, \quad \Im\mathcal{W}_\varphi(\mho_{j,i}^k) = \Im\mathcal{W}^1_\varphi(\Lambda_{j,i}^k) = \partial_{\varphi_i}\Delta I_j^k(\varphi).$$

Moreover as we said before $\Delta I_j^k = 2G_j^{0k}$.
Now we have set up all the necessary formalism to study the cancellations in the series for the vector $\Delta I_j^k(\varphi)$ and its Jacobian matrix $\partial_{\varphi_i}\Delta I_j^k(\varphi)$. The cancellations occur because the applications $\Im\mathcal{W}^1$ or equivalently $\Im\mathcal{W}$ (defined on $\mathbb{V}(\mathcal{A})$) are clearly not injective so that apparently different trees can give the same contribution.
We have introduced all this formalism on the trees to be able to identify cancellations directly in the formal space of trees $\mathbb{V}$. We have considered trees modulo isomorphism, now we add identities due to the dynamics.

Definition 2.7. *Given two trees A,B in* $\overset{0}{\mathcal{T}}$ *we set*

$$A = B \leftrightarrow A - B \in \ker \Im\mathcal{W}^1,$$

notice that isomorphic trees are equal.
This equality can hold for all initial data φ *or only for some special values in the latter case we will set* $A = B(\varphi = \bar{\varphi})$. *The same reasoning can be done with the operator* $\Im\mathcal{W}$ *in* $\mathcal{A}$.

Remark 2.8. *Notice that by our definition of equivalent trees adding a fruit of order* k *type* i *and angle* j *in the free node* v *of a tree* $A \in \overset{0}{\mathcal{T}}$ *is equivalent to adding a mark* $x_j^{[i]}(t)\partial_j^v$ *to the node* v *and multiplying by the* η *and* φ *dependent function* G_j^{ik}.

The cancellations between trees are due to the symmetries of the Q_j and O_j operators that we evidenced in Proposition1.16, we will write them again schematically:

a) The operators Q_j and O_j are symmetric; given F and $G \in H$

$$\Im F(t)Q_j(G(\tau)) = \Im G(t)Q_j(F(\tau))$$
$$\Im F(t)O_j(G(\tau)) = \Im G(t)O_j(F(\tau))$$

b) The operator Q_j preserves the parity; moreover $\Im f = 0$ if $f \in H$ is odd.

c)Given two continuous functions $F, G \in H$ if $\pi_P FG \neq c$ holds then:

$$\Im^T G(t)d_tF(t) = F(T)G(T)) - \Im^T F(t)d_tG(t).$$

d) By energy conservation the stable and unstable manifolds are on the same energy level.

Each of these properties brings some cancellations, we will first check those coming from property (b), as they are the simplest ones:

Lemma 2.9. *for each* j, k *and for any even function* $f(t)$*:*

$$\mathbb{V}(\mathcal{A}^k_{(j,f)}) \in \ker \Im\mathcal{W}_{\varphi=0}.$$

In the same way

$$\mathbb{V}(\mathcal{T}^k_{(j,f)}) \in \ker \Im\mathcal{W}^1_{\varphi=0}.$$

Proof. By Proposition1.16 (b) we only need to prove that $\mathcal{W}_{\varphi=0}(A)$ is odd for all $A \in \mathcal{A}^k_{(j,f)}$.

We proceed by induction

$$\mathcal{W}(\mho^1_{(j,f)}) = f(t)\nabla^{e_j} f^1(\tilde{\omega}t, q_0(t))$$

which is odd as f and f^1 are even. If $k > 1$ then call $l(A) > 0$ the number of subtrees of level one. $\mathcal{W}(A)$ is the product of $l(A)$ odd functions times f^1 derived $l(A)+1$ times so it is odd.
Then, for each j, k, the function

$$G^{h_\alpha}_{j_\alpha,k_\alpha}(\psi = 0) = \Im\mathcal{W}(x^i_j \mho^k_j) = 0$$

as it is the integral of an odd function. So in $\overset{0}{\mathcal{T}}$ all the trees with fruits have zero value. Finally if $A \in \overset{0}{\mathcal{T}}$ is fruitless then $\mathcal{W}^1(A)$ is odd. □

Theorem 2.10. *[homoclinic intersection] The stable and unstable manifold intersect at* $q = \pi, \psi = 0$.

Proof. the distance between stable and unstable manifold at $q = \pi, \psi = \varphi$ is:

$$\sum_{j=1}^{n}\sum_{k=1}^{\infty}(\eta)^k |\Delta I^k_j(\varphi)| = \sum_{j=1}^{n}\sum_{k=1}^{\infty}(\eta)^k |a_j \Im\mathcal{W}^1_\varphi(\Lambda^k_j)|.$$

□

Another important feature for identifying cancellations is the symmetry with respect to changes of the first node.

Lemma 2.11. *By Proposition1.16(a) we have:*

$$\begin{aligned} &\forall A \in \overset{r}{\mathcal{T}},\ \forall v \in A: \quad P(A,v) - A \in \ker \Im\mathcal{W}^1_\varphi \\ &\forall A \in \overset{r}{\mathcal{T}}_{(j,f)(i,h)}: \quad P_1(A) - A \in \ker \Im\mathcal{W}^1_\varphi \end{aligned} \tag{2.6}$$

for the same reasons:

$$\begin{aligned} &\forall A \in \overset{r}{\mathcal{A}},\ \forall v \in A: \quad P(A,v) - A \in \ker \Im\mathcal{W}_\varphi \\ &\forall A \in \overset{r}{\mathcal{A}}_{(j,f)(i,h)}: \quad P_1(A) - A \in \ker \Im\mathcal{W}_\varphi. \end{aligned} \tag{2.7}$$

Proof. Notice that given a tree A and one of its nodes v if $w \in \mathcal{P}(v_A, v)$ then:

$$P(A,v) = P(P(A,w),v),$$

so that we only need to prove the assertion for $v \in s(v_A)$. Given $A \in \overset{r}{\mathcal{T}}$ and $v \in s(v_A)$ such that $j_v = j$ we compare: $\Im\mathcal{W}^1(A)$ and $\Im\mathcal{W}^1(B)$ with $B = P(A, v)$, so B has first node v (no label j_v) and a node v_A in $s(v)$ with $j_{v_A} = j$.

$$\Im\mathcal{W}^1(A) = (-\eta)^{\delta_{v_A}} \Im\nabla^{\sum_j m_{v_A}(j)e_j} f^{\delta_{v_A}} \prod_{\substack{w\in s(v_A)\\ w\neq v}} Q_{j_w}[\mathcal{W}^1(A^{\geq w})]$$
$$Q_j[(-\eta)^{\delta_v}\nabla^{\sum_j m_v(j)e_j} f^{\delta_v} \prod_{w_1\in s(v)} \mathcal{W}^1(A^{\geq w_1})],$$

which by the symmetry of Q_j is equal to

$$\Im\nabla^{\sum_j m_v(j)e_j}(-\eta)^{\delta_v} f^{\delta_v} \prod_{w_1\in s(v)} \mathcal{W}^1(A^{\geq w_1}) Q_j[(-\eta)^{\delta_{v_A}}\nabla^{\sum_j m_{v_A}(j)e_j} f^{\delta_{v_A}} \prod_{\substack{w\in s(v_A)\\ w\neq v}} Q_{j_w}[\mathcal{W}^1(A^{\geq w})]].$$

This is the value of B, namely, both in A and in B, $m_v(i)$ with $i \neq j$ is the number of elements in $(s(v), \mathcal{M}(v), \mathcal{F}(v))$ having label i and $m_v(j) - 1$ is the number of elements in $(s(v), \mathcal{M}(v), \mathcal{F}(v))$ having label j.

□

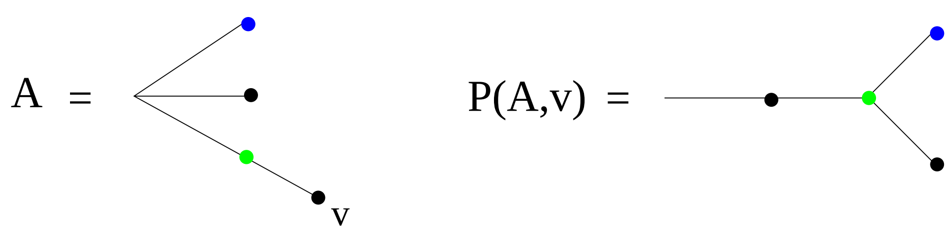


Figure 2.2: An example of trees that are equivalent by changing the first node

Example 2.12. let A be the tree in Figure2.2:

$$\mathcal{W}^1_\varphi(A) = \nabla^{e_1+e_2+e_0} f^1(\tau_0) Q_1(\tau_0,\tau_1)[\nabla^{e_1} f^1(\tau_1)] Q_0(\tau_0,\tau_2)[\nabla^{e_0} f^1(\tau_2)]$$
$$Q_2(\tau_0,\tau_3)\big[\nabla^{e_2+e_0} f^1(\tau_3) Q_0(\tau_3,\tau_4)[\nabla^{e_0} f^1(\tau_4)]\big]$$

while

$$\mathcal{W}^1_\varphi(P(A,v)) = \nabla^{e_0} f^1(\tau_0) Q_0(\tau_0,\tau_1)\Big[\nabla^{e_2+e_0} f^1(\tau_1) Q_2(\tau_1,\tau_2)\big[\nabla^{e_1+e_2+e_0} f^1(\tau_2)$$
$$Q_1(\tau_2,\tau_3)[\nabla^{e_1} f^1(\tau_3)] Q_0(\tau_2,\tau_4)[\nabla^{e_0} f^1(\tau_4)]\big]\Big]$$

so we apply repeatedly the Proposition 1.16(b).

We have seen that many trees in $\overset{0}{\mathcal{T}}$ are equivalent; we will concentrate on relations for the vectors $\mho_{i,j}$ and $\Lambda_{i,j}$. Let us summarize some properties of the coefficients $c(A,v)$.

Consider $A \in \overset{0}{\mathcal{T}}$ and let $\mathtt{A}$ be the rootless tree associated to A. By the Lagrange theorem if v_A is the first node of A and $\mathcal{S}(\mathtt{A}, v_A)$ is the stabilizer of v_A in $\mathtt{A}$ then

$$|\mathcal{S}(A)| = |\mathcal{S}(\mathtt{A}, v_A)| = |\mathcal{S}(\mathtt{A})|$$

as v_A is the only uncolored node of $\mathtt{A}$.

Lemma 2.13. • *(i) let $[v]$ be the cosets of v by the action of $\mathcal{S}(A)$ and $m(v) = |[v]|$:*

$$\sum_v \partial_i^v \frac{1}{\mathcal{S}(A)} A = \sum_{[v]} \frac{m[v]}{|\mathcal{S}(\mathtt{A})|} \partial_i^v A$$

the sum $[v]$ means choosing a term from each coset to obtain summands that are all different.

The coefficient $c(A,v) \equiv \frac{m[v]}{|\mathcal{S}(\mathtt{A})|}$ is the cardinality of the subgroup $\mathcal{S}(v_A, v)$ of $\mathcal{S}(\mathtt{A})$ that fixes v_A and v;

- *(ii) This subgroup fixes all the nodes of the path $\mathcal{P}(v_A, v)$ and so does not depend on the labels of the nodes on the path. So given a tree A and a node v*

$$c(A,v) = c(P(A,v), v_A)$$

Proof. • (i) we group the identical terms in $\sum_v c(A)\partial_i^v A$ corresponding to nodes in the same coset $[v]$ of A so we have $m[v]$ terms for each coset $[v]$.

- (ii) first we note that $\mathcal{S}(\mathtt{A})$ sends adjoint nodes in adjoint nodes so for each permutation $\sigma \in \S(v_A, v)$ and for each v_i in the path $\mathcal{P}_{v_A,v}$ (that have length m) $\sigma(v_i) = w_i$ is adjoint to $w_{i+1} = \sigma(v_{i+1})$. Now as by definition σ fixes $w_0 = v_0 = v_A$ and $w_m = v_m = v$ the list of nodes $\{w_i\}_{i=0}^m$ is a path joining v_A to v. In a tree the paths are unique so $w_i = v_i$ for each $i \leq m$.

□

This Lemma and Remarks 1.43 imply the following identities on the vectors $\Lambda_{i,j}$ and $\mho_{i,j}$. we write them explicitly only for Λ.

Proposition 2.14. *For $i,j \neq 0$, $f,h \in H$ and for each k the following equality holds: $\Lambda^k_{(i,f)\,(j,h)} = \Lambda^k_{(j,h)\,(i,f)}$.*

Proof. We have seen that $\mathcal{A}^k_{(i,f)\,(j,h)} \leftrightarrow \mathcal{A}^k_{(j,h)\,(i,f)}$ and that the trees in correspondence have the same value. We only need to prove that the corresponding summands in $\Lambda^k_{(i,f)\,(j,h)}$ and $\Lambda^k_{(j,h)\,(i,f)}$, have the same coefficient; this follows from Lemma2.13(ii). Namely given A in $\mathcal{A}^k_{(i,f)\,(j,h)}$ then $\mathcal{S}(A)$ is the stabilizer of the two marked nodes of A, then $c(A) = c(P_1(A))$. □

Proposition 2.15. *Given any two functions in H: $h(t)$ and $f(t)$, for each k and for each $i = 1, n$ we have:*

$$\Lambda^k_{(0,h)\,(i,f)} = \Lambda^k_{(i,f)\,(0,h)} + L_h(\Lambda^k_{(i,f)})$$

Proof. As in the preceding Proposition $\overset{r}{\Lambda}{}^k_{(0,h)\,(i,f)} = \overset{r}{\Lambda}{}^k_{(i,f)\,(0,h)}$.
To prove that $\Lambda^{k0}_{(0,h)\,(i,f)} = L_h(\Lambda^k_{(i,f)})$ we show that the coefficient of corresponding summands are the same (we have seen that to each element of $\mathcal{A}^{k0}_{(0,h),(i,f)}$ there corresponds an unique summand of $L_h(\Lambda^k_{(i,f)})$). Given a summand (tree A) of $\Lambda^{k0}_{(0,h)\,(i,f)}$ marked in the node v, its coefficient $c(A)$ is the inverse of the cardinality of the stabilizer of v_A, v. If v_1 is the only[2] node following v_A and not in $\mathcal{P}(v_A, v)$ then $|\mathcal{S}(A)|$ is as well the cardinality of stabilizer of the the path joining v_1 to v (which passes by v_A by definition). So it has the same coefficient as $P_1\bar{g}_{\alpha_h}(B, v_1)$ where B is the tree , having first node v, such that $P_1\bar{g}_{\alpha_h}(B, v_1) = A$. $\square$

[2]it is unique as $A \in \mathcal{A}^{k0}_{0_h\,(i,f)}$

Chapter 3

Basic estimates on tree expansions

We prove upper bounds on the value of trees of order k. In particular our bounds on $\mathcal{W}^1(A)$ are exponentially small for all $A \in \mathcal{A}$ (we will call these the "analytic bounds"). Upper bounds on $\mathcal{W}^1(T)$ for $T \in \mathcal{T}$ are derived more or less in the same way as in [G1]; notice however that we do not request that $f(\psi, q$ is a trigonometric polynomial. We also consider bounds on the values $\mathbb{V}^1$ of fruitless trees, which will be useful in Chapter 5. Moreover in Section 3.2 we will prove some technical lemmas on asymptotic power series which will be useful in Chapters 4 and 5.

In Chapter 2 we have introduced a tree representation for the series expansion of $\psi_j(\varphi, \eta)$ and $I_j(\varphi, \eta)$. The KAM theorem 1.2 guarantees the convergence of this two series and of the splitting matrix. So we can consider the series:

$$\Lambda_j = \sum_{k \geq 1} \Lambda_j^k, \quad \text{and the functions} \quad \mathcal{V}_\varphi^1(\Lambda_j) \quad \text{and} \quad \Im \mathcal{W}_\varphi^1(\Lambda_j),$$

are well defined smooth functions of η by the KAM theorem. We would like to consider series of the type:

$$\sum_{\alpha \in I} c(A_\alpha) A_\alpha,$$

where I is a numerable set and the $A_\alpha \in \mathcal{T}$. For such series we have no guarantee of the convergence of the corresponding values. We will consider them as formal series and write identities between the formal series which are true term by term. Such identities will be written as $A \sim B$.

In this chapter we prove that such formal series are polynomial asymptotic series in η, ε.

Definition 3.1. *A formal power series*

$$x = \sum_k (\eta)^k x_k(\varepsilon)$$

is polynomially asymptotic in η, ε if there exists a neighborhood of $\varepsilon = 0$ where for any $q \in \mathbb{N}$ there exists $p(q)$ such that:

$$x_k(\varepsilon) \leq \varepsilon^{-p(q)k}, \quad \forall k \leq \varepsilon^{-q}, \ \forall \varepsilon \neq 0.$$

3.1 Upper bounds on the values of trees

Given a fruitless tree $A \in \overset{m}{\mathcal{A}}$ of order k (so with at most $2k-1$ nodes), its value through $\Im \mathcal{W}^1_\varphi$ is of the form:

$$2\Im \prod_{v>v_0} (\Im_+^{\tau_w} + \Im_-^{\tau_w})(-\frac{1}{2})^{N(A)}(\eta)^{\delta_{v_0}} \nabla^{\sum_j m_{v_0}(j)e_j} f^{\delta_{v_0}}$$
$$\prod_{v>v_0} (\eta)^{\delta_v} \nabla^{\sum_j m_v(j)e_j} f^{\delta_v} w(\tau_w, \tau_v) \quad \text{(a)}$$

Its value $\mathcal{V}^1$ is:

$$\prod_{v\geq v_0} (\Im_+^{\tau_w} + \Im_-^{\tau_w})(\eta)^{\delta_v}(-\frac{1}{2})^{N(A)} \nabla^{\sum_j m_v(j)e_j} f^{\delta_v} w(\tau_w, \tau_v), \quad \text{(b)}$$

where w is the node preceding v and by convention: $\tau_{w_0} = t$.

Remember that, setting $x = e^{-|t|}$

$$w_j(t,\tau) = \sigma(t)x_j^1(t)x_j^0(\tau) - \sigma(\tau)x_j^0(t)x_j^1(\tau))$$
$$x_j^1 = \begin{cases} |t| & j \neq 0 \\ \frac{|t|x}{x^2+1} - \frac{1}{4}(x - x^{-1}) & j = 0 \end{cases} \qquad x_j^0 = \begin{cases} 1 & j \neq 0 \\ \frac{2x}{x^2+1} & j = 0. \end{cases} \quad (3.1)$$

And that the operators $\Im$ and $\Im_\rho^t$ are: $\Im = \Im^{0^-} - \Im^{0^+}$ and

$$\Im_+^t = \begin{cases} \Im^t & \text{if } t \geq 0 \\ \Im^{0^+} - \Im^{0^-} + \Im^t & \text{if } t \leq 0 \end{cases},$$

same for $\Im_-^t$.

We expand f^1 in Fourier series in the rotator angles,

$$f^1(\psi, q) = \sum_{|\nu|\leq N} e^{i\nu\cdot\psi} f_\nu(q),$$

so that each node has one more label $\nu_v \in \mathbb{Z}^n$. Be will represent as $A(\nu)$ a tree A with labels ν_v such that such that

$$\sum_{v\in A} \nu_v = \nu.$$

As A is fruitless $\mathcal{V}^1(A)$ depends on the initial data via the function $e^{i\varphi\cdot\nu}$.

In each node v with $\delta = 1$ we have as factor the function $d^{n_v} f_{\nu_v}(q(t))$ where $n_v = m_v(0)$. Moreover as $q(t) = 4\arctan(e^t)$ then $f_\nu(q(t)) = F_\nu(e^t)$ with $F(y)$ analytic in some strip around $y \geq 0$.

To find upper bounds on the trees one needs very few assumptions on the perturbating function f[1], we will consider some (not minimal) hypothesis that guarantee that the value of an integral of type a) on fruitless trees of total frequency ν and order k, with initial data $\varphi \in \mathbb{T}^n_{s_0}$ are bounded by

$$e^{s_0|\nu|}(k!)^{c_1}[P(\varepsilon, \varepsilon^{-1})]^k e^{-\frac{D}{\sqrt{\varepsilon}}|\omega\cdot\nu|}.$$

Where D is defined in Definition 1.28, $P(\varepsilon, \varepsilon^{-1})$ is a polynomial and we will fix s_0 of order one.

We prove t dependent bounds for the analytic integrals (b); this bounds will be useful in Chapter 5. We consider them here only because the proof is parallel to that of integrals (a). Notice however that in this context there is no guarantee that the values $\mathcal{V}^1$ of fruitless trees are bounded for $t \to \infty$ as such trees have no dynamical meaning. The bounds on integral (a) assure that the formal tree series we will consider in Chapter 4 are all asymptotic series.

The functions $f_\nu(q)$ are such that $F_\nu(e^t) \in H_0(a, D)$ (remember that a, D are those of definition 1.28). Naturally by our analyticity assumptions $f_\nu(q(t))$ is limited for $|t| \to \infty$ in $|\text{Im } t| < 2\pi$.
Notice that if $D < \pi/2$ the image of $C(a, D)$ via $q(t)$ is a compact region and that there exists ν such that $F_\nu(e^t)$ has singularities on the lines

$$|\text{Im } t| = D \text{ and } |\text{Re } t| = a.$$

Moreover as the image of $\mathbb{R} \times [-\pi/2, \pi/2]$ through $e^{iq(t)}$ is the Riemann sphere there must be a singularity $|\text{ Re } t| \leq \pi/2$.

Definition 3.2. *We consider the subset of $H_0(a, D)$:*

$$B(a, D) := \{f(\psi(t), q(t)) \in H_0(a, D) \; : \; \max_{t \in C(2a, D-\sqrt{\varepsilon})} |F_\nu(e^t)| \leq M\sqrt{\varepsilon}^{-p}\} \tag{3.2}$$

for some $p \in \mathbb{N}_0$.

In Appendix A.1 we will give various examples of functions $f(q, \psi)$, with essential singularities in q and satisfying this condition (even with $p = 0$).

Proposition 3.3. *(i)The functions $d_q^k f_\nu(q(t)) = F_\nu^k(e^t)$ are all in $H_0(a, D)$ if f is so moreover if f is in $B(a, D)$ then so are the F_ν^k and:*

$$\max_{t \in C(3a, D-2\sqrt{\varepsilon})} |F_\nu^k(e^t)| \leq k! M\sqrt{\varepsilon}^{-(p+k)}$$

[1]In Appendix A.6 we will prove non-optimal upper bounds for the tree expansion of the homoclinic trajectory for Hamiltonian (*)

Proof. The assertion is equivalent to proving that for any finite b and $d \neq 0, \pi/2$:

$$\min_{\substack{t,\tau \in C(b,d) \\ |t|,|\tau|<b \\ \mathrm{Im}\,(t-\tau)=k\ll 1}} |q(t)-q(\tau)| \geq A(b,d)k.$$

A direct computation of the minimum gives

$$A(b,d) \geq He^{-b} \tag{3.3}$$

for any ε independent d and big enough b. Then the image of $C(3a, D-2\sqrt{\varepsilon})$ through $q(t)$ is compact, contained in the image of $C(2a, D-\sqrt{\varepsilon})$ and the distance between the frontiers of the two sets is greater or equal to $e^{-2a}\sqrt{\varepsilon}$. We can use Cauchy estimates on the derivatives $\partial_0^k f_\nu(q(t)) = F_\nu^k(e^t)$.
We can prove 3.3 geometrically by noticing that, provided that b is big enough, the minimum distance $|q(t)-q(\tau)|$ is attained on the border $|t| = |\tau| = b$ (whose image in the variable q is a circle around $q=0$, for large enough b).

This is clearly seen in Picture3.1; to prove it one notices that $q(t)$ is convex, moreover if $t(\bar{q}), \tau(\bar{q})$ are such that Re $q(t) =$ Re $q(\tau) = \bar{q} \in (0,\pi]$ then the function $Im(q(t(\bar{q})) - q(\tau(\bar{q}))$ is strictly increasing in $(0,\pi]$. By triangulation this implies that the minimum distance is on the border i.e. on the image in$q$ space of $|t| = b$ which for large enough b is a circle around $q = 0$.

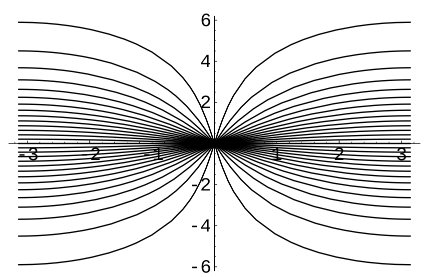


Figure 3.1:

Analogous reasonings can be applied to a generalized pendulum.
A more direct proof, valid only for functions having $D \neq \pi/2$, is the following. The functions $d_q^k f_\nu(q(t))$ are all limited, so we can bound them by $k!C^k$, with $C = O_\varepsilon(1)$, in the ε-independent domains $|\ \mathrm{Im}\ t| < 2\pi$, $|\ \mathrm{Re}\ t| > 3a$.
In the rectangles $|\ \mathrm{Im}\ t| < d - 2\sqrt{\varepsilon}$ $(d < \pi/2)$, $|\ \mathrm{Re}\ t| < 3a$ the application $t \to q(t)$ is conformal, let us call the inverse $T(q) = \log\tan(q/4)$. Then if[2] $g(t) = f(q(t))$:

$$d_q^k f_\nu(q) = d_q^k g \circ T = \sum_{\{p_h\}_k} \frac{1}{\prod_h p_h!} d_t^{\sum p_h} g(t)|_{t=T(q)} (\frac{d_q^h T(q)}{h!})^{p_h},$$

as $T(q)$ is (in a limited ε independent domain) not ε close to its singularities we set $\frac{d_q^h T(q)}{h!} \leq C^h$ for some order one C, then we bound $d_t^{\sum_h p_h} g(t)$ with $\varepsilon^{-(p+\sum_h p_h)/2}(\sum_h p_h)!$.
Finally we bound the sum:

$$\varepsilon^{-p/2} \sum_{\{p_h\}_k} \frac{(\sum_h p_h)! \varepsilon^{-(\sum_h p_h)/2}}{\prod_h p_h!} \leq \varepsilon^{-(p+k)/2} \sum_{\{p_h\}_k} \frac{(\sum_h p_h)!}{\prod_h p_h!} \leq C^k \varepsilon^{-(p+k)/2} 2^k k!$$

[2]As in Chapter 1 the symbol $\{p_h\}_k$ is a list of non-negative numbers p_h, $h \in \mathbb{N}$ such that $\sum_{h\geq 1} hp_h = k$.

as the sum in the middle term is the order k derivative, computed in zero, of $f \circ f$ where $f(x) = \frac{x}{1-x}$ for $x \in \mathbb{R}$. Notice that this proof holds true allso for a generalized pendulum[3]. □

If we restrict our attention to rational functions $F_\nu(e^t)$ and call t^i_ν their poles in $|\mathrm{Im}t| \le \pi$ (all with $\mathrm{Im}t \neq 0$) then:

$$D = \ \mathrm{Min}\ _{\nu, i \in [1, N(\nu)]} |\ \mathrm{Im}\ (t^i_\nu)|; \quad a = \max_{\nu, i\in[1,N(\nu)]} |\ \mathrm{Re}\ (t^i_\nu)|. \tag{3.4}$$

Moreover the following proposition holds.

Corollary 3.4. *(i) The functions $\partial^k_0 f_\nu(q(t)) = F^k_\nu(e^t)$ are all limited rational functions of e^t, whose poles are the same as those of $F^0_\nu(e^t)$. (ii) If the order of the pole y^i_ν is p^i_ν for F^0_ν then it is $p^i_\nu + k$ for F^k_ν (except for $\pm i\frac{\pi}{2}$ where it is always p^i_ν).*

Proof. (i) First we recall that limited rational functions can be decomposed in "partial fractions"(see [RU]) as:

$$F_\nu(y) = C + \sum_{i=1,\cdot,N_\nu} P_i(\frac{1}{y - y^i_\nu}),$$

where the polynomials P_i have no constant coefficient. Then as $f_\nu(q(t)) = F_\nu(e^t)$, we have

$$d_q f_\nu(q(t)) \cdot \dot{q}(t) = d_t f_\nu(q(t)) = d_t F_\nu(e^t),$$

and so $F^1_\nu(y) = \frac{1+y^2}{y} y d_y F_\nu(y)$. Now $d_y F(y) = \sum_{i=1,\cdot,N_\nu} P'_i(\frac{1}{y-y^i_\nu})$ is a sum of polynomials of degree greater or equal to two, so $(1+y^2)P'_i(\frac{1}{y-y^i_\nu})$ is limited and $F^1_\nu(y)$ admits the same kind of representation as $F(y)$ (it has obviously the same poles). For $k > 1$ we proceed recursively. (ii) The order of the pole y^i_ν is the degree of the corresponding polynomial.

□

Having fixed $\nu = \sum_v \nu_v$, in integral (a) we shift the integration to $\mathbb{R} + i\sigma(\omega_\nu)d$ where $d < D$, $\omega_\nu = \frac{\omega}{\sqrt{\varepsilon}} \cdot \nu$ and $\sigma(x)$ is the sign of x . As the functions are all analytic in $|\mathrm{Im}(t) \le d$ the integral (a) is unchanged.

In integral (b) we consider complex values of the time $t + id$ with $t, d \in \mathbb{R}$. Then we use Lemma 1.12 (ii) to shift the integration on the nodes.
Notice that in integral (a) we cannot choose the sign of the shift in the single node integrals and so we need to work in the (symmetric) domains $H(a, D)$ to guarantee the indifference of extending in the lower or upper half-plane. To simplify the notation we set

$$\sigma(\omega_\nu) = + \quad \text{and define} \quad E(d, \nu) = e^{-|\omega_\nu|d}.$$

[3]In the Appendix A.1 we will show that the only functions $f(q(t), \psi(t))$satisfying the bounds 3.2 and having only isolated singularities on$|\ \mathrm{Im}t| = \pi/2$ are rational functions, which obviously satisfy Proposition 3.3.

If A has k nodes with $\delta = 1$ let $\{\nu_v\}_\nu^k$ be the lists of k vectors $\nu_v \in \mathbb{Z}^n$ such that $\sum \nu_v = \nu$. The value of $A(\nu)$ (tree $A \in \overset{m}{\mathcal{A}}$ with total frequency ν) in integral (a) is:

$$(-\frac{1}{2})^{N(A)} e^{i\nu\cdot\varphi} E(d,\nu) \sum_{\{\nu_v\}_\nu^k} [\prod_{\substack{s=1,\ldots,n \\ \delta_v=1\,, v\geq v_0}} (i\nu_{v\,s})^{m_v(s)}]$$

$$\oint \frac{dR_{v_0}}{2i\pi R_{v_0}} \int_{-\infty}^{\infty} d\tau_{v_0} e^{-\sigma(\tau_{v_0})R_{v_0}} [d^{n_{v_0}} f_{\nu_{v_0}}^{\delta_v}(q(\tau_{v_0}+id))] e^{i\omega_v \tau_{v_0}}$$

$$\prod_{v>v_0} \oint \frac{dR_v}{2i\pi R_v} (\int_{-\infty}^{\tau_w} d\tau_v + \int_{\infty}^{\tau_w} d\tau_v) e^{-\sigma(\tau_v)R_v(\tau_v+id)} w_{j_v}(\tau_w+id, \tau_v+id)$$

$$\prod_{v\geq v_0} [d^{n_v} f_{\nu_v}^{\delta}(q(\tau_v+id))] e^{i\omega_v\tau_v}; \quad \text{(a)}$$

naturally $f_\nu^0 = 0$ for all non zero ν.
The same tree in integral (b), has value:

$$e^{\omega_\nu d}(-\frac{1}{2})^{N(A)} e^{i\nu\cdot\varphi} e^{d\omega_\nu} \sum_{\{\nu_v\}_\nu^k} [\prod_{\substack{s=1,\ldots,n \\ \delta_v=1\,, v\geq v_0}} (i\nu_{v\,s})^{m_v(s)}] \prod_{v\geq v_0} \oint \frac{dR_v}{2i\pi R_v}$$

$$(\int_{-\infty}^{\tau_w} d\tau_v + \int_{\infty}^{\tau_w} d\tau_v) e^{-\sigma(\tau_v)R_v(\tau_v+id)} w_{j_v}(\tau_w+id, \tau_v+id)$$

$$\prod_{v\geq v_0} [d^{n_v} f_{\nu_v}^{\delta_v}(q(\tau_v+id))] e^{i\omega_v\tau_v}. \quad \text{(b)}$$

As usual w is the node preceding v, $m_v(s)$ is the number of nodes in the list $v, s(v)$ with label $j = s$, $n(v)$ the number of those with label $j = 0$ and $\omega_v = \omega_{\nu_v}$, finally $\tau_{w_0} = t$.
The residues in R are introduced following the definitions of $\Im^t$ for complex values of t given in Subsection 1.1.3. The factors $(i\nu_{v\,s})^{m_v^s}$ come only from nodes with $\delta_v = 1$ so their product is bounded by[4] N^{2k}. Now we want estimates on the integrals that depend only on the order k; we start by splitting the sums in monomials.

1) Split $w_j(\tau_w + id, \tau_v + id)$ into 6 terms if $j = 0$ or 2 terms if $j \neq 0$: we obtain 6^{3k-1} terms. Each of this terms is of the form

$$\tau_v^h x_v^{-l} y(x_v) \tau_w^{h'} x_w^{-l'} y'(x_w),$$

where $x_v = e^{-|\tau_v|}$, $0 \leq h, h', l', l \leq 1$ and both $y(x)$, $y'(x)$ are analytic in $|x| \leq 1$ (we will call this the limited x dependent part of the Wronskian).

[4]In Appendix A.6, we will deal with functions $f(\psi, q)$ which are not trigonometric polynomials in ψ, the same reasoning could be applied in this Chapter, so removing the extra hypothesis on f, we consider only trigonometric polynomials only for simplicity.

2) Separate $\int_{-\infty}^{\tau_w} d\tau_v + \int_{\infty}^{\tau_w} d\tau_v$, and $\Im d\tau_{v_0}$ in integral (a). We get other 2^k terms like:

$$\prod_{v\geq v_0}\oint \frac{dR_v}{2i\pi R_v}\Big(\int_{\rho_v\infty}^{\tau_w} d\tau_v e^{-\sigma(\tau_v)R_v(\tau_v+id)}e^{i\omega_v\tau_v}(\tau_v)^{h_v}x^{l_v}\prod_{j=1}^{|s(v)|+2} y_v^j(x_v)\Big).$$

where $0\leq l_v, h_v \leq |s(v)|+1$. Notice that ρ_v is not the sign of τ_v but an extra label. The functions y_v^j are chosen in the following way:

(i) one of the y_v^j is either $\cos(q(\tau_v+id))$, $\sin(q(\tau_v+id))$ or one of the $F_{\nu_v}^k$.

(ii) one is the limited x_v dependent part of a term from the Wronskian at the node v.

(iii) for each node v' following v there is one function y_v^j which is the x_v dependent part of a term coming from the Wronskian $w(\tau_v, \tau_{v'})$.

Notice that the functions y are by definition all in $H(a,D)$ and respect condition 3.2.

3) Given a node $v\in s(v_0)$ split the integral $\int_{\rho_v\infty}^{\tau_{v_0}} d\tau_v$ as $\int_{\rho_v\infty}^{0} d\tau_v - \int_{\rho_{v_0}\infty}^{0} d\tau_v + \int_{\rho_{v_0}\infty}^{\tau_{v_0}} d\tau_v$ and proceed recursively for all nodes (other 3^{2k+1} terms). We consider first the contributions from the term with $\int_{\rho_{v_0}\infty}^{\tau_w} d\tau_v$ for all nodes (the others will be expressed as products of the same kind of integrals) .

Set $\rho_{v_0}=-1$, we want to estimate:

$$I_-(A) = \prod_{v\geq v_0}\oint \frac{dR_v}{2i\pi R_v}\Big(\int_{-\infty}^{\tau_w} d\tau_v e^{R_v(\tau_v+id)}e^{i\omega_v\tau_v}(\tau_v)^{h_v}x_v^{-l_v}\prod_{j=1}^{|s(v)|+2} y_j^v(\tau_v). \qquad (3.5)$$

Finally for integral (a) we split the first integral $\int_{-\infty}^{0} = \int_{-\infty}^{-a_0} + \int_{-a_0}^{0}$. $a_0>0$ is suitably large ($a_0=2a$).
In integral (b) we split $\int_{-\infty}^{t} = \int_{-\infty}^{-a_0} + \int_{-a_0}^{t}$ for $|t|\leq a_0$ and maintain $\int_{-\infty}^{t}$ otherwise. We consider the first term and expand the functions y_j^v as Taylor series in $x_v = e^{\tau_v}$ (the sign plus comes from the fact that we are considering only $t\leq -a_0<0$).

Remark 3.5. *The mapping $t\to e^t$ maps the region*

$$\{Re(t)<0, \quad 0\leq Im(t)\leq 2\pi i\}$$

in the unitary ball $|x|\leq 1$ and the half-lines $t+iy$ with $t\leq 0$ and $0\leq y\leq 2\pi$ going to $-\infty$ in the radiuses , of angle y, going to zero. Conversely the mapping $t\to e^{-t}$ maps the region

$$\{\ Re(t)>0, \quad 0\leq Im(t)\leq 2\pi i\}$$

in the unitary ball $|x|\leq 1$ and the half-lines $t+iy$ with $t\geq 0$ $0\leq y\leq 2\pi$ (going to ∞) in the radiuses of angle $-y$(going to zero). Notice that by our symmetry assumptions the image of $H(a,D)$ is the same in both mappings; moreover the y_j^v coming from the f_ν^k are all analytic in $x=0$, in the ball $|x|<e^{-a}$ and in all the section $Arg(x)<D$. The y_j^v coming from f^0 have a double pole in $\pm i$ and those coming from the Wronskian have simple poles in $\pm i$.

We set $y_j^v(\tau_v) = \sum_{r=0} y_j^{v\,r} x^r$ and $C_{\{r_v\}} = \prod_v y_j^{v\,r_v}$. The integral is

$$I_m^{a_0} = \text{Res} \sum_{\{r_v\}} C_{\{r_v\}} \prod_v \frac{\partial^{n_v}}{\partial E_v^{h_v}} \prod_v (\int_{-\infty}^{\tau_w} d\tau_v e^{R_v(\tau_v+id)+E_v\tau_v} e^{i\omega_v\tau_v} x_v^{r_v} d\tau_v \tag{3.6}$$

with $w_0 = -a_0$. Starting from the end-nodes we now perform the integrals in $d\tau_v$ then the derivatives in E_v and finally the residues in R_v, we do this first for all the end-nodes and then proceed to the inner nodes, hierarchically .

Proposition 3.6. *Integral (3.6) produces the bounds*

$$I_m^{a_0} \le \varepsilon^{-\alpha m} (m!)^{2\tau+2} C_1^k \prod_v [\prod_{j=1}^{|s(v)|+2} (\sum_h |y_j^{v\,h}| x_0^h) (\frac{t^s}{y_0})^{2k};$$

m is the number of nodes ($\le 2k-1$) ,$|s(v)|$ the number of nodes following v and C_1 is some order one constant.
In integral a) $y_0 = e^{-a_0}$ and $s = 0$; while in integral b) $s = 0$, $y_0 = e^{-a_0}$ if $|t| \le a_0$ and $s = 1$, $y_0 = e^{-|t|-d}$ otherwise.
Finally $\alpha \le \frac{1}{2}$ is defined in Chapter 1 and τ is the diofantine exponent of $\frac{\omega}{\sqrt{\varepsilon}}$ up to order K:

$$\varepsilon^{-\frac{1}{2}} |\omega \cdot n| > \varepsilon^{\alpha} \gamma |n|^{-\tau} \quad \textit{for some } \gamma = O_\varepsilon(1) \textit{ and for all } |n| < KN.$$

If we choose $a_0 > a$ the series are all convergent (by the analyticity of the y_j's in x_0).

We choose $x_0 = \frac{e^{-a}}{4}$ and estimate the coefficients of the Taylor series in the ball $|x| \le \frac{e^{-a}}{2}$:

$$\sum_{k=0}^{\infty} |y_j^{v,k}| x_0^k \le 2 \max_{|x| \le 2x_0} (y_j^v).$$

Proposition 3.6.

$$\text{The integral} \quad \int_{-\infty}^{t} x^K e^{iA\tau} e^{B\tau} = \frac{x^K e^{(iA+B)t}}{K+B+iA}$$

so the E_v derivatives in the end-node v give 2^{h_v} terms of the form:

$$h_1^v! \frac{x_w^{r_v} e^{idR_v} e^{(i\omega_v+R_v)\tau_w}}{r_v + R_v + i\omega_v} (\tau_w)^{h_2^v} \quad h_1^v + h_2^v = h_v. \tag{3.7}$$

The residue of R_v^{-1} times (3.7) is (3.7) if $|r_v| + |\omega_v| \ne 0$ and

$$\frac{h_2^v!}{(h_2^v+1)!} (\tau_w)^{h_1^v} (\tau_w + id)^{h_2^v+1} \quad \text{if } |r_v| + |\omega_v| = 0.$$

Developing the binomial we obtain other 2^{h_v+1} terms all of the type:

$$G^{h_v+1}\bar{m}!x_w^{r_v}e^{i\omega_v\tau_w}(\tau_w)^{\tilde{h}_v}.$$

The constant G is the maximum between one $(r_v \neq 0)$, $(\min_{|\nu|\le N}|\omega\cdot\nu|)^{-1}$ or $(\frac{\pi}{2})$ (we use that $d < \frac{\pi}{2}$). After integrating all the end-nodes following a node w we can integrate in $d\tau_w$ a sum of $2^{2\sum_{v\in s(w)}h_v+1}$ terms of the type:

$$G^{\bar{h}}\bar{h}!x_w^{\tilde{r}_w}e^{i\Omega_v\tau_w}(\tau_w)^{\hat{h}}$$

where $\tilde{r}_v = \sum_{v\in s(w)} r_v$, $\Omega_v = \sum_{v\in s(w)}\omega_v$ and $\bar{h}+\hat{h} \le \sum_{v\in s(w)} h_v + 1$. We have proved that the integrals derivatives and residues correspond to calculating the integrands in (3.6) at the limiting point (a_0 or t), ignoring the oscillating factors $e^{i\Omega a_0}$, substituting the Taylor coefficients with their moduli and multiplying by a factor bounded by:

$$2^{6k-3}(k!)^4 \max_{0<|\nu|<mN}(|\omega\cdot\nu|)^{-2\tau(2k-1)} \le C^k(k!)^{4\tau+4}.$$

x_0 is equal to e^{-a_0} in integral a) and is $x_0 = e^{-|t|}$ in integral b). If $|t| \le a_0$ then $x_0 \le e^{-a_0}$. The factor $\frac{1}{y_0^{2k}}$ comes from the divergent terms $x_0^{-\sum_v l_v}$ evaluated at the limiting point. The term $|t|^{2k}$ in integral b) can be bounded by ${a_0}^{2k}$ if $|t| \le a_0$. □

We now consider the "left out part" $\int_{-a_0}^t d\tau_{v_0}$ (we will set $t = 0$ in integral (a)). Let v_1 be a node of level one.
We break the integral $\Im^{\tau_{v_0}} d\tau_{v_1}$ as $\Im^{-a_0} d\tau_{v_1} + \int_{-a_0}^{\tau_{v_0}} d\tau_{v_1}$. If we choose the first term and m_1 is the number of nodes of $A^{\ge v_1}$, the integral on $A^{\ge v_1}$ can be bounded by $I_{m_1}^{a_0}$ and we are left with the problem of bounding the "left out part" $\int_{-a_0}^t d\tau_{v_0}$ on the remaining subtree $A^{/v_1}$. We repeat the procedure hierarchically and we end up with 2^m terms of the form:

$$I_{m_1}^{a_0}\cdots I_{m_p}^{a_0}\prod_{v\in\vartheta}\int_{-a_0}^{\tau_w} d\tau_v \mathcal{W}^1(\vartheta)$$

where the subtree ϑ has $\tilde{m}$ nodes and $\tilde{m}+\sum m_j = m$. We bound the last integral by the maximum of the integrand for integral (a) and for $|t| < a_0$ in integral (b) we obtain

$$(Ca_0)^{\tilde{m}}\prod_{v\in\vartheta,i}\max_{\tau_v\in[0,-a_0]}|y_j^v(\tau_v)|.$$

In integral (b) for $|t| > a_0$ we obtain

$$(Ce^{|t|}|t|)^{\tilde{m}}\prod_{v\in\vartheta,i}\max_{\tau_v\in[0,-a_0]}|y_j^v(\tau_v)|.$$

Let us now examine the $3^m - 1$ integrals left aside in the analysis of item 3). Starting from the end-nodes we cut off all the subtrees ϑ that contribute a definite integral $\Im_\rho^0$.

Such integrals are of the type $I_\rho(\vartheta_i)$ that we have already considered . We are left with an integral again of the type $I_{\rho_0}(\vartheta_0)$ where ϑ_0 is the tree deprived of the ϑ_i. The total number of nodes of the ϑ_i $i = 0, \cdots, h$ is m.
Now we only have to compute the maxima of the $|y_j^v(\tau_v)|$ that means the maxima of the moduli of $G_1 = \frac{1}{\cosh(t)}$, $G_2 = e^{-t}\sinh(t)$, $G_3 = 1 - \frac{2}{\cosh^2(t)}$, $G_4 = \frac{sinh(t)}{\cosh^2(t)}$ and of all the F_ν^k in the regions $\mathrm{Re}(t) > a+1$, $0 \leq \mathrm{Im}(t) \leq 2\pi$ and on $\mathrm{Im}(t) = d$.
To bound the functions F_ν^k we go back to the variable $q = \mathrm{arctg}(e^t)$ so $F_\nu^k \to d_q^k f_\nu(q)$. The maxima are then taken in a compact region $\subset \mathbb{T} \times i\mathbb{R}$ where the f_ν have no singularities, and which is contained in the image of $H(a,d)$ which is compact as $d < D \leq \pi/2$. Let us consider the integral (a) and set $d = D - \sqrt{\varepsilon}$, this means that in some of the considered functions we are going $\sqrt{\varepsilon}$ close to the singularity with $\mathrm{Im}(t_\nu^i) = D$ [5]. As we are not interested in optimality, we will estimate the maximum of G_1 with $\frac{1}{\sqrt{\varepsilon}}$ that of G_2 with a constant , and G_3, G_4 by $\frac{1}{\varepsilon}$.

Lemma 3.7. *The functions G_i contribute at most a factor ε^{-k-2k_0+1} where k_0 is the number of nodes with $\delta_v = 0$.*

Proof. There are $k_0 \leq k-1$ nodes with $\delta_v = 0$ contributing either G_3 or G_4, then each of the $k + k_0 - 1$ nodes $v \neq v_0$, carries a summand of

$$\max_{t \in C(D-2\sqrt{\varepsilon},2a)}(|x_j^0|) \max_{t \in H(2a,D-2\sqrt{\varepsilon})}(|x_j^1|)$$

from the Wronskian so either G_1^2 or G_1G_2. □

The functions F_ν^n appear exactly k times. Moreover $\sum_{i=1}^k n_{v_i}$ counts each node with $\delta_v = 1$ plus all its successive nodes. as each node with $\delta_v = 0$ has $s(v) \geq 2$

$$\sum_{i=1}^k n_{v_i} \leq \sum_v n_v - 3k_0 = 2k - k_0 - 1$$

We can bound the maxima of the F_ν^n in $H(2a, D - 2\sqrt{\varepsilon})$ via Proposition 3.3 so we have a factor $\sqrt{\varepsilon}^{-(p+2)k+k_0}$

Finally we notice that $E(D - 2\sqrt{\varepsilon}, \nu) \sim E(D, \nu)$ and we sum on all the trees of order k. This implies the following proposition:

Proposition 3.8. *We obtain the following bound on the order k of fruitless trees with initial data φ such that $|Im\ \varphi| \leq s_0$:*

$$[\sum_{A \in \mathcal{A}(0F)} c(A) \prod_v n_v!] \sum_{|\nu| \leq kN} e^{s_0|\nu|} C_1^k (k!)^{c_1} N^{2k} E(D,\nu) \sqrt{\varepsilon}^{-(p+5)k+5}, \tag{3.8}$$

where C is an ε independent constant and $c_1 = 2\tau + 2$.

[5] We approach all the singularities simultaneously only if $D = \frac{i\pi}{2}$. This fact can be used to give better bounds on integral (a); we will give some examples in Section 6.2.

Integral (b) is bounded by:

$$e^{\omega_\nu d} e^{(2k+1)(|t|+d)} \sum_{|\nu| \leq kN} e^{s_0|\nu|} C_1^k (k!)^{c_1} N^{2k} \sqrt{\varepsilon}^{-(p+5)k+3}. \tag{3.9}$$

The extra factor $e^{|t|+d}\varepsilon^{-1}$ *comes from* $w(t, \tau_{v_0})$.

The bound (3.9) is much overestimated (as explained in [GGM4]). In particular if $|t|, d$ are of order η, as D is η independent, the maxima of F_ν^h and of the G_i are not taken in a region near their singularities and so are ε- independent, for small enough ε. Moreover as $\omega_\nu \leq \varepsilon^{-3/2}$ for all $k \leq \varepsilon^{-1}$ then if $\eta < \varepsilon^{3/2}$ the factor $e^{\omega_\nu d}$ is small. In this case can use the following bound on analytic contributions to $\psi_j^k(t)$ with $t \in \mathbb{C}$ and $|t| = O(\eta)$:

$$(kN)^n e^{s_0 kN} C_1^k (k!)^{c_1} N^{2k}. \tag{3.10}$$

Remark 3.9. *Consider for each k a finite sum of integrals of type (b) which is known a priori to be bounded in t. This is possible only if all the integrals carrying divergent terms t^h or $e^{|t|}$ cancel. Then we can bound such finite sums by*

$$e^{\omega_\nu d} \sum_{|\nu| \leq kN} e^{s_0|\nu|} C_1^k (k!)^{c_1} N^{2k} \frac{\sqrt{\varepsilon}^3}{\sqrt{\varepsilon}^{(p+5)k}}.$$

We will generally consider formal power series on trees whose coefficients are the $c(A)$ of the preceding section. The following bound can be useful:

Lemma 3.10. *Given a tree $A \in \mathcal{A}_1$ let $\mathcal{S}(A)$ be its symmetry group and $n(v)$ be the number of nodes w in the list $v, s(v)$ such that $j_w = 0$. The following bounds hold:*

$$T_i(k) = \sum_{A \in \mathcal{A}_i^k} \frac{1}{|\mathcal{S}(A)|} \leq (4n)^k.$$

$$N_i(k) = \sum_{A \in \mathcal{A}_i^k} \frac{1}{|\mathcal{S}(A)|} \prod_{v \in A} n(v)! \leq (4n)^k.$$

A proof of this assertions is in Appendix A.2. Now let us see how one can apply the bounds (3.8) to trees with markings and with fruits.
If we want to consider formal power series on marked trees we only need to remember that for any $h(t) \in B(a, D)$ applying the linear function

$$D_{(j,h)}(A) \equiv \sum_{v \in A} h(\tau_v) \partial_j^v A,$$

is equivalent to multiplying by

$$Nk \max_{t \in C(2a, D - \sqrt{\varepsilon})} |h(t)|$$

if $j \neq 0$ and by

$$\frac{k}{\sqrt{\varepsilon}} \max_{t \in C(2a, D-\sqrt{\varepsilon})} |h(t)|$$

if $j = 0$. If $h(t)$ is not in H_0 then we set $d = D = 0$ (we will call this non-analytic bounds).

As we have seen the value of a fruit is

$$\mathcal{V}^1(\mathcal{F}_j^{i,k} = \frac{1}{2} x_j^{[i]}(t) \Im \mathcal{W}^1(x_j^i \partial_j^{v_0} \Lambda^k).$$

Moreover by Remark 2.8 adding a fruit of order k type i and angle j in the free node v of a tree $A \in \overset{0}{\mathcal{T}}$ is equivalent to adding a mark $x_j^{[i]}(t)\partial_j^v$ to the node v of A and multiplying by the η and φ dependent function $\frac{1}{2}\Im \mathcal{W}_\varphi^1(x_j^i \partial_j^{v_0} \Lambda^k)$. This is the sum of $2n^k$ values of trees with fruits (and with a marking $x_j^i \partial_j$ in the first node), so we repeat the procedure and cut away the fruits.

So we have $(2n)^k$ lists of l (at most $2k-1$) marked trees without fruits. t The value $\Im \mathcal{W}^1$ of a list is the product of the values $\Im \mathcal{W}^1$ of the trees and the value of a tree with fruits is the sum of the values of lists of trees obtained. As the tree values depend only on the order the sum is $(2n)^k$ times the value of a list.

We can apply the analytic bounds only to those trees whose markings are all analytic ($\partial_{j_{v_l}} x^0_{j_{v_l}}$). All the trees with carry a mark $\partial_{j_{v_l}} x^1_{j_{v_l}}$ are bounded via the non-analytic integrals ($d = 0$) Notice that ,in our bounds, each marking with $j = 0$ gives a factor bounded by $\frac{N}{\sqrt{\varepsilon}}$ and that there are exactly $2l$ markings.

Lemma 3.11. *The bound (3.8)implies the following bound for trees with fruits:*

$$[\sum_{A \in \mathcal{A}_j^k} c(A) \prod_v n_v!] \sum_{|\nu| \leq kN} e^{\ Im\ \varphi|\nu|} (2nC_1)^k (k!)^{c_1} N^{4k} \sqrt{\varepsilon}^{-(p+5)k+3}. \tag{3.11}$$

Proof. We have decomposed a tree with fruits A in $(2n)^k$ lists of marked trees $A_1, \ldots, \mathcal{A}_l$ each of order k_i such that $\sum_{i=1}^l k_i = k$ and bearing a total of $2l$ markings. The value of a list is:

$$\prod_{v \in \{A_i\}} n_v! \sum_{|\nu| \leq kN} e^{s_0|\nu|} (2nC_1)^k (k!)^{c_1} N^{4k} \sqrt{\varepsilon}^{-(p+5)\sum_{i=1}^l k_i + 5l - 2l}$$

□

.

Theorem 3.12. *The bounds (3.11) and (3.10) imply that the values of fruitless tree power series expansions of definition 2.6:*

$$\mho_{(j,f)} = \sum_{k \geq 1} \mho^k_{(j,f)}, \quad \mho_{(j,f),(i,h)} = \sum_{k \geq 1} \mho^k_{(j,f),(i,h)}, \ldots,$$

via $\Im\mathcal{W}$ *and* $\mathcal{V}^1_\varphi$, *for* $|\ Im\ \varphi| \leq s_0$ *and* $|t| = O(\eta)$, *are asymptotic power series in* η *and* ε.
Moreover, for $k \leq \varepsilon^{-1}$ *and* $\eta < \sqrt{\varepsilon}^{p+5+2c_1}$, *the value of trees of order* k *is bounded from above by* c^k *with* $c \ll 1$

Proof. Let us first consider the value of the fruitless tree power series expansion $\mho_j$ through $\Im\mathcal{W}_\varphi$[6]. In each node, $v \neq v_0$, we apply an operator O_j so we can divide O_j in tree terms (applying a label $i = 0, 1, b$ respectively for R^0_j, R^1_j and Q_j). Then we cut off the terms due to the operators R^i_j.
We obtain the lists of trees described above and we can use the bound (3.11). This implies that the series $\mho_j$ are asymptotic moreover for $k \leq \varepsilon^{-1}$ $k! \leq \varepsilon^{-k}$ and so if $\eta < \sqrt{\varepsilon}^{p+5+2c_1}$ is small enough then (3.11) is bounded by c^k with $c \ll 1$.
Applying the bounds (3.10) is the same only easier as one considers directly fruitless trees.
If we want to consider formal power series on marked trees we only need to remember that for any $h(t) \in B(a, D)$ applying the linear function

$$D_{(j,h)}(A) \equiv \sum_{v \in A} h(\tau_v)\partial^v_j A,$$

is equivalent to multiplying by

$$Nk \max_{t \in C(2a, D-\sqrt{\varepsilon})} |h(t)|$$

if $j \neq 0$ and by

$$\frac{k}{\sqrt{\varepsilon}} \max_{t \in C(2a, D-\sqrt{\varepsilon})} |h(t)|$$

if $j = 0$. □

Now we define a generality criterium. From now on a Proposition is said to be true “in general ” if it is true for (possibly fixed) functions f and for all $|\eta| \leq \varepsilon_0^p$, $|\varepsilon| \leq \varepsilon_0$ for some non zero ε_0.

Corollary 3.13. *(i)In general the values through* $\Im\mathcal{W}^1_\varphi$ *of non analytic trees, or of fruitless trees with total frequency* ν *such that* $\nu_F = 0$, *of order* $k \leq \varepsilon^{-1}$, *are of the type:*

$$P(\varepsilon, \varepsilon^{-1})^k \quad \textit{where} \;\; P \;\; \textit{is a polynomial.}$$

(ii) A formal power series of definition 2.6 whose summands are all fruitless trees with[7] $\nu_F \neq 0$ *is asymptotic under the same conditions of Theorem 3.12; moreover its terms of order*[8] $k < (\sqrt{\varepsilon})^{-(\frac{1}{\tau_F}-2b)}$ *in* η *are all bounded from above by*[9]*:*

$$P(\varepsilon, \varepsilon^{-1})^k O_\varepsilon(e^{-\frac{c}{\varepsilon^{b\tau_F}}}).$$

[6]Notice that we are not distinguishing between analytic and non analytic terms.
[7]For instance $\mho_{i\,j}$ with i or $j \leq m$.
[8]remember that τ_F is the diofantine exponent of ω_1.
[9]We will derive much better bounds for systems with one fast frequency

Proof. (i) We are not interested in shifting the integration in the complex plane, so all the integrands of integral a) can be bounded with ε independent constants. Then as we consider trees of order $k \leq \varepsilon^{-1}$, one can bound the factors $k!$ with ε^{-k}.
(ii) We are simply using the bounds (3.8). Fixed $k < (\varepsilon)^{-(\frac{1}{\tau_F}-2b)}$ $(b < \frac{1}{2\tau_F})$ then the frequencies that are accessible at order k are such that $|\nu| \leq Nk$. Moreover ω_1 is diofantine:

$$|\omega_1 \cdot \nu_1| \geq \gamma_F |\nu_1|^{-\tau_F}\,, \quad \text{with } \nu_1 \in \mathbb{Z}^m$$

and so for $\nu \leq kN$:

$$E(D,\nu) \leq e^{-\frac{|\omega_1\cdot\nu_1|}{\sqrt{\varepsilon}}} e^{|\omega_2||\nu_2|} \leq C^k e^{-(\frac{ck^{-\tau_F}}{\sqrt{\varepsilon}})}.$$

Consequently

$$\max_{\nu \leq (\varepsilon)^{-(\frac{1}{\tau_F}-2b)}} E(D,\nu) = O_\varepsilon(e^{-\frac{1}{\varepsilon^{b\tau_F}}}).$$

□

3.2 Identities for asymptotic power series.

We will prove some simple classical identities, true for asymptotic power series, which will be useful in the following sections.

Lemma 3.14. *(i) The sum and product of asymptotic power series is still an asymptotic power series. The division by an asymptotic power series $x(\eta,\varepsilon)$ such that $x(0,\varepsilon) \neq 0$ is still asymptotic. The integration and derivation of an asymptotic power series on the parameter η is still an asymptotic power series.*

(ii) Consider two formal power series that satisfy the formal relation $AB \sim C$ and such that

$$A = \sum_{k=0}^{\infty} (\eta)^k A_k \quad \textit{with } |A_k| \sim (\frac{c}{\eta_0})^k \quad \textit{for all } k \leq K,$$

with $c \ll 1$; same for B and C. Then their order K truncations $A^{\leq K}$, $B^{\leq K}$ satisfy the relation

$$A^{\leq K} B^{\leq K} = C^{\leq K} + o(c^K),$$

for all $\eta \leq \eta_0$.

(ii) Consider a function $f(x)$ analytic in a domain D and

$$x(\eta) = \sum_{k=0}^{K} (\eta)^k x_k \quad \textit{with } |x_k| \sim (\frac{c}{\eta_0})^k,$$

for all $k \leq K$ and for some $c \ll 1$, such that $x(\eta) \in D$ for all $\|eta| \leq \eta_0$. The following property holds:

$$f(x) - \sum_{k=0}^{K} (\eta)^k [f(\sum_{h=0}^{k} (\eta)^h x_h)]_k = o(c^K),$$

for all $|\eta| \leq \eta_0/2$.

Proof. (i) consider two asymptotic power series

$$A = \sum_{k=0}^{\infty} (\eta)^k a_k(\varepsilon) \quad B = \sum_{k=0}^{\infty} (\eta)^k b_k(\varepsilon)$$

with $a_0 \neq 0$ and such that $\max(|a_k|, |b_k|) \leq \varepsilon^{-pk}$ for all $k \leq K = \varepsilon^{-q}$. Their product and sum is obviously asymptotic. Moreover:

$$A^{-1} \sim \sum_{k=0}^{\infty} r_k(\eta)^k \sim \frac{1}{a_0} \frac{1}{1 + \frac{1}{a_0} \sum_{h=1}^{\infty} a_h(\eta)^h}$$

this is an analytic function of $x = \frac{A}{a_0} - 1$ provided that $|x| > 1$. Now for any truncation of A of order $K = \varepsilon^{-q}$ this condition is verified and we can find the coefficients r_k $(k \leq K)$ as finite combinations of the a_i with $i \leq k$.
(ii)This says simply that:

$$A^{\leq K} B^{\leq K} = \sum_{k=0}^{K} (\eta)^k \sum_{h=0}^{k} A_k B_{k-h} + (\eta)^K \sum_{k=1}^{K} \sum_{b=1}^{k} (\eta)^b A_k B_{K+b-k} = C^{\leq K} + (\eta)^K \sum_{k=1}^{K} \sum_{b=1}^{k} (\eta)^b A_k B_{K+b-k}$$

where

$$(\eta)^K \sum_{k=1}^{K} \sum_{b=1}^{k} (\eta)^b A_k B_{K+b-k} \leq 2K(c)^{K+1}.$$

(iii) $f(\eta) = f(\sum_{k=0}^{K} (\eta)^k x_k)$ is an analytic function of η for $\eta \leq \eta_0$ So its Taylor expansion at order K has the property:

$$f(\eta) = \sum_{k=0}^{K} \frac{(\eta)^k}{k!} f^{(k)} + \frac{f^{(K+1)}(\eta')}{(K+1)!} (\eta)^{K+1}.$$

Finally we apply Cauchy estimates on $f^{(K+1)}(\eta')$ in $|\eta| \leq \eta_0/2$. □

Lemma 3.15. *The equation $A + B \sim C + D$ where A, B, C, D are formal power series such that $|a_k|, c_k$ are at most polynomially small in ε while $b_k, d_k = O_\varepsilon(\varepsilon^\infty)$ for all $k \leq \varepsilon^{-q}$ is in general equivalent to the two equations:*

$$a \sim c \;\; b \sim d \quad \textit{for all} \;\; k \leq \varepsilon^{-q}.$$

Proof. We are simply saying that in general it is not possible that

$$P(\varepsilon, \varepsilon^{-1}) = f(\varepsilon) \;\; \text{where} \;\; f \;\; \text{is a trancendental function.}$$

□

This leads to the following relation for matrix formal series. Consider three matrix formal power series $D \in \text{Mat } n \times n$ and $X.Y \in \text{Mat } n \times h$ and let $D(K)$, $X(K)$, $Y(K)$ be their truncations to order $K = \varepsilon^{-q}$:

$$D(K) = \sum_{k=0}^{K} (\eta)^k D_k(\varepsilon),\ X(K) = \sum_{k=0}^{K} (\eta)^k X_k(\varepsilon),\ Y(K) = \sum_{k=0}^{K} (\eta)^k Y_k(\varepsilon).$$

Suppose that $D(K)$ is symmetric and that for $\eta \leq \varepsilon^p$:

$$\text{set } \sup_{i,j} D_{ij,k} = |D_k|, \text{ and } |D_k||X_k| \leq (\frac{c}{\eta})^k, \quad |Y_k| \leq C(\varepsilon)(\frac{c}{\eta})^k$$

for some $c \ll 1$ and $C(\varepsilon) = O_\varepsilon(\varepsilon^\infty)$. Moreover suppose that X_0 has an $h \times h$ non zero minor and $X_{ij\,0} = O_\varepsilon(1)$.

Lemma 3.16. *(i) Suppose that the expansions of D and X admit a decomposition $D_k = D_{1\,k} + D_{2\,k}$, $X_k = X_{1\,k} + X_{2\,k}$, all truncations of an asymptotic series; moreover $(D_{1\,k})_{ij}$ and $(X_{1\,k})_{ij}$ at most polynomially small in ε, while $(D_{2\,k})_{ij}$ $(X_{2\,k})_{ij}$ are $O_\varepsilon(C(\varepsilon))$. Then the formal power series relation $DX \sim Y$ is equivalent to*

$$D_1X_1 \sim 0 \quad D_1X_2 + D_2X_1 + D_2X_2 \sim Y \quad \textit{for all } k \leq \varepsilon^{-q}.$$

(ii) The formal power series relation

$$DX \sim Y,$$

implies that $D(K)$ has in general at least h eigenvalues $\lambda \leq O_\varepsilon(C(\varepsilon))$ for all $|\eta| \leq \varepsilon^{-p}$. If $D_{1\,0} + D_{2\,0}$ is non singular, the eigenvalues are $\lambda = O_\varepsilon(C(\varepsilon))$
(iii) Moreover if D is a convergent series in η with convergence radius $\eta_0 = \varepsilon^{p_1}$ then D as well has an eigenvalue $\lambda \leq O_\varepsilon(\max(C(\varepsilon), (c)^{\varepsilon^{-q}})$ for all $|\eta| < \varepsilon^{\max(p,p_1)}$.

Proof. (i) Is a direct consequence of Lemma 3.15.
(ii) Lemma 3.14 (ii) implies that:

$$D_1(K)X_1(K) = R_1\,, \quad \text{with } |R_1| \leq o(c^K)$$

where $K = \varepsilon^{-q}$, D_1 is (in general) symmetric and the columns of X_1 are independent and of order one. Let us set $D_1(K)$ in diagonal form $\lambda_1(\eta), \ldots \lambda_n(\eta)$; correspondingly X_1' still has independent and order one columns. This means that for each $j = 1, \ldots, h$ there exists $i(j)$ such that $(X_{1\,0})_{i(j)\,j} \neq 0$. Then the equation $\lambda_{i(j)} X_{1\,i(j)\,j} \sim 0$ implies that $\lambda_{i(j)\,k} = 0$ for all $k \leq K$.

As $D(K)$ is a $C(\varepsilon)$-small perturbation of $D_1(K) + D_{2\,0}$, classical perturbation theory guarantees the existence of at least h eigenvalues[10] of order $\leq O_\varepsilon(C(\varepsilon))$
(iii) We simply note that as D is convergent then $D = D(K) + o((\frac{\eta}{\eta_0})^K)$ with $\eta_0 = \varepsilon^{p_1}$. □

[10]The small eigenvalues are exactly h if for instance

$$D_{1\,0} + D_{2\,0} = \begin{pmatrix} Id_{n-h} & 0 \\ 0 & C(\varepsilon) Id_h \end{pmatrix}.$$

Chapter 4

Upper bounds on homoclinic splittings I

We prove that determinant of the splitting matrix is exponentially small in ε. The techniques are those of [GGM1] and we discuss them first for completely isochronous systems and then we generalize to Hamiltonian (*) with $F(q) = \cos q - 1$.

Notice that such bounds probably can be derived using the methods proposed in [LMS] (where the authors consider the case $m = n$). This would be a quick (and intrinsic) proof of the exponential smallness. Notice however that the bounds so obtained are generally not optimal as one has to set the Hamiltonian system in normal form and consequently looses the information on the nature of the singularities of f. We have seen in Chapter 3 that the singularities of f fix the parameter D of the bound (3.8), so that the exponentially small term in the value of a fruitless tree of total frequency ν is $E(D,\nu)$. In Chapter 6 we will prove that, for functions $f(\psi,q)$ which are trigonometric polynomials in ψ and rational in e^{iq}, the exponentially small term $E(D,\nu)$ is optimal as the (computable) first order of the splitting matrix has (under some non degeneracy assumptions as discussed in Lemma 3.16) exactly m exponentially small eigenvalues:

$$\lambda_i = \sum_{\nu \le N} P_{i,\nu}(\varepsilon,\varepsilon^{-1}) E(D,\nu).$$

4.1 Cancellations and splitting determinants

We use the tree formalism of Chapter 2 to find formal identities for the splitting vector and the splitting matrix. Then we apply the various Lemmas on asymptotic power series of Subsection 3.2 to prove $O_\varepsilon(\varepsilon^\infty)$ upper bounds on the splitting determinant. This is a generalization of the strategy proposed in [GGM1] for partially anisochronous systems with three degrees of freedom. It is based on the existence of linear formal power series relations (like those of Lemma 3.16 for the splitting matrix.
Such linear relations are discussed in Subsections 4.1.2 and 4.1.3; Subsection 4.1.1 is dedicated to proving that the stable-unstable manifolds are Lagrangian; we do not

need this property to prove the exponentially small bounds we report the proof only for completeness.

4.1.1 The generating function of the splitting

Using Observation 1.43 and the Lemmata 2.13 we can verify that the $n+1$ dimensional manifold $\Delta I_j(\varphi, q, \eta) = \sum_k (\eta)^k \Delta I_j^k(\varphi, q)$ is Lagrangian. In particular we have that:

Theorem 4.1 (Eliasson, Gallavotti). *The splitting vector $\Delta I_j^k(\varphi)$ is the derivative with respect to the angle φ_j, $j = 1, \ldots, n$ of a function $S(\varphi)$ called generating function. $S(\varphi)$ is the value $\Im \circ \mathcal{W}$ of of the tree vector:*

$$\sum_{B \in \hat{\mathcal{A}}^k} \frac{B}{N(B)|\mathcal{S}(B)|}.$$

Proving the theorem is equivalent to proving for each k, j the relation:

$$\mho_j^k = \sum_{A \in \mathcal{A}_j^k} \frac{A}{|\mathcal{S}(A)|} = \sum_{B \in \hat{\mathcal{A}}^k} \frac{D_j(B)}{N(B)|\mathcal{S}(B)|} \tag{4.1}$$

where $N(B)$ is as usual the number of nodes in B and $D_j = D_j(h(t) = 1)$ is defined in (1.34).
The theorem is equivalent to this last relation ((4.1)) as

$$\Delta I_j^k(t = 0) = \Im \circ \mathcal{W}(\mho_j^k)$$

and as we are considering fruitless trees:

$$\Im \circ \mathcal{W}_\varphi\Big(\sum_{B \in \hat{\mathcal{A}}^k} \sum_v \partial_j^v \frac{B}{N(B)|\mathcal{S}(B)|}\Big) = \partial_{\varphi_j} \Im \circ \mathcal{W}_\varphi \sum_{B \in \hat{\mathcal{A}}^k} \frac{B}{N(B)|\mathcal{S}(B)|}.$$

We prove relation (4.1) simply by translating it in a relation between trees with two markings:

Proof. For each $A \in \mathcal{A}_j^k$ we consider $N(A)$ copies A_v of A, each having an evidenced node v; now as $j \neq 0$ $\mathcal{A}_j^k \equiv \overset{r}{\mathcal{A}}{}_j^k$. For each coset $[v]$ we have $m[v]$ identical copies, we will name them $A_{[v]}$; we have:

$$\sum_{A \in \mathcal{A}_j^k} \frac{A}{|\mathcal{S}(A)|} = \sum_{A \in \mathcal{A}_j^k} \frac{1}{N(A)} \sum_{[v]\, \delta_v = 1} \frac{m[v] A_{[v]}}{|\mathcal{S}(A)|} =$$

$$\sum_{A \in \mathcal{A}_j^{\overset{r}{k}}} \frac{1}{N(P(A, v))} \sum_{[v]} \frac{m[v] P(A, v)}{|\mathcal{S}(A)|}$$

Then by Lemma2.13(ii):

$$\frac{m[v]P(A,v)}{|\mathcal{S}(A)|} = \frac{m[v_A]\partial_j^{v_A} B}{|\mathcal{S}(B)|}$$

where $B \in \hat{\mathcal{A}}$ is the tree (first node v) such that $\partial_j^{v_A} B = P(A,v)$.

□

Corollary 4.2. *With the same technique one can prove that*

$$\sum_{A\in\hat{\mathcal{A}}^k} \frac{A}{N(A)|\mathcal{S}(A)|} = \frac{1}{k} \sum_{B\in\hat{\mathcal{A}}^k:\delta_{v_B}=1} \frac{B}{|\mathcal{S}(B)|},$$

this is the representation of the generating function given in [G2]. This representation shows that the generating function is a function of the homoclinic trajectories $\psi_j(\varphi, t)$.

Proof. For each tree in $\hat{\mathcal{A}}^k$ we consider k copies each with one node $\delta_v = 1$ in evidence, conversely for each tree in $B \in \hat{\mathcal{A}}^k : \delta_{v_B} = 1$ we consider $N(B)$ copies each with one node in evidence. The corresponding coefficients are two points stabilizers and so are the same for corresponding trees on the left and right hand side. Now calculating the value of the generating function, and summing over $k = 1, \infty$, we obtain (simply via the definitions of the values of trees):

$$S(\varphi,\eta) = \sum_{k=1}^{\infty} \frac{(\eta)^k}{k} \Im[f(\sum_{h<k} (\eta)^h \psi^h(t,\varphi))]_{k-1} \sim \int_0^{\eta} d\tilde{\eta} \Im f(\psi(t,\tilde{\eta},\varphi)).$$

□

We will not be interested in proving that this is a true (not formal) relation. To do so one simply needs to show that all the involved functions have dynamical meaning and so their series expansion in η is a-priori convergent.

Remark 4.3. *The generating function is a function on rootless trees; call $\mathtt{A}^1$ the cosets of $\hat{\mathcal{A}}$ with respect to the usual equivalence relation : $A_1 \sim A_2$ if there exists v in A_2 such that $A_1 = P(A_2, v)$.*
Let $K \in \mathtt{A}^1$ and consider a representative A: then there are $N(A)$ trees in the coset K all with $N(A)$ nodes. As the trees in K have the same value the value of K is well defined, the generating function is:

$$S(\varphi) = \Im \circ \mathcal{W}_{\varphi}(\sum_{K\in\mathtt{A}^1} C(K)K)$$

and the coefficient $C(K) = \sum_{A\in K} \frac{1}{N(A)|\mathcal{S}(A)|}$.

Corollary 4.4. *The fruits can be written in terms of the generating function (at least as trees):*

$$G_j^i = a_j \Im \circ \mathcal{W}_{\varphi}[\sum_{B\in\hat{\mathcal{A}}^k} \frac{1}{N(B)|\mathcal{S}(B)|} \left[D_j(x_j^i)[B] + \delta_{j0} L_{x_0^i}(B)\right)\right]$$

4.1.2 Cancellations due to energy conservation

We consider the cancellations due to energy conservation i.e. the fact that the S/U manifolds are on the same energy level. These cancellations are best seen directly on the values of trees an in a non-perturbative setting; then if needed they can be translated in cancellations on the trees. This cancellations were first noticed in [G1]. Let us set

$$H_\eta(I_\eta(t,\varphi), p_\eta(t,\varphi), \psi_\eta(t,\varphi), q_\eta(t,\varphi)) \equiv E_\eta \equiv \sum_h (\eta)^h E^h,$$

where by the KAM results reported in Subsection 1.1.1, E_η is analytic in η near $\eta = 0$ and is independent of $\sigma(t)$. Recalling that $\psi_\eta(0^\sigma,\varphi) = \varphi$ and $q_\eta(0^\sigma,\varphi) = \pi$ we find:

$$I_\eta(0^+,\varphi)\cdot AI_\eta(0^+,\varphi) + (p_\eta(0^+,\varphi))^2 + 2\eta f(\varphi,\pi) = 2E_\eta =$$

$$(I_\eta(0^-,\varphi))^2 + (p_\eta(0^-,\varphi))^2 + 2\eta f(\varphi,\pi),$$

now we derive in φ_j with $j = 1,\dots,n$ and compute at the homoclinic point[1] $I(0^+,\varphi = 0) = I(0^-,\varphi = 0)$:

$$A\frac{\partial}{\partial\varphi_j}(\Delta I_\eta(\varphi))|_{\varphi=0}\cdot(2I_\eta(0,\varphi=0) = \frac{\partial}{\partial\varphi_j}(\Delta p_\eta(\varphi))|_{\varphi=0}(2p_\eta(0,\varphi=0).$$

Now let us write this perturbatively (i.e. in terms of trees); by the boundedness condition

$$p^k(0^\sigma,\varphi) = \Im^{0^\sigma}\circ\mathcal{W}_\varphi\mho^k_{(0,x_0^0)}.$$

We are on the lower branch of the separatrix so $p_0(\pi) = -2$ and $I_0(\varphi) = A^{-1}\tilde{\omega}$; now call $\mho_{(j,h)} = \sum_{k\ge1}\mho^k_{(j,h)}$, let Δ be the splitting matrix and for $j = 0, n$ set[2]:

$$a_j I_j^{(1)} = \Im^{0^\sigma}\circ\mathcal{W}_{\varphi=0}\mho_{(j,x_j^0)} = \Im^{0^\sigma}\circ\mathcal{W}^1_{\varphi=0}\mho_{(j,x_j^0)};$$

$$d_0 = \Im\circ\mathcal{W}_{\varphi=0}\mho_{(0,x_0^0)\,j}$$

(this is the φ gradient of Δp_η). Finally call $I^{(1)} = \{I_j^{(1)}\}_{j=1}^n$.

Proposition 4.5. *The splitting matrix satisfies the following equation*

$$\Delta(\tilde{\omega} + AI^{(1)}) = -d_0(-2 + I_0^{(1)}).$$

This means that we can tie the behavior of some fruits to that of the splitting matrix.

[1] Clearly at the homoclinic point p^+ and p^- coincide as well

[2] Note that as we are at the homoclinic point the only non zero contributions come from fruitless trees

Remark 4.6. *There are quite a few cancellations cancellations (on fruitless trees) coming from the symmetry of $\Im$ via integration by parts in the time variable. This cancellations are a simple generalization of the results in [GGM1]. Nevertheless they imply some heavy computations and are only actually not needed to prove exponential smallness for the splitting determinant, therefore we will state this results in Appendix A.7*[3].

4.1.3 Relation between trees with and without fruits

We have seen that trees with K fruits are homogeneous functions of degree K in the $G_j^{l\,h}$ with $l = 0, 1$, $j = 0, n$. We know as well that $G_j^{l\,h}(\varphi = 0) = 0$. We want to estimate the matrix

$$det(\Delta) \;\text{ where } \Delta_{i,j} = \partial_j \sum_k (\eta)^k \Delta I_i^k(\varphi = 0))$$

so it should be clear that it is useful to group trees by their degree in $G_j^{l\,h}$ rather than in η. We then decompose $\overset{0}{\mathcal{T}} = \mathcal{A} \oplus \mathcal{A}(1F) \oplus \cdots$ and add up the degrees in η (this are the formal power series discussed in Chapter 3):

$$\mho_j = \sum_k \mho_j^k \quad G_j^l = \sum_k G_j^{l\,k} \quad \Lambda_{(j,h)} = \sum_k \Lambda_{(j,h)}^k \quad G_j^l = \Im \mathcal{W}^1 \Lambda_{(j,x_j^{[l]})} = \sum_k G_j^{l\,k}$$

etc... The cancellations described in Chapter 2 are obviously still true in the sense of formal power series.

Remark 4.7. *To pass from $\mathcal{A}_1(kF)$ to $\mathcal{A}_1((k+1)F)$ one can apply the fruit adding linear functions discussed in Section 1.2.*

The problem is that in general $A \in \mathcal{A}(kF)$ and $\sum_h D_j^{ih}(A)$ (or $\sum_h B_j^{ih}(A)$) do not have the same symmetry group and so we cannot translate this relation on the $\Lambda^{kF} \to \Lambda^{k+1F}$ except in the case of $k = 0$.

We use Remark 2.8 to write a tree with one fruit in v as a tree with a mark x_j^i times the fruit function G_j^i The fruit adding functions become special mark adding functions T

Definition 4.8.

$$D_j(x_j^l) = D_j^l, \quad L_{x_0^l} = L^l, \qquad F^l(A) = x_0^l(y)\partial_0^y f^\alpha(A)$$

$$F_h^l(A) = h(y)x_0^l(y)\partial_{00}^y f^\alpha(A) \quad F_{x_0^m}^l = F^{l\,m}$$

where as usual α is $\overset{\delta=0}{\bigcirc}$.

[3]Notice that the cancellation mechanism that we illustrate in the Appendix A.7 is exactly the same used in [GGM1], the only difference is that here the cancellations can be seen directly on the trees and so the notation is more compact.

Lemma 4.9. *In $\mathcal{A}_1^{1F}$ we have:*

$$\Lambda_{(j,h)}^{1F} = \sum_{l=0,1} \sum_{m=0}^{n} G_m^l \left\{ D_m^{[l]}(\mho_{(j,h)}) + \delta_{m0}[L^{[l]}(\mho_{(j,h)}) + \delta_{j0} F_h^{[l]}(\mho_0)] \right\} \tag{4.2}$$

Proof. By Lemma 1.39 we only need to prove that the summands in the two sides of the relation have the same symmetry coefficient. This is true as the symmetry group of trees with one fruit is the subgroup of the symmetries, of the corresponding fruitless tree, that fixes the node where we will attach the fruit.

As both D_j^l and L^l act as sum on the nodes one can write them as sums on the cosets $[v]$:

$$D_j^l = \sum_v x_j^l \partial_j^v = \sum_{[v]} m[v] x_j^l \partial_j^v \,; \cdots$$

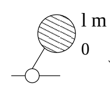
so that the summands of (4.2) are all different.

Consider a tree A in $\Lambda_{(j,h)}^{1F}$ carrying a fruit of type l label r and order m in the node v; we will call B the tree obtained by removing the fruit. If $B \in \mathcal{A}$ (it respects the grammar) then

$$A = G_r^{l\,m} x_r^{[l]} m[v] \partial_r^v B \text{ and } |\mathcal{S}(A)| = \frac{|Sa(B)|}{m[v]}.$$

If $B \notin \mathcal{A}$, consider again A. If $v \neq v_0$, then $r = 0$, $\delta_v = 0$ and there is a unique node following v, call it y. Then

$$A = G_0^{l\,m} m[y] g_{\alpha_{x_0^{[l]}}}(B', y)$$

for some $B' \in \mathcal{A}$ (B' is simply A without the piece 

).

Again we have $|\mathcal{S}(A)| = \frac{|\mathcal{S}(B')|}{m[v]}$, namely as y is the only node following v then v is fixed by $\mathcal{S}(A)$.

We now consider $v = v_0$, $B \notin \mathcal{A}$, this means that:

$$A = G_0^{l\,m} F_h^{[l]}(B'')$$

and there is an only node of level 1; by Remark1.27: $|\mathcal{S}(A)| = |\mathcal{S}(B'')|$. □

Definition 4.10. *In the following we will be interested in trees marked only with the functions x_j^i so we will contract the notations:*

$$\mho_{i\,j}^{l\,m} = \mho_{(i,x_i^l)\,(j,x_j^m)} \,; \cdots$$

4.1.4 Formal power series relations involving the splitting matrix

Proposition 4.11. *The splitting matrix depends only on the trees with zero or one fruit.*

Proof. Using Lemma4.9 we write:

$$\partial_j \Delta I_i(t=0,\varphi=0) = a_i^{-1}\Im\circ\mathcal{W}_0^1(\mho_{ij}) + \partial_j\{\Im\circ\mathcal{W}_0^1\{\sum_{l=0,1}\sum_{r=0}^{n} G_r^l[D_r^{[l]}(\mho_i)+$$

$$+\delta_{r0}(L^{[l]}(\mho_i)]\} + \text{ terms of order } \geq 2 \text{ in } G_r^l\}$$

so as $G_j^l(\varphi=0)=0$ (see the proof of Lemma 2.9) the terms of order ≥ 2 in G_j^l don't give contributions to the derivative:

$$\Delta_{ij} = \Im\circ\mathcal{W}_0^1[\mho_{ij} + \sum_{l=0,1}\sum_{r=0}^{n}\partial_j(G_r^l)[D_r^{[l]}(\mho_i) + \delta_{r0}(L^{[l]}(\mho_i)]$$

□

Similarly the value of

$$\partial_j G_r^l = \frac{1}{2}a_r\partial_j\Im\circ\mathcal{W}_0^1(\Lambda_r^l)$$

with $r=0,\dots,n$ and $j=1,\dots,n$ depends on trees with at most one fruit:

$$2G_{rj}^l = 2\partial_j G_r^l = a_r\Im\circ\mathcal{W}_0^1[\mho_{ij}^l + \sum_{m=0,1}\sum_{h=0}^{n} G_{hj}^m[\mho_{rh}^{l\,[m]} + \delta_{h0}(L^{[m]}(\mho_r^l) + \delta_{j0}F^{l\,[m]}(\mho_0))].$$

This is a linear relation that we can express in matrix form as:

$$G = \mathcal{A}(\mathcal{O}\mathcal{M}G + J)$$

where G is a $2n+2\times n$ matrix with entries:

$$G_{ij} = \begin{cases} \partial_j G_{i-1}^0 & \text{if } i=1,\dots,n+1 \\ \partial_j G_{i-n-2}^1 & \text{if } i=n+2,\dots,2n+2\,. \end{cases}$$

The matrix J is again $2n+2\times n$ with entries:

$$J_{ij} = \begin{cases} \Im\circ\mathcal{W}_0^1\mho_{i-1,j}^0 & \text{if } i=1,\dots,n+1 \\ \Im\circ\mathcal{W}_0^1\mho_{i-n-2,j}^1 & \text{if } i=n+2,\dots,2n+2 \end{cases}$$

$\mathcal{O}$ and $\mathcal{A}$ are $2n+2\times 2n+2$ matrices:

$$\mathcal{O} = \begin{vmatrix} 0_{n+1} & Id_{n+1} \\ Id_{n+1} & 0_{n+1} \end{vmatrix} \quad \mathcal{A} = \frac{1}{2}\begin{vmatrix} 1 & & & 0 \\ & A & & \\ & & 1 & \\ 0 & & & A \end{vmatrix}$$

where A is the diagonal matrix with eigenvalues a_j $j=1,\dots,n$. Finally $\mathcal{M}$ is a $2n+2\times 2n+2$ matrix with entries:

$$\mathcal{M}_{ij} = \begin{cases} \Im \circ \mathcal{W}_0^1(\mho^{1\,1}_{i-1,\,j-1}+ \delta_{j\,1}(L^1(\mho^1_{i-1})+ \delta_{i\,1}F^{1\,1}(\mho_0)) & \text{if } \begin{array}{l} i=1,\dots,n+1 \\ j=1,\dots,n+1 \end{array} \\ \Im \circ \mathcal{W}_0^1\mho^{0\,1}_{i-n-2,\,j-1}+ \delta_{j\,1}(L^1(\mho^0_{i-n-2})+ \delta_{i\,n+2}F^{0\,1}(\mho_0)) & \text{if } \begin{array}{l} i=n+2,\dots,2n+2 \\ j=1,\dots,n+1 \end{array} \\ \Im \circ \mathcal{W}_0^1(\mho^{0\,0}_{i-1,\,0}+ \delta_{j\,n+2}(L^0(\mho^0_{i-n-2})+ \delta_{i\,n+2}F^{0\,0}(\mho_0)) & \text{if } \begin{array}{l} i=n+2,\dots,2n+2 \\ j=n+2,\dots,2n+2 \end{array} \\ \Im \circ \mathcal{W}_0^1\mho^{1\,0}_{i-1,\,j-n-2}+ \delta_{j\,n+2}(L^0(\mho^1_{i-1})+ \delta_{i\,1}F^{0\,1}(\mho_0)) & \text{if } \begin{array}{l} i=1,\dots,n+1 \\ j=n+2,\dots,2n+2 \end{array} \end{cases}$$

Proposition 4.12. *The matrix $\mathcal{M}$ is symmetric.*

Proof. If $i,j \neq 1, n+1$, then $M_{ij} = M_{ji}$ is equivalent to Proposition2.14.
Same if $i,j = 1$ or $n+1$, then the symmetry condition is:

$$\mho^{0,1}_{0\,0} + L^0(\mho^1_0) = \mho^{1,0}_{0\,0} + L^1(\mho^0_0)$$

and so equivalent to the symmetry of the operators Q_j (see Propositions2.14,2.15).
Lastly if $i = 1, n+1$ and $j \neq 1, n+1$ the condition is:

$$\mho^{l,m}_{0\,j} = \mho^{m,l}_{j\,0} + L^l(\mho^m_j)$$

that is Proposition2.15 for trees without fruits. □

It can be useful to evidence the block structure of the matrix $\mathcal{M}$:

$$\mathcal{M} = \left|\begin{array}{c|c|c|c} a_{00} & u_{00}^t & a_{01} & u_{01}^t \\ \hline u_{00} & M_{00} & v_{10} & M_{01}^t \\ \hline a_{01} & v_{10}^t & a_{11} & u_{11}^t \\ \hline u_{01} & M_{01} & u_{11} & M_{11} \end{array}\right| \qquad G[\,J\,] = \left|\begin{array}{c} g_0^t \\ \hline G^0 \\ \hline g_1^t \\ \hline G^1 \end{array}\right|$$

$\forall i,j \in [1,\dots,n]$ $(M_{11})_{ij} = \Im\circ\mathcal{W}_0^1\mho^{00}_{i\,j}$, $(M_{00})_{ij} = \Im\circ\mathcal{W}_0^1\mho^{11}_{i\,j}$, $(M_{01})_{ij} = \Im\circ\mathcal{W}_0^1\mho^{10}_{i\,j}$, $(u_{11})_j = \Im\circ\mathcal{W}_0^1\mho^{00}_{0\,j}$, $(u_0)_j = \Im\circ\mathcal{W}_0^1\mho^{11}_{0\,j}$, $(u_{01})_j = \Im\circ\mathcal{W}_0^1\mho^{10}_{0\,j}$, $(v_{01})_j = \Im\circ\mathcal{W}_0^1\mho^{01}_{0\,j}$.

Remark 4.13. *The definitions of $\mathcal{M}$ and J imply that $J = \mathcal{OMT}$ where T is the $2n+2 \times n$ matrix :*

$$T = \left|\begin{array}{c} 0_{n+1} \\ 0 \\ Id_n \end{array}\right|$$

Again from the definitions of Δ and G we have that $A\Delta = 2G_0 = 2T\mathcal{O}G$

In the preceding Subsection we proved that

$$2\Delta(\tilde{\omega} + AI^{(1)}) \sim -g_0(-2 + I_0^{(1)})$$

where $\omega + AI$ has norm of order $O_\varepsilon(\varepsilon^{-\frac{1}{2}})$ and (if there are slow frequencies) $-g_0(-2+I_0)$ has norm $O_\varepsilon(\varepsilon^P)$ for some P (see Chapter 3; remember that this are all convergent series). Independently in Appendix A.8 we have proven that $\mathcal{M}$ is degenerate and satisfies the equation:

$$\mathcal{M}Y_1 = B \text{ where } Y_1^t = (\overbrace{0}^{n+1}, \overbrace{2}^{1}, \overbrace{-\tilde{\omega}}^{n})$$

and $B^t = 2(\overbrace{I_0^{(1)}(t=0)}^{1}, \overbrace{I^{(1)}(t=0)}^{n}, \overbrace{0}^{n+1})$. This relation gives a constraint on G_0 (as $(\mathcal{O} - \mathcal{M})G = \mathcal{M}T$) that coincides with the one given by the energy conservation.

Proposition 4.14. *The matrix $G_0 = \frac{1}{2}A\Delta$ satisfies the relation:*

$$G_0(A^{-1}\tilde{\omega} + I^{(1)}) = -g_0(-2 + I_0^{(1)}),$$

notice that $g_0 = \frac{1}{2}d_0$.

Proof. It is obvious from the energy conservation see Proposition4.5, now we derive it from the degeneracy of $\mathcal{M}$. We use PropositionA.24 and the relations between $\mathcal{M}$ J and G:

$$(\mathcal{O} - \mathcal{A}\mathcal{M})G = \mathcal{A}\mathcal{M}T \rightarrow (\mathcal{A}^{-1}Y_1)^t\mathcal{O}G - (Y_1^t\mathcal{M})G = Y_1^t\mathcal{M}T$$

this implies that

$$-\omega^t A^{-1}G_0 + g_0^t(2 - I_0^{(1)}) - (I^{(1)})^t G_0 = 0$$

we have used $\mathcal{O}^2 = I$, $Y_1^t\mathcal{M} = B^t$ and $B^tT = 0$. □

So one can see that the (hard) cancellations due to the integration by parts are only needed to check the consistency of our equations.

We want to estimate $\det G_0$; now for the first time we consider the existence of a fast time scale, we will say $j \in F$ if $\tilde{\omega}_j = \frac{\omega_j}{\sqrt{\varepsilon}}$is fast (resp. S). We concentrate only on the m fast variables ψ_1, ψ_m by applying the $2n+2 \times m$ matrix $Y_2^t = (\underbrace{0}_{n+2}, Id_m, \underbrace{0}_{n-m})$.

Then we set

$$\mathcal{M}Y_2 = \mathcal{M}_F = \begin{vmatrix} m_0^t \\ M^0 \\ m_1^t \\ M^1 \end{vmatrix}.$$

Lemma 4.15. *The order $k < C\varepsilon^{-\frac{1}{2(1+\tau_F)}}$ truncations of the matrix M_1 and the vector m_1 are of order $O_\varepsilon(e^{-c\varepsilon^{-\frac{1}{2(1+\tau_F)}}})$.*

Proof. The entries of this matrices are all derivatives $\partial_j L$, where $j \in F$ is a fast variable, and L is the integral of a function in $\tilde{H}_0$. We are considering the analytic parts of $\mathcal{M}$, u_{11} and M_{11}, and $\mathcal{M}$ is a function on fruitless trees. So we only have contributions from fruitless trees with analytic markings and non-zero total frequency in the fast direction $\nu_F \neq 0$ and we can apply Corollary 3.13 (ii). □

Proposition 4.16. *There exist two matrix formal power series X and Y, in $Mat(n \times m)$ such that their $K \leq \varepsilon^{-\frac{1}{2(\tau_F+1)}}$ truncation is $|X| = O_\varepsilon(1)$ and $|Y| = O_\varepsilon(e^{-c\varepsilon^{-\frac{1}{2(1+\tau_F)}}})$, for $|\eta| \leq \varepsilon^P$ and $\varepsilon \leq \varepsilon_0 \neq 0$ (here we choose $P = (p+5)/2 + c_1$, following Theorem 3.12, in Section 4.2 we will use less restrictive hyptheses on P). Moreover this series satisfy the formal equation:*

$$G_0 X \sim Y.$$

Proof. We can prove this only formally, i.e. the convergence of the vectors we define is not guaranteed; the bounds on the truncations are assured by the computations of Chapter 3.

$$Y_2^t(\mathcal{O} - \mathcal{AM})G \sim Y_2^t \mathcal{AM} T \to \bar{X}^t G_0 - m_0 \cdot g_0^t - A_F M_0^t G_0 - m_1 g_1^t - A_F M_1^t G_1 = A_F M_1^t$$

where $\bar{X} = (0, Id_m)$ in an $n \times m$ matrix and A_F is the $m \times m$ diagonal matrix $a_1, \ldots, a_m$. Substituting relation (4.14) we have:

$$G_0(\bar{X} - M_0 A_F - \frac{1}{-2 + I_0^{(1)}}(A^{-1}\tilde{\omega} + I^{(1)}) \cdot m_0) \sim g_1 \cdot m_1^t + (Id_n + G_1) M_1 A_F.$$

Notice that the problem in proving the convergence is not so much in the convergence of G_1 or g_1 (that have dynamical meaning) as in that of proving convergence for the "bare" parts M_0, M_1, m_0 and m_1. This all is done, for systems with three degrees of freedom, in [GGM4][4]. □

Now we can apply Lemma 3.16 to G_0 or equivalently to Δ; we have proven that the determinant of the splitting matrix is bounded from above by some constant of order $O_\varepsilon(e^{-c/\varepsilon^{\frac{1}{2(1+\tau_F)}}})$ where c is a suitable constant of order one. Consider the splitting matrix truncated at some order $k \leq C\varepsilon^{-\frac{1}{2(1+\tau_F)}}$ call it $\Delta^{\leq k}$, we can write it as a sum of matrices $\Delta^{\leq k} = \Delta_1^{\leq k} + \Delta_2^{\leq k}$ where Δ_2 contains all the exponentially small terms (coming from analytic integrals with non-zero fast mode as discussed in the end of Section 3.1). Both Δ_1 and Δ_2 are well defined as asymptotic series. We can divide X

[4] We will discuss the article [GGM4] in detail in Section 6.2.

as well in analytic (X_2) and non analytic (X_1) terms, both asymptotic power series. So we apply Lemma 3.16(ii) which states that:

Corollary 4.17. *In general, the matrix* $\Delta_{\bar{1}}^{\leq k}$ *has (at least)* m *zero eigenvalues* $o(\eta^k)$*-close to* $Span(X_{\bar{1}}^{\leq k})$*. Moreover the determinant of the splitting matrix is bounded from above by:*

$$|\det \Delta| \leq O_\varepsilon(e^{-c/\varepsilon^{\frac{1}{2(1+\tau_F)}}}),$$

for some order one c*.*

Moreover in Section 6.2 we will use the following statement:

Corollary 4.18. *The splitting matrix satisfies the following equation:*

$$G_0(Id_n - \frac{1}{-2+I_0^{(1)}}(A^{-1}\tilde{\omega} + I^{(1)})u_{01}^t + M_{01}^t A) = -u_{11}g_1u_{11}^t - G_1AM_{11} + AM_{11}.$$

Proof. We insert Proposition 4.14 in the last n lines of the linear equation: $G = \mathcal{A}(\mathcal{OM}G + J)$. ☐

4.2 Extension to partially isochronous systems

In this Section we will summarize the (few) modifications that are necessary to apply our techniques to partially isochronous systems.

We consider the following Hamiltonian:

$$\frac{(I, A(\varepsilon)I)}{2} + \bar{\omega}\cdot J + \frac{p^2}{2} + \varepsilon(\cos(q)-1) + \mu f(\psi,\phi,q). \tag{4.3}$$

As in Section 1.1 $I \in \mathbb{R}^n, \psi \in \mathbb{T}^n$, $p \in \mathbb{R}, q \in \mathbb{T}$ and we have coupled our systems with N clocks of frequency $\bar{\omega} \in \mathbb{R}^N$. The action angle variables of the clocks are $J \in \mathbb{R}^N, \phi \in \mathbb{T}^N$.
A is the diagonal matrix with eigenvalues a_i described in Chapter 1.
The system 4.3 is integrable for $\varepsilon \neq 0$, $\mu = 0$. It represents a list of n uncoupled rotators, N clocks and a pendulum. We will denote the frequency of the rotators (which determines the initial data $I(0)$) by ω so that:

$$I(t) = I(0) = A^{-1}\omega\,,\ \ \psi(t) = \psi(0) + \omega t$$
$$J(t) = J(0)\,,\ \ \phi(t) = \phi(0) + \bar{\omega} t.$$

The clocks ϕ_j are not changed by turning on the perturbation in μ. As in the previous sections we will look for S/U trajectories converging exponentially to a quasi-periodic function with diofantine frequency $\Omega = (\omega, \bar{\omega})$. So we will fix the initial data of the rotators as in Section 1.1. As usual we divide our frequency vector $\Omega = (\omega, \bar{\omega})$ in slow and fast frequencies and call m the total number of fast frequencies in the $n+N$ vector

Ω. Notice that the clock frequencies can be indifferently slow or fast. It is well known that for diofantine values of Ω one can apply a local KAM scheme, equivalent to that of Theorem 1.1, to construct the local S/U manifolds for the Hamiltonian 4.3. As usual we apply the canonical change of variables 1.10 and set the unperturbed Lyapunov exponent to one. Consequently the characteristic frequency is $\tilde{\Omega} = \varepsilon^{-\frac{1}{2}}\Omega$.
We use the Hamiltonian flow to extend the local manifolds. The extended S/U manifolds are graphs on the angles. As in the previous section we consider them at the Poincaré section $q = \pi$. To avoid using many variables we will set the initial data[5] $\psi, \phi = \varphi \in \mathbb{T}^{n+N}$ and set:

$$J_i^{\pm}(\varphi, \pi, \eta) = I_{n+i}^{\pm}(\varphi, \pi, \eta) \quad \text{for } i = 1, \ldots, n.$$

Now we construct the S/U manifolds perturbatively exactly as in subsection 1.1.2. The Hamilton equations are

$$\begin{aligned} \dot{I}_j \;&=\; -(\eta) f_{\psi_j}(\psi, \phi, q), \qquad & \dot{\psi}_j &= a_j I_j, \text{ for } j = 1, \ldots, n, \\ \dot{J}_i &= -(\eta) f_{\phi_i}(\psi, \phi, q), & \dot{\phi}_i &= \bar{\omega}_i, \text{ for } i = 1, \ldots, N, \\ \dot{p} &= \sin(q) - (\eta) f_q(\psi, q)\,, & \dot{q} &= p, \end{aligned} \tag{4.4}$$

We insert in the Hamilton equations the convergent power series representation:

$$I_j(t, \varphi, \eta) = \sum_{k=0}^{\infty} (\eta)^k I_j^k(t, \varphi) \quad \psi_j(t, \varphi, \eta) = \sum_{k=0}^{\infty} (\eta)^k \psi_j^k(t, \varphi) \quad \text{for } j = 1, \ldots, n$$

$$J_i(t, \varphi, \eta) = \sum_{k=0}^{\infty} (\eta)^k I_{n+i}^k(t, \varphi) \quad \phi_i(t, \varphi, \eta) = \phi_i + \frac{\bar{\omega}}{\sqrt{\varepsilon}} t \quad \text{for } i = 1, \ldots, n$$

finally

$$p(t, \varphi, \eta) = \sum_{k=0}^{\infty} p^k(t, \varphi) \quad q(t, \varphi, \eta) = q^0(t) + \sum_{k=1}^{\infty} (\eta)^k \psi_0^k(t, \varphi)$$

we obtain, for $k > 0$, the hierarchy of linear non-homogeneous equations:

$$\begin{aligned} \dot{I}_j^k =& F_j^k(\{\psi_i^h\}_{\substack{i=0 \\ h<k}}^{n})\,, \qquad \text{for } j = 1, \ldots, n+N \\ \dot{\psi}_i^k =& a_j I_i^k\,, \qquad\qquad\quad \text{for } i = 1, \ldots, n \\ \dot{p}^k =& (\cos q^0)\psi_0^k + F_0^k(\{\psi_i^h\}_{\substack{i=0 \\ h<k}}^{n})\,, \quad \dot{\psi}_0^k = p^k \end{aligned}$$

with

$$F_j^k = -[f_j(\sum_{h=1}^{k-1} (\eta)^h \psi^h, \phi + \frac{\bar{\omega}}{\sqrt{\varepsilon}} t)]_{k-1} + \delta_{j0}[\sin(\sum_{h=1}^{k-1} (\eta)^h \psi_0^h)]_k.$$

Using the whisker calculus developed in subsection 1.1.3 we find:

$$I_i^k(t) = \Im^t(F_i^k) \qquad \psi_j^k(t) = a_j O_j^t(F_j^k)$$

with $i = 1, \ldots, n+N$ and $j = 0, \ldots, n$.

[5] Our convention will be that the rotator angles are $\varphi_1, \ldots, \varphi_n$ and the clocks are $\varphi_{n+1}, \ldots, \varphi_{n+N}$ so the fast variables are not ordered sequentially but will be $\varphi_{i_1}, \ldots, \varphi_{i_m}$.

4.2.1 Tree representation

Passing to tree representation is now easy (and identical to what done in Chapter 2). We have seen that $\phi_j^k = 0$ for $k \neq 0$ so the labels j of nodes $v \neq v_0$ will still have values in $0, \ldots, n$ and the vector space $\mathbb{V}(\mathcal{A}_j)$ such that $\mathcal{V}(\mho_j^k) = \psi_j^k$ for $j = 0, n$, is unchanged.

We have seen however that markings represent derivatives on the node function $f^{\delta_v}(\psi_0, \ldots, \psi_{n+N}$ so the set of trees adapted to this dynamics is generated by $\overset{m}{\mathcal{A}}$ via the usual fruit adding functions (with type label $i = 0, 1$, order label $\delta \in \mathcal{N}$ and angle $j = 0, n$) and via mark adding functions:

$$h(t, v)\partial_J^v \quad \text{with} \;\; J = 0, n+N.$$

We will improperly call this spaces $\mathcal{A}$ and $\overset{0}{\mathcal{T}}$ as well. The function $\mathcal{W}$ is defined on $\mathcal{A}$ exactly as in Section 2.1.1, leading to the relation:

$$\Delta I_j^k = \Im \circ \mathcal{W}(\partial_j \mho^k)\,, \quad \text{for} \;\; j = 0, n+N.$$

The same holds for $\mathcal{V}^1$ and $\mathcal{W}^1$. As an example we write down explicitly the function $\Psi_\varphi(A)$ for $A \in \overset{0}{\mathcal{T}}$:

$$\Psi_\varphi(A) = (-\frac{1}{2})^{N(A)}[(\eta)^{\delta_{v_0}} \prod_{\substack{v \in \overset{0}{A} \\ v \neq v_0}} (\eta)^{\delta_v} a_{j_v}] \nabla^{\sum_{j=0}^{n+N} n_{v_A}(j) e_j} f^{\delta_{v_0}} \prod_{\alpha \in \mathcal{F}(v_A)} x_{j_\alpha}^{[i_\alpha]} \prod_{\beta \in \mathcal{M}(v_A)} h_\beta(v_A, \tau_{v_0})$$
$$\prod_{\substack{v \in \overset{0}{A} \\ v > v_0}} \nabla^{\sum_{j=0}^{n+N} n_v(j) e_j} f^{\delta_v} \prod_{\alpha \in \mathcal{F}(v)} x_{j_\alpha}^{[i_\alpha]} \prod_{\beta \in \mathcal{M}(v)} h_\beta(v, \tau_v) w_{j_v}(\tau_w, \tau_v) \cdot \prod_{\alpha \in \mathcal{F}(A)} G_{j(\alpha)}^{o(\alpha), i(\alpha)}$$

where $N(A)$ is the number of free nodes, $\mathcal{F}(v)$ are the fruits stemming from v, $\mathcal{M}(v)$ is the list of markings of the node v and finally $n_v(j)$ is the number of elements in $\{v, s_0(v), \mathcal{F}(v), \mathcal{M}(v)\}$ having angle label equal to j. Remember that j_v, $j_\alpha = 0, \ldots, n$, while the angle-markings are $J_v = 0, n+N$.

The energy conservation for the system ((4.3)) leads to the relation:

$$\Delta(\tilde{\Omega} + A^1 I^{(1)}) = -d_0(-2 + I_0^{(1)}),$$

where as in subsection 4.1.2 , Δ is the $n+N \times n+N$ splitting matrix, $d_0 = \nabla_\varphi \Delta p|_{\varphi=0}$, $I_j^{(1)} = \sum_{k=1}^\infty I_j^k(t=0, \varphi=0)$ and A^1 is an $n+N \times n+N$ matrix so defined:

$$A^1 = \begin{vmatrix} A & 0 \\ 0 & 0 \end{vmatrix}.$$

Remark 4.19. *We can repeat the procedure proposed in this Subsection for any Hamiltonian ((*)) such that $p^2/2 + F(q) = 0$ is the separatrix of a generalized pendulum (see the introduction). We only have to use the Wronskian matrix of the generalized pendulum in equations (1.14) and consequently change the functions x_0^i in the definition of O_0. The qualitative behavior in unchanged.*

We can find bounds similar to those of Corollary 3.13 (ii) for the fruitless trees of the expansion of Hamiltonian (*). However if we do not impose Condition 3.2 to $f(\psi(t), q(t))$ we do not find optimal bounds as we cannot get near to the singularities. We set that $F(q)$ is analytic in $|\operatorname{Im} q| \leq r_1$ and $f(\psi, q)$ is analytic in $|\operatorname{Im} q| \leq r_1$, $|\operatorname{Im} \psi_i| \leq r_1$. This means that:

$$f(\psi, q) = \sum_{\nu, h \in \mathbb{Z}^{n+1}} f_{\nu,h} e^{i(\nu \cdot \psi + hq)}, \quad \text{with} \quad |f_{\nu,h}| \leq C e^{-r_1(|\nu|+|h|)}.$$

Theorem 4.20. *The contribution of fruitless trees of total harmonic ν is bounded by:*

$$e^{-r_1|\nu|} C_1^k (k!)^{c_1} E(c_2, \nu)(\varepsilon)^{-k},$$

where C, c_2 are ε independent constants and $c_1 = 2\tau + 2$.

Proof. The proof is identical to that of Proposition 3.8, if $f(\psi, q)$ is trigonometric in ψ. The only difference is that in the proper integrals we do not go ε-close to the singularities so such terms are not divergent in ε (the factor ε^{-k} comes from small denominators). The proof for general analytic functions $f(\psi, q)$ is not difficult but quite long; we will report it in the Appendix A.6. □

4.2.2 Formal power series relations involving the splitting matrix

The linear non-perturbative equation (4.2) is unchanged ,so :

$$\Delta_{ij} = \Im \mathcal{W}_0^1 [\mho_{ij} + \sum_{l=0,1} \sum_{k=0,n} \partial_j (G_k^l) [D_k^{[l]}(\mho_i) + \delta_{k0}(L^{[l]}(\mho_i)],$$

and the derivatives of the fruits are ($j \in [0, \ldots, n+N]$):

$$G_{k,j}^l = \partial_j G_k^l = \frac{1}{2} a_j \Im \mathcal{W}_0^1 [\mho_{k,j}^l + \sum_{m=0,1} \sum_{h=0,n} G_{h\,j}^m [\mho_{k,h}^{l\,[m]} + \delta_{h0}(L^{[m]}(\mho_k^l) + \delta_{j0} F^{l,[m]}(\mho_0))]$$

this are linear relations :

$$G = \mathcal{A}(\mathcal{O}\mathcal{M}G + J) \quad \Delta = (\mathcal{N}^t G + J_1). \tag{4.5}$$

G is now a $2n+2 \times n+N$ matrix with entries:

$$G_{ij} = \begin{cases} \partial_j G_{i-1}^0 & \text{if } i = 1, \ldots, n+1 \\ \partial_j G_{i-n-2}^1 & \text{if } i = n+2, \ldots, 2n+2 \end{cases}$$

The matrix J is again $2n+2 \times n+N$ with entries:

$$J_{ij} = \begin{cases} \Im \mathcal{W}_0^1 \mho_{i-1,j}^0 & \text{if } i = 1, \ldots, n+1 \\ \Im \mathcal{W}_0^1 \mho_{i-n-2,j}^1 & \text{if } i = n+2, \ldots, 2n+2 \end{cases}$$

$\mathcal{O}$, $\mathcal{A}$ and $\mathcal{M}$ are the $2n+2 \times 2n+2$ matrices defined in the preceding subsection:

$$\mathcal{O} = \left| \begin{array}{cc} 0_{n+1} & Id_{n+1} \\ Id_{n+1} & 0_{n+1} \end{array} \right| \quad \mathcal{A} = \left| \begin{array}{cccc} 1 & & & 0 \\ & A & & \\ & & 1 & \\ 0 & & & A \end{array} \right|$$

where A is the diagonal matrix with eigenvalues a_j $j = 1, \ldots, n$.

$$\mathcal{M} = \left| \begin{array}{c|c|c|c} a_{00} & u_{00}^t & a_{01} & u_{01}^t \\ \hline u_{00} & M_{00} & v_{10} & M_{01}^t \\ \hline a_{01} & v_{10}^t & a_{11} & u_{11}^t \\ \hline u_{01} & M_{01} & u_{11} & M_{11} \end{array} \right| \qquad G[\,J\,] = \left| \begin{array}{c} g_0^t \\ \hline G^0 \\ \hline g_1^t \\ \hline G^1 \end{array} \right|$$

$$\forall i, j \in [1, \ldots, n] \;\; (M_{11})_{ij} = \Im\mathcal{W}_0^1\mho_{i,j}^{00} \;,\; (M_{00})_{ij} = \Im\mathcal{W}_0^1\mho_{i,j}^{11} \;,\; (M_{01})_{ij} = \Im\mathcal{W}_0^1\mho_{i,j}^{10} \;,$$

$$(u_{11})_j = \Im\mathcal{W}_0^1\mho_{0,j}^{00} \;,\; (u_0)_j = \Im\mathcal{W}_0^1\mho_{0,j}^{11} \;,\; (u_{01})_j = \Im\mathcal{W}_0^1\mho_{0,j}^{10} \;,\; (v_{01})_j = \Im\mathcal{W}_0^1\mho_{0,j}^{01}.$$

Δ is the $n+N \times n+N$ splitting matrix; J_1 is again $n+N \times n+N$

$$J_{1\,i\,j} = \Im\mathcal{W}_0^1\mho_{i,j}^{00}$$

finally $\mathcal{N}$ is $2n+2 \times n+N$, we represent it in block structure as:

$$\mathcal{N} = \left| \begin{array}{c} n_0^t \\ \hline N^0 \\ \hline n_1^t \\ \hline N^1 \end{array} \right|$$

where:

$$N_{ij}^1 = \Im\mathcal{W}_0^1\mho_{i,j}^{00} \;,\; N_{ij}^0 = \Im\mathcal{W}_0^1\mho_{i,j}^{01} \;,\; (n_1)_i = \Im\mathcal{W}_0^1\mho_{0,i}^{00} \;,\; (u_0)_j = \Im\mathcal{W}_0^1\mho_{0,i}^{10},$$

with $i \in (1, \ldots, n+N), j \in (1, \ldots, n)$.

As in the previous Section we consider the $n+N \times m$ matrix Y_F such that Y_F^t is the canonical projection on the fast components; we apply this projection to the second relation in ((4.5)); then we use the energy conservation and the relation:

$$G_0 = \mathcal{I}\Delta = \left| \begin{array}{cc} Id_n & 0 \\ 0 & 0 \end{array} \right| \Delta.$$

We obtain:

$$Y_F^t\Delta = (Y_F^t\mathcal{N}^t)G + Y_F^t J_1 = (N_F^1)G_1 + n_{1\,F}g_1^t + (N_F^0)^t G_0 + n_{0\,F}g_0^t + J_{1\,F} =$$

$$= (N_F^1)^t G_1 + n_{1\,F}g_1 + (N_F^0)^t\mathcal{I}\Delta + \frac{1}{-2+I_0^{(1)}} n_{0\,F}(\Omega + A^1 I^{(1)})^t\Delta + J_{1\,F}.$$

We have again found m independent vectors X such that (at least formally) $\Delta X = Y$ with $|X| = O_\varepsilon(1)$ and $|Y| = O(e^{\frac{a}{\sqrt{\varepsilon}}})$:

$$X = Y_F + \mathcal{I}N_F^0 + \frac{1}{-2+I_0^{(1)}}(\Omega + A^1 I^{(1)})n_{0\,F}$$

$$Y = (J_{1,F} + (N_F^1)^t G_1 + n_{1\,F}g_1)^t.$$

This and Theorem 4.20 imply that:

Theorem 4.21. *The Hamiltonian (*) , considered in the domains V_m defined in the Introduction, has an homoclinic point at $q = \pi, \psi = 0$. The order $k < C\varepsilon^{-(\frac{1}{2(\tau_F+1)})}$ term of the splitting determinant in such point is bounded from above by:*

$$(\varepsilon C_1)^k (k!)^{c_1}\varepsilon^{-k}e^{-c_2/\varepsilon^{\frac{1}{2(1+\tau_F)}}},$$

where C, c_2 are ε independent constants and $c_1 = 2\tau + 2$.

Proof. We can adapt Corollary 3.13 (ii) to find exponentially small upper bounds for $\det\Delta$. Namely we set

$$\sum_{\nu:|\nu_F|\neq 0} e^{-r_1|\nu|}E(C_2,\nu) \leq \sum_{\nu:|\nu_F|\neq 0} \exp(-\sum_{j=m+1}^{n}|\nu_j|(r_1 - \varepsilon^\alpha|\omega_2|c_2)e^{-(r_1|\nu_F| + \frac{|\omega_1||\nu_F|^{-\tau_F}}{\sqrt{\varepsilon}})}.$$

Now if $\alpha > 0$ then $r_1 - \varepsilon^\alpha|\omega_2|c_2 > 0$,while if $\alpha = 0$ we consider this a condition on c_2. So we can sum on the slow frequencies ν_j with $j > m$. Finally we split the sum over the fast frequencies in $|\nu_F| \leq \varepsilon^{-(\frac{1}{2(\tau_F+1)})}$ (where $\frac{|\omega_1||\nu_F|^{-\tau_F}}{\sqrt{\varepsilon}}$ dominates) plus a remainder (where $|r_1||\nu_F|$ dominates). □

Finally this implies Theorem 0.6 provided that $\nu \leq \varepsilon^{1+\frac{\tau+1}{\tau_F+1}}$ As the splitting determinant is smaller than:

$$\sum_{k<K}[\det\Delta]_k + (\mu 7\mu_0)^K,$$

and we can choose $K = C\varepsilon^{-(\frac{1}{2(\tau_F+1)})}$.

Chapter 5

Upper bounds on homoclinic splittings II

Following [BB1], we construct recursively a transformation $\vartheta : \mathbb{T}^n_s \ni \varphi \to \alpha \in \mathbb{T}^n_s$ *such that in the induced symplectic coordinates the generating function of the splitting (which we prove is* $S \circ \vartheta$*) is the integral* $\Im$ *of a function* $F(\alpha, t) \in H_0$ *plus a remainder of order* η^K *with* $K = O(\varepsilon^{-B})$ *with* $B = -\frac{1}{\tau_F} + b$*. This implies that the splitting determinant, i.e. the determinant of the Hessian of* S*, is* $O_\varepsilon(\varepsilon^{-b\tau_F})$*. So this section provides a possibly simpler proof of the upper bounds on the splitting determinant. Moreover the existence of* ϑ *implies a stronger condition, which is useful to prove "fast " diffusion*[1]*. For each* $\alpha \in \mathbb{T}^n_s$ *the Hessian matrix of* $S \circ \vartheta$ *has the following block structure:*

$$M(\alpha) = \left| \begin{array}{c|c} M_F & N_F \\ \hline & \\ N_F^t & M_S \end{array} \right| \tag{5.1}$$

where M_F *is an* $m \times m$ *matrix whose entries are* $O_\varepsilon(\varepsilon^\infty)$*,* N_F *is a* $n - m \times m$ *matrix whose entries are* $O_\varepsilon(\varepsilon^\infty)$ *and* M_S *contains terms which are polynomial in* $\varepsilon, \varepsilon^{-1}$*.*

As in the preceding section we use tree techniques, so we will give constructive proofs of our assertions, nevertheless the strategy of this Chapter shadows quite faithfully [BB1]. Namely we will study an auxiliary problem:

$$\ddot{\Phi}_i = \delta_{i0} \sin(\Phi_0) - (\eta) a_i \partial_i f(\Phi) + A_i(\eta) g_i(t) \tag{5.2}$$

where the $g_i(t)$ are prefixed functions in H_0. We will look for exponentially quasi-periodic "solutions" of this system. There are two main differences:

1. As usual the tree techniques can be easily applied to anisochronous systems, so our results apply to Hamiltonian (4.3).

[1]We will not prove fast Arnold diffusion in this thesis, so this Chapter should be seen as an alternative (possibly more intrinsic) way of proving exponentially small homoclinic splitting

2. On the other hand it is quite difficult to prove the convergence of Lindstedt series. The auxiliary problem is not Hamiltonian so there is no guarantee that the quasi-periodic "solutions" of this system exist. Although it should be possible to prove convergence using the techniques of [GGM4], the procedure is not easy.

To avoid this we will consider order $(\eta)^K$ truncations of the solutions, with $K = \varepsilon^{-b}$. In the next subsection we will use the results of Chapter 3 to explain why this is sufficient. Let us first remind a simple variation property of the generating function through changes of coordinates on $\mathbb{T}^n$.

Proposition 5.1. *Given an analytic transformation $\vartheta : \mathbb{T}^n \to \mathbb{T}^n$, let ϑ_* be the corresponding symplectic transformation lifted to the cotangent bundle. The generating function of the splitting in the coordinates $I', \psi' = \vartheta_*(I, \psi), p' = p, q' = q$ at the Poincaré section $q' = \pi$ is $S' = S \circ \vartheta^{-1}$.*

Proof. Given $\vartheta : \mathbb{T}^n \to \mathbb{T}^n$ we consider the prescribed symplectic transformation:

$$\psi' = \vartheta(\psi) \quad I' = J(\vartheta)^{-t}$$

$$q' = q \qquad\qquad p' = p$$

this is the canonical lift to phase space of $\vartheta \times 1 : \mathbb{T}^{n+1} \to \mathbb{T}^{n+1}$. As the pendulum angle q is unchanged and the Poincaré section is the same $q = q' = \pi$ the two coordinate systems describe the same S/U manifolds so:

$$J^{\pm}(\psi', \pi) = \left[J(\vartheta)|_{\vartheta^{-1}(\psi')}\right]^{-t} I^{\pm}(\vartheta^{-1}(\psi'), \pi)$$

By the definition of the generating function we have

$$\Delta J(\psi', \pi) = \partial_{\psi'_j} S'(\psi') = \left[J(\vartheta)|_{\vartheta^{-1}(\psi')}\right]^{-t} \partial_{\psi_j} S(\psi)|_{\psi=\vartheta^{-1}}.$$

□

5.1 Moving Poincaré sections

Following the ideas in [BB1] we will study an "auxiliary" system of $K(n+1)$ linear non-homogeneous ODE's whose solutions we will call $\Phi_j^h(t)$ with $h \leq K$ and $j = 0, \ldots, n$. The idea is to choose the "auxiliary" system and the initial data (depending on a parameter $\alpha \in \mathbb{T}_d^n$) so that $\Phi_j(\alpha, t) \in H_0$. Then we will define a function $\tilde{S}(\alpha)$ and we will find sufficient conditions on the "auxiliary" system such that there exists a (real) analytic transformation $\vartheta : \mathbb{T}_d^n \to \mathbb{T}_d^n$ with $S = \tilde{S} \circ \vartheta$.
The "auxiliary system" is of the type $(0 < h \leq K,\ j = 0, \ldots, n)$:

$$\Phi_j^{(0)}(t) = \psi_j^{(0)}(t)$$

$$\ddot{\Phi}_i^{(k)} = a_i \left[F_i^{(k)}(\{\Phi_j^{(h)}\}_{h=0,j=0}^{k-1,n}) + A_i^k g_i(t)\right] \ i = 1, \ldots, n$$

$$\ddot{\Phi}_0^{(k)} = \cos(q_0(t))\Phi_0^{(k)} + F_0^{(k)}(\{\Phi_j^{(h)}\}_{h=0,j=0}^{k-1,n}) + A_0^k g_0(t)$$

This is the order K Taylor expansion of the equation 5.2.

We have modified the forcing terms by the functions $A_i(\eta)g_i(t)$ where $g(t) \in H_0$ is an even function tending exponentially to a quasi-periodic function with zero average for $|t| \to \infty$.

The initial data on $\Phi_j^{(k)}$ are for the moment free and the only restriction is that the functions $\Phi_j^{(k)}(t)$ tend exponentially to a quasi-periodic function as $|t| \to \infty$. For $\alpha \in \mathbb{T}^n$, we set $\Phi_j^{(h)}(t=0) = \Phi_j^{(h)}(\alpha)$ for $h > 0$, while $\Phi_j^{(0)}(\alpha) = \alpha$ for $j \neq 0$ and $\Phi_0^{(0)}(\alpha) = \pi$ (the initial data are $\eta-$close to (α, π)). We can repeat the procedure used in Subsection 1.1.4 to determine the $\Phi_j^k(\alpha, t)$ recursively (the required asymptotic behavior is the same). The only difference is in the initial data; this implies that $\Phi_j^{(k)}$ have the form:

$$\Phi_j^k(\alpha, t) = x_0^0(t)\Phi_j^{(h)}(\alpha) + a_j O_j^t[F_j^k(\{\Phi_i^{(h)}(\alpha,\tau)\}_{h=0,j=0}^{k-1,n}) + A_j^k g_j(\tau)] \tag{5.3}$$

Correspondingly:

$$\dot{\Phi}_j^k(t) = \Im^t[F_j^k(\{\Phi_j^{(h)}\}_{h=0,j=0}^{k-1,n})) + A_i^k g_i(t)]$$

Remember that we are using the formalism of subsection 1.1.4 where we did not need any convergence property on the series $\sum_k (\eta)^k \Phi_j^k$ to recursively establish the boundedness of the $\Phi_j^k(\alpha, t)$.

Proposition 5.2. *If the functions g_i repsect the property:*

$$\Im x_i^0 g_i \neq 0,$$

for each α we can fix $A_j^k(\alpha)$ and $\Phi_j^{(k)}(\alpha)$ so that $\Phi_j^k(\alpha, t) \in H_0$.

Proof. We proceed by induction using the fact that $F_j^0(\Phi_i^0)$ is in H_0 and that $F_j^k(x_1, \cdots, x_m) \in H_0$ if x_i are in H_0. Suppose that $\Phi_i^{(h)}(\alpha, \tau) \in H_0$, for all $i = 0, \ldots, n$ and $h < k$:

$$\begin{aligned}\Phi_j^{(k)}(\alpha, t) = x_j^0(t)\Phi_j^{(k)}(\alpha) + a_j \Big[Q_j^t \big[F_j^k(\{\Phi_i^{(h)}(\alpha,\tau)\}_{h=0,i=0}^{k-1,n}) + A_j^k g_j(\tau)\big] + \\ \frac{1}{2} x_j^0(t) \Im\big[x_j^1(\tau)\big(F_j^k(\{\Phi_i^{(h)}(\alpha,\tau)\}_{h=0,i=0}^{k-1,n}) + A_j^k g_j(\tau)\big)\big] + \\ \frac{1}{2} x_j^1(t) \Im\big[x_j^0(\tau)\big(F_j^k(\{\Phi_i^{(h)}(\alpha,\tau)\}_{h=0,i=0}^{k-1,n}) + A_j^k g_j(\tau)\big)\big]\Big].\end{aligned}$$

If we choose

$$A_j^k(\alpha) = -\frac{\Im[x_j^0(\tau)(F_j^k(\{\Phi_i^{(h)}(\alpha,\tau)\}_{h=0,i=0}^{k-1,n})]}{\Im x_j^0 g_j(\tau)} \tag{5.4}$$

and

$$\Phi_j^{(h)}(\alpha) = -\Im\big[x_j^1(\tau)\big(F_j^k(\{\Phi_i^{(h)}(\alpha,\tau)\}_{h=0,i=0}^{k-1,n}) + A_j^k(\alpha) g_j(\tau)\big)\big] \tag{5.5}$$

the non analytic terms cancel and we have that:

$$\Phi_j^{(k)}(\alpha,t)=Q_j^t[F_j^k(\{\Phi_i^{(h)}(\alpha,\tau)\}_{h=0,i=0}^{k-1,n})+A_j^k g_j(\tau)]$$

so $\Phi_j^{(k)}(\alpha,t)$ is in H_0 as $F_j^k(\{\Phi_i^{(h)}(\alpha,\tau)\})\in H_0$ e $Q:H_0\to H_0$. □

Notice that $A_j^h(\alpha)$ is now the integral of a function in H_0 and that (obviously) $\dot{\Phi}_j^h(t)$ is in H_0 as well. For simplicity we will normalize the g_i setting

$$\Im x_i^0 g_i=1.$$

In this Chapter we will always consider truncated series:

$$A_i(\alpha)=\sum_{k=1}^{K}(\eta)^k A_i^k(\alpha),\Phi_i(t,\alpha)=\sum_{k=0}^{K}(\eta)^k\Phi_i^k(t,\alpha)\dots\dots,$$

with $K=\varepsilon^{-\frac{1}{2\tau_F}}\equiv\varepsilon^{-q}$. However the relations we will find are all formal series relations on the corresponding complete series. We will express the A_j^k and $\Phi_i^k(t,\alpha)$ as values of finite sums of fruitless trees (see the next subsection).
This means that we can use the bounds on fruitless trees discussed in Chapter 3.

Lemma 5.3. *Provided that $f(\psi,q)$ is analytic in some $H(a,D)$ and respects the bounds 3.2 then:*
(i) $\Phi_j^k\in H_0$ respects the bounds of Remark 3.9. Moreover if $|t|=O(\eta)$ it respects the bound 3.10.
(ii) A_j (and all the values $\Im\mathcal{W}^1$ of the trees we will describe in the following subsection) is the truncation of an asymptotic power series in η,ε.

Proof. (i) To apply Remark 3.9 we only need to remember that $\Phi_j^k\in H_0$ is bounded by construction. We will see in the next subsection that Φ_j^k is a finite sum of values of analytic trees.
(ii) The A_j^k are the integral $\Im$ of functions in H_0 so their tree representation will be through analytic trees which can be bounded by 3.8. □

We will repeatedly use Lemma 3.14 to write formal power series identities as identities between the order K truncations plus a known (smooth in η and α) remainder of order $o(e^{-K})$. We will say that the identity is true up to order $O(\eta^K)$.

As seen for system (1) the energy conservation implies that (Subsection 4.1.2) the value of $A_0^k(\alpha)$ is related to the $A_j^k(\alpha)$ with $j\neq 0$. For compactness we will state this relation in terms of the sum

$$A_i(\alpha)=\sum_{k=1}^{K}(\eta)^k A_i^k(\alpha),\Phi_i(t,\alpha)=\sum_{k=0}^{K}(\eta)^k\Phi_i^k(t,\alpha)\dots$$

Proposition 5.4. *For each value of α we define*

$$k_i=(-\frac{1}{2})^{\delta_{i\,0}}\Im\dot{g}_i(t)(\Phi_i(\alpha,t)-\Phi_i^0(\alpha,t))=O(\eta),$$

for $i = 0, \dots, n$. We have:

$$2A_0(\alpha) = \frac{1}{1+k_0}\sum_{j=1}^{n} A_j(\alpha)(\omega_j + k_j) \tag{5.6}$$

up to order $O(\eta^K)$.

Proof. Our auxiliary system is the order K truncation of the expansion of system 5.2. This means that, by Lemma 3.14, for small enough η, ε, the function $\Phi_j(\alpha, t) \in H_0$ solves the equation 5.2 up to order $O(\eta^K)$:

$$\ddot{\Phi}_i = \delta_{i0}\sin(\Phi_0) - (\eta)a_i\partial_i f(\Phi) + A_i(\eta)g_i(t) + (\eta)^K F_R(\alpha, t). \tag{5.7}$$

The function F_R is analytic and bounded in $t \in \mathbb{R} \times (-id, id)$, $|\ \mathrm{Im}\ \alpha| \le s_0$. Remark 3.9 and Lemma 3.14 (iii) ensure that $|F_R(\alpha, t)| \le C^K \varepsilon^{-pK}$ for some $p \in \mathbb{N}$. The energy conservation for system 5.2 leads to:

$$\sum_{j=0,\dots,n} \frac{\dot{\Phi}_j^2(\alpha,t)}{2} + \cos(\Phi_0(\alpha,t)) - (\eta)f(\Phi(\alpha,t)) - \sum_j A_j(\alpha)\Im^t\dot{\Phi}_j(\alpha,t)g_j + G_R(\alpha,t) = \text{cost},$$

the function G_R has the same properties as F_R. $\dot{\Phi}_j(t)$ is continuous and $\dot{\Phi}_j^0 = \omega_j$ for $j \neq 0$, $\dot{\Phi}_j^0 = -2x_0^0$, so we obtain:

$$\textstyle\sum_{j=0}^{n} A_j \Im\dot{\Phi}_j g_j = 0 = \sum_{k=0}^{K}(\eta)^k\big\{\sum_{i=1}^{n}\big[\omega_i A_i^k \Im(g_i) + \sum_{h<k} A_i^h \Im\dot{g}_i\Phi_i^{k-h}\big] -$$

$$\textstyle 2A_0^k(\alpha)\cdot\Im(x_0^0 g_0) + \sum_{h<k} A_0^h \Im\dot{g}_0\Phi_0^{k-h}\big\}.$$

This is a formal power series relation:

$$\sum_{i=1}^{n}\omega_i A_i \Im(g_i) - 2A_0\Im(x_0^0 g_0) \sim \sum_{i=1}^{n} A_i\Im\dot{g}_i(\Phi_i - \Phi_i^0) + A_0\Im\dot{g}_0(\Phi_0 - \Phi_0^0).$$

We use the normalization $\Im g_i x_i^0 = 1$. By the boundedness of the Φ_i^k the

$$k_i = (\frac{1}{2})^{\delta_{0i}}\Im\dot{g}_i(\Phi_i - \Phi_i^0)$$

are smooth functions of α and are of order η, so $(1+k_0)^{-1}$ is a well defined asymptotic series for small enough η. Passing to the order K truncation we obtain the desired equality. Notice that the remainder is a known smooth function of η and α by Lemma 3.14(ii). □

5.2 Tree representations of the auxiliary dynamics

In describing the tree representation of $\Phi_j^{(h)}(\alpha)$ we will use the fact that the structure is the same as in the tree representation of system (1). Therefore we will not repeat the proofs (which are identical) but simply cite the Theorems we are adapting.

Definition 5.5. *Let $\mathcal{T}$ be the set (of equivalence classes) of marked rooted labeled trees with fruits such that:*
a) each node carries the labels $j = 0, \dots, n$ $\delta = 0, 1$ and $k \geq 0$.
b) The labels respect the following grammar:

$$\delta_v = 0\,,\; k_v = 0 \textit{ implies that } j_v = 0\;,\; s(v) \geq 2;$$

$$k_v > 0 \textit{ implies that } \delta_v = 0\; s(v) = 0.$$

By definition we will call fruits the nodes with $k_v > 0$. The markings are the same as in Section 1.2 (i.e. an angle marking $J_v = 0, \dots, n$ and a function marking $h(t, v) \in H_0$). As usual we will consider the vector space $\mathbb{V}(\mathcal{T})$ generated by $\mathcal{T}$ on $\mathbb{Q}$. We can redefine all the subspaces of Definition 1.31.

The order of a tree is now: $o(A) = \sum_v \delta_v + k_v$ and we can express $\mathbb{V}(\mathcal{T})$ as a direct sum of finite dimensional spaces of prescribed order. As in the preceding Section we will call $\mathcal{S}(B)$ the symmetry group of a tree $B \in \mathcal{T}$.

Given these definition we can set (see identity 2.1) $\alpha(\delta, k) = \overset{j,\delta,k}{\bullet}$, $\partial_j^{v_0}\alpha(\delta, k) \in \mathcal{T}_j$ and:

$$\mathcal{V}^1(\partial_j^{v_0}\alpha(\delta,k)) = \begin{cases} (\eta) a_j Q_j(\nabla^{e_j} f^1)] & \text{if } k = 0, \delta = 1 \\ \\ a_j(\eta)^k A_j^k Q_j(g_j)] & \text{if } k \neq 0, \delta = 0 \end{cases}$$

We can repeat what done in subsection 2.1.1 and set Φ_j^k in correspondence with elements Λ_j^k of $\mathbb{V}(\overset{0}{\mathcal{T}}{}_j^k j)$. Then we can restate Proposition 2.2:

Proposition 5.6. *For each j, k $\mathcal{V}^1(\Lambda_j^k) = \Phi_j^k(\alpha, t)$ where:*

$$\Lambda_j^k = \sum_{A \in \overset{0}{\mathcal{T}}{}_j^k} \frac{1}{|\mathcal{S}(A)|} A \equiv \sum_{A \in \overset{0}{\mathcal{T}}{}_j^k} c(A)$$

Now, as in Subsection 2.1.2, to each tree $A \in \overset{0}{\mathcal{T}}$ (possibly marked) we associate a value $\mathcal{W}^1$.

Given a tree in with no marks on the first node we add the marks $j_1, \dots, j_l$, $h_1(t)$,

$\ldots, h_l(t)$ and set:

$$\mathcal{W}^1(\prod_{i=1}^{l} h_i(v_A,t)\partial_{j_i}^{v_A}\alpha(1,0)) = \prod_i h_i(v_A,t)\nabla^{\sum_i e_{j_i}} f^1,$$

$$\mathcal{W}^1(\prod_{i=1}^{l} \partial_{j_i}^{v_A} h_i(v_A,t)A) = \prod_i h_i(v_A,t)(\nabla^{\vec{m}+\sum_i e_{j_i}} f^\delta) \prod_{v\in s(v_A)\cap \overset{0}{A}} a_{j_v} Q_{j_v}[\mathcal{W}^1(A^{\geq v})] \prod_{v\in \mathcal{F}(v_A)}$$

$$a_{j_v} Q_{j_v}[g_{j_v}] A_{j_v}^{k_v},$$

$$A_j^k = -\Im x_j^0 \mathcal{W}^1(\overset{0}{\Lambda}{}_j^k),$$

where $\vec{m}_i$ is the number of nodes in $s(v_0)$ having $j_v = i$. As in Section 2.1.2 $\overset{0}{\Lambda}_j$ is $\Lambda_j - \mathcal{F}_j$.

In $\overset{0}{\mathcal{T}}$ we can define the change of first node. Notice that Proposition 2.11 is still true.

Remark 5.7. *A tree with a fruit $k_v \neq 0$ j_v is equivalent to the tree deprived of the fruit, marked $Q_{j_v}[g_{j_v}]$ on the node w preceding v, and multiplied by $a_{j_v} A_{j_v}^{k_v}$. Notice that the dependence on the initial data is contained only in $A_{j_v}^{k_v}(\alpha)$ and that this markings are always in H_0.*

This means that we can use the analytic bounds 3.8 to bound the values of trees in $\overset{0}{\mathcal{T}}$ (even those with "fruits" i.e. nodes with $k_v > 0$). On the other hand as $h_j(t) = Q_j^t(g_j)$ we can apply proposition 1.16 (i) and the change of first node also to the nodes with $k_v > 0$ except that we never obtain trees $\in \overset{0}{\mathcal{T}}$.

Remark 5.8. *The A_j^k and Φ_j^k are trigonometric polynomials in α.*

Lemma 5.9. *We can restate Propositions2.14 and 2.15, if we consider,*
1) For $i, j \neq 0$, for all k and for any $f, h \in H_0$ we have the identity:

$$\Lambda^k_{(i,f)\,(j,h)} = \Lambda^k_{(j,h)\,(i,f)}.$$

2) If $i \neq 0$, for all k and for any $f, h \in H_0$ we have the identity:

$$\Lambda^k_{(i,f)\,(0,h)} + L_h(\Lambda^k_{(i,f)}) = \Lambda^k_{(0,h)\,(i,f)}.$$

the linear operator L_h is defined in Section1.2.2.

Definition 5.10. *We call $\overset{r}{\mathcal{T}}$ the subset of $\overset{0}{\mathcal{T}}$ of trees that stay in $\overset{0}{\mathcal{T}}$ by applying the change of first node*[2]*.*

[2] our convention is not to consider trees with only one node $k_v > 0$ in $\mathcal{T}^r$ as these are not proper nodes

Definition 5.11. *We consider a function $\tilde{S}(\alpha)$ that is similar to the generating function $S(\varphi)$. We call $\hat{\mathcal{T}}$ the set of trees in $\overset{r}{\mathcal{T}}$ having no markings; recall that for any $B \in \overset{0}{\mathcal{T}}$, $N(B)$ is the number of free nodes of B (i.e. such that $k_v = 0$) and $c(B) = \frac{1}{\mathcal{S}_{v_B}(B)}$:*

$$\tilde{S}(\alpha) = \Im \mathcal{W}^1_\alpha \{ \sum_{B \in \hat{\mathcal{T}}} \frac{c(B)}{N(B)} B \}$$

as usual restricted to trees of order $\leq K$.

Lemma 5.12. *The following identity holds:*

$$A_j(\alpha) = \Im \mathcal{W}^1_\alpha \sum_{B \in \hat{\mathcal{T}}} \frac{c(B)}{N(B)} \sum_{v \in B} \partial^v_j B$$

Proof. The proof is identical to that of Theorem4.1. $\square$

Clearly this is different from $A_j(\alpha) = \partial_j \tilde{S}(\alpha)$, which is generally false as we are considering trees with α dependent fruits. Nevertheless the $A_j(\alpha)$ are linear functions of $\partial_\alpha \tilde{S}(\alpha)$ as we will see in the following proposition. Let us first prove a technical Lemma.

Lemma 5.13. *Let $\hat{\mathcal{T}}^k$ be the subspace of $\overset{r}{\mathcal{T}}$ of trees of order k:*

$$\Im \mathcal{W}^1_\alpha [\sum_{B \in \hat{\mathcal{T}}^k} \frac{c(B)}{N(B)} \sum_{v : k_v > 0} \frac{d_{\alpha_j} A^{k_v}_{j_v}}{A^{k_v}_{j_v}} B] = \sum_{i,h<k} d_{\alpha_j} A^h_i \Im g_i (\Phi^{k-h}_i - A_i Q_i g_i)$$

Proof. We fix our attention on the nodes v of B with $k_v > 0$; we have

$$\sum_{B \in \hat{\mathcal{T}}^k} \frac{c(B)}{N(B)} \sum_{v : k_v > 0} \frac{d_{\alpha_j} A^{k_v}_{j_v}}{A^{k_v}_{j_v}} B = \sum_{B \in \hat{\mathcal{T}}^k} \frac{c(B)}{N(B)} \sum_{[v] : k_v > 0} m[v] \frac{d_{\alpha_j} A^{k_v}_{j_v}}{A^{k_v}_{j_v}} B. \tag{5.8}$$

Now we shift the first node in v (a representative of the coset $[v]$), we obtain $A(v) \notin \mathcal{B}$ whose first node v has $k_v > 0$, moreover in $A(v)$: $s(v) = v_1$ e $j_{v_1} = j_v$[3]; we call the set of trees of this form $\mathcal{D}^{k,k_v}$. Notice that any tree in $B \in \mathcal{T}^{k-k_v}_j$ with at least a node $v : k_v = 0$ is equal to $A^{\geq v_1}$ for some tree $A(v) \in \mathcal{D}^{k,k_v}$; moreover the value of $A(v)$ is:

$$\Im g_{j_v} a_{j_v} Q^t_{j_v} \mathcal{W}^1 [A^{\geq v_1}].$$

Now [4] $\Phi^{k-h}_j - a_j A^{k-h}_j Q_j g_j$, is a sum of trees in $\mathcal{T}^{k-h}_j$ with at least one node $v : k_v = 0$, and we have:

$$\sum_{i,h<k} \Im g_i (\Phi^{k-h}_i - a_i A^{k-h}_i Q_i g_i) = \Im \mathcal{W}^1_\alpha (\sum_{A \in \mathcal{D}^{k,h}} c(A) A) \tag{5.9}$$

[3] remember that $\mathcal{P}(A, v)$ shifts the labels j_w of the nodes in the path v_A, v towards the first node v_A, see Definition 1.40

[4] we subtract the only tree with all the node labels $k_v > 0$

as in $\mathcal{D}^{k,h}$ $c(A) = c(A^{\geq v_1})$.
In the right hand side of relation 5.9 we consider $N(A)$ copies of each tree A in each we evidence a node v with $k_v = 0$

$$\sum_{A \in \mathcal{D}^{k,h}} c(A)A = \sum_{A \in \mathcal{D}^{k,h}} \frac{c(A)}{N(A)} \sum_{[v]:k_v=0} m[v]A.$$

□

Proposition 5.14. *for each α we have:*

$$\partial_{\alpha_j} \tilde{S}^k(\alpha) = A_j^k + \sum_{h<k, i=0,\dots,n} M_{j\,i}^h \Im(g_i \Phi_i^{k-h}) - \sum_{h<k, i=0,\dots,n} M_{j\,i}^h A_i^{k-h} \Im(g_i Q_i(g_i))$$

where $M_{i\,j} = \partial_i A_j(\alpha)$ is an $n \times n+1$ matrix.

Proof. To prove this assertion we first consider the relation

$$\partial_{\alpha_j} \tilde{S}^k(\alpha) = \Im \mathcal{W}_\alpha^1 [\sum_{B \in \hat{\mathcal{T}}} \frac{c(B)}{N(B)} \sum_{v \in B} \partial_j^v B] + \Im \mathcal{W}_\alpha^1 [\sum_{B \in \hat{\mathcal{T}}} \frac{c(B)}{N(B)} \sum_{v:k_v>0} \frac{\partial_{\alpha_j} A_{j_v}^{k_v}}{A_{j_v}^{k_v}} B]. \tag{5.10}$$

Then by Lemma5.12 the first sum is $\Im g_j A_j^k$, in the second sum we set $M_{j\,j_v}^{k_v} = \partial_{\alpha_j} A_{j_v}^{k_v}$ and apply Lemma 5.13. □

Now we want an homogeneous linear equation relating $\nabla \tilde{S}$ to $A = \{A_j\}_{j=1}^n$ of the type

$$A = (1 - \mathcal{M})^{-1} \nabla \tilde{S}(\alpha),$$

where $\mathcal{M}$ is an $n \times n$ matrix of order η. In order to have such an identity we have to impose conditions on the functions g_j. There are (at least) two possible and nearly equivalent choices. One has a clearer dynamical meaning (and is the condition proposed in [BB1]) and leads to possibly more explicit formulas; the second on the other hand can be easily implemented by a computer, moreover it is obvious that there exist functions $g_i \in H_0$ satisfying the latter condition so one does not have to verify the existence. We will describe both conditions and use the second one.

$$(a) \quad \Im[g_i(\Phi_i(\alpha) - \Phi_\alpha^0)] = 0 \text{ for each } i = 0, \dots, n. \tag{5.11}$$

With this restriction Proposition 5.14 states that:

$$\partial_{\alpha_j} \tilde{S}(\alpha) = A_j - \sum_{i=1,\dots,n} M_{j\,i} A_i \Im(g_i Q_i(g_i)) -$$

$$\frac{1}{2} \frac{M_{j\,0}}{1 + k_0} \sum_{i=1,\dots,n} (\omega_i + k_i) A_i \Im(g_0 Q_0(g_0))$$

up to order $(\eta)^K$. We write this relation compactly as:

$$\nabla\tilde{S}(\alpha) = (1-\mathcal{M})A \text{ where } \mathcal{M}_{ij} = M_{ij} + \frac{M_{i\,0}}{1+k_0}(\omega_j + k_j).$$

Notice that $|\mathcal{M}| = O(\eta)$ so that $1-\mathcal{M}$ is invertible.
The second condition (which is easier to verify) is the following, as usual we call $\mathcal{A}$ the set of fruitless trees,

$$\mho_i^k = \sum_{B\in\mathcal{A}_i^k} c(B)B,$$

the value of this sum does not depend on the choice of the g_i and we ask that[5]:

$$(b) \qquad \Im[g_i(t)\mathcal{V}^1(\sum_{k=1}^{K}(\eta)^k\mho_i^k] = 0 \text{ for each } i. \tag{5.12}$$

This means that $\Im g_i(\Phi_i(\alpha,t)-\Phi_i^0(\alpha,t))$ depends only on trees with at least one node $k_v > 0$ and so:

$$\Im g_i(\Phi_i(\alpha,t)-\Phi_i^0(\alpha,t)) = \sum_{j=1,\ldots,n} A_j(\alpha,\eta)C_{ij}(\alpha,\eta)$$

up to order $(\eta)^K$.
We define

$$C_{ij}(\alpha,\eta) = \delta_{ij}a_i\Im g_i Q_i(g_i) + \sum_{h=1}^{K}(\eta)^h\Im g_i C_{ij}^{(h)}(\alpha,t).$$

The functions $C_{ij}^{(h)}$ are $\mathcal{V}_\alpha^1$ applied to $\overset{0}{\Lambda}{}_i^h$ deprived of one "fruit" with label j[6]. We substitute in Proposition 5.14 and find

$$\partial_{\alpha_j}\tilde{S}(\alpha) = A_j - \sum_{\substack{i=1,\ldots,n\\ l=0,\ldots,n}} M_{j\,l}C_{li}A_i - \frac{1}{2(1+k_0)}\sum_{l=0,\ldots,n} M_{j\,l}C_{l0}\sum_{i=1}^{n}A_i(\omega_i+k_i)$$

which is the required linear relation; in this case:

$$\mathcal{M}_{ij} = \sum_{l=0,\ldots,n} M_{il}C_{lj} - \frac{1}{2(1+k_0)}\sum_{l=0,\ldots,n} M_{i\,l}C_{l0}(\omega_j+k_j).$$

Proposition 5.15. *The generating function can be written in compact notation as:*

$$\tilde{S}(\alpha,\eta) = \int_0^\eta \{\Im f(\Phi(\alpha,t)) + \sum_{i=0,\ldots,n}\partial_{\eta'}(A_i)[\Im(g_i(\Phi_i(\alpha,t)-\Phi_i^0(\alpha,t))]d\eta'$$

[5] for any non zero ε this is a finite set of orthogonality conditions

[6] We discussed in Lemma 4.9 the problem of taking away a fruit from a tree without changing its combinatorial coefficient, for trees with more than one fruit it is not easy to describe the needed linear function, but it is clear that it is well defined , so we will not go in any details.

$$- \sum_{i=0,\dots,n} A_i^2 \Im(g_i Q_i g_i); \tag{5.13}$$

up to order $O(\eta^K)$.

Proof. The proof is identical to that of Corollary 4.2; In $\tilde{S}^k$, for each tree A consider k copies:

$$\tilde{S}^k = \sum_{B \in \hat{\mathcal{T}}^k} \frac{c(B)B}{N(B)} = \frac{1}{k} \sum_{B \in \hat{\mathcal{T}}^k} \Big(\sum_{\substack{[v]:k_v=0 \\ \delta_v=1}} \frac{c(B)m[v]B}{N(B)} + \sum_{[v]:k_v>0} \frac{c(B)m[v]k_v B}{N(B)} \Big)$$

The first sum in the right hand side is equivalent to the first term in the right hand side of 5.13, as in Corollary 4.2. Finally we can apply Lemma 5.13 to the second term:

$$\Im \mathcal{W}^1 \Big[\sum_{[v]:k_v>0} \frac{c(B)m[v]k_v B}{N(B)} \Big] = \sum_{h=1}^{k-1} \sum_{i=0}^{n} h A_i^h \Im(g_i(\Phi_i^{k-h} - A^{k-h} Q_i g_i)).$$

Now we consider two formal power series:

$$A = \sum_{h=1}^{\infty} (\eta)^h A^{(h)} \quad B = \sum_{h=1}^{\infty} (\eta)^h B^{(h)},$$

the following equality holds:

$$\int_0^\eta d\eta' B \partial_\eta A \sim \sum_{k=2}^{\infty} \frac{(\eta)^k}{k} \sum_{h=1}^{k-1} h A^h B^{k-h}.$$

Finally if we chose $A = B$ we obtain that:

$$\frac{1}{2} A^2 \sim \sum_{k=2}^{\infty} \frac{(\eta)^k}{k} \sum_{h=1}^{k-1} h A^h B^{k-h}.$$

Notice that condition (a) would give a cleaner expression for the generating function. As usual the remainder is a known smooth function of η, α. □

We can gather the results in the following Theorem:

Theorem 5.16. *Given* $n+1$ *functions* $g_j(t)$ *respecting condition 5.12, We can fix the initial data and the functions* $A_j(\alpha)$ *so that for each* $\alpha \in \mathbb{T}^n$*: (1) The order* K *solution of equation 5.2,* $\Phi(\alpha, t) \in H_0$*, it is a polynomial in* η *and a trigonometric function in* α*.*

(2) There exists an order K *generating function* $\tilde{S}(\alpha)$*, again polynomial in* η *and trigonometric function in* α*. This function is of order* $O(\eta)$ *together with its* α *derivatives; moreover it respects Proposition 5.15 (always to order* $(\eta)^K$*) and is the integral*

$\Im$ of a function in H_0. (3) The coefficients $A_j(\alpha)$ again polynomial in η and trigonometric functions in α are related to each other and to the generating function by the identities (valid up to order K):

$$2A_0(\alpha) = \frac{1}{1+k_0}\sum_{j=1}^{n} A_j(\alpha)(\omega_j + k_j)$$

$$A = (1+\mathcal{M})^{-1}\nabla\tilde{S}$$

5.3 Connection between the auxiliary dynamics and the splitting

Theorem 5.17. *There exists an analytic change of coordinates[7] $\vartheta : \mathbb{T}^n_{s_1} \to \mathbb{T}^n_{s_1}$ such that $\tilde{S}(\vartheta(\varphi)) = S(\varphi) + o(\eta^K)$.*

We follow closely the strategy of [BB1]. First we move along the trajectory for a time t_α such that $\Phi_0(t_\alpha) = \pi$.

Lemma 5.18. *For each $\alpha \in \mathbb{T}^n_{s_1}$ there exists $t_{\alpha,\eta}$ analytic in $\alpha \in \mathbb{T}^n_{s_1}$ and $|\eta| \le \eta_0$ such that:*

$$\Phi_0(\alpha, t(\alpha,\eta), \eta) = \pi \quad t(\alpha, 0) = 0$$

Proof. We apply the implicit function theorem knowing that

$$\Phi_0(\alpha, 0, 0) = \pi \quad \dot{\Phi}_0(\alpha, 0, 0) = -2.$$

By our bounds 3.10 we have that

$$\sup_{\substack{|\eta|\le\eta_0 \\ \alpha\in\mathbb{T}^n_s}} |\Phi_0(\alpha, 0, \eta) - \pi| \le \eta_0 C,$$

so $|t(\alpha,\eta)| \le C\eta_0$. Then we verify:

$$\sup_{\substack{|\eta|\le\eta_0 \\ |t|\le\eta_0 C\ ,\ \alpha\in\mathbb{T}^n_s}} |1 + 2\dot{\Phi}_0(\alpha, t, \eta)| = \sum_{k=1}^{K}(\eta)^k \sup_{\substack{|\eta|\le\eta_0 \\ |t|\le\eta_0 C\ ,\ \alpha\in\mathbb{T}^n_s}} |\Phi_0^k(\alpha, t)| \le \frac{1}{2}.$$

Notice that $\nabla_\alpha t(\alpha, 0) = 0$ so $\nabla_\alpha t(\alpha,\eta) = O(\eta)$. □

Lemma 5.19. *Now consider the application $\mathbb{T}^n_{s_1} \to \mathbb{T}^n_{s_1}$:*

$$\varphi_j = \Phi_j(\alpha, t(\alpha,\eta), \eta) \tag{5.14}$$

for sufficiently small values of η this is a diffeomorphism of $\mathbb{T}^n_{s_1}$, η close to the identity.

[7] remember that for $\alpha \in \mathbb{T}^n_s$ we mean the thickening of the torus of length s: $\mathbb{T}^n_s = \mathbb{T}^n \times (-is, is)$

Proof. Let us write relation 5.14 as: $\varphi_j(\alpha,\eta)=\alpha+\vartheta_1(\alpha,\eta)$ with $\vartheta_1(\alpha,0)=0$; precisely:

$$\vartheta_1(\alpha,\eta)=\omega t(\alpha,\eta)+\sum_{k=1}^{K}(\eta)^k\Phi_j^k(\alpha,t(\alpha,\eta)).$$

The relation is invertible locally as:

$$\varphi_j(\alpha,0)=\alpha_j \quad \partial_i\Phi_j(\alpha,0,0)=\delta_{ij}.$$

The function we obtain is a diffeomorphism of $\mathbb{T}^n_d$ provided that

$$\sup_{\substack{|\eta|\le\eta_0\\ \alpha\in\mathbb{T}^n_s}}|\omega_j\nabla_\alpha t(\alpha,\eta)\big(1+\sum_{k=1}^{K}(\eta)^k\dot\Phi_j^k(\alpha,t(\alpha,\eta))\big)+\nabla_\alpha\sum_{k=1}^{K}(\eta)^k\Phi_j^k(\alpha,t(\alpha,\eta))|<\frac{1}{2}.$$

This holds true as $|\nabla_\alpha t(\alpha,\eta)|=O(\eta)$. Moreover $|\dot\Phi_j^k(\alpha,t)|$ and $|\nabla_\alpha\Phi_j^k(\alpha,t)|$ are bounded by 3.10 for $|t|\le C\eta$. □

Now we invert the relation $\varphi_j(\alpha,\eta)=\alpha+\vartheta_1(\alpha,\eta)$ we call the inverse $\vartheta_2(\varphi)$ and $t(\vartheta_2(\varphi),\eta)\equiv t_\varphi$.

Consider the equation:

$$\ddot\Psi_j(t)=f_j(\Psi(t))+\delta_{j0}\sin(\Psi_0(t))+A_j(\vartheta_2(\varphi))g(t+t_\varphi)+(\eta)^KF_R(\vartheta_{2(\varphi)},t+t_\varphi) \quad (5.15)$$

with initial data $\Psi_j(0)=\varphi$ if $j\neq 0$ and $\Psi_0(0)=\pi$. The function F_R is defined in 5.7. This equation admits an order $(\eta)^K$ solution (we call it $\Psi_j(\varphi,t)$) which is the truncation of an asymptotic power series in η,ε. So for $\eta\le\varepsilon^p$ the solution is η close to the separatrix of the pendulum and exponentially quasi-periodic.
We can solve equations 5.15 perturbatively and, as the initial data are now η independent, we obtain:

$$\Psi_j^k(t)=O_j[F_j^k(\Psi^h)+A_j^kg_j].$$

Lemma 5.20. *The asymptotic conditions determine the solution uniquely so* $\Psi(\varphi,t)=\Phi(\vartheta_2(\varphi),t+t_\varphi)$ *up to order* $(\eta)^K$.

Proof. $\Phi_j(\vartheta_2(\varphi),t+t_\varphi)$ and $\Psi_j(t)$ coincide at $t=0$ by definition. Moreover they solve the same equation up to order $O(\eta^K)$. Namely as seen in expression 5.7 there exists $G_R(\alpha,t)$ such that

$$\ddot\Psi_j(t)=f_j(\Psi(t))+\delta_{j0}\sin(\Psi_0(t))+A_j(\vartheta_2(\varphi))g(t+t_\varphi)+(\eta)^KG_R(\vartheta_2(\varphi),t+t_\varphi)$$

where G_R is bounded and exponentially quasi-periodic with $<G_R>=0$. So $H(\varphi,t)=\Phi_j(\vartheta_2\varphi,t+t_\varphi)-\Psi_j(t)$ is a bounded and exponentially quasi periodic solution of

$$\ddot H(\varphi,t)=(\eta)^KG_R(\vartheta_{2(\varphi)},t+t_\varphi).$$

By the results of Subsections 1.1.3 and 1.1.4 $H(\varphi,t)=O(\eta^K)$ for $t\in\mathbb{R}\times(-id,id)$, $|$ Im $\alpha|\le s_0$. Remark 3.9 and Lemma 3.14 (iii) ensure that $|H_R(\alpha,t)|\le(C^K\varepsilon^{-pK}$ for some $p\in\mathbb{N}$. Moreover $H(\alpha,t)$ is analytic in α for $|$ Im $\alpha|\le\frac{1}{2}s_0$ □

We can represent the series expansion in term of trees; in this case we have the non analytic operators O_j and so we do not think of the nodes $k_v > 0$ as fruits and apply the operators $\mathcal{V}$ and $\mathcal{W}$. Notice that the nodes with $k_v > 0$ now have value $a_j A_j^k(\vartheta_2(\varphi)) O_j^t(g(t+t_\varphi)$.

Lemma 5.21. *The generating function*

$$\bar{S}(\varphi) = \Im \mathcal{W}_\varphi \{ \sum_{B \in \mathcal{T}^r} \frac{c(B)}{N(B)} B \}, \tag{5.16}$$

satisfies the relation:

$$\bar{S}(\varphi) = \tilde{S}(\vartheta_2(\varphi)) + \sum_{i=1}^{n} A_i^2 \Im [g_i [Q_i^{t+t_\varphi}(g_i) - O_i(g_i)]] + O(\eta^K).$$

Proof. We use Proposition 5.15 which can be obviously restated for Ψ_j as:

$$\bar{S}(\varphi,\eta) = \int_0^\eta \{ \Im f(\Psi(\varphi,t)) + \sum_{i=0}^{n} (\partial_{\eta'} A_i) [\Im (g_i(\Psi_i(\varphi,t) - \Psi_i^0(\varphi,t))] d\eta'$$

$$- \sum_{i=0}^{n} A_i^2 \Im (g_i O_i g_i);$$

then we apply Lemma 5.20 and $\Psi_i^0 = \Phi_i^0$ obtaining

$$\int_0^\eta \{ \Im f(\Phi(\vartheta_2(\varphi),t)) + (\partial_{\eta'} A_i) [\Im (g_i(\Phi_i(\vartheta_2(\varphi),t) - \Phi_i^0(\vartheta_2(\varphi),t))] d\eta' -$$

$$A_i^2 \Im (g_i Q_i g_i)] + \sum_{i=0}^{n} A_i^2 \Im (g_i(Q_i g_i - O_i g_i)) + L_R(\varphi),$$

where L_R is analytic in φ for $|\text{ Im } \varphi| \leq \frac{s_0}{4}$.

□

To avoid confusion with the complex norm, $|v| = \sqrt{\sum_i v_i \bar{v}_i}$, we will define for all $v \in \mathbb{C}^n$ $q(v) = \sum_{i=1}^n v_i^2$.

Lemma 5.22. *Following [BB1] we prove that for all $\varphi \in \mathbb{T}^n_{s_0}$:*

$$|\tilde{S}(\vartheta_2(\varphi)) - S(\varphi) + o(\eta^K)| \leq C |q(\nabla \tilde{S})|. \tag{5.17}$$

Proof. By our definitions $\bar{S}(\varphi) - S(\varphi)$ is the (value $\Im \mathcal{W}$ of the) sum of trees A with at least one node $v : k_v > 0$, weighted by $\frac{c(A)}{N(A)}$, so it is at least linear in the A_j^h for $j = 1, \ldots, n$, $h = 1, K-1$; as usual we call fruitless trees the trees of order zero in A_j etc.. The linear term in the A_j^h is sum of trees with only one fruit j, h:

$$\text{Lin} = \sum_{B \in \hat{\mathcal{T}}^{1F}} c(B) B.$$

As there is only one fruit (node v_1) the coefficient $c(B) = \frac{1}{\mathcal{S}(B)}$ is the order of the stabilizer of the path $v_B\ v_1$. We can shift the first node to v_1 without changing the combinatorial coefficient and apply 5.13)

$$\Im \mathcal{W}^1(\text{Lin}) = \sum_{j=0,\dots,n} A_j \Im g_j(t + l_\varphi) \Psi_j^{0F}(t) = A_j(\Im g(\tau) \mathcal{V}^1(\mho_j(\tau)) + A_j(\Im g(\tau) \mathcal{V}^1(\Lambda_j^{1F}(\tau))$$

plus higher order terms in A_j. By condition 5.12 the linear part is zero. Then

$$\bar{S}(\varphi) - S(\varphi) = \sum_{k=3,K} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{l=1}^{k-2} \sum_{h=1}^{k-l-1} A_i^l A_j^h \tilde{C}_{ij}^{k-l-h}$$

where $\tilde{C}_{ij}^k$ is a sum of trees, deprived of two fruits with labels i, j, and whose order without these fruits is k. Now we substitute $\bar{S}$ with $\tilde{S}$ using Lemma 5.21. Finally we substitute $(\Im g - \mathcal{M})^{-1} \nabla \tilde{S} = A + o(\eta^K)$.

As usual in equation 5.17 we can explicitly compute the remainder which is an analytic function of η and φ for $\varphi \in \mathbb{T}^n_{s_1}$ with for instance $s_1 = \frac{1}{4} s_0$. □

Now finally we can prove the theorem and construct the transformation $\vartheta : \mathbb{T}^n_d \to \mathbb{T}^n_d$ $(d < s_0/4)$ sending § in $\tilde{S}$. This is almost identical to the proof of Theorem 4.1 in [BB1].

Proof of Theorem 5.17. We want to find ϑ such that

$$\tilde{S}(\vartheta_2(\varphi) + \vartheta(\varphi)) - S(\varphi) = f_R(\eta, \varphi) = o(\eta^K), \tag{5.18}$$

for some function $f_R(\eta, \varphi)$ analytic in $\eta \le \eta_0$ and $\varphi \in \mathbb{T}^n_d$. Note that if $\nabla \tilde{S}(\alpha) = 0$ then the equation is solved by $\vartheta = 0$. In general we look for a solution of the form

$$\vartheta(\varphi) = \nabla \tilde{S}(\alpha)|_{\alpha = \vartheta_2(\varphi)} y \tag{5.19}$$

where y is a scalar parameter and from now on we will write $\nabla \tilde{S}(\vartheta_2(\varphi))$ instead of $\nabla \tilde{S}(\alpha)|_{\alpha=\vartheta_2(\varphi)}$. Then we can write[8]:

$$\tilde{S}(\vartheta_2(\varphi) + v) = \tilde{S}(\vartheta_2(\varphi)) + \nabla \tilde{S}(\vartheta_2(\varphi)) v + (v, R(\varphi, \eta, v) v \tag{5.20}$$

where the matrix $R(\varphi, \eta, v)$ is such that:

$$\Big(v, R(\varphi, \eta, v) v\Big) = \tilde{S}(\vartheta_1(\varphi) + v) - \tilde{S}(\vartheta_1(\varphi)) - \nabla \tilde{S}(\vartheta_1(\varphi)) v.$$

[8]The operator (a, b) with $a, b \in \mathbb{C}^n$ is the real scalar product

$$(a, b) = \sum_{i=1}^{n} a_i b_i$$

Substituting 5.19 in 5.20, we find that 5.18 is equivalent to:

$$\bar{S}(\varphi) - S(\varphi) + q\big(\nabla\tilde{S}(\vartheta_2(\varphi))\big)y+$$

$$(\nabla\tilde{S}(\vartheta_2(\varphi)), R(\varphi,\eta,\nabla\tilde{S}(\vartheta_1(\varphi))y)\nabla\tilde{S}(\vartheta_1(\varphi)))y^2 = o(\eta^K),$$

and finally to

$$\frac{\bar{S}(\varphi) - S(\varphi) + o(\eta^K)}{q\big(\nabla\tilde{S}(\vartheta_2(\varphi))\big)} = y + R_1(\varphi,\eta,\nabla\tilde{S}(\vartheta_2(\varphi))y)y^2 \tag{5.21}$$

where

$$R_1(\varphi,\eta,\nabla\tilde{S}(\vartheta_2(\varphi))y) = \frac{\Big(\nabla\tilde{S}(\vartheta_2(\varphi)), R(\varphi,\eta,\nabla\tilde{S}(\vartheta_2(\varphi))y)\nabla\tilde{S}(\vartheta_2(\varphi))\Big)}{q\big(\nabla\tilde{S}(\vartheta_2(\varphi))\big)}$$

is smooth and satisfies $R_1(\varphi,\eta,y) = O(\eta)$ and $\partial_y R_1(\varphi,\eta,y) = O(\frac{\eta}{|y|})$ for all $\varphi \in \mathbb{T}^n_s$. Now we fix the $o(\eta^K$ term f_R to be equal to the remainder (which is a known analytic function of η and φ) of expression 5.17 so that the norm of the left hand side of relation 5.21 is bounded from above by C.

By the contraction mapping theorem, for η small enough, for all $u \in \mathbb{R}$ such that $|u| < 2C$, there exists a unique solution $y = g(\eta,\varphi,u)$ of the equation

$$u = y + R_1(\varphi,\eta,\nabla\tilde{S}(\vartheta_2(\varphi))y)y^2,$$

such that $|y| < 3C$. Moreover, The function $g(\eta,\varphi,u)$ so defined is smooth and analytic in $\varphi \in \mathbb{T}^n_d$, $|\eta \leq \eta_0$ as $\nabla\tilde{S}(\vartheta_2(\varphi))$ is so.
Setting

$$\vartheta'(\varphi) := g(\eta,\varphi,\frac{\tilde{S}(\vartheta_2(\varphi)) - S(\varphi) + o(\eta^K)}{|\nabla\tilde{S}(\vartheta_2(\varphi))|^2})\nabla\tilde{S}(\vartheta_1(\varphi)) \tag{5.22}$$

if $\nabla\tilde{S}(\vartheta_1(\varphi)) \neq 0$ and $\vartheta'(\varphi) = 0$ if $\nabla\tilde{S}(\vartheta_2(\varphi)) = 0$, we get a continuous function $\vartheta'(\varphi)$ which satisfies 5.18 and such that $|\vartheta'(\varphi)| \leq 3C|\nabla\tilde{S}(\vartheta_2(\varphi))| = O(\eta)$ and $\nabla\vartheta'(\varphi) = O(\eta)$.
To complete the proof, we remark that if $f, g : U \to \mathbb{C}$ are analytic in U open subset of $\mathbb{C}^m$ and g is not identically zero and $f = O(g)$ locally in U, then $\frac{f}{g}$ (which a-priori is defined only where $g \neq 0$) has an analytic extension defined in the whole set U. Namely on each locally irreducible hyper-surface on which g is zero, also f is zero with vanishing order al least equal to that of g. So applying standard results of complex analysis (see for instance [R]) we obtain our claim.
Hence

$$\frac{\tilde{S}(\vartheta_2(\varphi)) - S(\varphi) + o(\eta^K)}{|\nabla\tilde{S}(\vartheta_2(\varphi))|^2},$$

(which is bounded by C in $\mathbb{T}^n_d$) is analytic in $\mathbb{T}^n_d$, so is g and finally ϑ. Moreover The transformation

$$\vartheta(\varphi) = \vartheta_2(\varphi) + \vartheta'(\varphi) = \varphi + \eta L(\varphi),$$

is a diffeomorphism in $\mathbb{T}^n_d$, with $d < \frac{s_0}{4}$ such that:

$$\sup_{\substack{|\eta|\le\eta_0 \\ \alpha,\varphi\in\mathbb{T}^n_s}} (2|\nabla_\alpha \vartheta_1| + |\vartheta'(\varphi)|) = O(\eta) \le C\eta < \frac{1}{2}. \tag{5.23}$$

□

Theorem 5.23. *(i) The splitting matrix Δ, which is the Hessian of $S(\varphi)$ at $\varphi = 0$, satisfies the relation:*

$$\Delta = (1 + \eta O)^t \tilde{\Delta}(1 + \eta O) + o(\eta^K) \tag{5.24}$$

where $\tilde{\Delta}$ is the Hessian of $\tilde{S}(\alpha)$ at $\alpha = 0$.
(ii) $\tilde{\Delta}$ has the block structure described in equation5.1.

Proof. Relation 5.24 is a direct consequence of Theorem 5.17. Namely as $\varphi = 0$ implies that also $\vartheta(\varphi) = 0$ we have[9] (by the parity of f):

$$\nabla_\alpha \tilde{S}|_{\alpha=0} = \nabla_\varphi \tilde{S}|_{\varphi=0} = 0$$

and consequently

$$J_\varphi \vartheta|_{\varphi=0} H(\tilde{S})|_{\alpha=0} J_\varphi \vartheta|_{\varphi=0} = H(S)|_{\varphi=0}.$$

Finally using relation 5.23 and the fact that ϑ is η-close to identity we obtain relation 5.24.
(ii) We use the analytic bounds 3.8, for $\tilde{S}$. $\tilde{\Delta} = H(\tilde{S})|_{\alpha=0}$ where $\tilde{S}$ is a trigonometric polynomial in α and the $\Im$ integral of a function in H_0 let us consider the Fourier series of $H(\tilde{S})$:

$$H(\tilde{S})_{ij} = \sum_{\nu\le KN} \nu_i \nu_j e^{i\nu\cdot\varphi} S(\nu),$$

where $S(\nu)$ is the sum of the values through the analytic integrals (a) of trees in $\overset{0}{\mathcal{T}}$ of order $\le K$ and total frequency ν. By The bounds 3.8 all the $S(\nu)$ with non zero fast component ν_F are $O_\varepsilon(\varepsilon^\infty)$ while those with zero fast component are polynomial in $\varepsilon, \varepsilon^{-1}$. □

[9] Given $f : \mathbb{R}^n \in \mathbb{R}^n$ we will call Jf the Jacobian and $H(f)$ the Hessian.

Chapter 6

Lower bounds on the splitting for systems with one fast frequency

We find lower bounds on the splitting determinant and on the eigenvalues of the splitting matrix, for systems with one fast frequency, such that $f(\psi, q)$ is a rational function of e^{iq} and satisfies suitable non degeneracy conditions. This can be done independently by using the results of Chapter 4 or of Chapter 5.

6.1 Basic lower estimates

In Section 4 we have proved that the splitting matrix Δ at the intersection point $\phi, \psi = 0, q = \pi$ can be written for any $K \leq \varepsilon^{-\frac{1}{2\tau_F}}$ as:

$\Delta = \Delta_1^{\leq K} + \Delta_2^{\leq K} + (\eta)^K \Delta^R = \sum_{k \leq K} (\eta)^k (\Delta_{1\,k} + \Delta_{2\,k}) + (\eta)^K \Delta^R$. We are interested in systems with one fast frequency so $\tau_F = 0$; we choose $K = \frac{C}{\sqrt{\varepsilon}}$ with $C \gg 1$.

All the entries of $\Delta_2^{\leq K}$ are exponentially small by definition, they will contain a factor that is the integral of some function analytic in a domain $H(b, d)$ with total fast mode $\nu_1 \neq 0$. Moreover $\det(\Delta_1^{\leq K}) = 0$ by Lemma 3.16. The remainder $(\eta)^K \Delta^R$ is bounded by:

$$|(\eta)^K \Delta_{ij}^R| \leq (\frac{\eta}{\eta_0})^K < C_2(\eta\eta_0^{-1})^K).$$

(with $\eta_0 = \varepsilon^{3/2}$ as seen in Appendix A.4).

Similarly in Chapter 5 We have proved that the splitting determinant is equal to the determinant of $\tilde{\Delta}$ times the determinant of $(1 + \eta O)^2$. And that $\tilde{\Delta}$ has the block structure 5.1.

We know from KAM theory that the series expansion for Δ is absolutely convergent for $|\eta| < \eta_0$. This means that the series expansion of the determinant:

$$\det \Delta = \eta^n \sum_{k<K} Q_k \eta^k + R_K$$

is absolutely convergent as well, and that each summand of Q_k contains at least a factor $(\Delta_{2\,k})_{ij}$ for some i, j and $h < k$.

We know that $X_1^{\leq K} = O_\varepsilon(1)$ is $O((\eta/\eta_0)^K)$ close to the eigenvector of $\Delta_1^{\leq K}$ with $o((\eta/\eta_0)^K)$ eigenvalue. Now we set $\Delta_1^{\leq k}$ in block form via an orthonormal change of variables:

$$\Delta_1^{\leq k} = \left| \begin{array}{c|c} \lambda_R & 0 \\ \hline 0 & \Delta_1' \end{array} \right| \qquad \text{with } \lambda_R = o((\eta/\eta_0)^K).$$

We are considering a system with one fast frequency so Δ_1' is an $n-1 \times n-1$ matrix.

Proposition 6.1. *(i) If* $\det \Delta_1' \neq 0$ *and* $\Delta_{2\,11}^{\leq K} \neq 0$*, the splitting determinant is given by the determinant of* Δ_1' *times the size of* $\Delta_2^{\leq K}$*. Precisely the bounds:*

$$\begin{aligned} &a\varepsilon^p \leq \det \Delta_1' \leq b\varepsilon^{-p}, \quad a\varepsilon^p e^{-\frac{c}{\sqrt{\varepsilon}}} \leq |\Delta_{2\,11}^{\leq K}| \leq b\varepsilon^{-p} e^{-\frac{c}{\sqrt{\varepsilon}}} \\ &|\Delta_{2\,ij}^{\leq K}| \leq b\varepsilon^{-p} e^{-\frac{c}{\sqrt{\varepsilon}}} \end{aligned} \tag{6.1}$$

imply that

$$a^2\varepsilon^{2p} e^{-\frac{c}{\sqrt{\varepsilon}}} \leq \det \Delta \leq b^2 \varepsilon^{-2p} e^{-\frac{c}{\sqrt{\varepsilon}}}.$$

(ii) If the eigenvalues of Δ_1' *are bounded by:*

$$a'\varepsilon^{p'} \leq |\lambda_i| \leq b'\varepsilon^{-p'} \quad \textit{for } \; i = 1, \ldots, n-1$$

then so are $n-1$ *eigenvalues of* Δ*. The remaining eigenvalue is bounded by:*

$$a''\varepsilon^{p'} e^{-\frac{c}{\sqrt{\varepsilon}}} \leq |\lambda_i| \leq b''\varepsilon^{-p'} e^{-\frac{c}{\sqrt{\varepsilon}}}.$$

Proof. $\Delta_2^{\leq K} + \Delta^R$ respects the same bounds (with possibly different constants a, b) as $\Delta_2^{\leq K}$ as the remainder

$$(\eta\eta_0^{-1})^K < \frac{1}{2}\min(a,b)\varepsilon^p e^{-\frac{c}{\sqrt{\varepsilon}}}$$

for small enough values of ε.
Moreover the bounds 6.1 imply that $\det \Delta$ is:

$$\det \Delta = \det \Delta_1' \Delta_{2\,11}^{\leq K} + Q$$

where Q contains at least two entries of $\Delta_2^{\leq K}$.

(ii) This is simply the fact that the determinant is the product of the eigenvalues. ☐

This decouples the problem in a polynomial and an exponentially small part. We will proceed in two steps:

1. Compute the first order of Δ_2, with the purpose of finding general lower bounds. Then use the upper bounds on analytic and non analytic integrals of order $k \geq 2$ of Chapter 3 to extend the lower bounds on all $\Delta_2^{\leq k}$. This gives us the size of the exponentially small eigenvalue.
2. Compute the non zero eigenvalues of $\Delta_1^{\leq k}$, via classical perturbation theory.

6.1.1 Lower bounds on the Melnikov integral

In this subsection we will use for the first time the restriction that $f(\psi, q)$ is a rational function of $\cos(q), \sin(q)$.

Let $f(q, \psi) := \mathbb{T}^{n+1} \to R$ have the usual Fourier expansion in the rotator angles:

$$f(q, \psi) = \sum_{|\nu| \leq N} f_\nu(q) e^{i\nu \cdot \psi}$$

where all the functions $f_\nu(q)$ are rational functions of $x = e^{iq}$ with no poles on the unit circle $(f_\nu(q) = H_\nu(x))$.
The parity of f leads to $f_\nu(-q) = f_{-\nu}(q)$ while the reality of f implies that $f_\nu\bar{(}q) = f_{-\nu}(q)$. Moreover f has zero mean value.
We are considering lower bounds on the first order of the expansion of the splitting matrix (these are all analytic integrals) so at first we will make no difference between slow and fast variables.

$$M_{ij} = \Im f_{\psi_i \psi_j}(q(t), \frac{\omega}{\sqrt{\varepsilon}} t),$$

M is the Melnikov term for the splitting matrix.
We substitute $x = (\frac{e^t - i}{e^t + i})^2$ (notice for each value of x there are two solutions e^t and $-e^{-t}$) in the $H_\nu(x)$ and we obtain for each value of ν a rational function of e^t (we call it $F_\nu(e^t)$). The parity condition is $F_\nu(y) = F_{-\nu}(-y) = F_{-\nu}(\frac{1}{y})$ the reality is[1] $\bar{F}_\nu(y) = F_{-\nu}(y)$ and so we have $\bar{F}_\nu(y) = F_\nu(-y) = F_\nu(\frac{1}{y})$. Notice that F has all the poles of f as function of x plus possibly poles at $e^t = \pm i$. The absence of poles on the unit circle $|x| = 1$ implies that there are no poles for real values of t; this and the fact that $x \to 1$ (exponentially) for $t \to \pm\infty$ imply that all the $F_\nu(e^t)$ are the sum of a constant function and a function $G_\nu(e^t)$ that is exponentially decreasing to zero for $t \to \pm\infty$. The Melnikov integral depends only on G:

$$\Im(\nu_i \nu_j) e^{i\nu \cdot \frac{\omega}{\sqrt{\varepsilon}} t} F_\nu(e^t) = (\nu_i \nu_j) \int_{-\infty}^{\infty} G_\nu(e^t) = (\nu_i \nu_j) I_\nu$$

as purely oscillating functions give no contribution to $\Im$ and G_ν is clearly $\mathcal{L}_1$.

Now if the functions $f_\nu(x)$ are not all polynomials (and so the function f is a trigonometric polynomial) then some of them must have poles for finite values of x. Let us call the poles x_ν^j $j = 1, \ldots, n_\nu$ and the the corresponding values of t (in $|Im(t)| \leq 2\pi$ and via the relation $x = (\frac{e^t - i}{e^t + i})^2$) $t_\nu^i \in \mathbb{C}$ $i = 1, 2N_\nu$ each with degree p_ν^i. The poles of $F(e^t)$ are the t_ν^i plus possibly $\pm i\frac{\pi}{2} + 2ik\pi$.

Lemma 6.2. *The poles of G_ν in $|Im(t)| \leq \pi$ come in groups of four, namely if the complex number $t_\nu^i = a_\nu^i + ib_\nu^i$ is a pole for G_ν then so are $-\bar{t}_\nu^i = -a_\nu^i + ib_\nu^i$, $-t_\nu^i + i\pi = -a_\nu^i - ib_\nu^i + i\pi$ and $-\bar{t}_\nu^i + i\pi$; correspondingly $G_{-\nu}(e^t)$ has poles $-t_\nu^i$, $\bar{t}_\nu^i$ etc...*

[1] By $\bar{F}(y)$ we mean the function having as coefficients the complex conjugates of the coefficients of F, so $\bar{F}(y) = \overline{F(\bar{y})}$.

Proof. Each x^j_ν has two preimages t^j_ν and $-t^j_\nu + i\pi$; moreover by the reality condition $F_\nu(y) = \bar{F}(-y)$, if $y = e^{t^j_\nu}$ is a pole so is $-\bar{y}$. □

We notice that $|z|F(e^z) \to 0$ uniformly for $z = t + id$ with $d \in \mathbb{R}$ fixed and $\mathbb{R} \ni t \to \infty$; moreover $F(e^t) = F(e^{t+2i\pi})$ so we consider the integral 6.1.1 plus the same shifted by $2i\pi$ and apply Jordan's lemma. Precisely if $\omega_\nu = \frac{\omega}{\sqrt{\varepsilon}} \cdot \nu > 0$ we shift by $2i\pi$ and if it is negative by $-2i\pi$, we will call σ_ν the sign of ω_ν. If $\sigma_\nu = +$ we will consider the poles in the principal domain $0 \le \text{Im } t \le 2\pi$ instead of $-\pi \le \text{Im } t \le \pi$ and vice versa for $\sigma_\nu = -$.

$$I_\nu(1 - e^{-2|\omega_\nu|\pi}) = 2\pi i\sigma_\nu \sum_{i=1}^{2N_\nu} \text{ Res }^{\sigma_\nu}(e^{i\omega_\nu t}G_\nu(e^t), t^\nu_i) \tag{6.2}$$

The apex $\sigma_\nu = \pm$ on the Residue indicates that the poles are set in the upper or lower half-plane. Let

$$\sum_{k=-p^i_\nu,\infty} g^{i,k}_\nu(t - t^i_\nu)^k$$

be the Laurent expansion of G_ν near the poles $t^i_\nu = a^i_\nu + ib^i_\nu$:

$$\text{Res }(e^{i\omega_\nu t}G_\nu(e^t), t^\nu_i) = e^{-\omega_\nu b^i_\nu} e^{i\omega_\nu a^i_\nu} \sum_{k=1,p^i_\nu} \frac{(i\omega_\nu)^{k-1}}{k-1!} g^{i,-k}_\nu \tag{6.3}$$

For each value of ν such that $\omega_\nu > 0$ we consider a pole t^i_ν such that $\text{Re}t \ge 0$ [2] and the corresponding pole $-t^i_\nu$ of $G_{-\nu}$. The contribution to the integral is

$$2\pi i(\text{ Res }^+(e^{i\omega_\nu t}G_\nu(e^t), t^\nu_i) - \text{ Res }^-(e^{-i\omega_\nu t}G_-\nu(e^t), -t^\nu_i) \tag{6.4}$$

and as $G_\nu(e^t) = G_{-\nu}(e^{-t})$ the Laurent expansion of $G_{-\nu}$ near the pole is

$$\sum_{k=-p^i_\nu,\infty} (-1)^k g^{i,k}_\nu(t + t^i_\nu)^k.$$

The sum 6.4 is:

$$e^{-|\omega_\nu||b^i_\nu|} e^{i|\omega_\nu||a^i_\nu|} 2 \sum_{k=1,p^i_\nu} \frac{(i|\omega_\nu|)^{k-1}}{k-1!} g^{i,-k}_\nu.$$

Now we consider the poles $-\bar{t}^i_\nu$ and $\bar{t}^i_\nu$ of G_ν and $G_-\nu$; the relation $\overline{G(y)} = G(\frac{1}{\bar{y}})$ implies that the Laurent series of G_ν in the point $-\bar{t}^i_\nu$ is:

$$\sum_{k=-p^i_\nu,\infty} (-1)^k \bar{g}^{i,k}_\nu(t + \bar{t}^i_\nu)^k$$

[2]by the symmetry relations there must be such a pole.

($g_\nu^{i,k}$ are the coefficients of the expansion near t_ν^i). So finally for each ν such that $\omega_\nu > 0$ and each couple of poles t_ν^i and $-\bar{t}_\nu^i$ of F_ν (and $-t_\nu^i$, $\bar{t}_\nu^i$ of $-\nu$)in the upper half-plane (resp. lower half-plane) we obtain the real value:

$$(I_\nu + I_{-\nu})(1 - e^{-2|\omega_\nu|\pi}) = \\ 4\pi i e^{-|\omega_\nu||b_\nu^i|} \sum_{k=1,p_\nu^i} \frac{(i\omega_\nu)^{k-1}}{k-1!}(e^{i|\omega_\nu||a_\nu^i|} g_\nu^{i\,k} + (-1)^k e^{-i|\omega_\nu||a_\nu^i|} \bar{g}_\nu^{i,k}) \quad (6.5)$$

One should notice that this formula holds also for functions with some non polar singularity for finite $t \in \mathbb{C}$; in the latter case we always obtain exponentially decreasing functions of ω_ν (as predicted by the Paley Wiener theorem) but we cannot give general formulas for the decreasing rate as the residues are no longer finite sums $k = 1, p_\nu^i$. Consider a ν such that $\nu_F \neq 0$ then if all the p_ν^j are finite the frequency ν contributes a term of order either zero or $e^{-c/\sqrt{\varepsilon}}$.

6.1.2 Systems with one fast frequency

Let us go back to systems with one fast frequency:

$$\omega_1 = O_\varepsilon(1) \;\; |\omega_2| = O_\varepsilon(\varepsilon^{(\frac{1}{2}+\alpha)}),$$

with $0 \leq \alpha \leq \frac{1}{2}$. On such systems we can give "general" lower bounds on the determinant of the splitting matrix provided that we impose some non-degeneracy conditions on the frequencies of f so that the hypothesis of Proposition 6.1 are verified.

Proposition 6.3. *The sum of the exponentially small terms of order $2 \leq k \leq K$ are bounded from above by:*

$$C_4 e^{-\omega_1 d_1 \frac{D}{\sqrt{\varepsilon}}} \sum_{1<k\leq K} (\eta)^k (\frac{C\sqrt{\varepsilon}^3}{\sqrt{\varepsilon}^P})^k) \quad (6.6)$$

d_1 is the divisor of the frequencies of f in the fast component ($j = 1$)(it is different from one only for functions $f(q, \psi)$ whose fast frequencies are not coprime see Appendix A.3) and $P = \max(p + 5, 4\tau_S + 4)$

Proof. We apply Corollary 3.8 discussed in Section 3.1. Namely, a tree with fruits carrying an analytic integral of total frequency ν is bounded from above by:

$$J_k(\nu) = [\sum_{A\in\mathcal{A}} c(A)] e^{\text{ Im } \varphi|\nu|} C_1^k (k!)^{c_1} N^{2k} E(D, \nu) \frac{\sqrt{\varepsilon}^3}{\sqrt{\varepsilon}^{(p+5)k}},$$

restricted to frequencies ν with non-zero fast component, $\nu_F \in \mathbb{Z}$. We are considering systems with one fast frequency so, if $K < \varepsilon^{-\frac{1}{2}+\alpha}$, $\tau_F = 0$ and $c_1 = 4 + 4\tau_S$.
We choose $K = c/\sqrt{\varepsilon}$, bound $E(D, \nu)$ with $2e^{-|\omega_1 \nu_F|\frac{D}{\sqrt{\varepsilon}}} e^{\varepsilon^\alpha |\omega_2||m|}$, the sum on fruitless trees of order k by $(2n)^k$ and finally $k!$ with $\frac{C_2}{\sqrt{\varepsilon}}^k$. Now we sum on the frequencies

$|\nu| \leq kN$ with non zero fast component accessible at order k. First we fix the value of ν_F and sum on the slow modes ($e^{\varepsilon^\alpha |\omega_2} = O_\varepsilon(1)$ even for $\alpha = 0$), we obtain a factor bounded by C^k for some order one C. Then if ν_F^k is the minimum non zero fast mode accessible at order k

$$\Delta_{2\,k} \leq \sum_{\nu_F \geq \nu_F^k} J_k(\nu_F) = \sqrt{\varepsilon}^3 (\frac{C}{\sqrt{\varepsilon}^{p+7+4\tau_S}})^k \sum_{\nu_F \geq \nu_F^k} e^{-|\omega_1 \nu_F| \frac{D}{\sqrt{\varepsilon}}} \leq$$
$$C_4 \sqrt{\varepsilon}^3 (\frac{C}{\sqrt{\varepsilon}^{p+7+2\tau_S}})^k e^{-|\omega_1 \nu_F^k| \frac{D}{\sqrt{\varepsilon}}}. \tag{6.7}$$

as the contributions to $\Delta_{2\,k}$ are by definition all of the form $J(\nu)$. Finally by the definition of the divisor in the fast direction $d_1 \leq \nu_F^k$ for all k. This leads to the proposed bound with $P = p + 5 + 4\tau_S + 4$ the better bound proposed rises from the observing that in each node we can have either a small denominator coming from the improper integrals (so $\varepsilon^{2\tau_S+2}$ or a term from the proper integral $\varepsilon^{(p+5)/2}$. □

If we fix $|\eta| \leq |\sqrt{\varepsilon}|^P$, we can add up the terms $2 \leq k \leq K$:

$$\Delta_2^{\leq K} \geq \eta M_2 - C\varepsilon^{3/2} (\frac{\eta}{|\sqrt{\varepsilon}|^P})^2 [e^{-\frac{|\omega_1 d_1| D}{\sqrt{\varepsilon}}}],$$

where M_2 is the fast (exponentially small) part of the matrix M.
Finally we consider the Melnikov term M_2; to have a simpler expression we consider at first only functions $f(q,\psi)$ such that the fast and slow variables ψ_F are partially decoupled[3] $f(q,\psi) = g_1(q,\psi_F) + g_2(q,\psi_S) + G(\psi, q)$.

Lemma 6.4. *The size of the first order of Δ_2 is (generally) greater than:*

$$M_2 \geq C_3 \varepsilon^{-c_3} e^{-|\omega_1| \frac{h_M}{\sqrt{\varepsilon}}}$$

where $h_M = \min_{\nu, t_\nu^i} \nu_F b_\nu^i$ evaluated on the frequencies $\nu_F \neq 0$ (ν_F is the fast component) and $c_3 = \frac{p_M}{2}$ where p_M is the order of the pole b_ν^i which realizes the minimum.

Proof. We use the results of Subsection6.1.1. The trees of order one are all analytic so $M_{2\,ij}$ is zero if i,j are both slow. In particular in equation 6.5 there are only contributions from frequencies ν such that $\nu_F \neq 0$. We write $\omega_\nu = \frac{\omega_1}{\sqrt{\varepsilon}} \nu_F + B$ where $B \leq CO_\varepsilon(1)$. We substitute all the oscillating terms and the $e^{-Bb_\nu^i}$in 6.5 with order one constants :

$$|M_{2\,ij}| = 4\pi \sum_{\substack{\nu\,:\,\nu_F \neq 0 \\ |\nu| \leq N}} |\nu_i \nu_j| \sum_{l=1,\ldots,n_\nu} \frac{e^{-\frac{|\omega_1 \nu_F||b_\nu^i|}{\sqrt{\varepsilon}}}}{1 - e^{-2\frac{|\omega_1 \nu_F|}{\sqrt{\varepsilon}}\pi}} C_\nu^i(\varepsilon)^{-p_\nu^i}.$$

[3]This is called a non-degeneracy condition in [BB1]; in this way $r_1 = 1$ and the first order matrix Δ_1^M has a $n-1 \times n-1$ minor whose entries are of order one in ε.

Moreover setting

$$b_M = \min_{\nu, t^i_\nu} \nu_F b^i_\nu,$$

all the summands are smaller or equal to $e^{\frac{-|\omega_1| n_M b_M}{\sqrt{\varepsilon}}} (\varepsilon)^{-p_M/2}$. □

Proposition 6.3 and Lemma 6.4 imply immediately the following.

Theorem 6.5. *The Melnikov integral M_2 dominates in the expansion of $\Delta_2^{\leq K}$, for $\eta < \min(\varepsilon^{P-c_3-3/2}, \varepsilon^{(P)/2}$, provided that the perturbing function verifies the condition $d_1 D \geq h_M$. In this case the entries $\Delta_{2i,j}$ such that $M_{2\,ij} \neq 0$ are bounded from below by $\frac{1}{2} M_{2\,ij}$.*

For example if the fast component of the frequencies of $f(q, \psi)$ contains the divisor d_1 (see Appendix A.3) and all the $G_\nu(e^t)$ have the same poles then the condition $d_1 D \geq h_M$ is automatically satisfied.

Even if these conditions are not verified one can give rules to determine the (ε independent) "possible" dominating order, by simple considerations on the mode vectors $\nu \in \mathbb{Z}^n$. In general, our candidate will be the first analytic integral (fruitless tree) whose total fast mode is d_1 and containing a node v such that F_{ν_v} has a pole with imaginary part equal to D. The value of this integral is still the Fourier transform of an exponentially decreasing function with known singularities (the same as those of $G(e^t)$) but the singularities are not (generally) polar any more and we cannot use the same estimates as for the Melnikov term.

Remark 6.6. *"Hopefully" the size of the exponentially small eigenvalue is $O[e^{-\frac{|\omega_1 d_1| D}{\sqrt{\varepsilon}}}]$ for $|\eta| \leq |\varepsilon|^P$.*

Proof. A term of order $[e^{-\frac{|\omega_1 d_1| D}{\sqrt{\varepsilon}}}]$ appears for the first time in a fruitless tree of order $k = m_1$ (m_1 is the minimal length in the fast direction, Appendix A.3) containing a node v such that F_{ν_v} has a pole with imaginary part equal to D.

The problem is that, as we have said in subsection 6.1.1, if $m_1 \neq 1$ then it is not necessarily true that the value of the tree is greater than $CP(\frac{1}{\varepsilon})[e^{-\frac{|\omega_1 d_1| D}{\sqrt{\varepsilon}}}]$ as the singularities are generally not polar. If the last inequality holds we can add up the trees of higher order using the upper bounds and the assertion is true. If the value is zero or not of the correct order then we consider the contributions to fruits of order $k = m_1$ coming from the same fruitless tree, if this is again zero (or not of the correct order) we pass to a higher order fruitless tree with the same fast mode [4] and so on. □

This ends the analysis of step 1. Now we compute the polynomial eigenvalues:

Lemma 6.7. *The matrix Δ_1 is of order η^{r_1}; the leading order has contributions only from analytic integrals (with zero total fast mode) and so has the first line and column (corresponding to the fast variable) equal to zero. So the non zero eigenvalues of $\Delta_1^{\leq K}$ are of the size of the eigenvalues of $\Delta_{\bar{1}, r_1}^{\leq K}$ if this matrix has rank $n-1$.*

[4]in the Appendix A.3 we have proven that each divisor is accessible for infinite k

Proof. By LemmaA.12 r_1 is the first order such that in the generating function there is an analytic integral with total zero fast component. The value of such integral is generally NOT exponentially small in ε as seen in Corollary[5] 3.13. The leading order of $\Delta_{\mathrm{I}}^{\leq K}$ is the Hessian of $S^{r_1}(\varphi)$ at zero. So it is clear that such integral gives contributions only to the slow components of the matrix. Finally classical perturbation theory ensures that the eigenvalues of $\Delta_{\mathrm{I}}^{\leq K}$ are η^{r_1+1} close to those of $\Delta_{\mathrm{I},r_1}^{\leq K}$. This provides upper and lower bounds on the non-zero eigenvalues. □

This finally leads to the following theorem on the splitting determinant for perturbating functions f rational in e^{iq} that contain their divisor in the fast direction and such that the F_ν have all the same poles :

Proposition 6.8. *For ε sufficiently small and for all $\eta < \min(\varepsilon^{P-c_3-3/2}, \varepsilon^{(P)/2}$ the splitting determinant is bounded (from above and below) by expression of the type $C(\varepsilon)e^{-\frac{|\omega_1 d_1|D}{\sqrt{\varepsilon}}}$ where $C(\varepsilon)$ is a rational function of ε.*

6.2 Examples of Melnikov dominance

In this section we will use a simplified version of [GGM4] and [Ge] to find improved lower bounds on the splitting determinant. We work on examples with three time scales and three degrees of freedom, it should be clear however that the technique is general (for systems with one fast frequency) so we point out the necessary generalizations. We first review the techniques of [Ge] which enable us to prefix the Lyapunov exponent, thus simplifying the expression of the $\Phi_j^k(\varphi,t)$. The article [GGM4] then proves the convergence of the Lindstedt series by showing the existence of compensations between seemingly divergent terms due to the small denominators. We will not go into the details of this (very interesting) technique as we only want to find better upper bounds for the terms of the series expansion of the splitting determinant of order $k \leq \varepsilon^{-\frac{1}{2}+\alpha}$.

In Subsection 6.2.1 we describe (an adapted version of) the techniques of [GGM4]; then in Subsection 6.2.2 we find appropriate bounds, similar to those of Chapter 3. Notice that the proofs would be simpler if we could assume that the splitting determinant is exponentially small wherever it is convergent by KAM theory.

6.2.1 Systems with prefixed Lyapunov exponent

As in [Ge] and in [GGM4] we consider the following Hamiltonian:

$$\frac{1}{2}(I^2 + \varepsilon J^2 + p^2) + (g + \eta G(\eta, g))(\cos(q) - 1) + \eta F(\psi, \phi, q), \tag{6.8}$$

I, ψ, J, ϕ and p, q are conjugate action angle variables. The characteristic frequency will be a diofantine vector: $\frac{\omega_1}{\sqrt{\varepsilon}}, \varepsilon^{\frac{1}{2}}\omega_2$. $G(\eta, g)$ is an, a priori unknown, analytic function

[5] Clearly it is possible that for some perturbating function the integral is zero or arbitrarily small, but this implies giving a relation between f and ε

of its arguments for $|\eta| \leq \eta_0$ and $|g| \leq g_0$. We will prove that G can be determined uniquely by imposing that the Lyapunov exponent of the separatrix is g (at least to order $k = \varepsilon^{-1}$). Finally the parameter g will be fixed (as a function of η) so that $g(\eta) + \eta G(\eta, g(\eta)) = 1$. Under such conditions the system 6.8 is of the type described in Section 1.1.

We can apply the theory developed in Section 1.1 to the Hamiltonian 6.8 so we will not repeat the procedure but simply write the equations for the time evolution on the separatrix.

First we will expand G in Taylor series in η:

$$G(\eta, g) = \sum_{\eta^0}^{\varepsilon^{-1}} (\eta)^k g_{k+1},$$

the equations for the separatrix are:

$$I_k(t) = \Im F_1^k\,, \quad \psi_k = O_1 F_1^k\,, \quad J_k(t) = \Im F_2^k\,, \quad \phi_k = \varepsilon O_1 F_2^k \equiv O_2 F_k^2\,,$$

$$q_k = O_0 (F_0^k + \textstyle\sum_{h=1}^k g_h [\cos(\sum_{j \leq k-h} (\eta)^j q_j)]_{k-h})\,,$$

for $k \geq 1$ while $q_0(t) = \arctan(e^{gt})$ and $\psi_0(t), \phi_0(t) = \varphi + \omega t$ (and we will call $\varphi = \psi(t=0), \phi(t=0)$ the η independent initial data of the rotators). As in Subsection 1.1.2 :

$$F_j^k = [\partial_j f(\psi(\eta), \phi(\eta), q(\eta))]_{k-1} + g\delta_{j0}[\sin(\sum_{h<k} (\eta)^h q_h)]_k,$$

where

$$\partial_1 = \partial_\psi \ \partial_2 = \partial_\phi, \ \partial_0 = \partial_q.$$

In the following we will write the operators O_j with $j = 0, 1$ as:

$$O_j(G) = \Im^t w_j(t, \tau) G(\tau) + x_j^0 \Im^0 \sigma(\tau) x_j^1 G, \tag{6.9}$$

it is easily seen that this is equivalent to 1.31.

We can represent the series expansion in terms of trees as in Chapter 2. The nodes $v \neq v_0$ will carry the labels: $j_v = 0, 1, 2$, $\sigma_v \in \mathbb{N}_0$, $\delta_v = 0, 1$, $\rho_v = 0, 1$, with the usual grammar:

$$\delta_v = 1 \to \sigma_v = 0\,, \quad \delta_v = \sigma_v = 0 \to s(v) \geq 2, \quad \delta_v = 0 \to j(v) = 0\,, \ j(v') = 0 \ \forall v' \in s(v)$$

The $\sigma_v = h$ represents the application of a "counter-term" g_h. As the g_h have degree h in η we redefine the order of a tree as:

$$O(A) = \sum_{v \in A} \delta_v + \sigma_v.$$

Finally the label $\rho_v = 0, 1$ represents the application respectively of the first and second summand in 6.9.

We expand the function $f(\varphi+\omega t, q_0(t)$ in its harmonics and apply an extra label $\nu_v \in \mathbb{Z}^2$ such that $|\nu_v| \le N$.
Following [GGM4] we will call leafs the subtrees stemming form a node v with $\rho_v = 1$, this terms (just like the fruits in Subsection 2.1.2) contribute by a fixed function x_j^0 , times a coefficient depending on the subtree $A^{\ge v}$. So one can factorize the value of a tree as a product of values of marked leafless trees which in [GGM4] are called amputated trees. Graphically we represent the leafs by drawing a line on their stalk and do not consider the nodes inside the leaf as nodes of the amputated tree; we will call $L(v)$ the list of nodes v' with $s(v'=0)$ attached to the node v .

Notice that with this notation all the nodes $v > v_0$ of an amputated tree have $\rho_v = 1$ so we can remove this label from all the nodes except the first.

Using the notations of Section 1.2 we will consider the set $\mathcal{T}$ of marked trees with leafs, $\mathcal{A}$ of trees without leafs, and the subsets of Definition 1.31. In particular we will be interested in :

$$\Lambda_j^k = \sum_{A\in\mathcal{T}_J^k} \frac{1}{|\mathcal{S}(A)|} A \text{ and } \mho_j = \sum_{A\in\mathcal{A}_j^k} \frac{1}{|\mathcal{S}(A)|} A.$$

notice that now $\mho$ represents the leafless contributions to the series expansion.

The value $\mathcal{W}(A)$ of a tree with a marking $h(t)\partial_l$ in v_0 and leafs $L(v_0)$ is defined recursively:

$$\begin{aligned} \mathcal{W}(h(v_0,t)\partial_l^{v_0}\mho^1, t) &= -h(v_A,t)\nabla^{e_l} f^1 \\ \mathcal{W}((h(v_A,t)\partial_l^{v_A} A, t) &= -g_{\sigma_{v_0}}^{\delta_{v_0}} a_{j_{v_0}} h(\tau_{v_0})(\nabla^{\vec{m}_{v_0}} f^{\delta_{v_0}}) \prod_{v\in s(v_0)} \mathcal{V}(A^{\ge v}), \end{aligned} \tag{6.10}$$

where $m_v(j)$ is the number of nodes v' in the list $s(v), \mathcal{M}(v)$ having label $j_{v'} = j$; $a_0 = a_1 = 1$, $a_2 = \varepsilon$, $g_0^1 = 1$ and $g_{\sigma_v}^0 = g_{\sigma_v}$. Finally for $A \in \mathbb{T}_*$:

$$\mathcal{V}(A) = \Im^{t(\rho_{v_0})} w_{j_{v_0}}(t(\rho_{v_0}), \tau_{v_0}) \mathcal{W}(A), \quad \text{with} \quad t(x) = \begin{cases} 0 & \text{if } x = 0 \\ t & \text{if } x = 1 \end{cases}$$

Remember that

$$w_j(t,\tau) = (t-\tau) \quad \text{if} \quad j = 1,2$$
$$w_0(t,\tau) = \frac{t-\tau}{\cosh(gt)\cosh(g\tau)} + \frac{\sinh(gt)}{\cosh(g\tau)} - \frac{\sinh(g\tau)}{\cosh(gt)}.$$

This definition can be extended to $\mathbb{V}(\mathcal{T})$ by linearity, as seen in Subsection 2.1.2, this implies that:

$$\psi_k(t) = \mathcal{V}(\Lambda_1^k) \quad \phi_k(t) = \mathcal{V}(\Lambda_2^k),$$

similar identities can be found for the actions, however for the moment we will concentrate on the angles. Notice that if $\rho_v = 0$ the value

$$\mathcal{V}(A^{\ge v}) = x_{j_v}^0(t) \Im x_{j_v}^1(\tau) \mathcal{W}(A^{\ge v}, \tau).$$

As for trees with fruits, a tree with a leaf in v is equivalent to the tree without the leaf (amputated), marked $x^0_{j_v}$ and multiplied by the value of the leaf (which is time-independent); we can write:

$$\mathcal{W}(A,t) = g^{\delta_{v_0}}_{\sigma_{v_0}} a_{j_{v_0}} (\nabla^{\vec{m}_{v_0}} f^{\delta_{v_0}}) \prod_{v\in s_0(v_0)} \mathcal{V}(A^{\geq v}) \prod_{v'\in L(v_0)} x^0_{j_{v'}}(\tau_{v_0}) \Im x^1_{j_{v'}} \mathcal{W}(A^{\geq v'}), \qquad (6.11)$$

Remark 6.9. *Notice that in an amputated tree the integrals are all ordered: namely the τ_v have all the same sign and*

$$|\tau_v| \leq |\tau_w|.$$

Definition 6.10. *Given an amputated (marked) tree A and a node v in A we will consider the total rotator harmonic of the subtree $A^{\geq v}$:*

$$\nu_T(v) = \sum_{w\geq v} \nu_w,$$

remember that the nodes inside a leaf are not nodes of the amputated tree.

Notice that in a leafless tree the total rotation $\nu_T(v_0)$ gives the dependence on the initial data φ. If the tree has leafs each with total rotation $\nu^L_T(i)$ the dependence on the initial data is

$$e^{i(\nu_T(v_0)+\sum_i \nu^L_T(i))\cdot\varphi}.$$

Lemma 6.11. *(i) Given a function $F(t)$ such that*

$$F(t) = \sum_{|\nu|\leq M} \sum_{k=0}^{\infty} f_{\nu,k} e^{i\omega\cdot\nu t} e^{-kg|t|}, \quad \text{with} \quad f_{\vec{0},0} = 0$$

then the integral

$$\Im^t(t-\tau)F(\tau)d\tau = \Im^t\Im^\tau F(\tau') = \sum_{|\nu|\leq M} \sum_{k=0}^{\infty} \frac{f_{\nu,k}}{(i\omega\cdot\nu+\sigma(t)kg)^2} e^{i\omega\cdot\nu t} e^{-kg|t|} \qquad (6.12)$$

(ii) Given a function $G(t)$ such that

$$G(t) = \sum_{|\nu|\leq M} \sum_{k=0}^{\infty} g_{\nu,k} e^{i\omega\cdot\nu t} e^{-kg|t|}, \quad \text{with} \quad g_{\vec{0},1} = 0$$

then

$$\Im^t\left(\frac{\sinh(gt)G(\tau)}{\cosh(g\tau)} - \frac{G(\tau)\sinh(g\tau)}{\cosh(gt)}\right) =$$

$$\Im^t(\cosh(g\tau)\Im^\tau \frac{G(\tau')}{\cosh(g\tau')} - \Im^t \frac{\sinh(g\tau)}{\cosh^2(g\tau)} \Im^\tau \sinh(g\tau')G(\tau') = H(t)$$

moreover the function $H(t)$ has the same properties as F:

$$H(t) = \sum_{|\nu|\leq M} \sum_{k=0}^{\infty} h_{\nu,k} e^{i\omega\cdot\nu t} e^{-kg|t|}, \quad \text{with} \quad h_{\vec{0},0} = 0.$$

Proof. (i) We are simply using the identity:

$$\Pi_P H = 0 \to H = \Im^t \dot{H}.$$

(ii) Same as (i) we have to prove that

$$\Pi_P \Im^t (\frac{\sinh(gt)G(\tau)}{\cosh(g\tau)} - \frac{G(\tau)\sinh(g\tau)}{\cosh(gt)}) = \Pi_P \sinh(gt)G = 0.$$

The last identity is obvious as $\Pi_P \sinh(gt)G = \frac{1}{2}g_{\vec{0},1}$. in the left hand side, we notice that the only constant terms can come from the constant terms of G so:

$$\Pi_P \Im^t (\frac{\sinh(gt)G(\tau)}{\cosh(g\tau)} - \frac{G(\tau)\sinh(g\tau)}{\cosh(gt)}) = g_{\vec{0},0} \Im^t (\frac{e^{g|t|}}{e^{-g|\tau|}} - \frac{e^{g|\tau|}}{e^{-g|t|}}) = 0.$$

□

Proposition 6.12. *The value of a tree with $\rho_{v_0} = 1$ is a limited function of t; moreover we can fix recursively the coefficients g_h $h \leq k$ (independently from the initial data φ) so that the value of a tree of order k with $s_{v_0} = 1$ can be expanded as:*

$$\mathcal{V}(A)[t] = \sum_{|\nu| \leq kN} \sum_{h=0}^{\infty} a_{\nu,h}(A) e^{i\omega\cdot\nu t} e^{-hg|t|}, \quad \textit{with} \sum_{A \in \mathcal{A}_j^k} c(A) a_{\vec{0},0}(A) = 0.$$

So as we are interested only in $\mathcal{V}(\Lambda_j)$ we can set $a_{\vec{0},0}(A) = 0$.

Proof. If $j_{v_0} = 1, 2$ the proof is obvious as

$$\psi_k = \Im^t \Im^\tau F_1^k + \Im^0 \Im^\tau F_1^k,$$

where F_1^k (and consequently $\Im^\tau F_1^k$) has no polynomial component as proven in Subsection 1.1.4.

For $j_{v_0} = 0$ we obtain the conditions:

$$\begin{aligned} \Pi_P \frac{1}{\cosh(gt)} (F_0^k + \sum_{h=1}^{k} g_h [\sin(\sum_{j=0}^{k-h} q_j)]_{k-h} = 0 \\ \Pi_P \sinh(gt)(F_0^k + \sum_{h=1}^{k} g_h [\sin(\sum_{j=0}^{k-h} q_j)]_{k-h} = 0 \end{aligned} \tag{6.13}$$

The first condition is always verified as the functions ψ_k, ϕ_k, q_k and $x_0^0 = \frac{1}{\cosh(gt)}$ are all limited and x_0^0 tends exponentially to zero. The second condition fixes the g_h recursively:

$$g_k \Pi_P (\tanh^2 t) = 2g_k = -\Pi_P (e^{g|t|}(F_0^k + \sum_{h=1}^{k-1} g_h [\sin(\sum_{j=0}^{k-h} q_j)]_{k-h})). \tag{6.14}$$

The latter identity makes sense only if the right hand side is φ and t independent:

$$\Pi_P(e^{g|t|}(F_0^k + \sum_{h=1}^{k-1} g_h[\sin(\sum_{j=0}^{k-h} q_j)]_{k-h})) = \Pi_P e^{g|t|}\mathcal{W}(\mho_0^k) = c$$

Let us proceed by induction: $\mathcal{W}(t, A)$ is a product of $(\prod_j \partial^{n(j)+l(j)} f^\delta$ and $\mathcal{V}(A^{\geq v})$ with $v \in s(v_0)$ so it is a limited function which can be expanded as:

$$\mathcal{W}(A, t) = \sum_{|\nu|\leq kN} \sum_{h=0}^{\infty} w_{\nu,h}(A) e^{i\omega\cdot\nu t} e^{-hg|t|}$$

naturally we cannot expect that $w_{\vec{0},0} = 0$.
As seen in Lemma 6.11(ii) the constant part of $e^{g|t|}\mathcal{W}(A, t)$ depends only on $w_{\vec{0},1}$; moreover $\mathcal{W}(\mho_0^k)$ is limited and adding leafs with $j_L = 0$ means multiplying by $\frac{1}{\cosh(gt)}$ which is exponentially decreasing. Therefore the only contributions to 6.14 come from trees with at most one leaf L with $j_L = 0$.

Let us now consider leafs with $j_L = 1, 2$. Given a tree A let us choose a leaf L and detach from the tree all the leafs identical to L; we will call B the corresponding tree without the leafs L and $\mathcal{B}$ the set of trees with no leafs identical to L. Adding k leafs L to the tree B is equivalent to applying k derivatives ∂_{j_L} to the nodes of B. Therefore if the total rotation of B is zero such derivative is zero as well.

The total zero momentum contributions from trees with one leaf L with $j_L = 0$ cancel with the corresponding counter-terms. To illustrate this cancellation let us Fourier expand $f(\psi, q)$ fully obtaining a "frequency" label ν_v, n_v. Now let us compare the zero momentum contributions of a tree A with a node v carrying a leaf and the corresponding tree A without the leaf (which appears in the counter-term. In the first case we consider the zero order (in e^{gt}) terms of the expansion

$$e^{inq(t)} = (\frac{(e^{gt} + i)^2}{e^{2gt} + 1})^{2n}$$

in all the nodes and have a $j = 0$ derivative in v. In the second we consider the order one term in e^{gt} in the node v and order zero term in all the others. The order one term is $4in_v$ so the ratio of two values is two (in the first tree there is a factor two coming from the e^{gt} expansion of $\cosh^{-1}(gt)$). This implies that the constant term of all trees A of order h carrying a leaf $j_L = 0$ are canceled by the tree with only one node $\sigma_v = h$ and the same leaf. □

This means that we can apply Lemma 6.11 to all the nodes so that

$$\Im^t w_j(t, \tau)\mathcal{W}(A, \tau) = \varepsilon^{j-1}\Im^t\Im^\tau\mathcal{W}(A, \tau') \quad \text{if} \quad j = 1, 2 \tag{6.15}$$

$$\Im^t w_0(t, \tau)\mathcal{W}(A, \tau) = \Im^t\Im^\tau \frac{\mathcal{W}(A, \tau')}{\cosh(g\tau')} + \Im^t(\cosh(g\tau)\Im^\tau \frac{\mathcal{W}(A, \tau')}{\cosh(g\tau')} +$$

$$\Im^t \frac{\sinh(g\tau)}{\cosh^2(g\tau)} \Im^\tau \sinh(g\tau')\mathcal{W}(A,\tau'),$$

in each node with $j_v = 0$ we choose one of the three terms and denote it with an extra label $p_v = 1, 2, 3$ in the nodes with $j_v = 1, 2$ we set $p_v = 1$.

This Proposition and the relation 6.14 show that the coefficients g_h are fixed uniquely; a direct consequence is that the value of the splitting vector and splitting matrix can be expressed via amputated trees such that for each node v the integrations $\Im^t$ are always on functions F with no constant component $f_{\vec{0},0}$. This is true for v_0 as well as $\Delta I = \Im F_1^k$ and F_1^k has no constant component. To complete this brief review of the articles [Ge] and [GGM4] let us conclude by stating the following property (proved in [Ge] and [GGM4]):

Corollary 6.13. *Fixing the g_h as in relation 6.14 implies that the Lyapunov exponent of the separatrix is g.*

Notice that all we have done in this Subsection does not depend on the number of degrees of freedom (and on the choice of the matrix A).

Remark 6.14. *To prove the convergence of the Lindstedt series it is necessary to show compensations between the " resonances" which are subtrees stemming from v having a purely oscillating term in $\mathcal{W}(A^{\geq v})$, such terms generate small denominators $(\omega \cdot \nu)^{-1}$. In our approximation however ($k \leq \varepsilon^{-1}$) we can approximate all the small denominators with ε^α so we will ignore the compensations.*

6.2.2 Improved bounds for three dimensional systems

In this subsection we will adapt Section 3.1 to the Hamiltonian:

$$\frac{1}{2}(I^2 + \varepsilon J^2 + p^2) + (g + \eta G(\eta, g)) \cos q + \eta f(q, \psi, \phi),$$

Where f is a trigonometric polynomial in ψ, ϕ. To fix a class of examples we will consider

$$f(\psi, \phi, q) = (\cos(\psi) + \cos(\phi)) f(q)$$

such that $f(q_0(t))$ is a rational function in e^{gt} tending to zero for $|t| \to \infty$ and with at least one pole with $g|\text{Im }t| < \pi/2$.
We will perform the computations for

$$f(q) = \frac{2}{\cos(q) + 3},$$

which has one simple pole in $g|t| = i\pi/4$.
Moreover we will consider an example where $f(q)$ is a trigonometric polynomial and find better bounds that those proposed in Section 3.1 and in [GGM3].

$$f(\psi, \phi, q) = (\cos(\psi) + \cos(\phi))(\cos(q) - 1)$$

In this subsection we will return to the resummation tree notation as we need to evidence the analytic and non analytic parts in the splitting determinant. The resummated trees will however carry the extra labels σ_v (counter-term label) and ν_v (rotation label).

Moreover we will use the fact that the 2×2 splitting matrix Δ satisfies Corollary 4.18:

$$\Delta U \sim V, \tag{6.16}$$

where U, V are 2×2 matrices and U is invertible. This means that

$$\det \Delta = \det(U^{\leq K})^{-1} \det V^{\leq K} + o(\eta^K).$$

The matrices U, V are those defined in Subsection 4.1.4; it is easily seen however that for systems with three degrees of freedom one can choose U and V in the following way[6] (U_i, V_i are the columns of U and V):

$$U_1 = \begin{vmatrix} 1 & 0 \\ 0 & \varepsilon^{-1} \end{vmatrix} \omega + \begin{vmatrix} I^1(t=0,\varphi=0) \\ J^1(t=0,\varphi=0) \end{vmatrix} \qquad U_2 = \begin{vmatrix} 1 \\ 0 \end{vmatrix} - M_0 - \frac{m_0}{-2+P_0^1} U_1$$

and

$$V_1 = g_0(-2+P_0)\,, \quad V_2 = g_1 m_1 + (\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + G_1) M_1.$$

U_1 is the value of the actions of the rotators at the homoclinic point and correspondingly $-2+P_0^1$ is the value of p at the homoclinic point.
G_0 and G_1 are the gradients in φ of the values of fruits respectively of type $0,1$ and label $j = 1, \dots, n$.
g_0 and g_1 are the gradients of of the values of fruits respectively of type $0,1$ and label $j = 0$. This matrices can be computed using the techniques of the preceding subsection, for instance:

$$(G_1)_{ij} = \Im x_i^0(t) \mathcal{W}(\Lambda_{i,j}, t) \ \text{ with } \ i,j \neq 0 \dots$$

Remark 6.15. *For systems with n degrees of freedom we consider the equation*

$$G_0(Id_n - \frac{1}{-2+I_0^{(1)}}(A^{-1}\tilde{\omega} + I^{(1)})u_{01}^t + M_{01}^t A) = -g_1 u_{11}^t - G_1 A M_{11} + A M_{11},$$

where the first column of M_{11} and of u_{11}^t are exponentially small.

Proposition 6.16. *The G_i and g_i can be bounded from above (up to order $k = \varepsilon^{-\frac{1}{2}}$)by:*

$$|(G_1^k)_{ij}|, |(g_1^k)_i|, |(g_0^k)_i| \leq (\eta)^k C^k \varepsilon^{-k}.$$

The following lemma will be useful in the proof.

Lemma 6.17. *A tree of order k with m nodes $\delta_v = 0$ can have at most $k - m$ small denominators ε^{-1} and $2m$ denominators $\varepsilon^{-\frac{1}{2}}$.*

[6] We are simply using Propositions 4.14 and 4.16 instead of Corollary 4.18

Proof. Each node with $j_v = 1, 2$ can carry a small denominator ε^{-1} coming from the purely oscillating terms of $\mathcal{W}(A, t)$

$$\Im^t \Im^\tau \sum_{|\nu| \leq Nk} w_{\nu,0} e^{i\omega\cdot\nu}$$

the result is again a purely oscillating term. If we have m nodes with $\delta_v = 0$ then at least $m+1$ of the k nodes with $\delta = 1$ have $j = 0$ so we have at most $k - m - 1$ small denominators due to nodes with $j = 1, 2$.

By the boundedness of $\mathcal{W}(A, t)$ the only purely oscillating terms for a node with $j_v = 0$ appear in

$$\Im^t \sinh(g\tau) \mathcal{W}(A, \tau)$$

if $w_{\nu,1} \neq 0$ and contribute

$$\frac{w_{\nu,1}}{i(\omega\cdot\nu)} \sum_{h=1}^{\infty} \frac{e^{i\omega\cdot\nu t} x^{2h-1}}{2h - 1 + i\omega\cdot\nu} \tag{6.17}$$

which is a function with no purely oscillating term. The purely oscillating contribution from the $\Im^t$ integral in

$$\Im^t \frac{\sinh(g\tau)}{\cosh(g\tau)^2} \Im^\tau \sinh(g\tau') \mathcal{W}(A, \tau') - \Im^t \cosh(\tau) \Im^\tau \frac{\mathcal{W}(A, \tau')}{\cosh(\tau')},$$

has again only one small denominator. So, even if each node carries two integrals (and thus two potentially small denominators), there can be only one factor $\varepsilon^{-\frac{1}{2}}$ for each node with $j = 0$. The (remaining) nodes with $j = 0$ are $2m + 1$ so the small denominators produce a factor bounded by ε^{-k}. □

Proof of Proposition 6.16. We proceed in two steps:
1) Given a tree A, split the integral $\Im$ on the first node in three terms:

$$\Im = \Im^{-2a} - \Im^{2a} + \int_{-2a}^{2a},$$

if we choose the third term we consider the nodes following v_0 and repeat the procedure on the external integral (each node carries two integrals as seen in 6.15):

$$\Im^t = \Im^{2\sigma(t)a} + \int_{\sigma(t)a}^{t} . \tag{6.18}$$

Each time we apply the first term of 6.18 we are cutting off the subtree $A^{\geq v}$.Let us call ϑ_i with $i = 1, \ldots, H$ the list of such trees and ϑ' what remains of A. We have at most 9^k terms of the type:

$$\prod_{i=0,H} \mathcal{V}(\vartheta_i, 2\sigma(t)a) \int_{-2a}^{2a} \prod_{v\in\vartheta'} \int_{\sigma(t)a}^{\tau_w} A_{p_v,j_v}(\tau_v) \Im^{\tau_v} B_{p_v,j_v} \Phi_0(\vartheta_j).$$

where

$$\Phi_\varphi(\vartheta) = (-1)^{N(\vartheta)}[\prod_{v\in\vartheta} a_{j_v}(\eta)^{\delta_v}] \prod_{v\geq v_0} \nabla^{\vec{m}_v} f^{\delta_v}, \tag{6.19}$$

$$A_{p,j}(t) = \begin{cases} 1 & \text{if } p=1 \\ \cosh(gt) & \text{if } p=2 \\ \frac{\sinh(gt)}{\cosh^2(gt)} & \text{if } p=3 \end{cases}$$

$$B_{p,j}(t) = \begin{cases} 1 & \text{if } p=1,\ j=1,2 \\ \frac{1}{\cosh(gt)} & \text{if } p=1,2\,,\ j=0 \\ \sinh(gt) & \text{if } p=3. \end{cases}$$

2) Given a tree of order k we compute $\mathcal{V}(A, 2\sigma(t)a)$ using Lemma 6.11, we expand all the functions f $\frac{1}{\cosh(gt)}$ etc. in series in $x = e^{|g|t}$ and remember that

$$\Im^t e^{\alpha t} = \frac{e^{\alpha t}}{\alpha}, \quad \alpha \in \mathbb{C}.$$

The small denominators can be bounded by

$$|\alpha| = |k + i\omega\cdot\nu_T(v)| \geq \begin{cases} 1 & \text{if } k\neq 0 \\ \sqrt{\varepsilon} & \text{if } k=0 \end{cases} \tag{6.20}$$

so ε small terms occur only in integrals of purely oscillating functions.

We apply proposition 3.6 bounding the denominators with ε^{-k} as seen in Lemma6.17. We bound the definite integrals with the maximum of the integrand and $a_j(\varepsilon)$ with one:

$$\prod_{v\in\vartheta} a_{j_v}(\eta)^{\delta_v} \max_{t\in(-2a,2a)}(|A_{p_v,j_v}||B_{p_v,j_v}|| \prod_{v\geq v_0} \partial_{j_v} \prod_{v'\in s(v)} \partial_{j_{v'}} f^{\delta_v}|),$$

this are all ε independent constants (we are not shifting the integration to complex t's). We obtain the following bounds on the $G_i$and g_i:

$$|G_1^k|, |g_1^k|, |G_0^k|, |g_0^k| \leq (\eta)^k C^k \varepsilon^{-k},$$

which comes from the small denominators. □

The factor $\det U$ has an η independent part equal to $\frac{\omega_2}{\varepsilon^{\frac{1}{2}}}$ plus a η-dependent correction which can be bounded by

$$C_2 \max(|I^1|, |J^1|, |(M_0)_i|) = O_\eta(\eta).$$

M_0 and m_0 are values of fruitless trees with two markings[7], one analytic which can be applied on any node and one non analytic on the first node; M_1 and m_1 (which are exponentially small) are values of fruitless trees with two analytic markings, moreover one of the markings has $j = 1$:

$$(M_1)_i = \Im \mathcal{W}_0^1 \mho_{i,1}^{0\,0}\,, \quad (m_1) = \Im \mathcal{W}_0^1 \mho_{1,0}^{0\,0}.$$

As we are now considering resummated trees, their value is obtained through the operator $\Im\mathcal{W}_\varphi^1$ and thus by applying $\Im_+^{\tau_w} + \Im_-^{\tau_w}$ in each node. Given a tree with total rotation ν_T we shift the integration for the analytic trees to $\mathbb{R} + i\sigma(\nu_T \cdot \omega)d$, as seen in Section 3.1. We are considering integrals of the type:

$$A^1(\nu) = (-\frac{1}{2})^N (A) E(d,\nu) \sum_{\{\nu_v\}_\nu^m} [\prod_{s=1,\dots,n\, v\geq v_0} (i\nu_{v\,s})^{m_v^s}](i\nu_{y\,j})$$

$$\oint \frac{dR_{v_0}}{2i\pi R_{v_0}} \int_{-\infty}^{\infty} d\tau_{v_0} e^{-\sigma(\tau_{v_0})R_{v_0} x_i^l(\tau_{v_0}+id)} d^{n(v_0)} f_{\nu_{v_0}}(q(\tau_{v_0}+id)) e^{i\omega_v \tau_{v_0}}$$

$$\prod_{v>v_0} \oint \frac{dR_v}{2i\pi R_v} (\int_{-\infty}^{\tau_w} d\tau_v + \int_{\infty}^{\tau_w} d\tau_v) x_j^0(\tau_y) e^{-\sigma(\tau_v)R_v(\tau_v+id)} w_{j_v}(\tau_w+id, \tau_v+id)$$

$$\prod_{v\geq v_0} d^{n(v)} f_{\nu_v}(q(\tau_v+id)) e^{i\omega_v\tau_v},$$

with two markings i, x_i^l in the node v_0 and j, x_j^0 in the node y. Clearly in the non analytic integrals ($l = 1$) we set $d = 0$.

To re-obtain the nested integrals $\prod_{v\geq v_0} \Im_w^\tau$ we remember that

$$\Im_+^t + \Im_-^t = \Im^t + \frac{1}{2}\sigma(t)\Im. \tag{6.21}$$

Moreover, as seen in 2.5 the value of a subtree stemming form a node v is

$$(\Im_+^{\tau_w} + \Im_-^{\tau_w}) w(\tau_v + id, \tau_w + id) \mathcal{W}_\varphi^1(A^{\geq v}, \tau_v + id),$$

if we fix the initial data at the homoclinic point $\varphi = 0$ we can group the value as sum of three contributions

$$\Im^{\tau_w+id} w(\tau_w+id, \tau_v+id)\mathcal{W}_{\varphi=0}^1(A^{\geq v}, \tau_v+id)+$$

$$\frac{1}{2} x_j^1(\tau_w+id) \Im x_j^0(\tau_v+id) \mathcal{W}_{\varphi=0}^1(A^{\geq v}, \tau_v+id)+$$

[7] Notice that, for systems with n degrees of freedom, we still should consider values of fruitless trees with two markings, one of which analytic.

$$\frac{1}{2}\sigma(\tau_w)x_j^0(\tau_w+id)\Im\sigma(\tau_v)x_j^1(\tau_v+id)O_1\mathcal{W}^1_{\varphi=0}(A^{\geq v},\tau_v+id),$$

so that the second and third summand are different from zero respectively if $\mathcal{W}^1_{\varphi=0}(A^{\geq v},t)$ is even or odd. Moreover the second and third summand act like fruits, namely the contribution of $A^{\geq v}$ to the value of A is a fixed function (resp. x_j^1 and $\sigma(t)x_j^0$) times a t-independent factor:

$$(E(\nu_T(A^{\geq v}),d))^{-1}\Im x_j^0(\tau)O_1\bar{\Psi}^1_{\varphi=0}(A^{\geq v},\tau)$$

or

$$(E(\nu_T(A^{\geq v}),d))^{-1}\Im\sigma(\tau)x_j^1(\tau)O_1\bar{\Psi}^1_{\varphi=0}(A^{\geq v},\tau).$$

The factor $(E(\nu_T(A^{\geq v}),d))^{-1}$ is there as we have shifted back the integration on the real axis (remember that $\sigma(t)x_j^1(t)$ is analytic).
We will represent the choice of one of the summands by applying the type labels $h_v=b,0,1$.

Lemma 6.18. *Given a fruitless tree $A\in\mathcal{A}_{i,j}$ let v_0 and v be its marked nodes; the only contributions to the M_l and m_l , $l=0,1$ are from trees such that*

$$w\in\mathcal{P}(v_0,v)\to h_w=b,0\quad w\notin\mathcal{P}(v_0,v)\to h_w=b,1.$$

Proof. Given a node $w\notin\mathcal{P}(v_0,v)$ suppose that $h_w=0$ and that we don't give the h label to the other nodes.
The contribution of $A^{\geq w}$ is $\Im x_j^0\mathcal{W}^1_{\varphi=0}(A^{\geq w})$ and $A^{\geq w}$ is fruitless and with one marking ∂_{j_w} in the first node w. This is the integral of an odd function and so it is zero. In the same way if $w\in\mathcal{P}(v_0,v)$ then $A^{\geq w})$ has two markings and so $\mathcal{W}^1_{\varphi=0}(A^{\geq w})$ is even and $\Im\sigma(\tau)x_j^1\mathcal{W}^1_{\varphi=0}(A^{\geq v})=0$. □

We have obtained a tree with "leafs" (i.e. markings x_j^0), it should be noticed that a label $h_v=1$ acts just like a leaf as it contributes $x_j^0(\tau_w)C(A^{\geq v})$, whereas the label $h_v=1$ is a proper marking on $A^{\setminus v}$. Now starting from the end-nodes we cut away the subtrees with labels $h_v=0,1$; the value of a tree A is then a product of values of amputated trees with leafs and with two markings (i.e. any number of markings x_j^0 (leafs) and at most two markings x_j^1). If we are considering M_1 and m_1 there is only one marking x_j^1.
Remember that, in M_1 and m_1, we have shifted the integration before dividing Q_j using 6.21. In the previous Subsection we imposed that the value $\mathcal{W}_\varphi$ of a tree with leafs A will have no contributions from the constant part of $\mathcal{V}(A^{\geq v})$ for all $v\in A$.

Lemma 6.19. *(ii) Given a tree A with at most two non-analytic marking x_j^1 in the first node v_0 and in a node v, the value of the tree is given by trees such that for each $v\in A$ $v\notin\mathcal{P}(v_0,v)$ then $\mathcal{V}(A^{\geq v})$ has no constant part.*

Proof. For any $w\notin\mathcal{P}(v_0,v)$, the contribution of $A^{\geq w}$ to the value of A is $\mathcal{V}(A^{\geq w},t+id)$, where $\mathcal{V}(A^{\geq w},t)$ has no constant component. On the other hand if v is marked x_j^1 and $w\in\mathcal{P}(v_0,v)$ then $A^{\geq w}$ is carries a marking which is not a leaf. □

Remark 6.20. *If we work directly on the exponentially small pieces of the splitting matrix D we can assume that no node carries constant components. Such components must cancel out after summing on all the possible contributions.*

Proposition 6.21. *(i) The contributions to M_0 and m_0 of order k are bounded by $k!C^k\varepsilon^{-3k/2}$.*
(ii)The contributions to M_1 and m_1 of order $k \le \varepsilon^{-\frac{1}{2}}$ and total rotation ν are bounded from above by:

$$E(D,\nu)C^k \max(\sqrt{\varepsilon}^{-(p+2)k+1}, k!\varepsilon^{-3k/2}).$$

Proof. (i) We want to evaluate an integral of the type:

$$\sum_{\{\nu_v\}_\nu^m}[\prod_{s=1,\dots,n\, v\ge v_0}(i\nu_{v\,s})^{m_v^s}](i\nu_{y\,j})\Im^0 x_i^l(\tau_{v_0})d^{n(v_0)}f_{\nu_{v_0}}(q(\tau_{v_0}))e^{i\omega_v\tau_{v_0}}$$

$$\prod_{v>v_0}\Im^{\tau_w}x_j^0(\tau_y)e^{-\sigma(\tau_v)R_v(\tau_v)}w_{j_v}(\tau_w,\tau_v)\prod_{v\ge v_0}d^{n(v)}f_{\nu_v}(q(\tau_v))e^{i\omega_v\tau_v},$$

first we apply Lemma 6.19 to evaluate the contributions of trees $A^{\ge w}$ such that $w \notin \mathcal{P}(v_0,v)$ and w follows a node $v' \in \mathcal{P}(v_0,v)$. As we are considering trees of order $k \le \varepsilon^{-\frac{1}{2}}$ the small denominators are controlled by Lemma 6.17 so we have a factor bounded by ε^{-k} ($\varepsilon^{-\frac{1}{2}k}$ for partially isochronous systems).
We repeat the procedure of Proposition 6.16 and split the integration as in expression 6.18[8] . As we are not shifting the integration near a complex singularity of $f(q(t))$ we can bound all the $|d^{n(v)}f|$ by an ε independent constant.
Having reached the nodes $v \in \mathcal{P}(v_0,v)$ we can have contributions from trees with zero total momentum. As $|\mathcal{P}(v_0,v)| \le k$ we still have to perform at most k integrations, k being reached only if v is an end-node.
This kind of bounds were discussed in Chapter 3, but the existence of counter terms will give us better bounds than the expected $(k!)^2\varepsilon^k$.let us first discuss $x_j^1 = t$.
The integrand at the first node of the path can have no constant component as it is t times a function with no constant component. So we can use double integrals:

$$\Im^t A(\tau)\Im^\tau B(\tau')e^{\alpha\tau'}$$

where the functions A, B are defined in (6.19). We remind that

$$\Im^t e^{\beta\tau}\Im^\tau \tau' e^{\alpha\tau'} = \frac{te^{(\alpha+\beta)t}}{\alpha(\alpha+\beta)} - (\frac{1}{\alpha(\alpha+\beta)^2} + \frac{1}{\alpha^2(\alpha+\beta)})e^{(\alpha+\beta)t}.$$

Let us consider three adjacent nodes $v_1 < v_2 < v_3$. The integrand in v_3 is $te^{\alpha t}$ with non zero $\alpha \in \mathbb{C}$. So we can apply the double integral above and obtain three terms contributing to the integrand in v_2. The linear term in t cannot produce constant

[8] notice that, as we cannot have subtrees $A^{\ge v}$ such that $\Pi_P\mathcal{V}(A^{\ge v}) \ne 0$, we do not have the factor $(k!)^2$ of Proposition 3.6

terms[9] The remaining, purely exponential, terms can instead produce constant factors. The integration of such constant factors produces a factor t^2 times some term with no constant factor as integrand of v_1. So recursively we can have a small denominator $(\sqrt{\varepsilon})^{-l}$ after passing l nodes with zero total momentum and so with no divisor. Notice that*the presence of the counter-terms implies that it is not possible to have two adjacent nodes both having zero total momentum.*
One can proceed in the same way for

$$x_0^1 = \frac{t}{\cosh gt} + \sinh gt,$$

namely we have the integrals (applied to $\mathcal{W}(A^{\geq v}, t)$):

$$\frac{1}{\cosh gt}\Im^t \sinh^2 g\tau\,, \quad \Im^t (t-\tau)\sinh g\tau\,, \quad \sinh gt \Im^t \frac{\sinh g\tau}{\cosh g\tau}\,, \quad \Im^t \frac{(t-\tau)\sinh g\tau}{\cosh gt \cosh g\tau}\,,$$
$$\frac{1}{\cosh gt}\Im^t \frac{\tau \sinh g\tau}{\cosh g\tau}\,, \quad \sinh gt \Im^t \frac{\tau}{\cosh^2 g\tau}\,, \quad \Im^t \frac{\tau(t-\tau)}{\cosh g\tau}\,, \quad \Im^t \frac{\tau(t-\tau)}{\cosh gt \cosh g\tau}.$$

Only the first five terms have constant terms as integrands, coming from purely oscillating factors of all the nodes $A^{\geq v}$. The result of such integrations however is either:

$$\frac{t^2}{\cosh gt}\,, \qquad \frac{t}{\cosh gt} \qquad t\sinh gt\,, \qquad t^2$$

and all (but the third) cannot have again constant terms if applied to a node $J = 1, 2$. Moreover $t \sinh gt$ can only produce a t^3 which we have already discussed. So we can have zero total momentum contributions from a chain of single nodes but in each step we can rise the t exponent only by one.The only exception is a possible

$$t^h \sinh gt \to t^{h+1} \sinh gt \to t^{h+2}$$

but this can only happen once in the whole path and then lead to a known purely polynomial term. In all the parts with no constant components (and carrying $(t-\tau)$) we can pass to double integrals so the t degree does not grow.
Finally this implies that we produce at most a factor $(k)!\varepsilon^{-k/2}$ (see the proof[10] of Proposition 6.16). Finally we consider the $|\partial_j^{m_v^j} f|$, and the proper integral parts which are ε independent and so can be ignored.
(ii) We want to evaluate an integral of the type:

$$E(\nu, d) \sum_{\{\nu_v\}_\nu^m} [\prod_{s=1,\ldots,n\, v\geq v_0} (i\nu_{v\,s})^{m_v^s}](i\nu_{y\,j})\Im^0 x_i^l(\tau_{v_0} + id) d^{n(v_0)} f_{\nu_{v_0}}(q(\tau_{v_0} + id))e^{i\omega_v \tau_{v_0}}$$

[9]It is simply t times the result of the integration of $e^{\alpha t}$, so the eventual constant terms in v_2 are canceled by the counter-terms.

[10]notice that we have shown that the sum of the non zero t exponents h_v is bounded by $k/2$ as the t exponent cannot grow (except once in the whole path) on a single node having the factor $(t-\tau)$. so the $(m!)^2$ becomes $k!$.

$$\prod_{v>v_0} \Im^{\tau_w} x_j^0(\tau_y+id)e^{-\sigma(\tau_v)R_v(\tau_v+id)}w_{j_v}(\tau_w+id,\tau_v+id)\prod_{v\geq v_0} d^{n(v)}f_{\nu_v}(q(\tau_v+id))e^{i\omega_v\tau_v},$$

we have small denominators contributing at most $\varepsilon^{-3k/2}$ and the factor $k!$ as in point (i). Finally we bound

$$|\partial_0^{n_v(0)}f(z)|_{|Rez|>2a,|Imz|\leq 2\pi} \leq n_v!C$$

where $C = |f(z)|_{|Rez|>2a,|Imz|\leq 2\pi}$ is an ε independent constant.
In evaluating the proper integrals, we notice that we do not get ε close to $g\pi/2$ which is the singularity of $\cos(q(t))$ and of the Wronskian. So we can bound these functions with ε independent constants in $C(D-\sqrt{\varepsilon},2a)$. We obtain:

$$E(D,\nu)C^k \sum_{A\in\mathcal{A}_j^k} \frac{N^2 k}{\mathcal{S}(A)} \prod_{v\in A,\delta_v=1} \int_0^{2|a|} dt|d^{n_v}f(q(t+id))|$$

to get better bounds on the integral we have to specify the function $f(q)$ so that we can bound the derivatives in q with some function whose primitive we can estimate more efficiently.
In general we can use the same bounds as in Section 3.1, bounding the integral with the maximum of the integrand, we obtain:

$$\prod_{v:\delta_v=1} n_v!\varepsilon^{\sum_{v:\delta_v=1}(p+n_v)} = \prod_{v:\delta_v=1} n_v!\varepsilon^{(p+2)k/2}, \tag{6.22}$$

as $\sum_{v:\delta_v=1} n_v = 2(k+m)-1-2m$ if m is the number of nodes with $\delta=0$.
Remember that (see Appendix A.2):

$$N(k,j) = \sum_{A\in\mathcal{A}_j^k} c(A)\prod_{v\in A} n_v! \leq (4n)^k.$$

□

Let us perform the computations for

$$f(q) = \frac{1}{\cos q + 3},$$

we can bound the absolute value of the order n derivative of this function by:

$$\frac{n!C^n}{|(\cos q+3)^{n+1}|},$$

then in the definite integrals we ignore the possible constant terms (leading to polynomial contributions), the functions coming from the Wronskian which are bounded by ε independent constants and obtain:

$$\int_{-2a}^{0} dt\Im^t|\partial_0^n f(q(\tau+i\pi/4-i\sqrt{\varepsilon}))| \leq n!C^n \int_{-2a}^{0} dt\Im^t\Big(\frac{\sqrt{(1+2\sqrt{\varepsilon})\cosh^(2g\tau)}}{\sqrt{\sinh^2(2g\tau)+4\varepsilon\cosh^2(2\tau)}}\Big)^{n+1}.$$

We bound the numerator with an ε independent constant and multiply and divide by $\cosh(2\tau)$ which does not vanish for $t \in [0, 2a]$. We obtain the integral:

$$\int_{-\sinh(2a)}^{0} dx \int_{-\infty}^{x} dy (\frac{1}{\sqrt{y^2+4\varepsilon}})^{n+1} = \varepsilon^{-(n-1)/2} \int_{-\infty}^{0} dx \int_{-\infty}^{x} dy (\frac{1}{\sqrt{y^2+1}})^{n+1},$$

if the integrals converge i.e. for $n > 1$. For $n = 0, 1$ we compute the primitive $\Im^t$ which is respectively :

$$\text{arcsinh } (x/\sqrt{\varepsilon}) \leq x/\sqrt{\varepsilon}\,, \text{ for } \varepsilon \ll 1\,, \quad \text{and } \sqrt{\varepsilon}^{-1} \arctan(x/\sqrt{\varepsilon}),$$

both functions are not divergent in $x = 0$ and so we can bound the definite integral in x by $C/\sqrt{\varepsilon}$. Now, as in 6.22, $\sum_{v:\delta_v=1}^{m} n(v) - 1 \leq k$. So the worst bounds come from the small denominator terms.

Let us now consider the example 2) which is a trigonometric polynomial. In this case we have to consider the divergence of the Wronskian and of the nodes with $\delta = 0$ in $t = ig\pi/2$. However we have an important simplification in the evaluation of the proper integrals $\int_0^1$(we chose $a = 1$). Let us set $g_0 = \cos(q_0(t))$ and $g_1 = \sin(q_0(t))$ the derivatives $\partial^{n_v(0)} f^\delta$ are either g_0 or g_1 so we remove the label δ and consider a new label $d = 0, 1$. The definite integrals are then:

$$\prod_v \int_0^1 dt (\Im^t \frac{|g_{d_v}(\tau+id)|}{|\cosh(g\tau+igd)|} + |\cosh(gt+igd)| \Im^t \frac{|g_{d_v}(\tau+id)|}{|\cosh(g\tau+igd)|} + \tag{6.23}$$

$$\frac{|\sinh(gt+igd)|}{|\cosh^2(gt+igd)|} \Im^t |g_{d_v}(\tau+id)| \sinh(g\tau+igd)).$$

Setting $d = i\pi/2 - \sqrt{\varepsilon}$, it is quite easy to find bounds on this integrals; for instance:

$$\int_0^1 \Im^t \frac{|\sin(q_0(\tau+id))|}{|\cosh(g\tau+igd)|} = \sqrt{2} \int_0^1 \Im^t \frac{\sqrt{\cos(2\sqrt{\varepsilon}) + \cosh(2t)}}{(-\cos(2\sqrt{\varepsilon}) + \cosh(2t))^{3/2}} \leq$$

$$C \int_0^1 \Im^t \frac{1}{(-\cos(2\sqrt{\varepsilon}) + \cosh(2t))^{3/2}} = \text{Csc}(\sqrt{\varepsilon})^2 \int_0^1 dt \frac{\sinh(t)}{\sqrt{1 + 2\sinh^2(t) - \cos(2\sqrt{\varepsilon})}} \leq C\varepsilon^{-1}$$

We can bound all the summands in expression 6.23 with $C\varepsilon^{-1}$. As there are at most $2k-1$ nodes the following proposition holds.

Proposition 6.22. *We can bound the terms of order k and total harmonic ν in M_1 and m_1 of example 1) and 2) respectively by:*

$$E(D,\nu) C^k k! \varepsilon^{-3/2k}\,, \qquad E(D,\nu) C^k \varepsilon^{-2k+1}.$$

This implies Melnikov dominance in the examples 1),2)for respectively:

$$\eta \le \varepsilon^2\,, \quad \eta \le \varepsilon^3.$$

Proof. The proposed bounds are obvious form what discussed above.

This proves that the formal power series involved in equation (6.16) are asymptotic for $\eta \le \varepsilon^2$. To prove Melnikov dominance one can work directly on the splitting matrix (see Remark 6.20) so this removes a factor $(k-1)|\varepsilon^{-k/2}$ in the bounds of example 1). Now we proceed as in Proposition 6.3, summing on the slow modes and on the terms of order higher than one (resp for $\eta \le \varepsilon^2$ and $\eta \le \varepsilon^3$. In example 1) we have a simple pole and in 2) a double one:

$$\mu\varepsilon^{-\frac{1}{2}} > \sqrt{\varepsilon}\mu^2\varepsilon^{-2}\,, \quad \mu\varepsilon^{-1} > \varepsilon\mu^2\varepsilon^{-4}.$$

□

6.3 D'Alembert-like problems

In the previous section we have refined the bounds on η that imply Melnikov dominance; we have found $\eta < \varepsilon^2$ for example (1) and $\eta < \varepsilon^3$ for example (2). Both values of η imply the convergence of the KAM construction as discussed in Appendix A.4. In this subsection we obtain still better bounds for Hamiltonians having a big uni-modal (quasi-monochromatic) perturbation, we work on the following class of examples:

$$\frac{1}{2}(\varepsilon J^2 + p^2) + I\frac{\omega_1}{\sqrt{\varepsilon}} + \cos q - 1 + \alpha A(\phi+\psi)B(q) + \eta f(\phi,\psi,q). \tag{6.24}$$

The functions $A(x)$, $B(x)$ are trigonometric polynomials of degree N; the function f is a trigonometric polynomial in ψ, ϕ and rational in e^{iq} with at least one pole for finite values of Imq. Finally α is a free parameter. Hamiltonians of the form 6.24 are interesting as they provide a "model" for the D'Alambert problem (see [CG] for a discussion of the D'Alambert problem). An Hamiltonian of the form6.24 (but where f is a trigonometric polynomial) is proposed in [GGM3].

The cited article contains a proof of the existence of stable/unstable manifolds provided that $\alpha\varepsilon^{-\frac{1}{2}} \ll 1$; moreover the gaps between persistent unstable tori are proved to be smaller than $e^{-C/\sqrt{\varepsilon}}$ for any order one C. The proof relies on the monochromaticity of $A(\phi+\psi)$ which permits us to perform a Poincaré Birchoff transformation on the Hamiltonian which reduces the α dependent part of 6.24 to size $\alpha\sqrt{\varepsilon}$. We report the details in Appendix A.5.To prove lower bounds on the splitting determinant for system 6.24 we must prove "Melnikov" dominance, which means computing the Melnikov integral and finding appropriate upper bounds on the terms of order h in α and k in η with $h+k>1$. The results of the previous section enable us to find such bounds provided that f is **not** *a trigonometric polynomial in q.*

6.3.1 Big uni-modal perturbations

Following the strategy proposed in the previous section we consider the Hamiltonian:

$$\frac{1}{2}(\varepsilon J^2+p^2)+I\frac{\omega_1}{\sqrt{\varepsilon}}+(G(\eta,\alpha,g))(\cos q-1)+\alpha A(\phi+\psi)B(q)+\eta f(\phi,\psi,q); \quad (6.25)$$

we will first fix the function $G(\eta, al, g)$ by Proposition 6.12 and then fix $g = g(\eta,\alpha)$ so that:

$$g(\eta,\alpha)+\eta G(\eta,\alpha,g(\eta,\alpha))=1.$$

Naturally the perturbation series of the homoclinic trajectory: $\phi(\eta,\alpha,\varphi,t)$, $q(\eta,\alpha,\varphi,t)$ will now be expanded both in η and α:

$$\phi(\eta,\alpha,\varphi,t)=\sum_{h,k=0}^{\infty}\alpha^h(\eta)^k\phi_{h,k}(\varphi,t)\dots$$

and this holds for $G(\eta,\alpha,g)$ as well

$$G(\eta,\alpha,g)=g+\sum_{h+k\geq 1}g_{h,k}.$$

The tree expansion of the homoclinic trajectory carries the following labels: the usual $j_v = 0$ or 2 $\rho_v = 0,1$ then $\delta_v = 0,1,2$ and $k_v, h_v \in \mathbb{N}_0$. The grammar is:

$$\delta_v=k_v=h_v=0\to\{|s(v)|\geq 2\,,j_v=j_{v'}=0\,\forall v'\in s(v)\}$$

$$\delta_v=1\to\{k_v=1\,,h_v=0\}\,,\quad \delta_v=2\to\{k_v=0\,,h_v=1\}.$$

Now we briefly repeat the procedure described in Subsection 6.2.1.
The order of a tree will now be given by two numbers (resp. the order in η and α of the corresponding values):

$$o_1(A)=\sum_{v\in A}h_v\,,\quad o_2(A)=\sum_{v\in A}k_v.$$

We define the vector space of "acceptable" trees of prefixed order $\mathbb{V}^{h,k}$ by defining its generators the set $\mathcal{A}^{h,k}$ of equivalence classes of "acceptable" trees of order (h,k). Now we proceed exactly as in Subsection 6.2.1, namely we add two labels: $\rho_v = 0,1$ and ν_v and we have the so-called trees with leafs. Then we consider marked trees (where leafs are particular markings) and define the value $\mathcal{W}$ of a tree, with a marking $h(t)\partial_l$ in v_0 and some leafs $L(v_0)$, as:

$$\mathcal{W}(A)[t]=\alpha^{h_v}(\eta)^{k_v}g^{\delta_v}_{h_v,k_v}h(\tau_{v_0})\nabla^{\vec{m}_v}f^{\delta_v}(\varphi+\omega t,q_0(t))\prod_{v\in s(v_0)}\mathcal{V}(A^{\geq v}),$$

$$\mathcal{V}(A)=\Im^{t(\rho_{v_0})}w_{j_{v_0}}(t(\rho_{v_0}),\tau_{v_0})\mathcal{W}(A)\,,\quad\text{for all } A\in\mathcal{A}_*. \quad (6.26)$$

As usual $m_v(j)$ is the number of nodes $v' \in s(v), \mathcal{M}(v)$ having $j_{v'} = j$, $g^2_{1,0} = g^1_{0,1} = 1$, $g^0_{h,k} = g_{h,k}$ and $m_v(j)$ is the number of nodes in the list $v, s(v)$ having label $j_v = j$. Moreover $f^1(\psi,\phi,q) = f(\psi,\phi,q)$, $f^{(2)}(\psi,\phi,q) = A(\phi+\psi)B(q)$ and $f^0(q) = \cos q$. Finally we define the vectors:

$$\Lambda_{k,h} = \sum_{A\in\mathcal{A}_{h,k}} c(A)A,$$

such that

$$\phi_{h,k} = \mathcal{V}(\partial_2^{v_0}\Lambda_{k,h})\,,\quad q_{h,k} = \mathcal{V}(\partial_0^{v_0}\Lambda_{k,h}).$$

As in Proposition 6.12 we fix the values of the parameters $g_{h,k}$:

$$g_{h,k} = \Pi_P e^{g|t|}\mathcal{W}(\partial_0\mho_{h,k} - \alpha_{h,k})$$

where $\alpha_{h,k}$ is the tree with only one node v $\delta_v = 0$, $h_v = h$, $k_v = k$.

This ensures that $\mathcal{V}(A)$ is a limited function of t with no constant term:

$$\mathcal{V}(A) = \sum_{|\nu|\le o(A)N}\sum_{l=0}^{\infty} v_{l,\nu}(A)\exp\big(-gl|t| + i(\varphi + \frac{\omega_1}{\sqrt{\varepsilon}}\nu_1 t + \sqrt{\varepsilon}\omega_2\nu_2 t)\big),$$

and $v_{0,\vec{0}}(A) = 0$ for all A.

We have shown that we can extend Subsection 6.2.1 to systems whose perturbation series involves two parameters. Therefore we can improve the bounds of Subsection 6.2.2 using the particular structure of the α perturbation. Notice that a naïf use of Subsection 6.2.2 produces the bounds:

$$|G_i^{k,h}|,|g_i^{k,h}| \le (\eta)^k\alpha^h(\frac{C}{\sqrt{\varepsilon}})^{k+h} \quad |M_0^{h,k}|,|m_0^{h,k}| \le (\eta)^k\alpha^h(k+h-1)!(\frac{C}{\varepsilon})^{3/2(k+h)}$$
$$|M_1^{h,k}|,|m_1^{h,k}| \le E(D,\nu)(\eta)^k\alpha^h C^{k+h}\max(\sqrt{\varepsilon}^{-(p+2)k+1},(k+h-1)!\varepsilon^{-3/2(k+h)}),$$

where D depends only of the function f and is defined in 1.28. Clearly this bounds do not imply Melnikov dominance for order one values of α.

Lemma 6.23. *All the divergent terms in h (i.e. $h!(\sqrt{\varepsilon})^{-3h}$) come from the estimates 6.20 of the small divisors namely in the estimates of $\mathcal{V}(A, 2\sigma(t)a)$ where a is defined in 1.28.*

Proof. The only divergent terms from the definite integrals are in M_1 and m_1. We have imposed that $B(x)$ be trigonometric polynomial and $f(\phi,\psi,q)$ non trigonometric in q with $D < \pi/2$. So in estimating the definite integrals in M_1 and m_1 we never reach the poles $(\pm\pi/2)$ of $B(q(t))$ and we can estimate

$$\max_{t\in C(D,a)}|B(q(t))| \le C = O_\varepsilon(1).$$

□

The estimate 6.20 is clearly not optimal, namely given a tree A with total harmonic $\nu_T = (\nu_1, \nu_2)$ at the homoclinic point $\varphi = 0$, then

$$\mathcal{W}(A) = \sum_{l=0}^{\infty} w_{l,\nu}(A) \exp\big(- gl|t| + i(\frac{\omega_1}{\sqrt{\varepsilon}}\nu_1 t + \sqrt{\varepsilon}\omega_2\nu_2 t)\big),$$

and $w_{0,\vec{0}}(A) = 0$. If we consider the term of order l in $e^{-g|t|}$ the small denominator is:

$$\alpha(l, \nu_T) = | - \sigma(t)gl + i(\frac{\omega_1\nu_1}{\sqrt{\varepsilon}} + \sqrt{\varepsilon}\omega_2\nu_2)| \geq \begin{cases} 1/\sqrt{\varepsilon} & \text{if } \nu_1 \neq 0 \\ 1 & \text{if } \nu_1 = 0\ , l \neq 0 \\ \sqrt{\varepsilon} & \text{if } \nu_1 = l = 0. \end{cases}$$

So, in each node v the denominator is big $(O_\varepsilon(\varepsilon^{-\frac{1}{2}}))$ provided that $\nu_T(A^{\geq v}) = \nu_1(v), \nu_2(v)$ has non zero fast component, $\nu_1(v) \neq 0$. Moreover each node (having non zero total momentum) having a non zero fast component produces a denominator of order ε while small denominators are at most $\sqrt{\varepsilon}$ as seen in the proof of Lemma 6.17.

Proposition 6.24. *For all $k \leq \varepsilon^{-\frac{1}{2}}, h \leq \varepsilon^{-1}$ following bounds hold:*

$$\begin{aligned} |G_i^{k,h}|, |g_i^{k,h}| \leq (\eta)^k \alpha^h C^h (\frac{C}{\sqrt{\varepsilon}})^k \quad |M_0^{h,k}|, |m_0^{h,k}| \leq (\eta)^k \alpha^h C^h (\frac{C}{\varepsilon})^k \\ |M_1^{h,k}|, |m_1^{h,k}| \leq E(D,\nu)(\eta)^k \alpha^h C^{k+h} \sqrt{\varepsilon}^{-(p+2)k+1}. \end{aligned} \tag{6.27}$$

Proof. We proceed in two steps:

1) Consider a tree A having no nodes with possible constant components. Starting from the end-nodes, let us "cut away" all the trees $A^{\geq w}$ such that $\nu_1(w) = 0$. We are left with a set of amputated trees A such that $\nu_1(v_A) = 0$ and $\nu_1(v) \neq 0$ for all the other nodes in A. The integration in each node produces a factor $\alpha^{-2}(l, \nu_T(v))$ and α is big for all the nodes $v > v_0$.We can suppose that $j_{v_0} = 0$ as the $j_v = 1$ have an extra small factor ε and so no small denominators. The node v_0 produces (in any case) at most a factor $\varepsilon^{-\frac{1}{2}}$. Given an amputated tree with k nodes $\delta = 1$, h nodes $\delta = 2$ and $m(\leq 2(k+h) - 1)$ nodes with $\delta = 0$ the small denominator term is at most:

$$\varepsilon^{(h+k+m-1)}\varepsilon^{-\frac{1}{2}} = \varepsilon^{k+h+m-3/2}.$$

In $A(\psi + \phi)$ the only frequencies accessible at order one are (n, n) with $|n| \leq N$ so an amputated tree with only one node v and $\delta = 2$, then $\nu_1(v) \neq 0$. This implies that if $h \neq 0$ then $k + h + m \geq 2$. So the trees whose nodes do not carry constant terms and with $h \neq 0$ nodes $\delta = 2$ can carry at most a factor $\varepsilon^{\frac{1}{2}}$ there are *no small denominators.* 2) In Proposition 6.21 we noticed that each tree A can have a path $\mathcal{P}$ (of length $\leq k + h + m$) of nodes which can have zero total momentum $\nu_T = 0, l_T = 0$. As such terms give rise at most to a factor ε^{-k} (for $k \leq \varepsilon^{-\frac{1}{2}}$), we can ignore the nodes with $j = 1$ as each carries a small factor ε.

Given a tree A (of order $\leq \varepsilon^{-\frac{1}{2}}$)with the nodes carrying momentum labels, we evidence the nodes in $\mathcal{P}$ and in particular we dash the branches going in the nodes with zero total momentum and call the subtrees obtained $\bar{A}_i$ with $i = 1, \ldots, R \geq 1$. We sign on the first and last node of the evidenced nodes in each subtree the initial and final t degree; the initial t degree $P_1(i)$ is smaller or equal to the number of cuts and is equal to the degree of the final node above it plus one. The final degree is $P_2(i) \leq P_1(i)$.

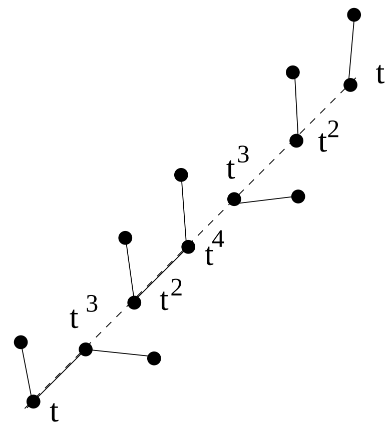


Figure 6.1:

If the path connecting the first and the last node of a subtree has length $l > 1$ and $P_1 - P_2 = r$ then the subtree produces a small denominator:

$$(\sqrt{\varepsilon})^{-(l-1+r)}\frac{l!}{(l-r)!}$$

supposing that all the internal nodes have $\nu_1 = 0$ (and $\nu_2 \neq 0$ by construction). Then, if an internal node has $\delta = 2$, one of its followers (either in $\mathcal{P}$ or not in $\mathcal{P}$), must be fast and so produce a term ε. As we are interested only canceling divergent terms for nodes with $\delta = 2$ we can suppose that all the l nodes have $\delta = 2$. Let us call A_i the subtree we have generated , we have a small denominator factor bounded by:

$$\sum_{A_i}(\varepsilon^{l(i)}(\sqrt{\varepsilon})^{-(l(i)-1+2r(i))}) \leq \varepsilon^{k/2-R} \leq O_\varepsilon(1)$$

as $R \leq k/2$. if we consider $x_j^1 = t$. If we consider a marking $\sinh gt$ there can be H subtrees of length 1 and zero total momentum. Each of these produces a factor ε (if $\delta_v = 2$ and $\nu_T = 0$ then there must be a fast node attached to v); so the small denominator term is bounded by:

$$\varepsilon^H \sqrt{\varepsilon}^{k-H-2R} \leq O_\varepsilon(1),$$

as now $R \leq H + (k-H)/2$. □

Theorem 6.25. *The Hamiltonian (6.24) is uniformly Arnold unstable in the domain:*

$$W(E) := \{H(I, J, \psi, \phi) = E\,, \quad O_\varepsilon(1) = b \leq |I|, |J| \leq a = O_\varepsilon(1)\}$$

for $E \in [E_1, E_2]$ with $E_1, E_2 = O_\varepsilon(1)$, provided that:
1) $\frac{\mu}{\sqrt{\varepsilon}} \leq \varepsilon^{p+2}$ and $\alpha \ll 1$ but still $O_\varepsilon(1)$. 2) $f(\psi, \phi, q)$ is sufficiently non degenerate; for instance we will suppose that $f_{e_1}(q), f_{e_2}(q)$ are not identically zero.

Proof. The bounds (6.27) imply that the contribution, to the entries of the splitting matrix, of a tree with fruits (of order $k, h < C\varepsilon^{-\frac{1}{2}}$) carrying an analytic integral of frequency ν is bounded from above by:

$$E(D,\nu)(\eta)^k\alpha^h C^{k+h}\sqrt{\varepsilon}^{-(p+2)k+1};$$

Moreover we know that $\det\Delta(\alpha,\mu=0)=0$ as the α perturbation is uni-modal. So we can explicitly compute the first relevant order of $\det\Delta$ which is either of first order in η and α or second order in η. Then we sum up the remainders in η and α using Proposition 6.3 $\square$

Chapter 7

Systems with more general unperturbed separatrices

In Section 6.2 we have given sufficient conditions for Melnikov dominance for systems with one fast time scale whose Hamiltonian is of the type 4.3. Now we would like to generalize the dependence of the q variable of the (η)-unperturbed pendulum. i.e. a system whose Hamiltonian is:

$$\frac{(I, A(\varepsilon)I)}{2} + \frac{p^2}{2} - F(q) + \eta f(\phi, q) \tag{7.1}$$

after the scaling change of variables of Remark1.10.
Naturally, in equation 7.1, we consider only periodic functions $F(q)$ which are analytic in a strip $|Im\ q \leq d$ and that do not modify the qualitative behavior of the unperturbed separatrix. We will impose the following conditions.

Condition 7.1. *$F(q)$ is even and analytic for $q \in \mathbb{T}_d$; moreover $F(q)$ verifies:*

1. *$q = \dot{q} = 0$ is an hyperbolic fixed point and the separatrix*

   $$\frac{\dot{q}^2}{2} - F(q) = 0$$

   contains only this fixed point. This holds true if:

   $$F(0) = F(2\pi) = 0, \quad F_q(0) = 0, \quad F_{qq}(0) = \lambda > 0, \quad F(q) > 0 \ \text{ for } \ q \neq 0, 2\pi.$$

2. *Moreover, on the separatrix, we can chose a sign for $\dot{q}$ and the equation of motion on the separatrix is:*

   $$\dot{q} = \pm\sqrt{2}\sqrt{F(q)} = G(q)$$

   where $G(q) \geq 0$ and $G(q) = 0$ if and only if $q = 0, 2\pi$. We will consider initial data $q(0) = \pi$.

Notice that the evenness of $F(q)$ implies that $\dot{q}(t)$ is an even function of t. This qualitative requests on F ensure the existence of a local Hyperbolic normal form for the " pendulum" near $q, p = 0$ and the convergence of the local KAM theorem 1.2. In the preceding Chapter we have considered " perturbative" examples which did not modify Melnikov dominance. In this Chapter we proceed in a completely non perturbative way; namely we give conditions on $G(q)$ sufficient to guarantee that the Melnikov integral dominates in equation 7.1 provided that f satisfies suitable non degeneracy conditions.
We look for functions $F(q)$ such that the time evolution on the separatrix $q(t)$ on a prefixed branch satisfies

Condition 7.2.

$$e^{iq(t)} = R(e^{-t}) \quad \textit{where } R(y) \textit{ is a rational function .} \tag{7.2}$$

Automatically the other branch of the separatrix satisfies:

$$e^{iq(t)} = R(e^{t}).$$

We will not try to classify the functions $F(q)$ satisfying Condition 7.2 but only give classes of examples. Then, in Sections 7.2 and 7.3, we will show that if $G(q)$ satisfies the condition 7.2 then one can prove for Hamiltonian 7.1 the same results as for Hamiltonian 4.3 (with the same techniques of Chapters 4 and 6).

7.1 Acceptable functions $F(q)$

Let us call S^1 the unitary circle in C $e^{iq(t)} = y \in S^1$, and let us call P the real axis plus the point at infinity. Both S^1 and P are circles on the Riemann sphere.

Lemma 7.3. *The only rational functions $w : P \to S^1$ such that $R(\infty) = 1$ are of the type*

$$R(z) = \frac{P(z)}{\bar{P}(z)}, \quad \textit{where } P(z) \textit{ is a polynomial with coefficents in } \mathbb{C},$$

$\bar{P}$ is the polynomial whose coefficients are the complex conjugate of the coefficients of P. The condition $R(\infty) = 1$ implies that the leading coefficient of P is real and so can be set to one both in P and in $\bar{P}$.

Proof. Our request is that for all $z \in P$, $|R(z)| = 1$ so we write that $R(z) = \frac{P(z)}{Q(z)}$ with P and Q of the same degree, with no common zeros and with the same leading coefficient. Then, without loss of generality, we can suppose both P and Q to be monic[1].Then if $\bar{P}$ ($\bar{Q}$) is the polynomial whose coefficients are the complex conjugate of the coefficients of P (Q) we have that

$$\frac{P(z)}{Q(z)} = \frac{\bar{Q}(z)}{\bar{P}(z)}$$

[1]we remind that a monic polynomial is a polynomial whose leading coefficient is one.

for all $z \in P$ and therefore for all $z \in \mathbb{C}$. All the polynomials involved are monic and the decomposition of rational functions in monic polynomials is unique so $Q(z) = \bar{P}(z)$. □

In particular this implies that there can be no real zeros of P.

Let us consider only polynomials $P(z)$ with zeros a_i having $|a_i| = 1$, with Im $a_i \neq 0$ and $i = 1, \cdots h$. We set $z = e^{-t}$ and

$$e^{iq(t)} = \prod_{i=1}^{h} \frac{(e^{-t} - a_i)}{(e^{-t} - \bar{a}_i)}$$

then modulo 2π

$$q(t) = -i \sum_{i=1}^{h} \log(\frac{(e^{-t} - a_i)}{(e^{-t} - \bar{a}_i)}) = 2 \sum_{i=1}^{h} \arctan[\frac{1}{\mathrm{Im} a_i}(e^{-t} - \mathrm{Re}\ a_i). \tag{7.3}$$

We derive the second term of relation 7.3 and obtain:

$$\dot{q}(t) = -2e^{-t} \sum_{i=1}^{h} \frac{\mathrm{Im}\ a_i}{(e^{-t} - \mathrm{Re}\ a_i)^2 + (\mathrm{Im}\ a_i)^2}. \tag{7.4}$$

We want to find conditions on $R(z)$ so that calling $D(z)$ the function such that $D(e^{-t}) = \frac{1}{2}\dot{q}^2(t)$, $D(z)$ can be expressed as an analytic function of $y = R(z)$ in some strip S^1_d (as usual S^1_d is an annular domain of width d around S^1).

Before stating a general proposition let us study a simpler (but still interesting) class of functions such that $D(z(y))$ can be explicitly computed.

Condition 7.4. *Consider the set of rational functions $R(z)$ such that:*

1. *$P(z)$ is monic and has degree two in z with zeros a_1 and a_2 such that $|a_i| = 1$..*

2. *$R(z) = R(-\frac{1}{z})$ and Im $a_i > 0$.*

This implies that the two zeros of $P(z)$ are a and $-\bar{a}$ for some a with $|a| = 1$. Moreover it can be easily verified that equation 7.3 parameterizes $[0, 2\pi)$ injectively and that $\dot{q}(t) < 0$ for all real t.

Remark 7.5. *(i) Systems satisfying Condition7.4 have $q(0) = \pi$ and $\dot{q}(t)$ an even function of t.*
(ii) For systems satisfying Condition7.4 $e^{iq(t)}$ has poles for purely imaginary values of t.

Proposition 7.6. *For all functions $R(z)$, satisfying Conditions7.4, there exists a unique function $H(y)$ which is a rational function of y with poles not in S^1 and such that $H(R(z)) = D(z)$.*

Proof. We will prove it by directly computing the function $H(y)$. The relation:

$$y = R(z) = \frac{(z-a)(z+\bar{a})}{(z-\bar{a})(z+a)},$$

implies that

$$(1-y)(z^2-1) - ibz(1+y) = 0 \quad \text{where} \quad ib = a - \bar{a}. \tag{7.5}$$

We compute the function $\dot{q}(t) = i\big[zd_z(\log R(z))\big]_{z=e^{-t}}$, by differentiating equation 7.5 with $y = R(z)$.

$$y'(1 - z^2 - ibz) = ib(1+y) - 2z(1-y), \tag{7.6}$$

we obtain that:

$$iz\frac{y'}{y} = i\frac{z}{y}\frac{ib(1+y) - 2z(1-y)}{(1-z^2-ibz)}.$$

We use relation 7.5 to simplify the denominator so we obtain:

$$\frac{z}{y}\frac{ib(1+y)-2z(1-y)}{-bz(\frac{1+y}{1-y}+1)} = \frac{1-y}{y}\frac{ib(1+y)-2z(1-y)}{-2b}.$$

The function we want to compute is $\frac{1}{2}\dot{q}^2$ so we square the last relation:

$$(\frac{y-1}{2by})^2(-b^2(1+y)^2 + 4(1-y)[(1-y)z^2 - bz(1+y)]$$

and we substitute again relation 7.5. As we said we obtain that $H(y)$ is a rational function of y:

$$\frac{1}{2}\left(\frac{y-1}{2by}\right)^2(-b^2(1+y)^2 + 4(1-y)^2).$$

□

$H(e^{iq})$ is a trigonometric polynomial; Setting $b = 2\text{Im } a = 2\beta$ an easy computation leads to

$$H(e^{iq}) = \frac{1}{2}(\frac{1}{\beta^2}(\cos q - 1)^2 + \sin^2 q).$$

In Figure 7.1 We show the graph of the separatrix

$$\frac{1}{2}p^2 = H(e^{iq})$$

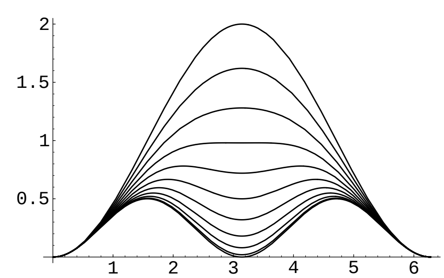


Figure 7.1:

in the phase plane p, q for various values of $0 \leq \beta \leq 1$. The limit value $\beta = 1$ is the pendulum, while the limit $\beta = 0$ is not in our class of functions as $F(\pi) = 0$. Notice that for all values of β the system has a critical point in $q = \pi$.

We have imposed that the zeros of $P(z)$ a and $-\bar{a}$ have positive imaginary part; this automatically forces $\dot{q}(t) \leq 0$.
Naturally this hypothesis is only for notational convention, to ensure that we are parameterizing the lower branch of the separatrix. If a and $-\bar{a}$ have negative imaginary part we only need to set $z = e^t$ in relation 7.3 to be on the lower branch of the separatrix.

Example 7.7. We conclude this simple example of functions $F(q)$ satisfying the conditions 7.2 by representing the phase curves of the Hamiltonian:

$$\frac{1}{2}p^2 - (\cos q - 1)^2 - \frac{1}{2}\sin^2 q$$

where $F(q) = H(e^{iq})$ with $\beta^2 = \frac{1}{2}$.

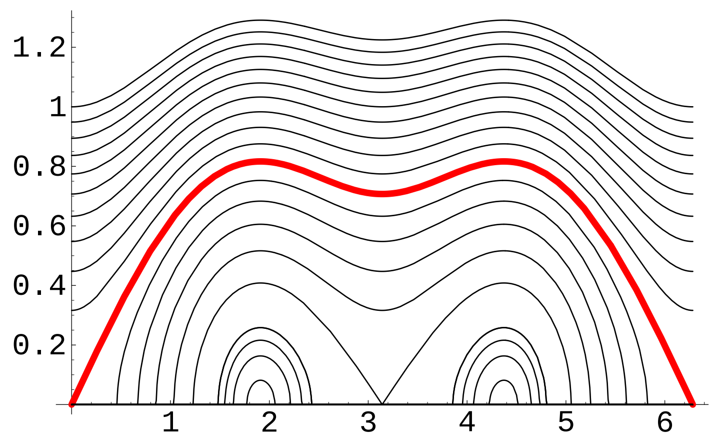


Figure 7.2: The separatrix is the line in red; notice that there are two stable fixed points and one unstable one (different from zero)

We will now consider the more difficult question of functions $R(z)$ such that $P(z)$ is of degree higher that two.

Condition 7.8. *We will restrict our attention to those functions $R(z)$ such that:*

1. $R(z) = R(-\frac{1}{z})$ *so the zeros of $P(z)$ come in couples a_i, $-\bar{a}_i$.*

2. *Let $\{a_, -\bar{a}_i\}_{i=1,h}$ be the list of zeros of $P(z)$ then:*

$$\sum_{i=1}^{h} \sigma(i) = 1 \quad where \;\; \sigma(i) = \sigma(\; Im \; a_i). \tag{7.7}$$

3. *The differential of $R(z)$ is different from zero on P.*

Lemma 7.9. *The Conditions 7.8 imply that the image of P through $R(z)$ is S^1 covered twice and precisely the preimage of each $y \in S^1$ is the couple $z, -\frac{1}{z}$.*

Proof. Let us factorize the application $R(z)$ as[2]

$$R(z) = \prod_{j=1,2h} R_j = \prod_{j=1,2h} \frac{z - b_j}{z - \bar{b}_j}$$

The image of P through each of the R_j is S_1 covered once. The winding number of a product is the algebraic sum of the winding numbers; it is easily seen that the winding number of each of the R_j is

$$\sigma(\text{ Im } b_i) = \sigma(\text{ Im } b_{i+h}) = \sigma(\text{ Im } a_i).$$

Finally as the differential of R in non zero on P then S_1 is covered by the image of P. □

Standard theorems on compact Riemann surfaces extend Lemma 7.9 to an annulus S^1_d.

Proposition 7.10. *There exists an annulus S^1_d such that if we call V the connected component of $R^{-1}(S^1_d)$ which contains P, the following properties hold:*
(i) $R : V \to S^1_d$ is a double covering of S^1_d.
(ii) V is invariant trough the application of $z \to -\frac{1}{z}$; moreover if $p, q \in V$

$$R(p) = R(q) \quad \Leftrightarrow \; q = -\frac{1}{p}.$$

To prove this statement we can use for instance in [F], Theorem 4.22:

it Suppose X and Y are locally compact spaces and $p : X \to Y$ is a proper[3] local homeomorphism. Then p is a covering map.

The map $R : V \to S^1_d$ is clearly a proper local homeomorphism if we choose d sufficiently small.

Corollary 7.11. *Given a function $D(z)$ holomorphic in a strip V' around P and such that:*

$$D(z) = D(-\frac{1}{z})$$

then there exists a function $H(y)$ holomorphic in a strip $S^1_{d'}$ such that in $V' \cup V$ $H(R(z)) = D(z)$.

---

[2]clearly $b_i = a_i$, $b_{i+h} = -\bar{a}_i$ for $i = 1, \ldots, h$.
[3]we remind that a map is proper if the preimage of each compact is compact

Proof. We fix d' so that the connected component of $R^{-1}(S^1_d)$ which contains P is contained in V'.
For any $y \in S^1_{d'}$ there exists an open set $A \in S^1_{d'}$ such there are two open sets B_1, B_2 that represent it in V'; moreover for all $z \in B_1$ $-1/z$ is in B_2 and vice-versa. This implies that the function $D(z)$ assumes the same values on the B_i and so can be lifted to A. Moreover the application $B_1 \to A$ is an isomorphism and so the lifted function is analytic. □

Finally we can state the main theorem of this Section:

Theorem 7.12. *Given any function $R(z)$ satisfying Conditions 7.8 there exists a unique Hamiltonian*

$$\frac{1}{2}p^2 - F(q)$$

satisfying the Conditions 7.1 and 7.2, such that $R(e^{-t}$ is the motion on the lower branch of the separatrix with initial data $q(0) = \pi$. The function $\dot{q}(t)$ on the separatrix is even.

Proof. Given $R(z)$ we only have to prove that the function $D(z)$ such that $D(R(e^{-t})) = \dot{q}^2/2$ respects the prescribed symmetry. We know by expression 7.4 that

$$D(z) = 2z^2\Big(\sum_{i=1}^{h}\big(\frac{\text{Im } a_i}{(z - \text{ Re } a_i)^2 + (\text{ Im } a_i)^2} + \frac{\text{Im } a_i}{(z + \text{ Re } a_i)^2 + (\text{ Im } a_i)^2}\big)\Big)^2,$$

so we can directly compute $D(-1/z)$ and check the identity. In the same way we check that each summand of $\dot{q}(t)$

$$-2e^{-t}\sum_{i=1}^{h}\frac{\text{Im } a_i}{(e^{-t} - \text{ Re } a_i)^2 + (\text{ Im } a_i)^2},$$

is even in t as $|a_i| = 1$ for all i. □

Let us show some examples of functions $R(z)$ satisfying all the conditions 7.8.

Lemma 7.13. *The function*

$$R_0(z) = \big(\frac{z-i}{z+i}\big)^4\frac{(z-a)(z+\bar{a})}{(z-\bar{a})(z+a)}, \quad P \Leftrightarrow S^1,$$

with $|a| = 1$ and $-\frac{1}{2} <$ Im $a < 0$, has non zero differential on P.

Proof. As we have seen in the proof of Lemma 7.9 the winding number of $R_0(z)$ is two; moreover $R_0(z) = R_0(-1/z)$.

We compute the logarithmic differential[4] :

$$d_z \log(R_0(z)) = \frac{8i}{z^2+1} + (a-\bar{a})\big(\frac{1}{z^2-(a+\bar{a})z+1} + \frac{1}{z^2+(a+\bar{a})z+1}\big).$$

[4]we should compute it as well in a neighborhood of the point at infinity; it should be obvious however that as $R(z^{-1}) = \frac{1}{R(z)}$ and the orientation is reversed then the logarithmic differential in a neighborhood of the point at infinity is equal to the logarithmic differential in a neighborhood of $z = 0$.

Now we set $a = \alpha + i\beta$ with $\beta < 0$ and $\alpha^2 + \beta^2 = 1$ and impose that the logarithmic derivative is non zero, this leads to:

$$8\,\beta^2\,z^2 + 2\left(-1 + z^2\right)^2 + \beta\left(1 + z^2\right)^2 \neq 0$$

which is equivalent to

$$-1 - 2\,\beta + \beta^2 + 2\,\beta^3 < 0.$$

This holds provided that:

$$-\frac{1}{2} < \beta < 0.$$

□

7.2 Computation of the Wronskian matrix

Consider an Hamiltonian of the type 7.1 with $F(q)$ satisfying the condition 5.11 and $f(q,\psi)$ a trigonometric polynomial in ψ and a rational function of e^{iq}.
We can repeat the procedure of the preceding Chapters to evaluate the Melnikov approximation of the splitting matrix and **prove Melnikov dominance for systems with one fast variable.**
We want to be able to repeat all the formal tree expansions and the bounds of section 3.1, to do this we have to compute a solution of the equation:

$$\dot{M} = \begin{vmatrix} 0 & F_q(q(t)) \\ 1 & 0 \end{vmatrix} M \quad \text{where} \;\; q(t) \text{ solves } \dot{q} = \sqrt{2F(q)}\,, \quad q(0) = \pi, \tag{7.8}$$

$M(t)$ is a 2×2 matrix and $M(0) = \mathrm{Id}$.
This is the fundamental solution of the linearized "pendulum" and has the role of the matrix W in subsection 1.1.2. We have to check that $M(t)$ is in $H_0(b,d)$ for some b, d; if this is true one can use the operator $\Im$ defined in subsection 1.1.3 to extend the integration. Then one can re-obtain the equations 1.30 for the perturbative expansion of the whiskers only with different functions x_0^i which nevertheless are $x_0^0 \in H_0(b,d)$ $x_1^0 \in H(b,d)$ and with the same parity properties.

There are classical methods to find the solution of the linearized equation equation 7.8. First we consider the solution $p(E,t), q(E,t)$ of the equations:

$$\begin{cases} \dot{p} = F_q(q) \\ \dot{q} \,=\, p \end{cases} \quad \frac{1}{2}p^2 + F(q) = E,$$

naturally $p(E,t) = \dot{q}(E,t)$ and $q(0,t) = q(t)$.
By simple substitution we see that the couples

$$\dot{p}(0,t)\,,\;\; \dot{q}(0,t) \quad \text{and} \quad \partial_E p(E,t)|_{E=0}\,,\;\; \partial_E q(E,t)|_{E=0}$$

are solutions of 7.8. Let us first consider $\dot q(t) = G(y(z))$; having fixed a dynamics $e^{iq(t)} = y = R(z)$ we know by equation 7.4 that

$$\dot q(t) = -2e^{-t}\sum_{i=1}^{h}\frac{\text{Im }a_i}{(e^{-t}-\text{ Re }a_i)^2+(\text{ Im }a_i)^2}.$$

Moreover by Proposition 7.6 $\dot q(t)$ is even[5] and has poles in $z = a_i, \bar a_i$; $\dot q$ is bounded for $|t| \to \infty$ and $|\text{Im}(t)| \le 2\pi$. So $\dot q$ is in $H_0(0, \bar d)$ with $\bar d = \min_i \arccos(|\text{ Re}(a_i)|)$ We know that $q(t) \neq 0$ for all $t \in \mathbb{R}$ so the vector

$$m_1(t) = \begin{pmatrix} \dot q(t)/\dot q(0) \\ \ddot q(t)/\dot q(0) \end{pmatrix} \text{ satisfyes the condition } m_1(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

To compute $\partial_E q(E,t)$ we derive the energy conservation relation and obtain:

$$\dot q(0,t)\partial_E \dot q(E,t)|_{E=0} - \ddot q(0,t)\partial_E q(E,t)|_{E=0} = 1;$$

by variation of constants we obtain:

$$q_E(t) = \partial_E q(E,t)|_{E=0} = \dot q(0,t)\int_0^t \frac{d\tau}{\dot q^2(0,\tau)} \tag{7.9}$$

which is well defines as $\dot q(t) \neq 0$ for all real t; moreover it is an odd function so $q_E(0) = 0$.Its derivative:

$$\dot q_E(t) = \ddot q(0,t)\int_0^t \frac{d\tau}{\dot q^2(0,\tau)} + \frac{1}{\dot q(t)}$$

is different from zero for $t = 0$; $\dot q_E(0) = \frac{1}{\dot q(0)}$. We notice that $\frac{1}{\dot q(t)}$ is a function in $H_0(b,d)$ for some b,d as it depends only on $z = e^{-t}$. So $q_E(t) \in H_0(b,D)$ for some b,d as the integration can be written as $\int_0^t = \Im^t - \Im^{0^{\sigma(t)}}$ which is closed on $H_0(b,d)$. Naturally $q_E(t)$ will not, in general, be a function only of e^{-t} and it will have non-polar singularities. This means that, to re obtain bounds like those of section3.1 ,one has to prove that $q_E(t)$ respects a condition like 3.2[6]. Naturally as we have seen $q_E(t)$ is not bounded (not even for the standard pendulum). We bound it exactly like we did in section3.1 to bound analytic trees. The term $\Im^0$ is a constant so we ignore it. The even function $Q = \frac{1}{\dot q(t)^2}$ has a double pole at $z = 0$ and at $z \to \infty$; moreover it has poles for finite values of z, coming in conjugated couples that we call $b_l, \bar b_l \notin \mathbb{R}$. If $\dot q$ has $2j$ poles (the $a_i, \bar a_i$) then $l = 1, 2(j-1)$. Naturally we have no guarantee that $|b_l| = 1$ so in general the poles are not purely imaginary when written in the t variable.

[5]and so naturally $\ddot q(t)$ is odd.

[6]remember that the condition 3.2 is satisfied by all rational functions of e^t times polynomial in t.

For $|t| > \max_l |\log|b_l|| = \tilde{b}$ we can write $Q(t)$ as function of $x = e^{-|t|}$(call it $\tilde{Q}(x)$) and expand it in a Laurent series around $x = 0$.

$$\tilde{Q}(x) = \sum_{k=-2}^{\infty} Q_k x^k$$

converges in the annulus $0 < |x| < e^{-\tilde{b}}$. When we apply the formal integration 1.18 to the expansion we obtain a purely polynomial term tQ_0. So for $|\text{Re } t| > \tilde{b}$ and $|\text{Im } t| \leq 2\pi$ the function $q_E(t)$ is

$$(\Im^0[\tilde{Q}(x)] - Q_0 t)\dot{q}(t) + \dot{q}(t) \cdot \Im^t[\tilde{Q}(x) - Q_0];$$

notice that the second summand is function only of x and has a simple pole in $x = 0$.

In the domain $M(\tilde{b}, \tilde{d}) :=,|\text{Re } t| \leq \tilde{b}$ and $|\text{ Im } t| < \tilde{d} = \min_i \arccos(|\text{ Re}(b_i)|)$ we simply bound the integral with the maximum of the integrand and obtain that:

$$\max_{t\in D(\tilde{b},\tilde{d})} q_E(t) \leq 2\tilde{b}\dot{q}(0) \max_{t\in D(\tilde{b},\tilde{d})} \dot{q}(t) \max_{t\in D(\tilde{b},\tilde{d})} \frac{1}{\dot{q}^2(t)}.$$

We have found a matrix M with all the properties of W defined in subsection1.1.2, namely it has the same parity and regularity properties, and the same qualitative asymptotic behavior. So we simply substitute

$$x_0^0(t) = \dot{q}(t)/\dot{q}(0) \quad x_0^1(t) = \sigma(t)\dot{q}(0)q_E(t)$$

in the definitions of the operators Q_j and we can perform all the symbolic tree expansions of Section1.2. We have to prove again proposition 1.16, to ensure the possibility of changing the first node. Then we use the bounds on x_j^i to re-derive the bounds on trees of Section 3.1. The cancellations of Chapter 4 depend only on the parity conditions and on the symmetry of the operators Q_j so they still hold true.

Finally we have to compute the Melnikov integral which implies the same computations of subsection6.1.1,provided that f is a trigonometric polynomial in ψ and rational in e^{iq}.

Example 7.14. Naturally it is pleasant to have an explicit expression for x_0^1 and actually it is not difficult to perform the integral 7.9. If we e consider the functions satisfying Condition 7.4, we can compute the x_0^i explicitly (we have used Mathematica to do the computations).

$$\begin{aligned} x_0^0 =& \frac{\dot{q}(t)}{\dot{q}(0)} = \frac{-2\,(-1+\alpha)\,(1+\alpha)\,e^t\,(1+e^{2\,t})}{1+(2-4\,\alpha^2)\,e^{2\,t}+e^{4\,t}} \\ \sigma(t)x_0^1 =& -\frac{e^{-t}}{4(\alpha^2-1)}\Big[\frac{-1+e^{6\,t}+e^{4\,t}\,(1+8\,\alpha^4+4\,t-16\,\alpha^2\,t)}{1+(2-4\,\alpha^2)\,e^{2\,t}+e^{4\,t})} - \\ & \frac{e^{2t}(1+8\,\alpha^4-4\,t+16\,\alpha^2\,t)}{1+(2-4\,\alpha^2)\,e^{2\,t}+e^{4\,t}}\Big], \end{aligned} \tag{7.10}$$

as usual $\alpha = \text{Re } a$.

7.3 Homoclinic splitting for the generalized pendulum

We show on an example the procedure for proving lower and upper bounds for systems with generalized pendulum. We consider the Hamiltonian:

$$\frac{1}{2}(I_1^2+I_2^2+p^2)+(\cos q-1)^2+\frac{1}{2}\sin^2 q+\eta(\cos(\psi_1)+\cos(\psi_2))\cos(2q),$$

this is a completely anisochronous system with three degrees of freedom. For $\eta=0$ the hyperbolic variables p,q are on a pendulum-like separatrix and precisely the dynamics is the one described in Example 7.7.

$$q(t)=2\arctan(\sqrt{2}e^{-t}+1)+2\arctan(\sqrt{2}e^{-t}-1).$$

For $\eta\neq 0$ we have the perturbative equations:

$$\begin{aligned} \dot I_i^k &= F_i^k\,, & \dot\psi_i^k &= I_i^k \\ \dot p_i^k &= [-\frac{1}{2}\sin(2q_0(t))+2\sin(q_0(t))]q^k+F_0^k\,, & \dot q^k &= p^k. \end{aligned} \tag{7.11}$$

Where as usual we set $\psi_0=q$ and

$$F_i^k=[\partial_{\psi_i}f(\sum_{h<k}(\eta)^h\vec\psi^h)]_{k-1}+\delta_{i0}[-\frac{1}{2}\sin(2\sum_{h<k}(\eta)^h\psi_0^h)+2\sin(\sum_{h<k}(\eta)^h\psi_0^h)]_k.$$

We have computed the Wronskian matrix of such dynamics in the previous subsection, see Equation 7.10

$$\begin{pmatrix} x_0^0 & \sigma(t)x_0^1 \\ \dot x_0^0 & \dot\sigma(t)x_0^1, \end{pmatrix}$$

where

$$x_0^0=\frac{e^t(1+e^{2t})}{1+e^{4t}}\,,\quad x_0^1=\frac{2\ (e^t+e^{3\,t})\ t}{1+e^{4\,t}}+\frac{e^{-t}+3\,e^t-3\,e^{3\,t}-e^{5\,t}}{2\ (1+e^{4\,t})}$$

notice that this matrix has the same parity, analyticity and asymptotic properties as the Wronskian of the linearized pendulum, studied in Subsection 1.1.2, so that the boundedness conditions on the solutions of 7.11 lead to the recursive equations:

$$I_j^k=\Im^t F_j^k\,,\quad \psi_j^k=O_jF_j^k,$$

in the operator O_0 the functions x_0^0 and x_0^1 come from the Wronskian of the generalized pendulum.

Now we consider the spaces $\mathbb{V}(\mathcal{A})$ and $\mathbb{V}(\overset{0}{\mathcal{T}})$ and associate to trees the values $\mathcal{V},\mathcal{W}$ and $\mathcal{V}_1,\mathcal{W}_1$ exactly as in Subsection 2.1.1 and 2.1.2. The only difference is the explicit expression of the functions x_0^i, which is irrelevant to the tree construction.

The upper bounds we derived in Chapter 3 depend only on the degree of the poles of the functions x_0^i and of the q dependent part of the perturbing function, in our case $\cos(2q)$. The x_0^i have simple poles in $e^{i\pi/4+k\pi/2}$ and

$$\cos(2q(t)) = \frac{-16\, e^{2\,t}\,(-1+e^{2\,t})^2}{(1+e^{4\,t})^2}$$

has double poles in $e^{i\pi/4+k\pi/2}$. As we are considering a system with three degrees of freedom we could use the improved bounds of Section 6.2, to do this, however, we should reformulate Proposition 6.12, which depends on the explicit expression of the Wronskian (not only on parity and analyticity properties). This is straightforward but lengthly so we will use the (much worst) bounds of Chapter 3.

Proposition 7.15. *The sum of terms of order higher than one and $\leq \varepsilon^{-\frac{1}{2}}$ in the splitting determinant are bounded from above by*[7]*:*

$$C\sqrt{\varepsilon}^3\left(\frac{\eta}{|\sqrt{\varepsilon}|^{p+7}}\right)^2[e^{-\frac{|\omega_1\pi|4}{\sqrt{\varepsilon}}}],$$

provided that $|\eta| \leq |\sqrt{\varepsilon}|^{p+7}$. p is the degree of the pole of $f(\psi(t), q(t))$ nearest to the real axis, so in this example $p = 2$.

Proof. It is a consequence of the upper bounds of Chapter 3 and of Propositions 6.3 and 6.1. ☐

The Melnikov integral for the splitting matrix is:

$$f_{11} = \int_{-\infty}^{\infty} \frac{16\, e^{2\,t}\,(-1+e^{2\,t})^2}{(1+e^{4\,t})^2}\,\cos(\frac{t\,\omega_1}{\sqrt{\varepsilon}})dt =$$

$$\frac{4\pi}{\sqrt{\varepsilon}}\operatorname{csch}(\frac{\pi\,\omega_1}{2\,\sqrt{\varepsilon}})\left(2\,\sqrt{\varepsilon}\,\sinh(\frac{\pi\,\omega_1}{4\,\sqrt{\varepsilon}}) - \cosh(\frac{\pi\,\omega_1}{4\,\sqrt{\varepsilon}})\,\omega_1\right)$$

which, for ε sufficiently small is dominated by $e^{\frac{\pi\,\omega_1}{4\,\sqrt{\varepsilon}}}$.

$$f_{i,j} = 0\,,\quad \text{for } i \neq j;\quad \text{and } f_{2,2} = \sqrt{\varepsilon}C(\varepsilon,\omega_2) \neq 0,$$

for some order one $C(\varepsilon, \omega_2)$.

Finally the splitting determinant is bounded from below by:

$$Ce^{\frac{\pi\,\omega_1}{4\,\sqrt{\varepsilon}}},\quad \text{if}\quad \eta\varepsilon^{3/2} < \varepsilon^9.$$

[7]we consider a three time scale system so $\tau_S = 0$

Chapter 8

Arnold diffusion

In this chapter we present a brief review of the procedure necessary to prove diffusion of the action variables, once given lower bounds on the splitting determinant. There are essentially three steps:
1) Prove the existence of heteroclinic intersections namely that for $\eta \leq \varepsilon^P$ *and given* $\omega, \omega_0 \in \Omega_\gamma$ *such that*

$$|\omega - \omega_0| \leq F(\varepsilon)$$

there exists $\bar{\phi}(\omega, \omega_0, \eta)$ *such that*[1]

$$I^-_\eta(\bar{\phi}(\omega, \omega_0, \eta), \omega, \rho(\omega)) = I^+_\eta(\bar{\phi}(\omega, \omega_0, \eta), \omega_0, \rho(\omega_0)).$$

Then the point

$$z_\eta(\omega, \omega_0) = I^-_\eta(\bar{\phi}(\omega, \omega_0, \eta), \omega, \rho(\omega)), \bar{\phi}(\omega, \omega_0, \eta), \pi$$

lies in[2]

$$W^-_\eta(\omega, \rho(\omega)) \cap W^+_\eta(\omega_0, \rho(\omega_0)) \cap \{q = \pi\}.$$

2)Compare the maximum distance for ω *and* ω_0 $F(\varepsilon)$ *with the size of the gaps of preserved tori, given by the Normal form theorem discussed in Appendix A.4.*
3) Prove the existence of a trajectory which "shadows" a chain of heteroclinic connections and has order one drift in the actions. We will not give any proof of this third step, but only cite some articles that contain the proofs of our claims.

Let us repeat some definitions (taken from [C]) already cited in the introduction.

Definition 8.1 (Heteroclinic chains). *A heteroclinic chain is a set of* $N \geq 1$ *trajectories* $z^1(t), \ldots, z^N(t)$ *together with* $N+1$ *different minimal sets*[3] $T_0, \ldots, T_N$ *such that for all* $1 \leq i \leq N$

$$\lim_{t \to -\infty} dist\ (z^i(t), T_{i-1}) = 0 = \lim_{t \to \infty} dist\ (z^i(t), T_i).$$

---

[1]See Section 1.1 for the definition of $\rho(\omega)$

[2]see Theorem 1.1.1 for the definitions of $W^{\pm}_\eta(\omega, rho)$

[3]A closed subset of the phase space is called minimal (with respect to a Hamiltonian flow ϕ^t_h) if it is non-empty, invariant for Φ^t_h and contains a dense orbit. In our case the minimal sets will be unstable tori $T(\omega_i)$ with $\omega_i \in \Omega_\gamma$.

Definition 8.2 (Transition chains). *A heteroclinic chain is called a transition chain if for any $r > 0$ there exists a trajectory $z(t)$ and a time $T > 0$ such that*

$$dist\ (z(0), T_0) \le r\,, \quad dist\ (z(T), T_N) \le r\,, \quad \sup_{0\le t\le T} dist\ (z(t), Z) < r$$

where Z is the closure of the union over i of the $\{z^i(t) : t \in \mathbb{R}\}$. The sets T_0 and T_N are said to be connected by as transition chain.

Definition 8.3 (Arnold instability). *Given $E \in \mathbb{R}$ consider an Hamiltonian $h(\varepsilon)$ (with Hamiltonian flow ϕ_h^t) such that $h(0)$ represents an integrable system.*
The system $(\phi_h^t, h^{-1}(E))$ is called Arnold unstable if there exist two positive numbers ε_0 and d_0 such that for all $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$ there exist (closed) invariant sets $T(\varepsilon)$, $T'(\varepsilon) \subset h^{-1}(E)$ satisfying the following conditions:

(i) $T(\varepsilon)$, $T'(\varepsilon)$ are continuous, at $\varepsilon = 0$, in the Hausdorff metric and if Π_I denotes the natural projection over the action variables then

$$\Pi_I T(0) = \{I\}\,, \quad \Pi_I T'(0) = \{I'\}\,, \quad \textit{with} \quad |I' - I| > d_0\,;$$

(ii) for each $0 < |\varepsilon| < \varepsilon_0$ $T((\varepsilon)$, $T'(\varepsilon)$ are connected by a transition chain.

8.1 Heteroclinic chains

In this section we deal with the first two steps of the proof of Arnold instability, namely the construction of heteroclinic chains.

8.1.1 Heteroclinic intersection for systems with one fast frequency

In the following we will consider systems with one fast frequency and in the a-priori stable variables of Hamiltonian (1.10). We can fix $\mu = \varepsilon^P$ and ensure Melnikov dominance, as discussed in the previous Sections. This means that we have lower and upper bounds on the splitting determinant (and on the eigenvalues of the splitting matrix) of the type:

$$a\varepsilon^p e^{-c\varepsilon^{-\frac{1}{2}}} \le \det \Delta^0(\omega) \le b\varepsilon^{-p} e^{-c\varepsilon^{-\frac{1}{2}}}.$$

The coefficients p, a, b, c depend on the perturbing function f.

We consider the function:

$$F(\varphi, \omega_0, \omega) = \tilde{I}_\mu^-(\varphi, \omega, \rho(\omega)) - \tilde{I}_\mu^+(\varphi, \omega_0, \rho(\omega_0)) \equiv c\sqrt{\varepsilon}\big(I_\mu^-(\varphi, \omega, \rho(\omega)) - I_\mu^+(\varphi, \omega_0, \rho(\omega_0))\big)$$

where $\omega, \omega_0 \in \Omega_\gamma$. Notice that

$$F(0, \omega_0, \omega_0) = 0\,, \quad \det \frac{\partial F}{\partial \varphi}(0, \omega_0, \omega_0) = 2^n \varepsilon^{n/2} \det \Delta^0(\omega_0).$$

Hence from the implicit function theorem there exists a function $\varphi(\omega,\omega_0,\varepsilon)$ for which

$$F_\mu(\varphi(\omega,\omega_0,\varepsilon),\omega,\omega_0)\equiv 0,$$

provided $|\omega-\omega_0|$ is small enough. Fixed ω_0 standard computations (see [C]) show that the smallness condition is:

$$|\omega-\omega_0|\le C\varepsilon^{-2p}e^{-2c\varepsilon^{-\frac{1}{2}}}.$$

To prove the existence of heteroclinic intersections we have to prove the existence of a chain of KAM tori at distances of order $B=O_\varepsilon(e^{-C\varepsilon^{-\frac{1}{2}}})$ for some $C>2c$, namely we have to adapt to our anisotropic setting (one fast and many slow time scales) the classical techniques discussed in detail in [C] or [CG].

Proposition 8.4. *There exists a list of Diophantine frequencies $\omega_1,\dots,\omega_h\in\Omega_\gamma$ such that:*

$$(i)\ \sqrt{\varepsilon}|\omega_i-\omega_{i+1}|\le e^{-C_1\varepsilon^{-\frac{1}{2}}}\quad (ii)\ \varepsilon^{-\frac{1}{2}}|\Pi_n(\omega_1-\omega_h)|\sim O_\varepsilon(1), \tag{8.1}$$

where Π_n is the projection on the n-th component. To each of the frequencies ω_i is associated a preserved unstable invariant torus of Hamiltonian 1.10, $\mathcal{T}(\omega_i,\rho_i)$ (with $\rho_i\in[-\frac{1}{2},\frac{1}{2}]$) of frequency $\sqrt{\varepsilon}\rho_i\omega_i$. The scaling factor ρ_i is chosen so that all the invariant tori are on the same energy surface, as explained in Remark??.

To prove the Proposition we proceed in two steps:

1. Define an appropriate set $\bar\Omega$ of Diophantine frequencies respecting condition 8.1.
2. Prove the existence of unstable KAM tori of frequency: $\sqrt{\varepsilon}\rho\omega$ for $\rho\in[-\frac{1}{2},\frac{1}{2}]$ and $\omega\in\bar\Omega$. We will only sketch the proof of this second point.

Definition 8.5. *Given an order one $C_1>2c$, set $A_1=e^{-C_1\varepsilon^{-\frac{1}{2}}}$ and consider the set:*

$$\bar\Omega:=\Big\{\omega\in\Omega:\begin{cases}(a)\quad \sqrt{\varepsilon}|\omega\cdot l|\ge\dfrac{A_1}{|l|^\tau}\quad\forall l\in\mathbb{Z}^n/\{0\}\quad: l_1\neq 0\\ (b)\quad \sqrt{\varepsilon}|\omega\cdot l|\ge\dfrac{\varepsilon^2}{|l|^\tau}\quad\forall l\in\mathbb{Z}^n/\{0\}\quad: l_1=0\end{cases}\Big\}.$$

As there is only one fast time scale the condition $\omega\in\Omega$ can be given only on the slow variables, while the fast variable is obtained by " energy conservation" $\omega\in\Sigma$ (Σ is the ellipsoid of Definition??), namely we consider a function $F:\mathbb{R}^{n-1}\to\Sigma$:

$$F(x):=\{\sqrt{2E-\sum_{i=2}^{n-1}x_i^2-\varepsilon^{-1}x_n^2},\ \ x_2,\dots,x_n\},$$

so that given $\beta=\frac{1}{2}+a$ ($\frac{1}{2}\le\beta\le 1$) and R,r,R_1,r_1,r_2, appropriate order one constants[4] and defining:

$$\tilde\Omega:=\{\tilde\omega\in\mathbb{R}^n:\tilde\omega\varepsilon^{-\frac{1}{2}}\in\Omega\}\ \text{ we have }\ \tilde\Omega=F(B(R,r)\cap M)$$

[4] This conditions automatically imply $\bar r\le\sqrt{\varepsilon}\omega_1\le\bar R$, notice that we are not using the same notation as in **??**, here ω_i is always the i'th component of ω.

where $B(R,r) \subset \mathbb{R}^{n-1}$ is the spherical shell[5] of radiuses $\varepsilon^\beta R, \varepsilon^\beta r$ and

$$M := \{\omega \in \mathbb{R}^{n-1} : \varepsilon r_1 \le \omega_n \le \varepsilon R_1\,,\ \ \omega_i > r_2\varepsilon^\beta\,,\ \ i = 2, \ldots, n-1\}.$$

As we always deal with $\tilde{\omega} = \sqrt{\varepsilon}\omega$ we will omit the tilde rescaling all the relations. The Jacobian of F in $B(R,r) \cap M$ is bounded from above and below by order one constants so that given a measurable set[6] $S \subset \Omega$ $\text{meas}(F^{-1}(S)) \sim \text{meas}(S)$.

Condition (b) naturally defines subsets of $B(R,r) \cap M$, moreover we can project the set respecting condition (a) on the subspace of the slow variables, call this set $\bar{\Omega}_4 \subset B(R,r) \cap M$.
Let us call $S(x)$ the $n-2$ dimensional sphere centered in the origin and of radius $\varepsilon^\beta x$. We take , $2r < R$ and consider $\bar{R}$ so that

$$R_1/2 < \bar{R} < R_1\,,\ \ \frac{r}{R} > \frac{r_1}{\bar{R}} \tag{8.2}$$

Definition 8.6. *Consider the sets*

$$S_2 := \big\{\omega \in S(R) : \varepsilon(R_1 - (R_1 - \bar{R})/4) \le \omega_n \le \varepsilon(\bar{R} + (R_1 - \bar{R})/4),\ \omega_i \ge r_2\varepsilon^\beta\,,\ \forall i \ne n\big\},$$
$$S_3 := \big\{\omega \in S(R) : \varepsilon R_1 \le \omega_n \le \varepsilon\bar{R}\,,\ \omega_i \ge r_2\varepsilon^\beta\,,\ \forall i \ne n\big\}.$$

$M \cap S(R) \supset S_3 \supset S_2$; and the sets all have measure of order $\varepsilon^{(n-3)\beta+1}$.

Given a set $X \in S(R)$ its cone $\mathcal{C}(X)$ is the set of semilines stemming from the origin and reaching points of X. We consider truncated cones $T(X) := \mathcal{C}(X) \cap B(R,r)$, and, for any $r < a < b < R$, $T_{a,b}(X) = T(X) \cap B(b,a)$.
Notice that by 8.2 if $X \in S_3$ then $T(X) \in M \cap B(R,r)$.

Remark 8.7. *Recall that given a measurable set $X \in S(R)$, the cone of X is measurable and meas $T(X) \sim \varepsilon^\beta$ meas (X), meas $T_{a,b}(X) \sim \varepsilon^\beta(b-a)$ meas (X).*

Definition 8.8. *Given $A_2 = e^{-C_2\varepsilon^{-\frac{1}{2}}}$ with $2c < C_2 < C_1$ and for all $s \in \mathbb{R}$, $1 < s < 4R/r$, we consider the sets:*

$$\begin{aligned}\bar{\Omega}_2(s) =& \{\omega \in B(R,r) :\ |\omega\cdot l| \ge \frac{s\varepsilon^2}{|l|^\tau}\ \ \forall l \in \mathbb{Z}^{n-1}/\{0\}\ \ |l| \le A_2^{-1}\},\\ \bar{\Omega}_3(s) =& \{\omega \in B(R,r) :\ |\omega\cdot l| \ge \frac{s\varepsilon^2}{|l|^\tau}\ \ \forall l \in \mathbb{Z}^{n-1}/\{0\}\}\end{aligned}$$

Remark 8.9. *Standard measure theoretic arguments imply that the sets $(\bar{\Omega}_i(s) \cap S(R))^C \cap S(R)$ all have measure of order $\varepsilon^{(n-3)\beta+2}$; this implies as well that $(\bar{\Omega}_i(s) \cap S_2)^C \cap S_2$ has measure of the same order and the same holds for intersections with S_3 and for $(\bar{\Omega}_2(s) \cap \bar{\Omega}_3(s) \cap S_2)^C \cap S_2$. We will repeatedly use such relations.*

---

[5] We call spherical shell of radiuses b, a the $n-1$ dimensional domain $\{x \in \mathbb{R}^{n-1} : a \le |x| \le b\}$.
[6] The symbol $\sim$ means that the two measures are of the same order in ε.

Lemma 8.10. *(i) Given a point $\omega \in \bar{\Omega}_2(2R/r) \cap S_2$ the whole solid ball $B_\rho(\omega)$ of center ω and radius $\rho = \varepsilon^2 A_2^{1+\tau}$ is contained in $\bar{\Omega}_2(R/r)$ and its intersection with $S(R)$ is contained in S_3.*

(ii) The whole truncated cone $T(\bar{\Omega}_2(R/r) \cap S_3)$ is in $\bar{\Omega}_2(1)$, same for $\bar{\Omega}_3$.

Proof. (i) First notice that any $n-2$ dimensional "ball" $B_\rho(x) \cap S(R) \in S_3$ if $x \in S_2$. Now consider $\omega \in \Omega_2(2R/r) \cap S_2$ and a vector $x \in \mathbb{R}^{n-1}$ on the unit sphere:

$$|(\omega + \rho x) \cdot l| \geq ||\omega \cdot l| - |l|\rho| \geq |\omega \cdot l|(|1 - \rho \frac{|l|}{|\omega \cdot l|}|), \quad \text{as} \quad \frac{|l|}{|\omega \cdot l|} \leq \frac{r|l|^{\tau+1}}{2R\varepsilon^2}$$

and $|l| \leq A_2$, setting $\rho = \varepsilon^2 A_2^{1+\tau}$ we have $0 < \rho \frac{|l|}{|\omega \cdot l|}|)\frac{1}{2}$.

(ii) Given a point $x \in \Omega_3(R/r) \cap S(R)$ (or in $x \in \Omega_2(R/r) \cap S(R)$) then $rx/R \in S(r)$ moreover for $r/R \leq t \leq 1$:

$$|tx \cdot l| = t|x \cdot l| \geq r/R \frac{R\varepsilon^2}{r|l|^\tau} = \frac{\varepsilon^2}{|l|^\tau}$$

□

Lemma 8.11. *The set $\bar{\Omega}_2(R/r) \cap S(R)$ is union of a finite number of disjoint convex domains. Each domain is contained in a $n-2$ dimensional "ball" of radius $C_3 \varepsilon^\beta A_2$ for an appropriately fixed order one C_3.*

Proof. $(\bar{\Omega}_2(R/r) \cap S(R)) \equiv$

$$S(R) \bigcap_{\substack{l \in \mathbb{Z}^{n-1} \\ |l| \leq A_2}} \Big(\{x \in \mathbb{R}^{n-1} \,:\, (x \cdot l) > \frac{R\varepsilon^2}{r|l|^\tau}\} \cup \{x \in \mathbb{R}^{n-1} \,:\, (x \cdot l) < -\frac{R\varepsilon^2}{r|l|^\tau}\} \Big),$$

now the intersection of sets such that each connected component is convex has the same property. Suppose, by contradiction, that there are points $x_1, x_2 \in \Omega_2(R/r) \cap S(R)$ such that the arc $\widehat{x_1 x_2}$ is all in $\Omega_2(R/r) \cap S(R)$ and has length grater than $2R^{-1}\sqrt{n}\varepsilon^\beta A_2$. Let $\langle x_1, x_2 \rangle$ be the plane generated by the vectors x_1, x_2, and on it consider the sector $\mathcal{S}$ of unit vectors orthogonal to $\widehat{x_1 x_2}$, this sector has angle $\vartheta = 2\sqrt{n}A_2$. The product space of $\langle x_1, x_2 \rangle^\perp$ with the sector $\mathcal{S}$ is a multi-cylinder in which there cannot be entire vectors $l \in \mathbb{Z}^{n-1}$ with $|l| \leq A_2^{-1}$.
Now we consider the intersection of the multi cylinder with the sphere $|x| = A_2^{-1} - 2\sqrt{n}$, on $\langle x_1, x_2 \rangle$ it is an arc of length greater than $2\sqrt{n}$ so that a ball of radius $\sqrt{n}$ is contained in the multi-cylinder. Now in each ball of radius $\sqrt{n}$ there is at least one entire vector. Namely let x be the center of the ball then $[x]$ (entire part of each component) is entire and $|x - [x]|_\infty \leq 1$. □

Let N be the number of connected domains of $\bar{\Omega}_2(R/r) \cap S(R)$ contained in S_3. Each domain contains an $n-2$ dimensional "ball" of radius $\rho = \varepsilon^2 A_2^{1+\tau}$, so that $N \leq A_2^{-(n-2)(\tau+1)} \varepsilon^{\beta(n-2)-2n+5}$.

Let us now consider the Cantor set $\bar{\Omega}_3(R/r) \cap S_3$, by Remark 8.9 we have that $(\bar{\Omega}_3(R/r) \cap S_3)^C \cap S_3$ has measure of order $\varepsilon^{(n-3)\beta+2}$. This implies that $\bar{\Omega}_3(R/r) \cap S_3 \cap \bar{\Omega}_2(R/r)$ is not empty and the measure of $(\bar{\Omega}_3(R/r) \cap S_3 \cap \bar{\Omega}_2(R/r))^C \cap S_3$ is of order $\varepsilon^{(n-3)\beta+2}$.

Lemma 8.12. *There exists a connected domain D of $\Omega_2(R/r) \cap S_3$ such that*

$$meas\,(D \cap \bar{\Omega}_3(R/r)) \geq A_2^{(n-2)(\tau+1)+1}.$$

Proof. Suppose the assertion to be false, then calling D_i $i = 1, \ldots, N$ the connected domains:

$$\text{meas } S_3 \sim \text{meas } (\bar{\Omega}_2(R/r) \cap S_3 \cap \bar{\Omega}_3(R/r)) = \sum_{i=1}^{N} \text{meas } (D_i \cap \bar{\Omega}_3) \leq A_2^{(n-2)(\tau+1)+1} N$$

which is absurd. □

Then we can use Lemma 8.10 (ii) and consider the truncated cone $T(D) \subset \bar{\Omega}_2(1)$, by Lemma 8.12 $P = T(D) \cap \bar{\Omega}_3(1)$ has measure of order $A_2^{(1+\tau)(n-2)+1}\varepsilon^{\beta}$; namely the the Cantor set P contains all radial segments having an endpoint in $D \cap \bar{\Omega}_3(R/r)$ and the other on $S(r)$.

Consider an $n-1$ dimensional ball of radius $\rho \sim \varepsilon^{\beta} A_2$ centered on a point $x \in D$ and which contains D (such ball exists by Lemma 8.12). Given $h = [\frac{2(R-r)}{3\rho R}]$, consider the points $x_i = t_i x$ with $t_i = 1 - 3/2 i\rho$ $h \geq i \in \mathbb{N}_0$ and let us cover $T(D)$ with a finite number of balls B_i of radius ρ and centered on points x_i.
Setting $\rho = 2C_3\varepsilon^{\beta}A_2$ we have that $B_i \cap B_j$ is empty if $|i-j| > 1$ and each $B_i \cap B_{i+1}$ contains a truncated cone $T_{a_i,b_i}(D)$ with $b_i - a_i \geq \rho/4$. We consider the sets $P_i = T_{a_i,b_i}(D) \cap \Omega_3(1)$, by Lemma 8.12 each P_i has measure of order $\varepsilon^{\beta} A_2^{(1+\tau)(n-2)+2}$.

Now we consider the Cantor set $\bar{\Omega}_4$ whose complementary set in $M \cap B(R, r)$ has measure of order $\varepsilon^{(n-2)\beta+1}A_1$. Its intersection with P_i has measure of order $\varepsilon^{\beta} A_2^{(1+\tau)(n-2)+2}$, provided that $A_1 < A_2^{(\tau+1)(n-2)+3}$. Consider a list $\omega_i \in P_i \cap \bar{\Omega}_4$; for each i we have that $\omega_i, \omega_{i+1} \in B_{i+1}$ so the list respects condition 8.1(i) moreover

$$\min_{y \in B_0} y_n \geq \bar{R} - 2C\varepsilon^{\beta}A_2 \quad \text{and} \quad \max_{y \in B_h} y_n \leq \frac{r}{R}R_1 + 2C\varepsilon^{\beta}A_2$$

for some order one C so the list respects condition 8.1(ii).

In the Appendix A.4 we have proved, generalizing similar results of [GGM3], that there exists a symplectic transformation, well defined in a region W of the phase space $(\tilde{I}, \psi)$, which sends Hamiltonian 1.10 in the local normal form:

$$\frac{1}{2}(J, AJ) + \sqrt{\varepsilon}G_1(PQ, \sqrt{\varepsilon}) + \mu g_1(\phi_S, J, P, Q) + \alpha f_1(\phi, J, P, Q) \tag{8.3}$$

where $\alpha = O_{\varepsilon}(e^{-C\varepsilon^{-\frac{1}{2}}})$ for any order one C. W is of order one in the actions both in the fast direction J_1 and in the degenerate one J_n, namely there exists points $w_1, w_2 \in$

W such that $|\Pi_{J_n}(w_1 - w_2)| = O_\varepsilon(1)$. We can then prove a KAM theorem for the Hamiltonian 8.3 for $\mu < \varepsilon^4$ with the frequencies ω in $\bar{\Omega}$ by choosing $(A_1)^2 \ll \alpha$. Roughly speaking KAM theorems are proved by performing an infinite sequence of symplectic transformations defined in a set of nested domains whose intersection is not trivial. Each approximation step reduces the order of the perturbation quadratically and is well defined provided an appropriate smallness condition is verified. Roughly speaking such condition is of the type:$\mu\gamma^{-2} \ll 1$ where μ is the small parameter and γ is the Diophantine constant of the frequency ω of the preserved torus. To apply this scheme to Hamiltonian 8.3 we first perform a finite number of approximation steps on the slow variables with J_1 as a parameter; the small denominators involved are $|\omega_S \cdot l|$ on which we have the stronger Diophantine condition so that the approximation scheme works provided that $\mu\varepsilon^{-4} \ll 1$. Eventually we will reduce the μ perturbation to order α and then continue with the classical KAM scheme on all the variables now the smallness condition is $\alpha A_1^{-2} \ll 1$.

Remark 8.13. *One could try as well to formulate a quantitative version of the implicit function theorem on rectangular domains like those* **??**. *Actually this is quite straightforward for isochronous systems and using the results of section 5.This would be a first step in proving fast Arnold diffusion for isochronous systems with three time scales, which is treated in detail in [BB2].*

8.2 Transition chains

In the preceding section we have proved that for $\eta \leq \varepsilon^P$ any two tori $T(\omega, \rho(\omega))$, $T(\omega_0, \rho(\omega_0))$ with $\omega, \omega_0 \in \bar{\Omega}_\gamma$ whose distance respects **??** are connected by a heteroclinic intersection. Then any two tori $T(\bar{\omega}, \rho(\bar{\omega}))$, $T(\omega_0, \rho(\omega_0))$ with $\omega_0, \bar{\omega} \in \bar{\Omega}_\gamma$ are connected by a heteroclinic chain composed of

$$n \sim |\bar{\omega} - \omega_0| e^{C\varepsilon^{-\frac{1}{2}}}$$

invariant tori.

Proposition 8.14. *The heteroclinic chain connecting two tori* $T(\bar{\omega}, \rho(\bar{\omega}))$, $T(\omega_0, \rho(\omega_0))$ *with* $\omega_0, \bar{\omega} \in \bar{\Omega}_\gamma$ *is a transition chain.*

The Hamiltonian 1.1.1 having one fast frequency and $\mu \leq \varepsilon^P$ *is Arnold unstable.*

This proposition is the adaptation to Hamiltonian * of [CV]. In particular it follows from the following Proposition of [CV].

Given the heteroclinic chain $\mathcal{T}^i$, z_i we call U_i a neighborhood of $\mathcal{T}^i$ where one can apply the Normal Form Theorem and then Theorem 1.2, we call $W^{s/u}_{i\,\mathrm{loc}}$ the local stable/unstable manifolds and w_i an intersection point

$$w_i \in W^u_i \cap W^s_{i+1} \cap U_i.$$

Finally we denote, for $i = 1, \ldots, N-1$, by Λ^s_i a connected $n+1$ sub-manifold of W^s_{i+1} contained in U_i and intersecting transversally $W^u_{i\,\mathrm{loc}}$ on the energy surface at $w_i \in U_i$.

Proposition 1. of [CV] *Given a neighborhood* B_{i-1} *of some*[7] $\xi_{i-1} \in \Lambda^s_{i-1} \cap (W^u_{i-1\,loc})^c$ *one can find* $\xi_i \in \Lambda^s_i \cap (W^u_{i\,loc})^c$*, a neighborhood* B_i *of* ξ_i *and a time* $T_i > 0$ *such that* $\phi^{-T_i} B_i \subset B_{i-1}$.

[7]The superscript ^{c} denotes the complementary set

Appendix A

A.1 Examples of functions in $B(b,d)$

We give examples of functions $F(e^t)$ having non polar singularities and respecting however condition 3.2 in the domains $C(b, D-\sqrt{\varepsilon})$ where D is the (by hypothesis non polar) singularity nearest to the real axis.
We will not consider the classification of such functions but only prove the existence of a class of functions $f(\psi, q)$ trigonometric in ψ and such that[1] $f(\varphi+\omega t, q(t))$ has a non polar singularity in D respecting condition 3.2. Notice that the only entire functions of q in this class are the trigonometric polynomials.

Let us first state an obvious property of the exponential function (which can be verified by direct computation):

Lemma A.1. *The function $f(z) = e^{\frac{a}{z-z_0}}$ with $a, z, z_0 \in \mathbb{C}$ has an essential singularity in z_0 and is bounded in the region:*

$$Re\ (z-z_0)\ Re\ a + \ Im\ (z-z_0)\ Im\ a = \ Re\ (\bar{a}(z-z_0)) \leq 0.$$

Now let us consider analytic functions $f(\psi, q) = f(\psi)g(q)$ where $f(\psi)$ is a trigonometric polynomial and g is even in q. Then $g(q) = G(\cos(q))$ with $G(x)$ real analytic for $x \in (-1, 1)$ and bounded. We want to find functions $G(x)$ such that $G(\cos(q(t))$ is bounded in some $C(b,d)$.
Let C_1 be $\mathbb{C}$ deprived of the half line Im $z = 0$, Re $z \geq -1$.

Theorem A.2. *For any $z_0 \in C_1$ there exists a function $G(z)$ such that $G(z)$ has singularities only in $z_0, \bar{z}_0$ and $G(\cos(q(t))$ is limited in $C(b, d(z_0)$ for some b.*

To prove the Theorem let us study the map

$$t \to z = \cos(q(t)) = 1 - \frac{2}{\cosh^2(t)}$$

[1] in this Appendix we will restrict our attention to $q(t)$ being the separatrix of the standard pendulum. We do this only to write down simple formulas but naturally the same reasoning hold for any $q(t)$ discussed in Chapter 7.

which is analytic for $t \in \mathbb{R} \times (-i\pi/2, i\pi/2)$. Moreover as the map is even we will consider only the domain $t \in \mathbb{R} \times [0, i\pi/2)$ whose image through $\cos(q(t))$ is $\mathbb{C}$ minus the half line Im $z = 0$ Re $z > 1$.
We will study the curves $z_d(s)$ in $\mathbb{C}$ for fixed $d \in (-i\pi/2, i\pi/2)$ which are the image of the lines $t = s + id$ with $s \in \mathbb{R}$.
The following statements can be easily verified by direct computation.

Lemma A.3. *For each $z_0 \in C_1$ there exists a unique $\pi/2 > d(z_0) \in \mathbb{R}^+$ such that the curve*

$$z(s, z_0) \equiv z_{d(z_0)}(s),$$

passes through z_0.
The $z_d(s)$ are all closed curves whose curvature is different from zero for all $s \in \mathbb{R}$.
The curves $z_d(s)$ are all symmetric with respect to the real axis and $z_d(s) = \bar{z}_d(-s)$.

For all $z_0 \in C_1$ such that Im $z_0 \geq 0$ let $L(z_0)$ $L(\bar{z}_0)$ be the lines tangent to the curve $z(s, z_0)$ in the points $z, \bar{z}$. The symmetry of the $z_d(s)$ implies that the equations of $L(z_0)$ $L(\bar{z}_0)$ are respectively:

$$Re(\alpha(z - z_0)) = 0\,, \quad Re(\bar{\alpha}(z - z_0)) = 0.$$

Moreover let $z^{\pm}(s, z_0)$ be the intersections of the curve respectively with the half planes Im $z \geq 0$ and Im $z \leq 0$.
Standard considerations on smooth curves with non zero curvature ensure that the following Lemma holds.

Lemma A.4. *The curve $z^+(s, z_0)$ (resp. $z^+(s, z_0)$) and a ball[2] $B_r(1)$ with r sufficiently small, are both all one one side of $L(z_0)$ (resp. $L(\bar{z}_0)$) and touch the line only in $z = z_0$ (resp. $z = \bar{z}_0$).*

Proof of Theorem A.2. Given $z_0 \in C_1$ let us suppose that

$$z^+(s, z_0) \subset \text{ Re } ((\alpha(z - z_0)) \leq 0\,, \;\; z^-(s, z_0) \subset \text{ Re } ((\bar{\alpha}(z - \bar{z}_0)) \geq 0,$$

This implies that the function:

$$G(z) = e^{\frac{\bar{\alpha}}{(z - z_0)}} + e^{\frac{\alpha}{(z - \bar{z}_0)}}, \tag{A.1}$$

is bounded inside $z(s, z_0)$. Moreover for any $x \in \mathbb{R}$ $G(z)$ is real. Clearly if $z_0 \in C_1$ is such that

$$z^+(s, z_0) \subset \text{ Re } ((\alpha(z - z_0)) \geq 0\,, \;\; z^-(s, z_0) \subset \text{ Re } ((\bar{\alpha}(z - \bar{z}_0)) \leq 0,$$

we will choose

$$G(z) = e^{-\frac{\bar{\alpha}}{(z - z_0)}} + e^{-\frac{\alpha}{(z - \bar{z}_0)}}.$$

□

[2] $B_r(z_0)$ is the ball of radius r centered in z_0.

Theorem A.2 shows that the condition $f \in B(a,d)$ does not imply that f is rational if $d \neq \pi/2$ on the other hand if $d = \pi/2$ then the following Proposition holds:

Proposition A.5. *Consider an analytic function $f(q)$ $(q \in \mathbb{T})$ such that $f(q(t)) \in B(a, \pi/2)$ and $f(qt)$ has isolated singularities. $f(q)$ is a rational function of e^{iq}.*

The image of Im$t = \pi/2$ through $t \to z = e^{iq(t)}$ is the half line Re $z \geq 0$, Im $z = 0$ and in general the image of $t + id$ with $t \in \mathbb{R}$ is plotted in Figures A.1and A.2.

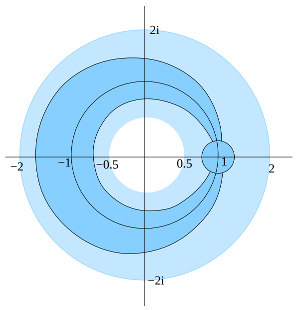


Figure A.1:

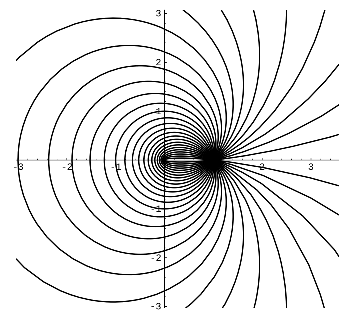


Figure A.2:

So the Proposition is equivalent to the following:

Proposition A.6. *Consider a single valued function $g(z)$ with $z \in \mathbb{C}$ analytic in $B/\{0\}$ where B is some ball centered in zero. Moreover suppose that for some (non zero) k:*

$$|g(z)||\ Imz|^k \leq C$$

in $B/\{0\}$ then $g(z)$ cannot have an essential singularity in $z = 0$.

Proof. **The proof of this Proposition is due to Prof. D'Ancona**

The proposed bounds have polar growth in the sectors $|Imz| \geq |Re(z)|$, now if we integrate $g(z)$ $k+1$ times and call $G_{-}(z)$ the $k+1$ primitive obtained by cutting away Re $t = 0$, Im $t \leq 0$ and starting from a point z_0 with Im $z_0 = 0$ Re z_0 far from the origin (but still in B)

$$|G(z)| \leq C'.$$

To prove this let us perform one integration on the path proposed in FigureA.3

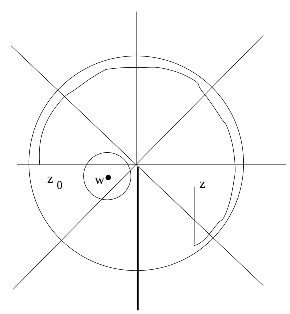


Figure A.3:

The part that is not close to the singularity is bounded by some constant while the integral on the line parallel to the imaginary axis gives the bound

$$|\int g(z)| \leq (\ \text{Im}\, z)^{k-1}.$$

The primitive G_+ obtained by choosing the cut Re $t = 0$, Im $t \geq 0$ differs from G_- by a polynomial of degree k. Now given a point w in $|Imz| < |Re(z)|$ we consider a circle $\mathcal{C}$ of radius $|w|/2$ centered in w; the circle does not intersect the imaginary axis so:

$$|f(w)| = |\int_{\mathcal{C}} \frac{G(z)}{(w-z)^{k+1}}| \leq \frac{C_1}{|w|^{k+1}},$$

independently of the chosen primitive.

□

A.2 Evaluation of the coefficients $T(k), N(k)$

We will prove that:

$$\begin{aligned} T_j(k) &= \sum_{A \in \mathcal{A}_j^k} \frac{1}{|\mathcal{S}(A)|} \leq (4n)^k; \\ N_j(k) &= \sum_{A \in \mathcal{A}_j^k} \frac{\prod_{v \in A:\, \delta_v = 1} n_v!}{|\mathcal{S}(A)|} \leq (4n)^k \end{aligned} \tag{A.2}$$

where n_v is the number of nodes in the list $v, s(v)$ that have label $j = 0$.
This are standard computations on trees and can be found for instance in [Bo] for trees without grammar. However here we present an easy and self contained proof of this statement for our set $\mathcal{A}_j^k$. We will rely some adaptations of the results of Section 1.2 which we will not prove again.
Given a real parameter α and two real analytic functions $f(x, y), g(y)$ with $x \in \mathbb{R}^n$ and $y \in \mathbb{R}$, such that

$$g(0) = 0 \;\; \nabla f(0) = 1, \;\; d_y g(0) = 1,$$

consider the equations:

$$x_i = \alpha \partial_{x_i} f(x, y) \;\; \text{for} \;\; i = 1, \ldots, n \;\; g(y) = \alpha \partial_y f(x, y). \tag{A.3}$$

This relation is invertible in some $|\alpha| \leq \alpha_0$ where the solution $x(\alpha), y(\alpha)$ is analytic in α.

We determine the series expansion

$$x_j(\alpha) = \sum_{h=1}^{\infty} x_j^{(h)} \alpha^h \quad \text{where} \;\; x_0 \equiv y,$$

recursively like in Subsection 1.1.4:

$$x_j^k = F_j^k \quad \text{with} \; F_j^k = [\partial_{x_j} \partial_{x_j} f(\sum_{h=1}^{k-1} \alpha^h x^{(h)})]_{k-1} - \delta_{j0}[g(\sum_{h=1}^{k-1} \alpha^h x_0^{(h)})]_k.$$

The theory we have developed in Chapter 1 implies that the series expansion can be represented by labeled trees and precisely:

$$x_j^k = \sum_{A \in \mathcal{A}_j^k} \frac{1}{|\mathcal{S}(A)|} \Phi(A),$$

where the value of a tree $\Phi(A)$ is:

$$\prod_{v \in A} \big(\prod_{v' \in s(v)} \partial_{j_{v'}} \big) (\partial_{j_v})^{\delta_v} f^{\delta_v} |_{x=0,y=0};$$

with $f^1 = f$ and $f^0 = -g$. Notice that f^0 appears only through its derivatives of order greater or equal than two.
To bound $T_j(k)$ we choose

$$f(x,y) = e^{\sum_{i=1}^n x_i + y}, \quad g(y) = 1 + 2y - e^y,$$

so that the value of any tree is one and $T_j(k) = x_j^k$. Now A.2 can be computed by estimates on the Taylor coefficients of $x(\alpha)$. An easy direct computation shows that relations A.3 can be inverted for

$$|\alpha| \le \frac{1}{4n} \quad \text{which implies A.2.}$$

To bound $N_j(k)$ we choose :

$$f(x,y) = \frac{e^{\sum_{i=1}^n x_i}}{1-y}, \quad g(y) = 1 + 2y - e^y.$$

Again easy computations show that the relations A.3 can be inverted for $|\alpha| \le \frac{1}{4n}$.

A.3 Notions on lattices in $\mathbb{Z}^n$

We briefly review some useful properties of lattices in $\mathbb{Z}^n$.
Let $v_1, \cdots, v_h$ be vectors in $\mathbb{Z}^n$; we will call V the $n \times h$ matrix whose columns are the v_j's and $K(V)$ the lattice spanned by the v_j's with coefficients in $\mathbb{Z}$.
Two vectors v_i $v_j \in \mathbb{Z}^n$ are independent if $av_i + bv_j \neq 0$ for all $a, b \in \mathbb{Z}^n$; clearly in any list of vectors in $\mathbb{Z}^n$ there are at most n independent ones, our vectors v_i will not be, in general, independent.
Nevertheless, for all $w \in K(V)$ we define the "coordinate sets":

$$A_w = \{a \in \mathbb{Z}^h : Va = w\}$$

these are the cosets of $\mathbb{Z}^h$ modulo the relations between the v_i.

On $K(V)$ we introduce the function[3]

$$|w| = \min_{a \in A_w} \sum_{i=1}^{h} |a_i|.$$

We will then use as coordinates of w any vector $a \in A_w$ that realizes the minimum of $|w|$.

Definition A.7. *The n positive numbers $d_1, \cdots, d_n$ such that*[4]

$$d_j = \ Mcd\ (v_{1\,j}, \cdots, v_{h\,j})$$

are the divisors of the lines of V.

Notice that the lattice $K(V) \subset \mathbb{Z}^n$ even if $h > n$.

Lemma A.8. *$K(V) \equiv \mathbb{Z}^n$ if and only if the determinants of the $n \times n$ minors are coprime. Given $w \in K(V)$ $w_j \geq d_j$ for each $j = 1, \ldots, n$; moreover for $h > 2$ for each j there exist infinite vectors $w \in K(V)$ such that $w_j = d_j$. Let us call*

$$W_j = \{w \in K(V) \ : \ w_j = d_j\}.$$

Proof. The first assertion is obvious: it is sufficient that one divisor $d_i \neq 1$ and the i coordinates of all the vectors in $K(V)$ are divided by d_i so that e_i is not in $K(V)$. The second assertion is a standard theorem on lattices in $\mathbb{Z}^n$, it is not immediate so we will not prove it. The third is almost the definition of Mcd: any linear combination (in $\mathbb{Z}$) of numbers all having a common divisor has the same divisor. If we consider k coprime numbers then there will be at least two having Mcd equal to one. This implies that there exists a unique linear combination of this two numbers that gives one. The required linear combination of all the numbers is obtained by adding any linear combination of all the numbers that gives zero. □

Definition A.9. *For each $j = 1, \ldots, n$ we define the projection m_j of $K(V)$ in the direction j as*

$$m_j = \min_{w \in W_j} |w|.$$

Let us now consider the lines of the matrix V (call them v^t)

Proposition A.10. *For each $j = 1, \ldots, n$ there exist vectors $u \in K(V)$ such that $u_j = 0$ and $u_i \neq 0$ for all the i such that $v_i^t \neq \alpha v_j^t$ with $\alpha \in \mathbb{Q}$. We call the set of such vectors U_j*

[3]It is easily seen that this is a well defined norm on $K(V)$.
[4]Mcd is the maximum common divisor.

Proof. Consider the sub-lattice orthogonal to v_j^t:

$$U_j^1 = \{y \,:\, v_j^t y = 0\},$$

now suppose that there exists v_i^t orthogonal to U_j^1 and which is not parallel to v_j^t in $\mathbb{Q}$. This is a contradiction as $v^{\perp\perp} = v$. □

Definition A.11. *For each $j = 1, \ldots, n$ we define the projection r_j of $K(V)$ orthogonal to the direction j as*

$$r_j = \min_{u \in U_j} |u|.$$

If we have a perturbating function with frequencies $\pm\nu_1, \cdots \pm \nu_k$, the lattice $K(V)$ where $v_i = \nu_i$, gives all the possible frequencies reached in the perturbation series. We are interested in how the various possible frequencies are reached and particularly at what order of the perturbation. Consider the following discrete-time dynamical system on $\mathbb{Z}^n$:

- At time one we have the list of vectors $V(1) \equiv \{v_i(1)\} = \{\pm\nu_i\}$
- At time l we have the list of vectors $V(l) \equiv \{v_i(l)\} = \sum_{j=1,l} v_{i_j}(1)$ (sum of l vectors of $\{v_i(1)\}$).

The vectors $\{v_i(l)\}$ are the possible values of the total frequency $(\sum_v \nu_v)$of a tree of order l.

This dynamical system never enters inside the rectangle centered in zero of length $2d_i$ in each direction i; nevertheless it touches each side of the rectangle infinite times.

Lemma A.12. *Let $\bar{l}(j)$ be the first time such that one of the vectors $\{v_i(l)\}$ has the j component equal to the divisor d_j and $\bar{o}(j)$ be the first time such that*

$$V(o(j)) \cap U_j \neq \{0\},$$

then , $\bar{l}(j) = m_j$ and $\bar{o}(j) = r_j$.

Proof. The two proofs are identical so we will consider only $\bar{l}(j)$.
The time $\bar{l}(j)$ exists and is finite, so consider the elements $\bar{v}$ of $V(\bar{l}(j))$ that have the j component equal to the divisor d_j.
$\bar{l}(j)$ is minimal, so if the sum expressing $\bar{v}$ contains ν_i then it does not contain $-\nu_i$ and vice-versa. This means that:

$$\bar{v} = \sum_{j=1,\bar{l}} v_{i_j}(1) = \sum_{i=1}^{k} k_i \nu_i \text{ with } k_i \in \mathbb{Z} \text{ and } \sum_i |k_i| = \bar{l}(j).$$

Now the vector $k \in \mathbb{Z}^h$ is in $A_{\bar{v}}$, if there existed $k' \in A_{\bar{v}}$ such that $\sum_i |h_i| < \bar{l}(j)$ this would contradict the minimality of $\bar{l}(j)$.

□

A.4 Normal form theorem

To obtain bounds on the convergence radius μ_0 of the KAM theorem 1.1 we perform a symplectic change of variables that brings Hamiltonian (*) in local "normal form". We will use the standard notations (see [Pö], [BG], or [CG], [GGM1]) and the existence of the fast time scale. For systems with one fast time scale this provides a symplectic change of variables defined in a region W such that $\Pi_I W = O_\varepsilon(1)$, that sends the perturbating terms depending on the fast angle to order $e^{-\frac{1}{\varepsilon^B}}$ for some $B(n) < 1$. This will be the basis for proving Arnold diffusion for systems with one fast variable. For completeness we state the theorem for m fast variables. The first step is to set the pendulum in local hyperbolic normal form (see [CG]), we obtain the local Hamiltonian:

$$\frac{1}{2}(I, AI) + \sqrt{\varepsilon}G(pq, \sqrt{\varepsilon}) + \mu f(p, q, \psi), \tag{A.4}$$

where the function $G(J, \sqrt{\varepsilon})$ is analytic for $|J| < \tilde{k}_0^2 \sim \sqrt{\varepsilon}$ and will be written as Taylor series: $G(J) = \sum_{k\geq 1} J^k G_k$.
The perturbating term $f(p, q, \psi)$ is a trigonometric polynomial of degree N in the rotator angles and an analytic function of $p, q \leq k_0$. So we consider the domain:

$$W(k_0, s_0) \equiv W_0 := \{|p|, |q| \leq k_0,\ I \in V_0(\varepsilon) \subset \mathbb{C}^n \psi \in \mathbb{T}^n \times (-is_0, is_0)\},$$

where $V_0(\varepsilon)$ is some n-rectangle contained in $D(\alpha, \delta)$ (i.e. such that $\Pi_{I_j} V_0(\varepsilon) = O(\frac{\omega_j}{a_j})$, see Chapter 1 for the definition of the sets $D(\alpha, \delta)$).
We write f in Taylor series:

$$f(p, q, \psi) = \sum f_{\nu,k,h} p^k q^h e^{i\nu\cdot\psi}.$$

For all $s < s_0$, $k < k_0$ we use the weighted norm:

$$|f|_{k,s} \equiv |f|_{W(k,s)} = \sum e^{s|\nu|} |f_{\nu,l,h}| k^{2(l+h)} e^{i\nu\cdot\psi}.$$

Definition A.13. *Given a sub-lattice $\Lambda \in \mathbb{Z}^n$ and a point set $D \in V_0(\varepsilon)$ we say that D is $K - \beta$ non-resonant modulo Λ if for all $I \in D$:*

$$|\omega(I) \cdot \nu| \geq \beta\ \ \forall \nu\ : \nu \notin \Lambda\ \cap\ |\nu| \leq K.$$

If Λ_0 is the lattice generated by the N frequencies ($\nu_i \in \mathbb{Z}^n$) of f, we set $\Lambda \in \Lambda_0$ to be the sub-lattice orthogonal to the fast components.
We choose a point set D in the following manner:
let P be the set of vectors $\omega \in \Omega$ (see Section 1.1) such that $|\omega_1 \cdot \nu_F| \geq \frac{\gamma}{|\nu_F|^{\tau_F}}$ for an order one γ.
Given $r_0 \in \mathbb{R}^+$, the domain $D(r_0)$ is a thickening of P such that $\forall I \in D(r_0)$ there exists $\omega \in P$ such that :

$$|AI - \omega| \leq \varepsilon^{\alpha+\frac{1}{2}} r_0$$

for $r_0 < R$; in the following we will set $b = \frac{1}{2} + \alpha$.

Lemma A.14. *$D_0 \equiv D(r_0)$ is $\beta - K$ non-resonant modulo Λ with*

$$K = \big(\frac{\gamma}{4R}\varepsilon^{-b}\big)^{\frac{1}{1+\tau_F}}, \quad \beta = (\gamma)^{\frac{1}{1+\tau_F}}(4R\varepsilon^b)^{\frac{\tau_F}{1+\tau_F}}.$$

Proof. Given $I \in D(r_0)$ $\omega(I) = AI$ is $\varepsilon^b r_0$-close to an $\omega \in P$ so

$$|\omega(I)\cdot\nu| \geq |\omega_1 \cdot \nu_F| - (\varepsilon^b|\omega_2||\nu| + \varepsilon^b r_0|\nu|)$$

with $r < |\omega_2| < R$ so we set:

$$\varepsilon^b|\omega_2|\gamma^{-1}|\nu|^{\tau_F+1}, \varepsilon^b r_0|\nu|\gamma^{-1}|\nu|^{\tau_F+1} < \frac{1}{4}.$$

□

We construct an analytic symplectic transformation (μ-close to identity) of the form:

$$Id + \mu S(I', p', \psi, q) = Id + \sum_{1<l\leq\frac{K}{N}} \mu^l \sum_{\nu\neq\Lambda}^{|\nu|\leq lN} S^{(l)}_{\nu,k,h} p'^k q^h e^{i\nu\cdot\psi},$$

that brings the Hamiltonian A.4 in the normal form[5]

$$(I', AI') + \sqrt{\varepsilon}G_1(pq, \sqrt{\varepsilon}) + \mu g_1(\psi'_S, I', p', q', \varepsilon, \mu) + \mu^{\frac{K}{N}} f_1(\psi', I', \varepsilon, \mu),$$

in a suitable domain $D'(r_1) \times \mathbb{T}^n_{s_1} \times B^2_{k_1}$, where

$$D'(r) = D(r) \cap \{I: \ \exists\omega \in P \ \text{ such that } \ |a_j I - j - \omega_j| \leq r_0 \varepsilon^{\delta_j - \alpha_j}\}.$$

The Hamilton-Jacobi equations are:

$$\begin{aligned} &\mu AI' \cdot S_\psi + \tfrac{1}{2}\mu^2 |AS_\psi|^2 + \sqrt{\varepsilon}G(qp' + \mu q S_q)) = \sqrt{\varepsilon}G_1(p'q + p'S_{p'}, \mu) + \\ &\mu g_1(\psi_S + \mu S_{I'}, I', p', q + \mu S_{p'}, \varepsilon, \mu) - \mu f(p' + S_q, q, \psi) + o(\mu^K) \end{aligned} \tag{A.5}$$

and we assume that we can find some domain $D'(r) \times \mathbb{T}^n_s \times B^2_k$ such that the functions in A.5 are evaluated inside their domain of analyticity. We will call Π_Λ the natural projection on functions NOT depending on the fast angles: $\Pi_\Lambda f(\psi, p, q) = g(\psi_S, p, q)$ and Π_J the natural projection on functions depending only on $J = pq$:

$$F = \sum F_{\nu,k,h} p^k q^h e^{i\nu\cdot\Phi} \ \ \Pi_J F = \sum F_{0,h,h}(pq)^h.$$

[5]The separation between the integrable G_1 and the non integrable g_1 is kept only because we will eventually set up a KAM scheme for the slow variables, so we need to estimate the size of the integrable part.

We are looking for a symplectic transformation such that $(\Pi_\Lambda)S = 0$, we will solve the Hamilton-Jacobi equations recursively and determine the functions $G_1(J,\mu) = \sum_{i\geq 0}\mu^i G_1(J;i)$ and $\mu g_1(\psi_S, I, p, q, \mu) = \sum_{i\geq 1}\mu^i g_1(\psi_S, I, p, q; i)$. The first order leads to[6]:

$$G_1(J,0) = G(J)\,, \quad G_1(J,1) = \frac{1}{\sqrt{\varepsilon}}\Pi_J f\,, \quad g_1(\psi_{1S}, I', p', q', 1) = (\Pi_\Lambda - \Pi_J)f\,,$$

$$S^{(1)}_{\nu,k,h} = -\frac{f_{\nu,k,h}}{i[I'\cdot\nu] + (k-h)\sqrt{\varepsilon}G_J(p'q)}.$$

The term $i[I'\cdot\nu] + (k-h)\sqrt{\varepsilon}G_J(0) = D(\nu,k,h)$ is the "small denominator" that in our case (i.e. up to order $\frac{K}{N}$) admits the lower bound $D(\nu,k,h) \geq \beta$ provided that $I' \in D'(r_0)$. The higher order terms are determined recursively; we set $\mu S^{<l} = \sum_{h=1}^{l-1}\mu^h S^{(h)}$ and $[f(\mu)]_l = \frac{1}{l!}\partial^l_\mu f|_{\mu=0}$.

$$G_1(J,l) = \tfrac{1}{\sqrt{\varepsilon}}\Pi_J[(\mu^2\tfrac{1}{2}|AS^{<l}_\psi|^2 + \sqrt{\varepsilon}G(qp' + \mu q S^{<l}_q)) - \sqrt{\varepsilon}G_1(p'q + p'S^{<l}_{p'}, \mu) -$$

$$-\mu g_1(\psi_S + \mu S^{<l}_{I'}, I', p, q + \mu S^{<l}_{p'}, \varepsilon, \mu) + \mu f(p' + S^{<l}_q, q, \psi)]_l)$$

the remaining resonant terms are in $\mu g_1 = \sum_{m=1}^{\infty}\mu^m g_1(\psi_s, I', p', q; m)$:

$$g_1(\psi_S, I', p', q; l) = (\Pi_\Lambda - \Pi_J)[(\tfrac{1}{2}\mu^2|AS^{<l}_\psi|^2 + \sqrt{\varepsilon}G(qp' + \mu q S^{<l}_q)) - \sqrt{\varepsilon}G_1(p'q + p'S^{<l}_{p'}, \mu) -$$

$$-\mu g_1(\psi_S + \mu S^{<l}_{I'}, I', p, q + \mu S^{<l}_{p'}, \varepsilon, \mu) + \mu f(p' + S^{<l}_q, q, \psi)]_l)$$

the terms of order μ^l and such that $\nu \neq \Lambda$ fix the value of $S^{(l)}_{\nu,k,h}$. We expand the Taylor series only in this expression. The symbol $\{k_i\}^r_k$ means the set of vectors in $\mathbb{N}^r$ such that $\sum_{i=1}^r k_i = k$, while $\{\nu_i\}^r_\nu$ is the set of r vectors in $\mathbb{Z}^n$ such that $\sum_{i=1}^r \nu_i = \nu$.

$$S^{(l)}_{\nu,k,h} = -\frac{1}{D(\nu,k,h)}[\sum_{\nu^{(1)}+\nu^{(2)}=\nu}\frac{1}{2}S^{(m)}_{\nu^{(1)},k_1,h_1}S^{(l-m)}_{\nu^{(2)},k-k_1,h-h_1}(\nu^{(1)}, A\nu^{(2)}) +$$

$$+\sqrt{\varepsilon}\sum_{r\geq 2}^{l}\sum_{\substack{\{k_i\}^r_k,\{h_i\}^r_h,\\ \{l_i\}^r_l,\{\nu_i\}^r_\nu}}(\frac{1}{r!}\partial^r_J G(p'q)\Pi^r_{i=1}S^{(l_i)}_{\nu_i,k_i,h_i}h_i +$$

$$\sum_{r\geq 1}\sum_{\substack{\{k_i\}^r_{k+r},\{h_i\}^r_h,\\ \{l_i\}^r_{l-1},\{\nu_i\}^r_{\nu_1}}}\frac{1}{r!}\partial^r_{p'}f(p',q,\Psi)\Pi^r_{i=1}k_i S^{(l_i)}_{\nu_i,k_i,h_i} -$$

$$\sqrt{\varepsilon}\sum_{m=0}^{l-2}\sum_{r\geq 2}^{l-m}\sum_{\substack{\{k_i\}^r_k,\{h_i\}^r_h,\\ \{l_i\}^r_{l-m},\{\nu_i\}^r_\nu}}(\frac{1}{r!}\partial^r_J G_1(p'q;m)\Pi^r_{i=1}S^{(l_i)}_{\nu_i,k_i,h_i}k_i -$$

[6]Notice that the pendulum and rotator terms cannot cancel each other, this is a consequence of the locality of our analysis.

$$\sum_{m=1}^{l-1} \sum_{\substack{k_a+k_b=k,\\ h_a+h_b=h}} \sum_{\substack{l_a+l_b=l-m,\\ \nu_a+\nu_b=\nu}} \sum_{r\geq 0}^{l-m} \sum_{\substack{s\geq 0,\\ r+s\geq 1}}^{l-m-r} \sum_{\substack{\{k_i\}^r_{k_a+r},\{h_i\}^r_{h_a},\\ \{l_i\}^r_{l_a},\{\nu_i\}^r_{\nu_a}}} \sum_{\substack{\{k_j\}^s_{k_b},\{h_j\}^s_{h_b},\\ \{l_i\}^s_{l_b},\{\nu_i\}^s_{\nu_b}}}$$

$$(\frac{1}{r!s!}\partial_q^r \partial_{\psi_S}^s g_1(\psi_S, I', p', q; m)\Pi_{i=1}^r k_i (S^{(l_i)}_{\nu_i,k_i,h_i} \Pi_{j=1}^s \nabla_I S^{(l_i)}_{\nu_j,k_j,h_j}$$

To avoid proliferation of symbols we will set:
$\max(|f|_0, |G|_0) = E_0$ and choose $r_0 > 1$ so that $r_0\varepsilon^{\delta_i-\alpha_i} \geq r_0\varepsilon \equiv \lambda_0 > k_0^2$. Finally we will call $b_j = \max(b, \delta_j - \alpha_j)$.

Proposition A.15. *Consider the nested domains:* $D_l \equiv D'(r_l) \times \mathbb{T}^n_{s_l} \times B^2_{k_l}$ *where* $r_l = \frac{1}{2}r_0 e^{-l\xi}$, $s_l = s_0(1 + l\xi)$ *and* $k_l = \frac{1}{2}k_0 e^{-l\xi}$*; the following bounds hold*[7]*:*

$$|S^{(l)}_{\nu,k,h}|_l \leq C_1(l-1)!B^{l-1} \quad |G_1(J,l)|_l \leq C_2(l-1)!B^{l-1}$$

$$|g_1(\psi_S, I', p', q; l)|_l \leq C_3(l-1)!B^{l-1}$$

with $C_1 = \frac{E_0}{\beta}$, $C_2 = C_3 = E_0$ *and* $B = c\frac{E_0^2}{\beta^2 k_0^4 \xi^2}$ *for some small enough order one c.*

Moreover the so defined transformation is a biholomorphism: $D_K \to D_0$ *provided that* $\xi = \frac{s_0}{4K}$, $\mu BK < 1$*. Thus the system can be written in normal form for*

$$\mu < \frac{\beta^2 k_0^4 \xi^2}{K^3} \tag{A.6}$$

in the domain $D(r) \times T^n_s \times B^2_k$*, with* $r = \frac{1}{2}r_0 e^{-s_0/4}$,$k = \frac{1}{2}k_0 e^{-s_0/4}$ s=$s_0/4$.

Remark A.16. *Notice that for systems with one fast time scale the domain P coincides with the whole $W(k, s_0/2)$ as all one dimensional vectors of norm one are diofantine with order one γ. Moreover in this case $\beta = O(1)$ as well so if we choose $K = \frac{c}{\sqrt{\varepsilon}}$, the bound on μ is $\mu \leq \varepsilon^{\frac{5}{2}}$.*

Remark A.17. *Notice that if we choose $K = O_\varepsilon(1)$ we can perform some steps of the normal form theorem for $\mu < \varepsilon$ so for order one $\eta = \mu/\varepsilon$.*

Proof. We proceed by induction, using the analyticity assumptions on G and f.

We will assume that the desired bounds hold for all $l < m$ and that $G_1(J, l)$ and $g_1(\psi_S, I', p', q, l)$ are analytic in D_{m-1}. This implies that the transformation

$$\begin{aligned} I = I' + \mu S^{<m}_\psi\,, &\quad \psi' = \psi + \mu S^{<m}_{I'}\,,\\ p = p' + \mu S^{<m}_q\,, &\quad q' = q + \mu S^{<m}_{p'} \end{aligned}$$

is well defined and $D_m \to D_0$ if

$$\max(|\mu S^{<m}_q|_m, \mu S^{<m}_{p'}|_m) \leq \frac{1}{4}k_m\,, \quad |\mu S^{<m}_{\psi_j}|_m \leq \frac{1}{4}r_m\varepsilon^{b_j}\,,$$
$$|\mu S^{<m}_{I'}|_s \leq \frac{1}{4}s_0\,, \quad |\mu S^{<m}_{\psi,I'}|_m < 1.$$

[7] By $|f|_l$ we mean $|f|_{D_l}$.

Substituting the bounds in this inequalities (and using Cauchy estimates for the derivatives) we obtain the constraint $\mu \max(\frac{8C_1}{k_0^2\xi}, \frac{8C_1}{\lambda_0\xi^2}) < 1$ provided that $\mu KB \leq \frac{1}{2}$. Having verified the analyticity of the transformation up to order m we use analytic bounds on G, G_1 and g_1 and the assumed bounds on the lower orders to bound $G_1(J;m)$ $S^{(m)}$ and $g_1(\psi_S, I', p', q; m)$. We repeatedly use the inequality:

$$\sum_{\{k_i\geq 1\}_{i=1}^{a}:\sum_i k_i=k} \prod_{i=1}^{a}(k_i-1)! \leq (k-1)!.$$

Let us first consider $S^{(m)}$, it is composed of five sums. In each we substitute the Cauchy estimates and the bounds coming from the inductive hypothesis.

(1) The sum of quadratic terms is bounded by $(k-1)!B^{k-1}\frac{C_1^2}{s_0^2\xi^2\beta B}$.

(2) The terms due to G are bounded by:

$$\frac{\sqrt{\varepsilon}E_0}{\beta}(m-1)!B^m \sum_{r\geq 2}(\frac{4C_1}{k_0^2\xi B})^r \leq \frac{8\sqrt{\varepsilon}E_0C_1^2}{k_0^4\xi^2\beta B}(m-1)!B^{m-1}$$

provided that $\frac{4C_1}{k_0^2\xi B} < \frac{1}{2}$.

(3) The terms due to f are bounded by:

$$\frac{E_0}{\beta}(m-1)!B^{m-1}\sum_{r\geq 1}(\frac{2C_1}{k_0^2\xi B})^r \leq \frac{4E_0C_1}{k_0^2\xi\beta B}(m-1)!B^{m-1}.$$

provided that $\frac{2C_1}{k_0^2\xi B} < \frac{1}{2}$.

(4) The terms due to G_1 has the same bound as (2) if we fix $C_2 = E_0$.

(5) If we fix $C_3 = E_0$ as well, the terms due to g_1 are bounded by:

$$\frac{E_0}{\beta}(m-1)!B^{m-1}\sum_{r\geq 0}\sum_{s\geq 0, r+s\geq 1}(\frac{2C_1}{k_0^2\xi B})^r(\frac{2C_1}{\lambda_0\xi B})^s \leq \frac{4C_1E_0}{\beta k_0^2\xi B}(m-1)!B^{m-1}$$

provided that $\frac{2C_1}{\lambda_0\xi B} \leq \frac{2C_1}{k_0^0\xi B} < \frac{1}{2}$.

This five bounds must be all set $< \frac{1}{5}C_1$. It is easily seen that, as $b \leq 1$ and $\lambda_0 \geq k_0^2$, all the desired bounds are implied by $\max(\frac{8C_1}{\lambda_0\xi^2}, \frac{8\sqrt{\varepsilon}E_0C_1}{k_0^4\xi^2\beta B} \leq \frac{1}{5}$. Now we discuss the bounds on G_1 and g_1. There are always the same five terms times a factor $\frac{\beta}{\sqrt{\varepsilon}}$ for G_1 and β for g_1. So all the bounds are verified if $, \frac{E_0C_1}{k_0^4\xi^2\beta B} \leq c << 1$. We fix $C_1 = \frac{E_0}{\beta}$ as this comes from the first order and $B = c\frac{E_0^2}{k_0^4\xi^2\beta^2}$. □

A.5 Fast averaging Theorem for uni-modal perturbations

In this Appendix we report Paragraph ¶7 of [GGM3]. We consider the Hamiltonian:

$$H = \frac{1}{2}(\varepsilon J^2 + p^2) + I\frac{\omega_1}{\sqrt{\varepsilon}} + \cos q - 1 + \alpha A(\phi+\psi)B(q) + \eta f(\phi,\psi,q),$$

for $\eta = 0$. $A(x)$ is a trigonometric polynomial with zero mean value:

$$A(x) = \sum_{0<|n|\le N} A_n e^{inx}.$$

The symplectic change of coordinates with generating function:

$$J'\phi + I'\psi + p'q - \alpha\sqrt{\varepsilon}B(q)\sum\sum_{0<|n|\le N}\frac{A_n}{in}e^{i(\phi+\psi)n},$$

is globally defined for $\alpha\sqrt{\varepsilon} \ll 1$ (on a domain slightly smaller that the domain of H) and in the new coordinates the size of the perturbation is $\alpha\sqrt{\varepsilon}$. Moreover the perturbation is still a monochromatic trigonometric polynomial with zero mean value[8]. Now we pass to local hyperbolic coordinates for the pendulum, let us call then x, y. The Hamiltonian is:

$$\frac{\varepsilon}{2}J^2 + \sqrt{\varepsilon}G(xy,\sqrt{\varepsilon}) + \sqrt{\varepsilon}\alpha F(\phi+\psi,x,y),$$

moreover, as F is uni-modal then the lattice generated by its frequencies $K(V)$is one-dimensional,so the sub-lattice Λ of frequencies $\nu \in K(V)$ orthogonal to the fast direction is $\{0\}$.

Now, following Remark A.17, for $\alpha\sqrt{\varepsilon} < 1$, we can apply an ε independent number of steps of the Normal Form Theorem of Appendix A.4 so that the Hamiltonian is of the form A.4 with a perturbation of order $\eta < \varepsilon^{3/2}$. Finally for $\eta \neq 0$ we apply the change of coordinates just described. We obtain a Hamiltonian of the type A.4 but with a perturbing function $f(\psi,\phi,x,y)$ which is not a trigonometric polynomial. So we truncate the Fourier series of the perturbating function at $|\nu| < N$ with $N = \frac{1}{2}\varepsilon^{-\frac{1}{2}}\sqrt{\log\varepsilon^{-1}}$. Finally in the normal form theorem we set $K = \varepsilon^{-\frac{1}{2}}\sqrt{(\log\varepsilon^{-1})^{-1}}$.

A.6 Proof of Theorem 4.20

We consider a tree $A \in \overset{m}{\mathcal{A}}$ with total frequency ν and consider in each node $v \in A$ the Fourier expansion of $d^{n_v} f^{\delta_v}_{\nu_v}(q)$:

$$f(\psi,q) = \sum_{\nu,l\in\mathbb{Z}^{n+1}} f_{\nu,l}e^{i(\psi\cdot\nu + l\,q)}, \quad |f_{\nu,l}| \le Ce^{-r_1(|\nu|+|l|)}.$$

[8] However it does not depend only on the angle variables any more

So the integral a) of Section 3.1, (evaluated at $\varphi = 0$, becomes:

$$(-\frac{1}{2})^N(A)E(d,\nu) \sum_{\{\nu_v\}_\nu^k,\{l_v\}} [\prod_{\substack{s=1,\dots,n\\ \delta_v=1,v\geq v_0}} (i\nu_{v\,s})^{m_v(s)} f_{\nu_v,l_v}] \prod_{v\geq v_0} (il_v)^{n_v}$$

$$\oint \frac{dR_{v_0}}{2i\pi R_{v_0}} \int_{-\infty}^{\infty} d\tau_{v_0} e^{-\sigma(\tau_{v_0})R_{v_0}} e^{il_{v_0}q(\tau_{v_0}+id)} e^{i\omega_{v_0}\tau_{v_0}}$$

$$\prod_{v>v_0} \oint \frac{dR_v}{2i\pi R_v} (\int_{-\infty}^{\tau_w} d\tau_v + \int_{\infty}^{\tau_w} d\tau_v) e^{-\sigma(\tau_v)R_v(\tau_v+id)} w_{j_v}(\tau_w+id,\tau_v+id)$$

$$\prod_{v\geq v_0 e^{il_v q(\tau_v+id)}} e^{i\omega_v\tau_v}. \quad \text{(a)}$$

As usual w is the node preceding v, $m_v(s)$ is the number of nodes in the list $v, s(v)$ with label $j = s$, $n(v)$ the number of those with label $j = 0$, $l_v \in \mathbb{N}_0$ and $\omega_v = \omega_{\nu_v} = (\omega \cdot \nu)\varepsilon^{-\frac{1}{2}}$.

Now we proceed exactly as in Section 3.1 and we apply Proposition 3.6; finally, in bounding the proper integrals we notice that we do not approach any singularity. We obtain the following bound[9]:

$$C^k(k!)^{2\tau+2}E(d,\nu) \sum_{\{\nu_v\}_\nu^k,\{l_v\}} [\prod_{\substack{s=1,\dots,n\\ \delta_v=1,v\geq v_0}} (i\nu_{v\,s})^{m_v(s)}](il_v)^{n_v}|f_{\nu_v,l_v}|(\max_{t\in H(a_0,d)} |\max(e^{iq(t)},e^{-iq(t)})|)^{|l_v|}.$$

Now we choose a_0 and $d = c_1$ so that

$$\max_{t\in H(a_0,c_1)} |\max(e^{iq(t)},e^{-iq(t)})| \leq e^{r_1/2}$$

(see Figure 1.3 for an example). Finally we apply the bounds on the Fourier coefficients[10] of $f(\psi, q)$, we obtain:

$$C^k(k!)^{2\tau+2}E(d,\nu)(\prod_{j=1}^{n} [\prod_{\substack{v\geq v_0\\ \delta_v=1}} d_{r_1}^{m_v(j)} \sum_{h=-\infty}^{\infty} x^h e^{-r_1|h|}]_{\nu_j}) \prod_{\substack{v\geq v_0\\ \delta_v=1}} d_{r_1}^{n_v} \sum_{h=-\infty}^{\infty} e^{-r_1|h|/2}, \qquad \text{(A.7)}$$

all computed in $x = 1$. As in Chapter 1 $[f(x)]_n$ denotes the term of order n in the Taylor expansion of $f(x)$ around $x = 0$. The series in expression A.7:

$$\sum_{h=-\infty}^{\infty} x^h e^{-r_1|h|}$$

[9] we are ignoring the nodes with $\delta = 0$ as they are clearly irrelevant

[10] Remember that we are considering a tree with total rotation ν

are absolutely convergent for $e^{-r_1} < |x| < e^{r_1}$ and so we can bound the term of order ν_j by:

$$C^k \prod_{v:\delta_v=1} m_v(j)! e^{-r_1|\nu_j|}.$$

Finally one can proceed as in Appendix A.2 To prove that:

$$\sum_{A\in\mathcal{A}_j^k} c(A) \prod_{v:\delta_v=1} \prod_{j=0}^{n} m_v(j) \leq (Cn)^k,$$

we simply choose

$$f(x_0,\ldots,x_n) = \prod_{j=0}^{n} \frac{1}{1-x_j}, \quad g(x_0) = 1 + 2y - e^y.$$

A.7 Cancellations due to integration by parts

This is a simple generalization of the results in [GGM1].
given $A \in \mathcal{A}_{(i,h)}$ consider a continuous function $h(t)$, remember that:

$$\mathcal{W}^1(A) = -\underbrace{\partial_{j_{v_0}} h(\tau_{v_0})}_{\text{marking}} \nabla^{\vec{m}(v)} f^{\delta_{v_0}}(\tau_{v_0}) \prod_{v_i\in s(v_0)} Q_{j_{v_i}}[\mathcal{W}^1(A^{\geq v_i})]$$

is a function of the time τ_{v_0}.

Lemma A.18. *Consider a function h in H_0 and a fruitless tree A , the function $h(\tau_0)O_1' \circ \bar{\Psi}_\varphi^1(A)$ can be extended to a homomorphic function on a predefined strip around the real axis $\tau_0 \in \mathbb{R}$. Moreover if $h(t) \in H$ is o continuous then:*

$$\Im \partial_t h(t) \mathcal{W}^1_{\phi_0}(A,t) = 0$$

Proof. The first assertion is simply the closure of H_0 under the action of Q_j (see Proposition 1.16(ii)); the second is equivalent to proving that for the continuous function $h(\tau_0)\mathcal{W}^1_\varphi(A)$ one has $\Im(\partial_{\tau_0}\pi_P h(\tau_0)\mathcal{W}^1_\varphi(A) = 0$ (π_P is the projection on polynomials). This is obvious as $\Im P = 0$ for any polynomial. □

Remark A.19. *As $f^\delta(t) = F(\psi_i(0) + \tilde{\omega}_i t, \psi_0(t)$ e $\dot{\psi}_0(t) = -2x_0^0(t)$ we have that:*

$$\partial_t f^\delta(t) = \sum_{j=1,\ldots,n} \tilde{\omega}_i \partial_{\psi_i} - 2x_0^0 \partial_{\psi_0}$$

Lemma A.20. *Given an odd function $G \in H_0$ the following relation holds:*

$$\partial_\tau Q_j(G) = Q_j[\partial_{\tau_y} G(\tau_y) + 2\delta_{j0} x_0^0(\tau_y) \partial_0^3 f^0(\tau_y) Q_0(G)]$$

Proof. if $j \neq 0$ one can verify the Lemma directly integrating by parts (notice that F_j^k has no constant components so that $\pi_P x_j^i G \neq c$ for any summand G coming from F_j^k), this is a heavy computation so we give an alternative proof.
We consider the vector $V = \begin{pmatrix} O_j(G) \\ \partial_t(O_j(G)) \end{pmatrix}$, by the definition of O_j it is a solution of $\dot{V} = L_j V + G$ where L_j is the 2×2 block of the matrix L (defined in Subsection 1.1.4) coming from the action-angle variables I_j, ψ_j (remember that $I_0 = p, \psi_0 = q$):

$$L_j = \begin{vmatrix} 0 & 1 \\ \delta_{j0} cos(q_0) & 0 \end{vmatrix}.$$

We derive with respect to t:

$$\ddot{V} = W_0 \dot{V} + (\dot{W}_0 V + \dot{G})$$

the first line of the solution $\dot{V}$ is

$$\partial_t(O_0(G)) = O_0(-\dot{q}(t)\partial_0^3 f^0(t) O_0(G) + \dot{G})$$

plus the first component of a solution of the homogeneous equation $t \to W(t)X$ that we determine via the initial data. $O_j^t(F)$ is zero for $t = 0$, and we have seen in Subsection 1.1.3 that the initial datum is determined by the boundedness condition $\partial_t(O_j(G))|_{t=0} = \Im^0(x_j^0 G)$ so:

$$\partial_t(O_j(G)) = O_j(2\delta_{j0} x_0^0(t)\partial_0^3 f^0(t) O_0(G) + \dot{G}) + x_j^0(t)\Im^0(x_j^0 G)$$

and as G is odd we can substitute $Q_j(G) = O_j(G)$.

Next we notice that the vectors $W^i = \begin{pmatrix} x_0^0(t) \\ \dot{x}_0^0(t) \end{pmatrix}, \begin{pmatrix} \sigma(t)x_0^1(t) \\ \sigma(t)\dot{x}_0^1(t) \end{pmatrix}$ are solutions of the system $\dot{W} = L_0 W$ so we apply the time derivative and obtain[11]:

$$\dot{x}_0^i = 2Q_0(x_0^i x_0^0 \partial_0^3 f^0(t)) + \delta_{i1}\sigma(t) x_0^0 \tag{A.8}$$

the last term is added to have the right behavior in $t = 0$ ($d_t \sigma(t) x_0^1|_0 = 1$).

$$O_j(2\delta_{j0} x_0^0(t)\partial_0^3 f^0(t) O_0(G) + \dot{G}) + x_j^0(t)\Im^0(x_j^0 G) = Q_j(2\delta_{j0} x_0^0(t)\partial_0^3 f^0(t) O_0(G) + \dot{G}) +$$

$$+ \frac{1}{2}\sum_i (x_j^i \Im x_j^{[i]}(2\delta_{j0} x_0^0(t)\partial_0^3 f^0(t) O_0(G) + \dot{G}) + x_j^0(t)\Im^0(x_j^0 G)$$

The last two sums cancel each other via relationA.8, for $j = 0$, and using the fact that if $j \neq 0$ then $\dot{x}_j^0 = 0$ and $\dot{x}_j^1 = \sigma(t)$. □

[11] we are using the fact that $O_0(\sigma(\tau)F) = \sigma(t)O_0(F)$

Proposition A.21. *Consider $A \in \mathcal{A}_{(i,h)}$ with $h(t) \in H$ continuous for $t \in \mathbb{R}$ (as h is in H it is analytic separately in $\mathbb{R}^+$ and $\mathbb{R}^-$):*

$$\Big(\sum_{j=1,\dots,n} (\tilde{\omega}_j \sum_{v \in A} \partial_j^v A) = 2 \sum_{v \in A} x_0^0(v) \partial_0^v A + 2L_{x_0^0}(A) - \dot{h}(t)\bar{A}) \Big)_{\varphi=0}$$

$\dot{h}(t)\bar{A}$ is simply the tree A marked $\dot{h}(\tau_0)$ instead of $h(\tau_0)$.

Proof. Let $\bar{A} = \frac{A}{h(v_0)}$ namely the tree A, marked with the function $h = 1$ we need to prove that:

$$\Im \partial_t \mathcal{W}_0^1(A) = \Im \mathcal{W}_0^1(\sum_{j=1,\dots,n} (\tilde{\omega}_j \sum_{v \in A} \partial_j^v A) - 2 \sum_{v \in A} x_0^0(v) \partial_0^v A - 2L_{x_0^0}(A) + \dot{h}(t)\bar{A}).$$

We know that:

$$\Im \partial_{\tau_0} \mathcal{W}_0^1(A) = \Im \dot{h}(\tau_0) \mathcal{W}_0^1(\bar{A}) + \Im h(\tau_0) \left\{ \partial_{\tau_0} \left[\mathcal{W}_0^1(\bar{A}) \right] \right\}$$

where the term in $\{\}$ parentheses is:

$$-(\partial_{\tau_0} \nabla^{\vec{m}(v_0)} f^{\delta_{v_0}}) \prod_{v_i \in S(v_0)} Q_{j_{v_i}}[\mathcal{W}^1(A^{\geq v_i})] +$$

$$+\nabla^{\vec{m}(v_0)} f^{\delta_{v_0}} \sum_{v_i \in S(v_0)} \mathcal{W}_0^1(A^{/v_i}) \partial_{\tau_0}[Q_{j_{v_i}} \mathcal{W}_0^1(A^{\geq v_i})]) \tag{A.9}$$

Now we set $\mathcal{W}_0^1(A^{\geq v_i}) = F$ (which is odd as $\varphi = 0$) and apply Lemma A.20 to $F \in H_0$:

$$\partial_{\tau_0} Q_{j_{v_i}}(F) = Q_{j_{v_i}}(\partial_{\tau_{v_i}} F) + 2\delta_{j0} Q_0(x_0^0(\tau_y) \partial_0^3 f^0(\tau_y) Q_0(F))$$

and

$$\partial_{\tau_{v_i}} F = \partial_{\tau_{v_i}}[\mathcal{W}_0^1(A^{\geq v_i})]$$

this is the same expression treated in A.9 calculated on trees of lower order. So we proceed recursively and obtain:

$$0 = \Im\{\dot{h}(\tau_0) \mathcal{W}_0^1(\bar{A}) + \sum_{v \in A} \mathcal{W}_0^1 \partial_{\tau_v}^v(A) +$$

$$2\delta_{j_v 0}(\mathcal{W}_0^1[A^{\setminus v}]) Q_0[x_0^0(\tau_y) \partial_0^3 f^0(\tau_y) Q_0(\mathcal{W}^1(A^{\geq v}))]\}$$

The symbol $\partial_{\tau_v}^v$ means a τ_v derivative applied to $f^{\delta_v}(\tau_v)$ so we can apply LemmaA.19. The third sum is $-2\mathcal{W}_0^1(L_{x_0^0}(A))$ (by the definition of $L_{x_0^0}(A)$). ☐

Corollary A.22. *In particular PropositionA.21 holds for $\mho_{(j,h)}^k$:*

$$\sum_{i=1,\dots,n} \tilde{\omega}_i \mho_{(j,h)\,i}^k = 2\mho_{(j,h)\,(0,x_0^0)}^k + 2L_{x_0^0}(\mho_{(j,h)}^k) + \dot{h}(t, v_0) \partial_j^{v_0} \mho^k$$

A.8 Properties of the matrix $\mathcal{M}$ due to cancellations

Lemma A.23. *The relation*

$$\dot{x}_0^l = 2Q_0(x_0^0 x_0^l \partial_0^3 f^0) + \delta_{l1}\sigma(t)x_0^0$$

implies that:

$$F^{l\,0}(\mho_0) = \frac{1}{2}(\dot{x}_0^l \mho_0 - \sigma(t)x_0^0)$$

Proof. In $F^{l\,0}(\mho_0)$ we change the first node to the only node v_1 of level one; then we substitute the leaf

$$Q_0(x_0^0 x_0^l \partial_0^3 f^0) = \frac{1}{2}(\dot{x}_0^l - \delta_{l1}\sigma(t)x_0^0)$$

□

Proposition A.24. *The matrix $\mathcal{M}$ verifies:*

$$\mathcal{M}Y_1 = A \quad where \quad Y_1^t = (\overbrace{0}^{n+1}, \overbrace{2}^{1}, \overbrace{-\vec{\omega}}^{n})$$

e $A^t = (\overbrace{I_0^+(t=0)}^{1} \overbrace{I^+(t=0)}^{n}, \overbrace{0}^{n+1})$*; remember that* $I_j^+(t=0, \varphi=0)$ $(j=0,\ldots,n)$ *is the initial datum in the actions at the homoclinic point* $\varphi = 0, \phi_0 = \pi$.

Proof. This is a consequence of Corollary A.22. let us the relation in components:

$$2M_{k\,n+2} = \sum_i \omega_i M_{k\,n+2+i} + (\sum_{j=1}^{n+1} \delta_{k,n+1+j})I^+_{k-n-2}$$

Translated in trees this is:

$$2\mho_{k\,0}^{l\,0} + 2L^0(\mho_k^l) + 2\delta_{k\,0}F^{l\,0}(\mho_0) = \sum_i \tilde{\omega}_i \mho_{k\,i}^{l\,0} - \delta_{l,1}(\sum_{j=0}^{n} \delta_{kj})\sigma(t)\mho_k^0 \tag{A.10}$$

here we simply used the definition of $\mathcal{M}$ and the identity:

$$I_j^+ = \frac{1}{2}(I_j^+ + I_j^-) = -\Im \mathcal{W} x_j^0 \sigma(\tau_0)\mho_j^0.$$

Expression A.10 can be derived from Proposition A.21 by setting $h_k^l = x_k^l$ for $l = 0, 1$ and $k = 0, \ldots, n$. as by Lemma A.23:

$$\dot{h}_k^l(\tau)\mho_k = \begin{cases} -2F^{0l}(\mho_{\boxed{0}}) + \sigma(t)\mho_{\boxed{0}}^0 & k=0 \\ 0 & l = 0 k = 1,\ldots,n \\ \sigma(\tau)\mho_{\boxed{k}}^0 & l = 1\; k = 1,\ldots,n \end{cases}.$$

□

Bibliography

[A1] Arnold V. I. *Small denominators and problems of stability of motion in classical and celestial mechanic*, Russ. math. Surveys,**18:6** (1963) pp. 85-191.

[A2] Arnold V. I. *Instability of dynamical systems with several degrees of freedom*, Sov. Math. Dokl. ,**5** (1964) pp. 581-5.

[BG] Benettin G. ,Gallavotti G. *Stability of motions near resonances in quasi-integrable systems*

[BB1] Berti M., Bolle P. *A functional analysis approach to Arnold diffusion*, Annales de l'IHP, section analyse non-lineaire, **18** 2001.

[BB2] Berti M., Bolle P. *Fast Arnold diffusion in systems with three time scales*,to appear in Discrete and Continuous Dynamical Systems.

[Be] Bessi U. *An approach to Arnold diffusion through the calculus of variations*, Nonlinear analysis T. M. A. , **26** (1996) pp. 1115 -35.

[BCV] Bessi U., Chierchia L., Valdinoci E. *Upper bounds on Arnold Diffusion Time via Mather Theory*, J. Math. Pures Appl. **80-1** (2001) pp. 105-129

[Bo] B. Bollobas:*Graph theory* Springer-verlag (Graduate texts in mathematics, 63), Berlin- Heidelberg-New York, 1979

[C] Chierchia L. *Arnold instability for nearly-integrable analytic Hamiltonian systems*, Proceedings of the workshop " Variational and local methods in the study of Hamiltonian systems", A. Ambrosetti, G.F. Dell'Antonio eds. Trieste, 1994.

[CF] Chierchia L., Falcolini C.:*A Direct Proof of a Theorem by Kolmogorov in Hamiltonian Systems*

[CG] Chierchia L., Gallavotti G.:*Drift and diffusion in phase space*, Annales de l'IHP, section Physique Théorique, **60** (1994) pp. 1-144; see also the *Erratum* **68** (1998), 135.

[CV] Chierchia L., Valdinoci E.:*A note on the construction of Hamiltonian trajectories along heteroclinic chains*,Forum Math.,12 **2** (2000), 247-254.

[Dy] Arnold V. I. *Dynamical Systems 3*, Encyclopedia of Math. Sci.**3** (1963)

[DGJS] Delshams A., Gelfreich V. G., Jorba V. G., Seara T. M. :*Exponentially small splitting of separatrices under fast quasi-periodic forcing*, Comm. Math. Phys., **150** (1997), 35-71.

[E] Eliasson L.H.: *Absolutely convergent series expansions for quasi periodic motions*, Reports Dept. of Math., , Univ. of Stockholm, Sweden, No.**2** (1988) 1-31

[F] Forster O.:*Lectures on Riemann surfaces* Springer-Verlag (Graduate texts in mathematics, 81), Berlin- Heidelberg-New York, 1980

[G1] Gallavotti G.:*Twistless KAM tori, quasi flat homoclinic intersections and other cancellations in the perturbation series of certain completely integrable Hamiltonian systems. A review*, Rev. Math. Phys., **6** (1994), 343-411.

[G2] Gallavotti G.:*Reminiscences on science at the I.H.E.S. A problem on homoclinic theory and a brief review*

[GGM1] Gallavotti G., Gentile G., Mastropietro V.: *Separatrix splitting for systems with three time scales*, Commun. Math. Phys.**202** (1999) 197-236.

[GGM2] Gallavotti G., Gentile G., Mastropietro V.: *Melnikov approximation dominance. Some examples*, Rev. Math. Phys. **11** (1999), 451-461.

[GGM3] Gallavotti G., Gentile G., Mastropietro V.: *Hamilton-Jacobi equation and existence of heteroclinic chains in three time scales systems*, Nonlinearity, **13** (2000), 323-340.

[GGM4] Gallavotti G., Gentile G., Mastropietro V.: *Lindstedt series and Hamilton-Jacobi equation for Hyperbolic tori in three time scales problems*

[Ge] Gentile G. *Whiskered tori with prefixed frequencies and Lyapunov spectrum* Dynam. Stability Systems **10** (1995), no. 3, 269–308.

[Gel] Gelfreich. V. *Melnikov method and exponentially small splitting of separatrices,* Physica D. **101** (1997), pp. 227-248.

[GR] Godsil C., Royle G.:*Algebraic Graph theory* Springer-Verlag (Graduate texts in mathematics, 207), Berlin- Heidelberg-New York, 2001

[GJ] Goulden I.P. Jackson D.M. :*Combinatorial enumeration.* Wyley Interscience Series in Discrete math., 1983.

[HM] Holmes P., Marsden J.: :*Melnikov method and Arnold diffusion for perturbations of integrable Hamiltonian systems*, J. Math. Phys. , **23** (1982), pp. 669-675.

[HMS] Holmes P., Marsden J., Scheurle J.: :*Exponentially small splitting of separatrices with applications to KAM theory and degenerate bifurcations*, Contemp. Math., **81** (1988),pp. 213-244.

[L] Lang S.:*Algebra*,Addison Wesley Publishing Co. 1993

[LMS] Lochak P., Marco J. P., Sauzin D.: *On the splitting of invariant manifolds in multidimensional Hamiltonian systems*, preprint Université Jussieu.

[M] Melnikov V. K..:*On the stability of the center for time periodic perturbations,* Trans. Moskow Math. Soc. **12**,1 (1963)

[Pö] Pöschel *Nekhoroshev estimates for quasi-convex Hamiltonian systems* math. Z. **213** (1993), 187-216.

[P] Poincaré H. *Les methodes nouvelles de la mécanique céleste* Vol. I-III Paris, 1892,93,99.

[R] Range R. M.: *Holomorphic functions and integral representations in several complex variables*, Springer Verlag (Graduate texts in mathematics, 108), 1986.

[RU] Rudin W.: *Real and complex analysis*

[S] Serre J.P. : *Trees* , Springer Verlag , 1980.

[V] Valdinoci E. : *Families of whiskered tori for a-priori stable/unstable Hamiltonian systems and construction of unstable orbits*, Math. Phys. Electr. **6** 2000