SISSA - SCUOLA
INTERNAZIONALE
SUPERIORE
STUDI AVANZATI
TRIESTE
Strada Costiera 11

ISAS - INTERNATIONAL SCHOOL
FOR ADVANCED STUDIES

Fast rotations of the rigid body and Hamiltonian perturbation theory

Thesis submitted for the degree of
"Doctor Philosophiæ"

CANDIDATE
Francesco Fassò

SUPERVISOR
Giancarlo Benettin

April 1991

TRIESTE

Fast rotations of the rigid body and Hamiltonian perturbation theory

Thesis submitted for the degree of

"Doctor Philosophiæ"

CANDIDATE

Francesco Fassò

SUPERVISOR

Giancarlo Benettin

April 1991

Contents

Introduction

A. The dynamics of a rigid body with a fixed point in a conservative force field is one of the classical problems of mechanics (see for instance [75, 7, 57]). It constitutes a Hamiltonian system with three degrees of freedom, so that the equations of motion are integrable by quadratures whenever there exist three independent integrals of motion in involution. Such integrals are known to exist only in few particular cases.

In the classical cases of Euler–Poinsot, Lagrange, and Kowalevskaya the three integrals exist, and are independent, in an *open and dense* subset of phase space. The first case is 'free' motion, namely under the influence of external forces with vanishing torque with respect to the fixed point (obviously, the reaction of the constraint is supposed to be ideal). The other two cases concern symmetric rigid bodies, having special mass distributions, in the constant gravity field (by symmetric we mean that the inertia ellipsoid relative to the suspension point is of revolution). Moreover, there are also cases in which the equations of motion are integrable, but only for a subset of initial conditions which is *closed and of zero measure* (examples can be found in [9,25]).

Among all problems of rigid body dynamics, the most studied one is certainly that of a 'heavy' (i.e. subject to a constant gravity) rigid body. In such a case, there always exists a second independent integral of motion, besides energy, namely the projection of the angular momentum along the direction of gravity. The 'third' integral has been longly searched for, and the very question of its existence was essentially open until recent times. For instance, Poincaré (who had extremely clear ideas about the non–integrability of nearly integrable systems) could not exclude its existence, and was very cautious about this question (see [67], vol. 1, pag. 255–259). Non–existence results for analytic integrals (of a certain kind) defined in open sets were later obtained by Husson [44] and, more recently, by Kozlov [52] and Ziglin [76] (see also [9, 43, 54]). Numerical evidence for chaotic motions of a heavy asymmetric body was given in [38].

The studies of rigid body dynamics have been spread over more than two centuries. In the absence of exact integrability results, they have been mainly addressed to establish, or exclude, the existence, in various cases, of special solutions (periodic motions, regular precessions)[1)], and to study their stability properties.

Another important approach has been the search for approximate descriptions, espe-

---

1) With the hope, as Klein and Sommerfeld say, that "by finding enough special cases we may some day be able to know more about the general solution of the problem". Leimanis observed about thirty years ago that such a hope "has not yet come true" (see [55]).

cially in connection with gyroscopic phenomena, and more generally with the *fast motions* of the bodies. Roughly speaking, a motion is 'fast' if the angular velocity ω is very large (in some natural units); so it can be characterized, for instance, by the fact that the kinetic energy is much larger than the potential energy. In such a case, one can be content with a description accurate up to terms of the order of some inverse power of ω.

It is then natural to set up a perturbative approach, in which one considers the potential force field as a small perturbation of the Hamiltonian describing the Euler–Poinsot system. This approach was quite common in celestial mechanics (see for instance chapter 6 of [11]). However, because of the well known difficulties due to resonances ('secular terms', 'small denominators'), such studies were generally restricted to first order expansions, and were not fully rigorous. The advantage of such a perturbative approach over other approximate methods (see for instance [70, 71, 72, 12]), lies in the fact that classical perturbation theory is concerned with the behaviour in time of the phase space functions which are integrals of motion for the unperturbed system, rather than with the behaviour of individual orbits. The point is that one can control the perturbed integrals of motion over time scales extremely longer than for the single perturbed trajectories.

In the early sixties, Arnold considered the problem of the fast motion of an asymmetric body in the constant gravity field (more generally, in an axially symmetric force field) by the rigorous methods of classical perturbation theory. Regarding the system as a perturbation of the Euler–Poinsot case, he made an application of the then new–born KAM theory. Basically, he succeeded in proving the integrability of the equations of motion in a subset of phase space which is *closed* but *of large relative Lebesgue measure*. The main consequence is a stability property of *all* fast motions.

B. The main purpose of the present thesis is to study the fast motion of a rigid body with a fixed point in an *arbitrary* (but analytic) conservative force field by the methods of Hamiltonian perturbation theory. However, at variance with Arnold, we will not use KAM theory. Indeed, the latter cannot be applied for a generic perturbation, because of the degeneracy of the Hamilton function of the Euler–Poinsot system (a problem which can be overcome in the special case considered by Arnold, using in an essential way the symmetry of the perturbation). Moreover, even if the perturbation had the very special property of 'removing the degeneracy', KAM theory would nevertheless assure stability only for the majority of initial conditions, and not for all of them.

Because of this, we shall base our analysis on what is now usually called the Nekhoroshev approach[63,65,66,15], in which one looks for results valid for all initial data, renouncing to have results valid for all times. In fact, by such a method one gets results for extremely long times, increasing exponentially with an inverse power of the 'small' parameter of the system; in the present case, this leads to times increasing exponentially fast with the (square root of the) angular velocity. For the sake of simplicity, we shall consider only the case of a symmetric body.

Our central result is the following: in the fast motion of a symmetric body about a fixed point, under the influence of an analytic conservative force field, both the modulus of the angular momentum and its projection on the inertia symmetry axis of the body vary at most of quantities $\mathcal{O}(1)$ (namely, independent of ω), for times $|t| \sim \exp(\sqrt{\omega})$ (the first of such two estimates is optimal). In addition, a detailed approximate description of the motions is obtained. Some partial results for a fast rotating body with no fixed point are also given.

These results do not constitute simply a corollary of Nekhoroshev's theorem. The main reason is that the action–angle variables of the free rigid body are not globally defined in the phase space. In fact, one can construct an atlas with two coordinate charts; but, because of the degeneracy of the unperturbed Hamiltonian, there remains the problem that one has no control on the time after which the system will leave the domain of each chart. As a consequence, the standard results of classical perturbation theory, which use canonical transformations to get suitable normal forms within each coordinate system, are not sufficient.

Indeed, this is a general problem, which is always met in perturbation theory, when one has to deal with systems which are degenerate and do not possess global action–angle variables (actually, degeneracy is a typical cause of non–existence of global action–angle variables). For instance, besides the rigid body, this problem is encountered in the Kepler system.

Because of its importance, we study such a problem in full generality. We show that, under fairly general conditions, the canonical transformations (and the corresponding normal forms) constructed 'locally' in each chart domain with the usual techniques of perturbation theory, are in fact the local representatives of a canonical transformation (and of a Hamiltonian system) which is 'globally' defined on the phase space. This study will be based on a theory of degenerate systems, given by Nekhoroshev [64], which is essentially a generalization of the Arnold–Liouville theorem.

Another problem concerns the quality of the estimates mentioned above. The existing, general proofs of Nekhoroshev's theorem would lead to very poor estimates for the variations of the angular momentum and for the times on which they are assured to hold. Obviously, better results can be obtained in specific cases, taking advantage of the peculiarities of the system. In the present case, a great effort has been paid to obtain the above mentioned optimal estimate $\mathcal{O}(1)$ for the variations of the angular momentum. To this end, use is made of a perturbative technique based on the construction of a normal form for the Hamiltonian vector field of the system, rather than for the Hamilton function alone (see [36]).

C. The thesis is organized as follows. Chapter 1 has an introductory character, and contains some general informations about rigid body and classical perturbation theory. Special emphasis is given to problems of globality for action–angle variables, and degeneracy. Chapter 2 is concerned with the global (i.e. chart independent) formulation of classical perturbation theory for degenerate systems. Chapter 3 is devoted to a preliminary, detailed study of the action–angle variables of the free symmetric rigid body. Chapter 4 deals with the construction, by perturbation theory, of the normal forms for the system;

these are then used in chapter 5 to study the fast motions of the rigid body.

D. Acknowledgments: I wish to express my gratitude to Giancarlo Benettin, who introduced me to the field of Hamiltonian perturbation theory, and constantly addressed and encouraged my work. I am also indebted to Prof. Gianfausto Dell'Antonio, for his guidance during my years at *SISSA*, and for many helpful discussions. I like to thank Luigi Galgani and Antonio Giorgilli, whose ideas deeply influenced my work. I discussed all the material of this thesis with Franco Cardin, to whom I wish to express my friendly gratitude. I have the pleasure of thanking Jean Bellissard and Ricardo Lima for their hospitality at the *Centre de Physique Theorique* of the *CNRS* at Luminy (Marseille) during the last year; I greatly profited of that stimulating environment to reconsider all the material of the thesis. My stage at Luminy has been supported by a fellowship of *Consiglio Nazionale delle Ricerche*. Finally, I wish to thank Dario Bambusi, who helped me in proving the estimates of proposition 4.5.

Chapter 1

The rigid body and classical perturbation theory: an overview

This chapter has an essentially introductory character: we give an overview of some relevant arguments. Consequently, many details and proofs are left for the next chapters. Problems related to action–angle variables will be of a certain importance in the sequel. Thus, we devote section 1 to recall some basic facts about such coordinates, and about completely integrable systems. Section 2 is an introduction to the 'free' motion of a rigid body with a fixed point, namely the Euler–Poinsot problem. Section 3 gives a survey of classical perturbation theory, while section 4 depicts its use in the study of the fast motions of a rigid body.

We notice that $\langle\ ,\ \rangle$ and $\times$ will always denote, respectively, the euclidean scalar product in $\mathbb{R}^n$ and the cross product in $\mathbb{R}^3$.

1.1 The Arnold–Liouville theorem and the action–angle variables

A. Action–angle variables. The key result about the existence of action–angle variables is the celebrated Arnold–Liouville theorem. We formulate here such a theorem in a form which is somehow stronger than Arnold's original statement [4], which is due essentially to Weinstein [74] and Nekhoroshev [64].

To this purpose, let us introduce some notations and terminology. Let $f = \{f_1, \ldots, f_m\}$ be m real functions defined on a manifold M. The subset R_f of M at each point of which the differentials $df_1, \ldots, df_m$ are linearly independent is foliated into the level surfaces described by the equations $f_j = \text{const}$, $j = 1, \ldots, m$. Each leaf of that foliation is an immersed manifold of codimension m. Let $\mathcal{L}_f(z)$ be the connected component containing

z of the leaf through the point z:

$$\mathcal{L}_f(z) \; = \; \text{the } z\text{-component of } \{z' \in M : f_j(z') = f_j(z), j = 1, \ldots, m\} \, .$$

Let us now restrict ourselves to the subset M_f of R_f which contains only *compact* leaves:

$$M_f \; = \; \{z \in M : df_1(z) \wedge \ldots \wedge df_m(z) \neq 0, \text{ and } \mathcal{L}_f(z) \text{ is compact}\} \, . \tag{1.1}$$

In the following, we shall call M_f the *properness* set of the function f.[2)] Obviously, M_f is an open subset of M. The foliation of M_f into compact connected leaves $\mathcal{L}_f$ will be denoted $\mathcal{F}_f$.

Recall now that two functions f, g on a symplectic manifold are said to be in involution if their Poisson bracket vanishes identically: $\{f, g\} = 0$. The Arnold–Liouville theorem gives properties of the foliation $\mathcal{F}_f$ in the case the functions $f_1, \ldots, f_m$, $m = \frac{1}{2} \dim M$, are pairwise in involution (i.e. $\{f_i, f_j\} = 0$ for all $i, j = 1, \ldots, m$).

Proposition 1.1 (Arnold–Liouville theorem) *Let $f = \{f_1, \ldots, f_m\}$ be m functions defined on a symplectic manifold M of dimension $2m$, which are pairwise in involution. Then:*

1. Let $\mathcal{L}$ be a leaf of the foliation $\mathcal{F}_f$ of M_f. Then there exist an open neighbourhood $U \subset M_f$ of $\mathcal{L}$, which is union of leaves of $\mathcal{F}_f$, and a diffeomorphism

$$z \mapsto (I_1(z), \ldots, I_m(z), \varphi_1(z), \ldots, \varphi_m(z)) \tag{1.2}$$

of U onto $B \times \mathbb{T}^m$, where B is an open set in $\mathbb{R}^m$, such that:

i) the symplectic two-form of M, restricted to U, is $\sum_j dI_j \wedge d\varphi_j$,

ii) in U, the leaves of the foliation $\mathcal{F}_f$ are described by $I_j = \text{const}$, $j = 1, \ldots, m$,

iii) in U, the functions $f_1, \ldots, f_m$ are invertible functions of $I_1, \ldots, I_m$.

2. Moreover, let $\mathcal{L}$ and $\mathcal{L}'$ be any two leaves of $\mathcal{F}_f$, and let (I, φ) and (I', φ') be coordinates as above, defined respectively in the neighbourhoods U and U'. Then, in every connected component of $U \cap U'$ (if not empty) one has

$$\begin{aligned} I' \; &= \; A^{-1} I \, + \, a \\ \varphi' \; &= \; A^t \varphi \, + \, \mathcal{F}(I) \end{aligned} \tag{1.3}$$

for some $A \in O(\mathbb{Z}, m)$, some $a \in \mathbb{R}^m$, and some function $I \mapsto \mathcal{F}(I) \in \mathbb{R}^m$.[3)]

A proof of the Arnold–Liouville theorem can be found in [64] (see also [7, 74, 35, 57]). In the appendix A at this chapter we give a proof of part *2*, which is simple but has a certain interest. Let us now make some comments about the Arnold–Liouville theorem:

---

2) Such a name is surely not very lucky. It is motivated by the fact that a *proper* map is a map such that the preimages of compact sets are compact.

3) $O(\mathbb{Z}, m)$ is the group of all the $m \times m$ matrices with integer entries and determinant ± 1. Notice that statements *ii)* and *iii)* above are not independent: we have inserted both for greater clarity.

i) $\mathcal{F}_f$ is a foliation of the properness set M_f into m–dimensional tori. Locally, in a neighbourhood of each leaf of the foliation, there exist symplectic coordinates I, φ adapted to the foliation, which will be called *action–angle variables of the foliation* $\mathcal{F}_f$.

ii) In principle, one could define as 'action–angle variables of the foliation $\mathcal{F}_f$' any set of local coordinates $I, \varphi(\text{mod}2\pi)$ which satisfy all the conditions of part *1* of proposition 1.1. It is easy to see that such coordinates are uniquely determined by $\mathcal{F}_f$, up to transformations of the form (1.3).

iii) Under the hypotheses of the theorem, M_f is also a fiber bundle with fiber $\mathbb{T}^n$, which we shall denote $M_f \xrightarrow{f} \mathcal{B}_f$. Indeed, statement *2* of the theorem shows that, by patching together the local action–angle coordinates systems of the neighbourhoods of the leaves of $\mathcal{F}_f$, one constructs an atlas for M_f which is a fiber bundle atlas. The base $\mathcal{B}_f$ of the bundle is locally defined by the projection $(I, \varphi) \mapsto I$. Thus, the (local) action variables I can be regarded as local coordinates on the basis $\mathcal{B}_f$.

iv) It is not assured by the theorem that one of the local systems of action–angle coordinates can be extended to cover all of M_f, namely that there exists an atlas of action–angle variables constituted by a single chart. When this happens, we say that M_f possesses *global action–angle variables*; otherwise, we say that M_f possesses action–angle variables. Obviously, M_f has global action–angle variables iff the bundle $M_f \xrightarrow{f} \mathcal{B}_f$ is trivial (i.e. diffeomorphic to $\mathcal{B}_f \times \mathbb{T}^n$).

Remark 1.1 We adopt here the traditional approach, and describe the foliation into tori by means of a set of indipendent functions in involution. Such approach is simple and convenient for applications. However, it has some shortcomings, which come from the obvious fact that the description of a foliation by functions is not univocal. Thus, we remark that the important object for all the theory is the foliation, not the functions which are used to describe it: in particular, the local action–angle variables depend only the foliations. Furthermore, some caution has to be posed in identifying the (maximal) domain in which the foliation is defined with the 'properness' set of the considered functions: a trivial reason is that the loose of linear independence of the differentials of these functions does not necessarily reflect a singularity of the foliation. A more important fact is that there exist foliations which cannot be described globally by a single set of functions. To avoid such problems, one could formulate the Arnold–Liouville theorem in terms of the geometric analog of maximal sets of functions in involution, namely Lagrangian foliations (see [74,33]).

B. Completely integrable systems. The Arnold–Liouville theorem, as formulated above, is purely a statement of symplectic geometry.[4] The interesting application of the

[4] It can be of some interest to notice that in every symplectic manifold of dimension $2m$ there exist m smooth functions which are pairwise in involution and have almost everywhere (with respect to the Lebesgue measure) linearly independent differentials (see [37]).

theorem is to the case in which the m independent functions in involution are integrals of motion of a Hamiltonian flow. This is the case of the so–called completely integrable systems. Usually[58,1,57], a Hamiltonian system on a symplectic manifold M of dimension $2m$ is said to be *completely integrable* if it possesses m smooth integrals of motion $f_1, \ldots, f_m$, which are pairwise in involution, and have differentials which are linearly independent in an open dense subset of M (one does not require the independence in all of phase space, so to consider, for instance, a harmonic oscillator or a pendulum as completely integrable in all of their phase space).

In the case of a completely integrable system, the Arnold–Liouville theorem assures that the 'properness' set M_f of the m integrals of motion in involution is fibrated in m–dimensional tori $\mathcal{L}_f(z)$, which are *invariant* under the flow.

Notice also that the Hamilton function h, being constant on each torus $\mathcal{L}_f(z)$, can be considered as defined on the basis $\mathcal{B}_f$ of the bundle $M_f \xrightarrow{f} \mathcal{B}_f$.

The description of the flow of a completely integrable system is especially significative, if referred to the local action–angle variables of the foliation into invariant tori: the local actions are integrals of motion, and the motions appear to be quasiperiodic. In fact, in the local action–angle variables (I, φ), the Hamilton function h has local representatives $\hat{h}$ (defined by $h = \hat{h} \circ I$) which are functions of the actions alone; thus, the equations of motion are (locally)

$$\dot{I}_i \;=\; 0\,, \qquad \dot{\varphi}_i \;=\; \omega_i(I) \qquad\qquad (i = 1, \ldots, m)\,. \tag{1.4}$$

We shall call the (local) action–angle variables relative to the foliation into invariant tori of a completely integrable system the *action–angle variables of the system*. We remark that they are not uniquely determined. Besides re–definition of the form (1.3), there can be a deeper cause of non–non–uniqueness: if the system is 'degenerate', in the sense that it possesses additional independent integrals, the foliation into m–dimensional invariant tori can be not uniquely defined. We shall come back later on this fact.

C. Global problems for action–angle variables. We consider now some questions of globality of the action–angle variables of a completely integrable system. We are interested in such problems because of their relations with perturbation theory, which will be discussed later (section 1.3D). In this subsection we make some general comments.

First of all we stress that, in general, action–angle variables do not exist everywhere in phase space: simple examples are the harmonic oscillator and the pendulum. One easily understands that the largest set in which action–angle variables can be defined (at least locally) is the union of all the invariant tori of the maximal dimension $m = \dim M$. Let M_* be such a set. One certainly has $M_* \supseteq M_f$ for any set f of independent integrals of motion in involution. In general, not all the phase space (or the subset of compact trajectories) is foliated into invariant tori of the maximal dimension: singularities of the foliations can exist, in the neighbourhood of which action–angle variables do not exist. Typical singularities are the isolated[5)] equilibria or periodic orbits, as well as the connections of hyperbolic equilibria, see figure 1.1. Some general results about the existence of such singularities can be found in [48, 37, 26].

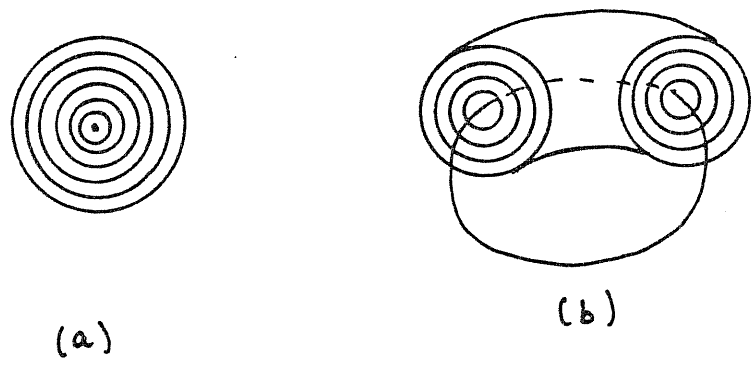

Figure 1.1

A second 'global' problem can be formulated as follows: does it exist a single action–angle coordinate system which covers all of a given subset of phase space, noticeably the maximal subset M_*? Such a problem has been first considered by Nekhoroshev [64], and then by others, in particular by Duistermaat [35]. As observed, this is the problem of the triviality of a bundle. A necessary and sufficient condition for the existence of global action–angle variables has been found[64,35,33], which seems to be applicable to specific cases. As an example, it has been shown that the spherical pendulum [35,31] and the Lagrange top [32] (see section 1.2E) do not possess global action–angle variables.[6)]

The non–existence of global action–angle variables is of no great importance as far as the dynamics of the completely integrable system is concerned. In fact, each chart domain is invariant under the flow. However, a perturbation of the flow may destroy such invariance. As better explained later, the non–existence of global action–angle variables does not constitute a real problem for perturbation theory, if it is not coupled to degeneracy.

---

5) Here, 'isolated' means that the considered equilibrium, periodic orbit or, more generally, invariant torus of lower dimension ($\leq m$) does not belong to a family which fill a torus of the maximal dimension m. For instance, a non–isolated lower dimensional torus (possibly a periodic orbit) is the closure of any resonant trajectory.

6) One can obtain some insight into the pendulum case by observing that the phase space of the plane pendulum (a cylinder) is divided by the separatrices into three *disjoint* connected components, in each of which global action–angle variables do exist; thus, one concludes that the union of these components possesses global action–angle variables, too. Such a conclusion is quite striking, since the phase curves corresponding to the oscillations are not homologous to those corresponding to the rotational motions. In the spherical pendulum, the homoclinic connections of the unstable hyperbolic equilibrium do not disconnect the phase space any more, and moreover the energy surfaces corresponding to oscillatory motions are not diffeomorphic to the rotational ones.

C. Degeneracy. A somehow different cause of non–existence of global action–angle variables is degeneracy. Roughly speaking, a completely integrable system is degenerate, if it has 'too many' integrals of motion, so that its Hamilton function does not depend on all the action variables. Precisely, we give the following:

Definition 1.1 *Let M be a $2m$–dimensional fiber bundle with fiber $\mathbb{T}^n$ and base $\mathcal{B}$, and let $h : \mathcal{B} \to \mathbb{R}$ be a smooth function. We say that:*

i) h is nondegenerate in B *if its differential $dh(b)$ is nonsingular at each point b of $\mathcal{B}$.*

ii) h is properly (or intrinsically) degenerate *if*

$$n \; := \; \sup_{b\in\mathcal{B}} \operatorname{rank} df(b) \; < \; m \; ;$$

the number n is called the number of frequencies *of h.*

As we now discuss, there are cases in which the non existence of global action–angle variables is deeply related to proper degeneracy. To this purpose, let us refer to a very simple case. Consider a Hamiltonian system on a $2m$–dimensional phase space M which possesses $m+1$ independent integrals of motion $f_1, \ldots, f_m, f_{m+1}$, and assume that there exist *two* distinct subsets of m pairwise in involution integrals, say $f' = \{f_1, \ldots, f_{m-1}, f_m\}$ and $f'' = \{f_1, \ldots, f_{m-1}, f_{m+1}\}$.

Let $M_{f'}$ and $M_{f''}$ be the corresponding properness sets. Then $M_{f'}$ is foliated into invariant tori of dimension m, and possesses action–angle variables. The same conclusion holds for $M_{f''}$. Let us also assume, for simplicity, that global action–angle variables do exist in each of these sets.[7] The interesting case is that in which both $M_{f'}$ and $M_{f''}$ are proper subsets of $M_{f'} \cup M_{f''}$.

Let us first consider the intersection of the two domains. Since f_m and f_{m+1} are independent, $M_{f'} \cap M_{f''}$ turns out to be foliated into invariant tori of dimension $m-1$ (which are the intersection of the tori the foliations corresponding to f' and f''). This shows in particular that the system is degenerate, with (at most) $n = m-1$ frequencies.

Consider now the union of the two properness sets. Certainly, action–angle variables can be constructed (at least locally) everywhere in $M_{f'} \cup M_{f''}$. However, one suspects that no global system of action–angle variables do exist therein.

We now discuss an important example.

Example: The Kepler system. The system is defined by the Hamilton function $h = \frac{1}{2}\langle p, p\rangle - \gamma^2 \langle q, q\rangle^{-1/2}$ on the phase space $(\mathbb{R}^3 \setminus \{0\}) \times \mathbb{R}^3$ (with symplectic two–form $dp \wedge dq$). The Hamilton function h and the angular momentum vector $m = q \times p$ are integrals of motions. The Hamilton function h, the modulus $G = \langle m, m\rangle^{1/2}$ of the angular momentum and the projection $J_z = \langle m, e_z\rangle$ of m along *any* direction fixed in space, with

[7] In the intersection of their domains the local action–angle variables corresponding to the two foliations need not to be related by (1.3).

unity vector e_z, are in involution. An elementary (but tedious) computation shows that the differentials dh, dG and dJ_z are everywhere linearly independent in $(\mathbb{R}^3 \setminus \{0\}) \times \mathbb{R}^3$, except in correspondence to all

* the circular orbits
* the rectilinear orbits passing through the origin
* the orbits lying in the plane orthogonal to e_z.

One sees that the properness set $M_z := M_{\{h,G,J_z\}}$ is the complement of the set $S_+ \cup S_* \cup S_z$, where S_+ is the subset of phase space in which $h \geq 0$ (the level surfaces of the three functions are there non–compact), S_* consists of circular orbits and rectilinear motions, and $S_z = \{(q,p) \in Q \times \mathbb{R}^3 : (p \times q) \times e_z = 0\}$.

A standard computation (see for instance [42]) shows that the action–angle variables corresponding to the choice h, G, J_z of the integrals of motion are globally defined in M_z. The three actions can be taken to be

$$I_1 = G\,, \qquad I_2 = \frac{\gamma^2}{\sqrt{-2h}}\,, \qquad I_3 = J_z\,. \tag{1.5}$$

Since h depends only on I_2, the system is properly degenerate, and has only one frequency.

It is now clear that action–angle variables can be constructed everywhere in the complement M_* of the 'singular' set $S_* \cup S_+$ introduced above. In fact, one has $M_* = M_{z'} \cup M_{z''}$ for any choice of $e_{z'}$ and $e_{z''}$. Furthermore, M_* is easily seen to be the larger set in which action–angle variables exist. However, no global set of action–angle variables manifestly exist in such a maximal set, because of the independence of the functions $J_{z'}$ and $J_{z''}$.

Another important example, very similar to the one above, is the 'free' rigid body with a fixed point, which we consider in the next section.

1.2 The Euler–Poinsot system

We give in this section some general informations about the 'free' motion of a rigid body about a fixed point, namely the 'Euler–Poinsot' problem ('free' means: under the influence of external forces having vanishing torque with respect to the fixed point).

A. The rigid body as a Hamiltonian system. First, we give a survey of the Hamiltonian description of the rigid body. For details, we demand to [57], or also to [7,1].

Consider two right–handed orthogonal reference frames having the body fixed point as common origin: a frame $\mathcal{B}_b = \{e_1, e_2, e_3\}$ attached to the body, and an inertial frame $\mathcal{B}_s = \{e_x, e_y, e_z\}$ fixed in space (we shall refer to them as to, respectively, the 'inertial' or 'spatial' and the 'body' or 'moving' frame); all the vectors $e_1, \ldots, e_z$ have unit euclidean norm. The configuration space of the rigid body with a fixed point is then the space of all the orientations of the two frames $\mathcal{B}_s$ and $\mathcal{B}_b$.

We obtain a particular matrix representation of the configuration space, by choosing a reference configuration. Let us take for it the configuration in which the two bases $\mathcal{B}_b$ and $\mathcal{B}_s$ coincide: $e_1 = e_x$, $e_2 = e_y$, $e_3 = e_z$. Then, each configuration is determined by the (unique) orthogonal matrix $\mathcal{R} \in SO(3)$ defined by $e_x = \mathcal{R}e_1$, $e_y = \mathcal{R}e_2$, $e_z = \mathcal{R}e_3$.

By well known facts about Lie groups, and since the Lie algebra $so(3)$ is isomorphic to $\mathbb{R}^3$ (with the cross product as Lie bracket), the tangent bundle $TSO(3)$ can be identified with $SO(3) \times \mathbb{R}^3$. Such an identification can be realized in two different ways, which correspond, respectively, to the description of the motion in the reference frame attached to the body, or in the inertial one. In the first case, the tangent vector $(\mathcal{R}, \dot{\mathcal{R}}) \in TSO(3)$ is identified with the point $(\mathcal{R}, w) \in SO(3) \times \mathbb{R}^3$, where w is the angular velocity vector in the frame attached to the body. In the 'inertial' description, instead, $(\mathcal{R}, \dot{\mathcal{R}})$ is identified with the point $(\mathcal{R}, w) \in SO(3) \times \mathbb{R}^3$, where now w is the angular velocity vector in the inertial frame.

In the present context, and especially for studying the properties of the integrals of motion, it is convenient to adopt the body description. Later on (subsection D) we shall instead turn to the inertial description. In order to avoid ambiguities, it would be advisable to employ different symbols to denote vectors of the two frames (as for instance in [7]). Unfortunately, this would involve significantly the notations. Thus, we shall not follow this usage. On the contrary, we shall tacitly identify all vectors with their representatives in the body base $\mathcal{B}_b$ until section 1.2C included, and with their representatives in the inertial base $\mathcal{B}_s$ after on.

Let A be the inertia operator of the body relative to its fixed point. We assume that the body has at least three noncollinear points, so that A is positive definite. We take, as usual, the basis vectors e_1, e_2, e_3 of the body frame $\mathcal{B}_b$ to be the eigenvectors of A, which are called *inertia axes* of the body. Thus, in the body base $\mathcal{B}_b$, A has diagonal matrix (which we continue to indicate with the same letter) $A = \mathrm{diag}(a_1, a_2, a_3)$; the eigenvalues a_j $(j = 1, 2, 3)$ are called the *principal inertia moments* of the body, relative to the fixed point.

The kinetic energy of the system is the quadratic form $(\mathcal{R}, w) \mapsto T(w)$ defined by

$$T(w) \;=\; \frac{1}{2}\,\langle w,\, Aw\rangle \;=\; \frac{1}{2}\big(a_1 w_1^2 + a_2 w_2^2 + a_3 w_3^2\big) \qquad (2.1)$$

where $w_j = \langle w,\, e_j\rangle$. $T(w)$ is the Lagrange function for the Euler Poinsot problem. The advantage of the 'body' description is obviously in the fact that the kinetic energy does not depend on the configuration $\mathcal{R}$.

We now turn to the Hamiltonian description. The Legendre transformation corresponding to the Lagrangian (2.1) is the mapping of $SO(3) \times \mathbb{R}^3$ onto itself defined by

$$(\mathcal{R}, w) \;\mapsto\; (\mathcal{R}, Aw)\;. \qquad (2.2)$$

The phase space is again $SO(3) \times \mathbb{R}^3$, the cotangent vector $m = Aw$ being the angular momentum vector in the body frame.[8)] The Hamilton function $(\mathcal{R}, m) \mapsto k(m)$ is given by

$$k(m) = \frac{1}{2}\langle m, A^{-1}m\rangle = \frac{1}{2}\left(\frac{m_1^2}{a_1} + \frac{m_2^2}{a_2} + \frac{m_3^3}{a_3}\right). \tag{2.3}$$

B. Integrals of motion. As is well known, the Hamilton function k and the angular momentum *in space* ($\mathcal{R}^t m$, with the notations above) are integrals of motion of the Euler–Poinsot system.[9)] In full analogy with the case of the Kepler system, the Hamilton function $k(m)$, the modulus $G(m) = \langle m, m\rangle^{1/2}$ and the projection $J_\zeta(\mathcal{R}, m) = \langle m, \mathcal{R}e_\zeta\rangle$ of the angular momentum along *any* direction e_ζ fixed in space are integrals of motion, and are pairwise in involution.

Let us first consider the general case of a body which has all the inertia moments distinct ('tri–axial' body; we shall understood that a_2 is the middle inertia moment). Since we are not aware of any precise statement about the properness set of the functions $k(m)$, $G(m)$ and $J_\zeta(\mathcal{R}, m)$, which we indicate Σ_ζ^T (T standing for tri–axial), we state here the following[10)]

Proposition 1.2 *The properness set Σ_ζ^T of the functions k, G, J_ζ is the complement of the set $S_*^T \cup S_\zeta$, where*

$$\begin{aligned} S_*^T &= \{(\mathcal{R}, m) \in SO(3) \times \mathbb{R}^3 : 2a_i k(m) = G(m)^2 \text{ for some } i = 1,2,3\} \\ S_\zeta &= \{(\mathcal{R}, m) \in SO(3) \times \mathbb{R}^3 : m \times \mathcal{R}e_\zeta = 0\}. \end{aligned} \tag{2.4}$$

The set Σ_ζ is the union of four connected components.

The proof of this proposition is deferred to the appendix B, at the end of the chapter.

One should notice that the 'singular' set S_*^T consists of:

---

8) In principle, the Euler–Poinsot system is a Hamiltonian system in the phase space $T^*SO(3)$, with its natural symplectic structure of the cotangent bundle. Here, an identification of $T^*SO(3)$ with $SO(3)\times\mathbb{R}^3$ has been obtained by identifying first $TSO(3)$ with $SO(3)\times\mathbb{R}^3$, and then each cotangent space with the corresponding tangent space by means of the euclidean structure of the latter one. Essentially, all these identifications allow one to consider as defined in one and the same space $\mathbb{R}^3$ (and then to draw in a same picture) the body, its angular velocity vector and its angular momentum vector (in fact, the latter vector is identified with the angular velocity vector of a body with unity inertia operator).

9) In the present formulation (in which we do not have an explicit expression for the symplectic two form of the system, and then we cannot explicitly write the Hamilton's equations) the constancy of m could be justified on the basis of general arguments of symmetry: see [7, 1, 57].

10) For instance, in [64] it is stated that there is also a singularity for $J_\zeta = 0$, and that the action–angle variables are defined in a set which is the union of eight (instead than four) connected components, while in [51] there appears a singularity for $L = 0$.

* the equilibria,
* the steady rotations about the three inertia axes,
* the four branches of the separatrices,

while S_ζ contains all the states with the angular momentum parallel to the axis e_ζ. Some insight on the singularities in the set S_*^T can be obtained from the (classical) figure 1.2a, which shows the intersection curves, in the cotangent space $m \in \mathbb{R}^3$, of two (nonzero) level surfaces of the functions $G(m)$ and $k(m)$. Notice in particular the steady rotations about the three inertia axes, and the 'separatrices' connecting the (unstable) rotations about the middle inertia axis.

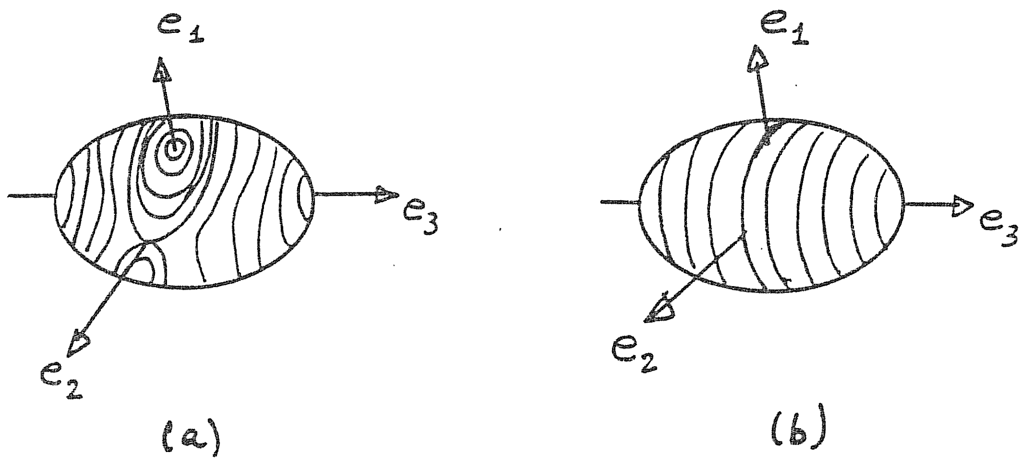

Figure 1.2

The action–angle variables for the (free) rigid body with a fixed point have been extensively studied (see [3, 64, 69, 51, 53]). Here, we merely present some basic facts. The three actions corresponding to the foliations of the functions k, G and J_ζ are globally defined in each connected components of Σ_ζ^T, and can be taken to be

$$I_1 = G\,, \qquad I_2 = I_2(k, G)\,, \qquad I_3 = J_\zeta\,, \tag{2.5}$$

where $I_2(k, G)$ is a certain (analytic) function. The Hamilton function is a function of only two of the action variables (G and I_2), so that the system is degenerate, with two frequencies.

Exactly as in the case of the Kepler system, action–angle variables do exist in the maximal set Σ_*^T, which is the complement of the set S_*^T, but they are not globally defined therein.

C. The symmetric case. A rigid body is said to be symmetric if two inertia moments coincide, say $a_1 = a_2$. In such a case, e_3 is the symmetry axis of the inertia ellipsoid (also called *gyroscopic* axis of the body), while e_1 and e_2 are not uniquely determined.

In the symmetric case, the Euler–Poinsot system has an additional integral of motion, namely the projection $L = \langle m, e_3 \rangle$ of the angular momentum along the gyroscopic axis e_3. The function $L(m)$ is pairwise in involution with all of the other integrals k, G, J_ζ (for any ζ) but it is not independent of them: since $A = \text{diag}(a_1, a_1, a_3)$, one immediately deduces from (2.3)

$$k = \frac{G^2 - L^2}{2a_1} + \frac{L^2}{2a_3}. \tag{2.6}$$

In the symmetric case, as independent integrals of motion in involution it is convenient to take G, L and J_ζ (for a reason which will be clear in a moment).

Proposition 1.3 *The properness set Σ_ζ of the functions G, L, J_z is connected, and is the complement of the set $S_* \cup S_\zeta$, where*

$$\begin{aligned} S_* &= \{(\mathcal{R}, m) \in SO(3) \times \mathbb{R}^3 : G(m) = |L(m)|\} \\ S_\zeta &= \{(\mathcal{R}, m) \in SO(3) \times \mathbb{R}^3 : m \times \mathcal{R}e_\zeta = 0\}. \end{aligned} \tag{2.7}$$

The proof of this proposition is quite similar to that of proposition 1.2, and is therefore omitted.

The singular set S_* contains now only

* the equilibria,
* the steady rotations about the symmetry inertia axes e_3.

The action–angle variables for the symmetric rigid body, in the set Σ_ζ, are well known. The actions can be taken to be G, L, J_ζ, while the corresponding angles g, l, j_ζ are illustrated in figure 1.3a (where the axis e_ζ is taken to be the axis e_z of the inertial frame; figure 1.3b shows, for comparison, the familiar Euler angles). As far as we know, the first clear introduction of these variables is due to Andoyer [2]. Later, Deprit proved in a simple way their canonicity [34]. For this reason, they are sometimes called in the recent literature *Deprit's variables*. We shall prefer to call them *Poinsot's variables*, because of their deep relation with Poinsot's description of motion (the angles g and l are angular coordinates on the two Poinsot cones). An exhaustive study of these coordinates (with particular attention to their singularities) will be accomplished in chapter 3.

Once again, action–angle variables exist in a maximal set Σ_*, which is the complement of S_*, but they are not global. It is of interest to notice that they are defined also in correspondence of the steady rotations about the degenerate inertia axes. The reason is that such periodic orbits are not isolated in phase space but, as one can easily argue, grouped together to form a three–dimensional invariant torus, which constitutes one of the leaves of the foliation into invariant tori. Some insight can be drawn from figure 1.2b, which is the analog of figure 1.2a. Notice in particular that the four heteroclinic connections of the tri–axial case are collapsed into the above family of periodic orbits (such a collapse is quite similar to that of the two separatrices of the pendulum as gravity goes to zero).[11]

---

11) The differentials of G, k, J_ζ are linearly dependent at each point of such orbits (they inherit

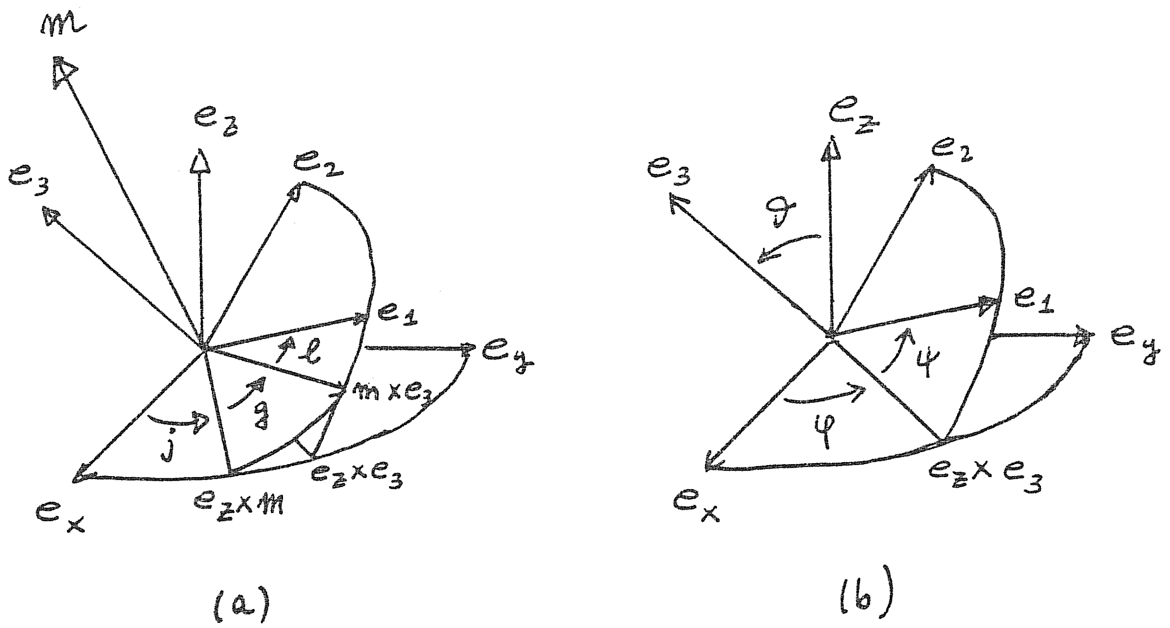

Figure 1.3

D. Poinsot's geometric description. We now turn to the description of the body motion in the inertial frame, i.e. in space. A deep insight into the motions of the free rigid body is furnished by Poinsot's classic geometric description.

Poinsot's theorem can be stated as follows: think to the inertia ellipsoid of the body relative to the fixed point as having its center in that point, and as being rigidly connected to the body; then, the inertia ellipsoid rolls without sliding on a fixed plane, orthogonal to the angular momentum vector (figure 1.4a). The point of contact draws a closed curve on the ellipsoid and another one on the plane. These two curves, called respectively *polhode* and *erpolhode*, are sections of the two *Poinsot cones*: the conical surfaces described by the angular velocity vector in the space and in the body.[12)]

Poinsot's description makes clear that the motions are quasiperiodic, with two frequencies. The motion can be thought as a composition of a rotation about the angular momentum direction, and a proper rotation of the body (about an instantaneous axis which is not fixed in the body, in the case of a tri–axial inertia ellipsoid). The corresponding frequencies have a very simple interpretation ([67], vol. 1, sect. 86):

$$\omega_1 = \frac{\alpha}{P}, \qquad \omega_2 = \frac{2\pi}{P}$$

---

from the tri–axial case the singularity on the separatrices), while this does not happen if L is considered in place of k, or of G.

12) The intersection curves of figure 1.2a look very similar to the polhodes, which are obtained with the dilatation $x_i \mapsto \sqrt{a_i}x_i$ of the coordinates along the three axis e_i.

where P is the period of the proper rotation, namely the time between two successive contacts of a same point of the polhode with the erpolhode, and α is the angle which is formed in the invariable plane by the two points of contact, and has the vertex in the foot of the angular momentum vector (see figure 1.4a).

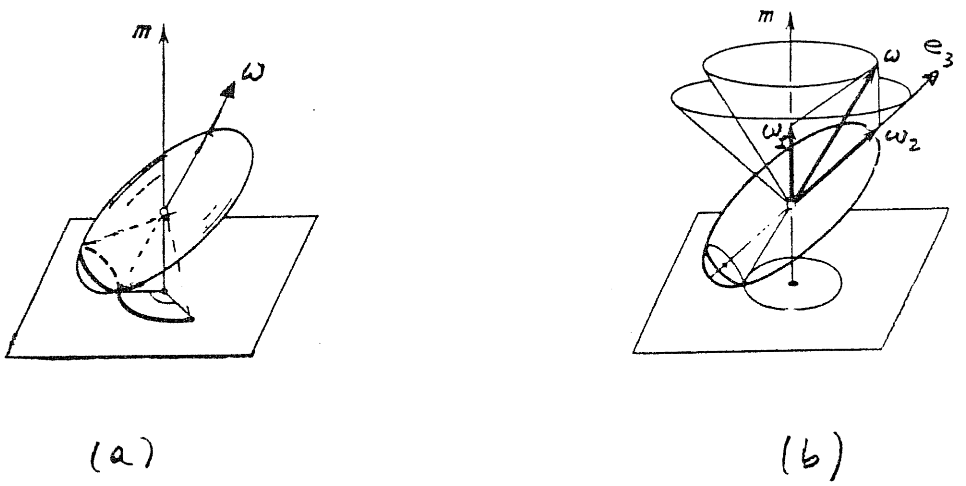

Figure 1.4

The description of the motion simplify somehow in the symmetric case (figure 1.4b). In such a case, polhode and erpolhode are circles. Moreover, as is well known[7]: the modulus $\langle w, w\rangle^{1/2}$ and the gyroscopic component $\langle w, e_3\rangle$ of the angular velocity w are constant; the vectors m, w, e_3 always lie in a plane; the angle between m and w, and that between w and e_3 are both constant (figure 1.4b). Using these facts, one easily verifies that the angular velocity w is given by

$$w = \omega_1 \mu + \omega_2 e_3 , \tag{2.8}$$

where μ is the unit vector in the direction of the angular momentum m and the two frequencies of the motion ω_1 and ω_2 are given by[13)]

$$\omega_1 = \frac{1}{a_1} G , \qquad \omega_2 = \frac{a_1 - a_3}{a_1 a_3} L . \tag{2.9}$$

Such a motion is called a 'regular precession' (see [56]). Equation (2.8) indicates that, during the motion, the vector w rotates uniformly on the fixed and on the moving circular Poinsot cones, with angular frequencies given, respectively, by ω_1 and ω_2. The motion is the composition of a uniform 'proper rotation' of the body about the symmetry axis

13) They are more simply obtained from the expression (2.6) of the Hamilton function, if one knows that G and L are the actions.

e_3 (with proper[14]) period $2\pi/\omega_2$), and by a uniform 'precession' of the gyroscopic axis e_3 about the angular momentum direction (with period $2\pi/\omega_1$).

E. The Lagrange top. We conclude this section by considering, for comparison, one of the other classical cases of integrability. As noticed in the introduction, in the motion of a 'heavy' rigid body there exist always two integrals of motion: the Hamilton function and the projection $J_z = \langle m, e_z \rangle$ of the angular momentum along the 'vertical' e_z, i.e. the direction of gravity. At difference with the Euler–Poinsot case, the components of m along any other axis $e_{z'} \neq e_z$ is now no longer constant: the force field has removed the degeneracy. The simplest case in which there exists a third independent (in an *open* subset of phase space) integral in involution with J_z, is that of a *Lagrange top* (or *heavy gyroscope*): the inertia ellipsoid of the body relative to the fixed point is symmetric, and the center of mass of the body lies on the inertia symmetry axis e_3. The third integral is the 'gyroscopic' component of m, $L = \langle m, e_3 \rangle$.

The motions of the Lagrange top are quasi–periodic with three frequencies: the body rotates about its symmetry axis e_3 ('proper rotation'), which in turn oscillates in the vertical plane ('nutation') and rotates about the vertical ('precession').[15]

1.3 Classical perturbation theory

A. Nearly integrable systems.[16] Classical perturbation theory deals with nearly integrable systems, i.e. Hamiltonian systems which differ 'little' form completely integrable ones. Employing action–angle variables, the Hamilton function of a nearly integrable system has the form, in the simplest case,

$$h(I,\varphi) = h_0(I) + \varepsilon f(I,\varphi) \tag{3.1}$$

where ε is a small parameter. We stress that ε should be considered a *local* parameter, defined by the condition that one has $\|h_0\| \approx \|f\|$ in the region of phase space of interest, where $\| \;\|$ is an appropriate norm (for instance the supremum norm or, better, some norm related to the derivatives of the functions) evaluated in the region of phase space of

---

14) Measured in a reference frame which rotates in space with angular velocity $\omega_2 \mu$

15) Notice that the terminology is different from that of the Euler–Poinsot case: the frequency of the 'precession' of the Lagrange top vanishes in the Euler–Poinsot case, while the 'nutation' of the Lagrange top is somehow similar to the 'precession' of the free body.

16) "Problème générale de la dynamique" (Poincaré [67]). "Such an approach may be open to criticism, but by means of it one can obtain some interesting results" (Kolmogorov [50]).

interest. Here, we disregard for the moment all the questions of globality of the action–angle variables, and consider the function (3.1) as defined in the phase space $B \times \mathbb{T}^n$, B being an open domain in $\mathbb{R}^n$.

The dynamics of nearly integrable systems is ultimately complex, being characterized by a deep coexistence of integrable and non–integrable behaviour. A significative example of a non–integrability result is the famous theorem by Poincaré on the non–existence of integrals of motion in nearly integrable systems[67,14,9]. Such a theorem properly applies to the case of a nondegenerate unperturbed Hamiltonian (see definition 1.1):

$$\det\left(\frac{\partial^2 h_0}{\partial I \partial I}(I)\right) \neq 0 \qquad \forall I \in B\ . \tag{3.2}$$

Roughly speaking, in such a case the theorem implies that the foliation of the phase space into invariant tori does not survive a (generic) perturbation, no matter how small it is. Nevertheless, as classical perturbation theory shows, some integrability properties still survive small perturbations (KAM theorem [49, 3, 4, 7, 61, 59, 68, 29]), and moreover the appearance of non–integrability behaviours may require extremely long time scales (Nekhoroshev's theorem [63, 65, 66, 15]). In fact, under suitable conditions:

i) not all the invariant tori are destroyed: on the contrary, the majority of them (in the sense of the Lebesgue measure) survive small perturbations, being only slightly deformed (KAM);

ii) the motions starting on a destroyed torus remain close to such a torus for extremely long times (growing faster than any power of $1/\varepsilon$: Nekhoroshev's theorem).

The main object of classical perturbation theory is to study the effect of perturbations on the long–time behaviour of the integrals of motion of the unperturbed system, more specifically of the action variables $I_1, \ldots, I_n$. The main tool of perturbation theory is canonical transformation theory, together with averaging techniques. We do not enter here into the details: a deep and exhaustive introduction is chapter 5 of [9]. Some indications will be given in the next, more technical chapters.

B. KAM theorem. As is well known, the invariant tori of a completely integrable system which survive a small perturbation, called *Kolmogorov tori*[9], evaluated in the region of phase space of interest, are those with strongly non–resonant frequencies $\omega = \frac{\partial h_0}{\partial I}$; for instance, those which satisfy a Diophantine–like condition of the form

$$|\langle \omega(I),\, \lambda \rangle| \ \geq\ \gamma(\varepsilon)\, \|\omega(I)\|\, \|\lambda\|^{-\tau} \qquad \forall \lambda \in \mathbb{Z}^n \setminus \{0\}\ , \tag{3.3}$$

where $\|\ \|$ is some norm, τ a constant and $\gamma(\varepsilon) \to 0$ as $\varepsilon \to 0$.[17)] A possible formulation of KAM theorem is the following[68,29]:

Proposition 1.4 (KAM) *Assume that $h(I,\varphi) = h_0 + \varepsilon f(I,\varphi)$ is analytic in a complex neighbourhood of $B \times \mathbb{T}^n$, and that h_0 is nondegenerate in B. Then, there exist:*

[17)] The estimates (3.5) below are obtained with $\gamma \sim \sqrt{\varepsilon}$.

i) a positive number ε^*, which depends on the domain B and on general properties of the functions h_0 and f,

ii) a subset $B_\varepsilon \subset B$ of large relative Lebesgue measure,

iii) a smooth canonical transformation $(I,\varphi) = \mathcal{C}_\varepsilon(J,\psi)$ of $B \times \mathbb{T}^n$, which is close to the identity, and

iv) a smooth nondegenerate completely integrable Hamiltonian $h'_\varepsilon(J)$ defined on $B \times \mathbb{T}^n$,

with the following properties: if $\varepsilon \leq \varepsilon^*$, then one has

$$\left(h_0 + \varepsilon f\right) \circ \mathcal{C}_\varepsilon \overset{B_\varepsilon}{=} h'_\varepsilon \,, \tag{3.4}$$

where $\overset{B_\varepsilon}{=}$ means equality of the functions at the two members, and of all their derivatives, whenever $J \in B_\varepsilon$. Moreover, B_ε, $\mathcal{C}_\varepsilon$ and h'_ε satisfy

$$\|J - I\| \leq \text{const}\sqrt{\frac{\varepsilon}{\varepsilon^*}} \tag{3.5a}$$

$$\|\psi - \varphi\| \leq \text{const}\sqrt{\frac{\varepsilon}{\varepsilon^*}} \tag{3.5b}$$

$$\mu(B \setminus B_\varepsilon) \leq \text{const}\sqrt{\frac{\varepsilon}{\varepsilon^*}}\,\mu(B)\,, \tag{3.5c}$$

$$h'_\varepsilon = h_0 + \varepsilon \overline{f} + o\left(\frac{\varepsilon}{\varepsilon^*}\right) \tag{3.5d}$$

where $\|\ \|$ is some norm in $\mathbb{R}^n$ and $\overline{f} = (2\pi)^{-n}\int_{\mathbb{T}^n} f d\varphi$. The constants appearing in (3.5) do not depend on ε, and are uniform in $B \times \mathbb{T}^n$.

The union $\mathcal{C}_\varepsilon^{-1}(B_\varepsilon \times \mathbb{T}^n)$ of the Kolmogorov tori is the *Kolmogorov set.* The Kolmogorov set, as constructed in the theorem, has large measure but, as is well known, its complement is open and dense. One can notice that, in the above formulation, KAM theorem is an integrability result (although on a nowhere dense set).

We are here mostly interested in the time variations of the action variables. It follows from proposition 1.4 that one has

$$\|I(t) - I(0)\| \leq \text{const}\sqrt{\frac{\varepsilon}{\varepsilon^*}} \qquad \forall t \in \mathbb{R}\,, \tag{3.6}$$

for the motion on Kolmogorov tori.

As is well known, there is a special case in which the estimate (3.6) holds for all the motions, not only for the 'majority' of them, so that KAM implies (topological) stability. In systems with $n = 2$ degrees of freedom the foliation into the (two–dimensional) Kolmogorov tori constitutes a topological obstruction to the variation of the actions, since they divide the (three–dimensional) energy surface in which are embedded. One can show that, if the unperturbed Hamiltonian h_0 satisfies the 'isoenergetic' nondegeneracy condition

$$\det\begin{pmatrix} \frac{\partial^2 h_0}{\partial I \partial I}(I) & \frac{\partial h_0}{\partial I}(I) \\ \frac{\partial h_0}{\partial I}(I) & 0 \end{pmatrix} \neq 0 \qquad \forall I \in B\,, \tag{3.7}$$

then the majority of the invariant tori on each energy surface will survive the perturbation. In such a case, as a corollary of proposition 1.4 one obtains the following

Proposition 1.5 (Arnold) *Assume that $h = h_0 + \varepsilon f(I,\varphi)$ is analytic in a complex neighbourhood of $B \times \mathbb{T}^2$, $B \subset \mathbb{R}^2$, and that h_0 satisfies the isoenergetic nondegeneracy condition (3.7). Then, if ε is sufficiently small, i.e. $\varepsilon \leq \varepsilon^*$ for some positive ε^* dependent on B, h_0 and f, one has*

$$\|I(t) - I(0)\| \leq \text{const}\sqrt{\frac{\varepsilon}{\varepsilon^*}} \tag{3.8}$$

for all $t \in \mathbb{R}$ and all $I(0) \in B$.

The estimate $\mathcal{O}(\sqrt{\varepsilon})$ in (3.8) cannot be improved (see [62]).

When the number of degrees of freedom exceeds two, such a topological obstruction does not exists any more: n–dimensional tori do not disconnect the $(2n-1)$–dimensional energy submanifolds, if $n \geq 3$. In other words, the complement of the Kolmogorov set is connected. Thus, for a set of initial conditions of small measure, but open and dense, the motions can in principle wander on the energy submanifold ('Arnold diffusion'). In a known, nontrivial although very special example[5,6] Arnold diffusion exists, although with an extremely slow average velocity, $\mathcal{O}(\exp - \sqrt{\varepsilon^*/\varepsilon})$. Nekhoroshev's theorem assures that, in generic situations, Arnold diffusion is a very slow phenomenon: $\|I(t) - I(0)\|$ is bounded by a power of ε, for times which grows exponentially fast with (a power of) $\varepsilon^*/\varepsilon$.

C. Nekhoroshev's theorem. At difference from KAM theorem, Nekhoroshev's theorem is a stability result for finite (though extremely long) times, but for all the initial conditions. For such a reason, Nekhoroshev's theorem seems to be more significant than KAM for physical applications (for a number of applications, see [13,17,19,41,20,21]). In its more general formulation, Nekhoroshev's theorem applies to analytic Hamiltonian perturbations of analytic, completely integrable Hamiltonians which satisfy a generic property called 'steepness' (see [65,45]). The simplest examples of steep functions are the convex functions and, more generally, the 'quasi–convex' functions. The latter are the functions whose level sets are boundaries of strictly convex sets, and have contact of the lowest order with the tangent plane. Precisely, one gives the following definition[65]:

Definition 1.2 *Let B and $B' \subset B$ be domains of $\mathbb{R}^n$. A smooth function $h : B \to \mathbb{R}$ is said to be* quasi–convex in B' *if for any $I \in B'$ there exists a constant $c(I)$, uniformly bounded away from zero and of constant sign in B', such that one has*

$$\frac{1}{c(I)}\left\langle \xi,\, \frac{\partial^2 h}{\partial I \partial I}(I)\xi \right\rangle \geq \langle \xi,\, \xi \rangle$$

for all $\xi \in \mathbb{R}^n$ such that $\left\langle \xi,\, \frac{\partial h}{\partial I}(I) \right\rangle = 0$.

For instance, the function $k(x_1, x_2) = x_1^2 + \eta x_2^2$, $(x_1, x_2) \in \mathbb{R}^2$, is convex if $\eta > 0$, quasi–convex in the domain $\{x_1 > |x_2|\}$ if $\eta < 0$.

It is convenient to enunciate Nekhoroshev's theorem with reference to a Hamiltonian ('with parameters', in Nekhoroshev's words[65]) of the form

$$h(I,\varphi,p,q) \;=\; h_0(I) \;+\; \varepsilon f(I,\varphi,p,q) \;, \tag{3.9}$$

defined for $I \in B$, a domain in $\mathbb{R}^n$, $\varphi \in \mathbb{T}^n$, and $(p,q) \in U$, a domain in $\mathbb{R}^{2(m-n)}$, for some $m \geq n > 0$ [18] (the symplectic two form is understood to be $\sum dI_j \wedge d\varphi_j + \sum dp_j \wedge dq_j$). Hamiltonians of the form (3.9) are encountered in two important cases.

First, consider a completely integrable system described by $h_0(I)$ in $B \times \mathbb{T}^n$, coupled by an interaction term $v(I,\varphi,p,q)$ to a system with $m-n$ degrees of freedom described by a certain Hamilton function $g(p,q)$. The Hamiltonian describing the interacting system has then the form

$$h(I,\varphi,p,q) \;=\; h_0(I) \;+\; g(p,q) \;+\; v(I,\varphi,p,q) \;, \tag{3.10}$$

and reduces to the form (3.9) in regions of phase space in which $\|h_0\| \gg \|g\|, \|v\|$. An example is a fast rotating rigid body, with no fixed point: in such a case h_0 is the kinetic energy of the motion relative to the center of mass (described by the action–angle coordinates I,φ), g the kinetic energy of the center of mass motion (described by the cartesian coordinates $(p,q) \in \mathbb{R}^6$), and v the potential energy of the external forces. Other cases are considered in [17, 19, 21].

Another important class of Hamiltonians of the form (3.9) is constituted by the small perturbations of properly degenerate completely integrable systems. A perturbation of a properly degenerate Hamilton function $h_0 = h_0(I_1,\ldots,I_n)$ has the form (3.9), if one interprets $I = (I_1,\ldots,I_n)$ and $\varphi \in \mathbb{T}^n$ as the action–angle variables corresponding to the nonzero frequencies, the p's as the other $m-n$ actions (on which h_0 does not depend) and $q(\mathrm{mod}2\pi) \in \mathbb{T}^{m-n}$ as the correspondent angles.

We now state Nekhoroshev's theorem:

Proposition 1.6 (Nekhoroshev) *Assume that the Hamilton function $h = h_0 + \varepsilon f$ as in (3.9) is analytic in a domain $B \times \mathbb{T}^{n-m} \times U$, that h_0 is quasi-convex (or more generally steep) in B and that ε is sufficiently small, i.e. $\varepsilon \leq \varepsilon^*$ for some $\varepsilon^*(B,h_0,f) > 0$. Then, for any motion $z(t) = \big(I(t),\varphi(t),p(t),q(t)\big)$ one has*

$$\|I(t) - I(0)\| \;\leq\; \left(\frac{\varepsilon}{\varepsilon^*}\right)^a \tag{3.11}$$

for all times $|t| \leq \min(T_{\rm esc}, T_0)$, where

$$T_0 \;=\; \mathcal{T} \exp\left[\left(\frac{\varepsilon^*}{\varepsilon}\right)^b\right] , \tag{3.12}$$

[18] Thus, the system has m degrees of freedom, but only n frequencies. We make the obvious convention that, if $m = n$, then the phase space reduces to $B \times \mathbb{T}^n$, with action–angle coordinates (I,φ).

while $T_{\rm esc}$ is the escape time of the solution from $B \times \mathbb{T}^n \times U$, and ε^*, a, b, $\mathcal{T}$ are positive constants, independent of ε.

Within the proof of the theorem, one can produce estimates for the various constants entering the statement. However, as is typical of perturbation theory, such estimates are rather dissatisfying. Especially bad is the dependence on the number of frequencies. The most important constants are the two exponents a and b, whose estimated values depend on the steepness properties of h_0, and are better for convex and quasi–convex functions. For example, if h_0 is convex, the following estimates have been produced[15]:

$$a = \frac{1}{16}\,, \qquad b = \mathcal{O}\left(\frac{1}{n^2}\right)\,. \tag{3.13}$$

Such values can be certainly somehow improved. However, a n–dependence of b seems to be unavoidable. Concerning a, it is perhaps possible to obtain $a = 1/4$, or even $a = 1/2$.[19)]

Proposition 1.6 allows immediate conclusions for nearly integrable Hamiltonians $h = h_0 + \varepsilon f$ of the form (3.1), namely without additional variables p, q. In such a case, one has clearly $T_{\rm esc} > T_0$ for all the motions $z(t) = (I(t), \varphi(t))$ with $I(0)$ not too near the border of B (by (3.11), a distance $\mathcal{O}(\varepsilon^a)$ is enough).

However, Nekhoroshev's theorem does not give any information about the motion of the 'additional' variables (p, q), when present. The escape of such variables from their own domain of definition U can prevent the confinement of the actions for exponentially long times. In some cases one can obtain some *a priori* control on such variables (typically by energy conservation) so to be sure that, at least for a large, significative set of initial conditions, they do not escape their own domain of definition for times $\leq T_0$. This situation is in fact rather common, if (p, q) are the canonical variables of an interacting subsystem[17,19,21]. The case of proper degeneracy can instead be quite different, as discussed in subsection E below.

Remark 1.2: The possibility of studying systems of the form (3.9) is one of the advantages of Nekhoroshev's theory over KAM theory, which cannot in general be applied to such systems. In the case of proper degeneracy, KAM theory works if the perturbation has very special properties, precisely if it 'removes the degeneracy'. Let, as above, $I = (I_1, ..., I_m)$ and $p = (p_1, ..., p_{m-n})$ be the action variables of the system, and $\varphi = (\varphi_1, ..., \varphi_m)$ and $q = (q_1, ..., q_{m-n})$ be the corresponding angle variables. Then, it is required that the Hamiltonian can be given the form $h_0(I) + \varepsilon f_0(I, p) + \varepsilon^2 f_1(I, p, \varphi, q; \varepsilon)$, with h_0 nondegenerate, v_0 nondegenerate as a function of the actions $p_1, \ldots, p_{m-n}$, and f_1 divisible by ε (in other words: by averaging the perturbation over the angles φ, the q coordinates disappear). For details, see [4,9]. By the way, the topological obstruction which led to the stability result of proposition 1.5 does not exist any more, as far as $m \geq 3$, even if $n = 2$.

D. Global problems. We now discuss the problems connected to questions of globality of the action–angle variables.

---

19) In the treatment of the rigid body of chapters 4 and 5, we shall actually obtain $a = 1/2$, but for the rather special case in which $n = 2$.

First, let us consider the singularities of the foliation into invariant tori, in the neighbourhood of which action–angle variables do not exist. Such singularities have a certain relevance for perturbation theory, because the perturbative study in a neighbourhood of them has to be performed by specific techniques (*not* using action–angle variables, of course). For instance, the study in the neighbourhood of isolated equilibria and periodic orbits (sometimes called 'limiting degeneracy' [4]) are classical topics in perturbation theory (Birkhoff's series, Poincaré sections; for the case of invariant tori of higher dimensions see [28]).[20)]

We now turn to the problems posed by the non–existence of global action–angle variables. Preliminary, notice that each action–angle chart domain is the union of invariant tori. Thus, the escape of the system from one of such domains can be caused only by the variation of the actions.

In connection with this problem, degeneracy plays a central role. In fact, as far as nondegenerate completely integrable systems are concerned, the theorems of perturbation theory indicate that all the action variables are essentially constant (at least on the time intervals of interest), so that it can be assured that each motion does not leave some chart domain (it is sufficient for this that the chart domains have sufficiently fat intersections).

However, things are very different in the case of degeneracy. The normal forms constructed by perturbation theory do not allow to control the variations of those action variables (previously denoted by p) which correspond to the null frequencies[†]. This situation is dangerous if, as in Kepler's or Euler–Poinsot's cases, the action variables which cannot be controlled are exactly those which are not globally defined. In that case, one can be unable to control the escape time of such actions from their own domains of definition. As a consequence, the statement of Nekhoroshev's theorem (as well as of KAM) becomes meaningless: $T_{\rm esc}$ could be an extremely short time, if compared with T_0. Moreover, one can even not to have any control at all on $T_{\rm esc}$.

This problem is encountered in the perturbative studies of the (non–planar) Kepler system, where one has to be sure that the angular momentum does not pass through (or too near to) the chosen z–axis. Even worse, in the n–body problem one has to check that this does not happen for any one of the planets. In Nekhoroshev's study of the n–body problem[65], a great care is posed in determining a set of initial conditions for which the motions do not fall into such singularities (for the time interval of interest). In much the same way, in the rigid body case one can assure in general that the angular momentum vector will not pass through (or too near to) a given spatial direction only for a subset of initial conditions (all those which are sufficiently far from resonances). We find such a fact unsatisfactory.

Motivated by this fact, we shall reconsider this problem from a general point of view. Chapter 2 is devoted to this problem. The source of the difficulties is in the fact that the

---

20) If the excluded neighbourhood of the singularity is small, and some mechanism (like conservation of energy) confine the motions in its interior, one can avoid a detail study in such a neighbourhood: this will be the case for the 'gyroscopic' rotations of the rigid body, see section 5.3.

† There is no 'fast' motion over which to average: it is like within a resonance.

normal forms of perturbation theory are constructed within a single coordinates system. It is then not clear whether it is possible to reconnect the descriptions obtained with different coordinate systems (the idea being simply that of changing chart, if the motion arrives near a singularity of the considered chart). The solution of this problem is obtained by showing that, under very mild conditions, the normal forms constructed in the different charts are the local representatives of a unique Hamiltonian system, intrinsically defined on the phase space.

A geometric formulation of perturbation theory has been studied by Moser [60] and Cushman [30], but only for the very special case of the perturbation of a 'completely degenerate' Hamiltonian h_0, which depends only on one of the action variables. In such a case (which is that of Kepler's system, but *not* of the n–body problem), all the orbits of the unperturbed system are periodic. The geometric formulation of perturbation theory is then essentially obtained by replacing the space averages with the corresponding time averages along the motions.

It turns out that such a simple method cannot be used in the case of systems having more than one frequencies. The appropriate setting for the study of the general case is a theory by Nekhoroshev [64] of degenerate integrable systems, which is essentially a generalization of the Arnold–Liouville theorem.

1.4 Fast motions of the rigid body and classical perturbation theory

A. Fast motions of the rigid body. We present in this section the basic ideas of the study of the fast motions of the rigid body with a fixed point by classical perturbation theory.

Let us preliminarily refer to the Euler angles φ, ψ, θ, which are defined in a well known way with reference to the two orthogonal frames introduced in section 1.2A (see figure 1.3b): φ is the angle between e_1 and $e_z \times e_3$, ψ is the angle between $e_z \times e_3$ and e_1, while $0 < \theta < \pi$ is the angle between e_z and e_3. The conjugate momenta are $p_\varphi = \langle m, e_z \rangle \equiv J_z$, $p_\psi = \langle m, e_3 \rangle$ and $p_\theta = \langle m, e_n \rangle$, where $e_n = (\sin\theta)^{-1} e_z \times e_3$ is the unit vector along the nodal line. The kinetic energy T is a function of the three momenta (on which it depends quadratically) and of the *two* angles ψ, θ. If the external forces acting on the body are positional and conservative, with potential energy $\mathcal{V}(\varphi, \psi, \theta)$, then the system is Hamiltonian (with respect to the symplectic two–form $dp_\varphi \wedge d\varphi + dp_\psi \wedge d\psi + dp_\theta \wedge d\theta$), with Hamilton function

$$T(p_\varphi, p_\psi, p_\theta, \psi, \theta) \; + \; \mathcal{V}(\varphi, \psi, \theta) \; . \tag{4.1}$$

An interesting (both for applications and in itself) problem of rigid body dynamics is that of the fast motion of a body, namely the motion with sufficiently large angular velocity. Roughly speaking, such motions can be characterized by the condition that the kinetic energy is much larger than the potential one: $|T| \gg |\mathcal{V}|$.[21)]

[21)] A more appropriate characterization requires the comparison of the derivatives of T and $\mathcal{V}$.

Such a problem can be naturally studied by perturbative techniques, since the Hamilton function $T + \mathcal{V}$ can be regarded as a 'small' perturbation of the Hamiltonian T of the Euler–Poinsot case. The 'small parameter' for the problem is

$$\varepsilon = \frac{\|v\|}{\|T\|} \tag{4.2}$$

where $\| \; \|$ is some appropriate norm.

The problem can be studied from two equivalent points of view. On one hand (as we shall do) one can work in regions of phase space in which the angular velocity w – more precisely some norm $\omega \equiv \|w\|$ of it – is large. In such regions, one has

$$p = \mathcal{O}(\omega^1)\,, \quad q = \mathcal{O}(\omega^0)\,, \quad T(p,q) = \mathcal{O}(\omega^2)\,, \quad \mathcal{V}(q) = \mathcal{O}(\omega^0)\,, \tag{4.3}$$

where $p = (p_\varphi, p_\psi, p_\theta)$, $q = (\varphi, \psi, \theta)$. On the other hand, one can rescale the coordinates, so to work in a region of phase space in which w, p and T are $\mathcal{O}(1)$, but the potential energy is small, precisely $\mathcal{O}(\omega^{-2})$. Indeed, consider the rescaling of variables

$$p = \lambda p'\,, \quad q = q'\,, \quad t = \lambda t' \qquad (\lambda > 0)\,. \tag{4.4}$$

Since T is quadratic in the $p's$ and $\mathcal{V}$ is independent of them, the transformation (4.4) conjugate (4.1) to

$$T(p', q') + \varepsilon \mathcal{V}(q')\,, \qquad \varepsilon = \lambda^{-2}\,. \tag{4.5}$$

If λ is proportional to ω, then one has $p' = \mathcal{O}(1)$ when $p = \mathcal{O}(\omega)$, so that (4.5) has the standard form for perturbation theory, the small parameter ε being proportional to ω^{-2}.

B. The action–angle variables. In order to apply perturbation theory, one needs to use action–angle variables. We disregard for the moment any questions of domains. We refer to section 1.2B (the only difference being that now we interpret all vectors as belonging to the inertial frame). The three actions can be taken to be (see (2.5))

$$I_1 = G\,, \qquad I_2 = I_2(T, G)\,, \qquad I_3 = J_z\,, \tag{4.6}$$

where $G = \langle m, m\rangle^{1/2}$, $J_z = \langle m, e_z\rangle$, and $I_2(T,G)$ is a certain (analytic) function. Like I_1 and I_3, I_2 is a positively homogeneous function of degree one of the angular momentum: $I_1(\lambda m) = \lambda I_2(m)$ for all $\lambda > 0$. Let φ_1, φ_2, φ_3 be the correspondent angles. The first two of them are angular coordinates on the fixed and, respectively, the moving Poinsot cone, while φ_3 is a rotation angle of m about the axis e_z.

As a function of the two actions I_1, I_2, the kinetic energy k has good properties of nondegeneracy. Precisely, one has the following

Proposition 1.7 (Arnold [3], Kozlov [51]) *In each connected component of the subset of phase space where the action–angle variables are defined, the function $(I_1, I_2) \mapsto k(I_1, I_2)$:*

i) is analytic, and positively homogeneous of degree two;

ii) satisfies the condition of isoenergetic nondegeneracy (3.7);
iii) is steep[22].

It is also of interest to consider the properties of the potential energy $\mathcal{V}$, as a function of the action–angle variables. Let us write $\mathcal{V}(\varphi,\psi,\theta) = v(I_1,I_2,I_3,\varphi_1,\varphi_2,\varphi_3)$, so that the Hamilton function is

$$k(I_1,I_2) + v(I_1,I_2,I_3,\varphi_1,\varphi_2,\varphi_3) . \tag{4.7}$$

It is easy to establish that

i) v is a homogeneous function of degree zero of the actions $I = (I_1,I_2,I_3)$:

$$v(\lambda I,\varphi) = v(I,\varphi) \qquad \forall\lambda > 0 ; \tag{4.8}$$

ii) if $\mathcal{V}$ does not depend on the Euler angle φ, then v does not depend on φ_3.

Finally, it is useful to reconsider the relation between the use of the large parameter ω and that of the small parameter ε in terms of the action–angle variables. The canonical transformation

$$I = \lambda I' , \quad \varphi = \varphi' , \quad t = \lambda t' , \tag{4.9}$$

conjugates the Hamilton function (4.7) to

$$k(I') + \varepsilon v(I',\varphi') , \qquad \varepsilon = \lambda^{-2} . \tag{4.10}$$

Let us now denote $\big(I(t;I_o,\varphi_o;\varepsilon),\varphi(t;I_0,\varphi_0;\varepsilon)\big)$ the solution at time t of the Hamilton equations for the Hamilton function (4.10), and $\big(I(t;I_o,\varphi_o;1),\varphi(t;I_0,\varphi_0;1)\big)$ that of (4.7), with initial datum (I_0,φ_0). The homogeneity properties of k and v imply $k \sim \lambda^2$, $v \sim \lambda^0$. Then by (4.9), one has

$$I(\lambda^{-1}t,\lambda I_o,\varphi_o;\lambda^2\varepsilon) = \lambda I(t;I_0,\varphi_0;\varepsilon) .$$

Consequently one has the relation

$$I(\lambda^{-1}t,\lambda I_o,\varphi_o;\lambda^2\varepsilon) - \lambda I_o = \lambda\,[I(t;I_0,\varphi_0;\varepsilon) - I_0] . \tag{4.11}$$

C. Arnold's results. As far as we know, the first rigorous study of the motion of a rigid body by Hamiltonian perturbation theory is due to Arnold[3]. As an application of KAM theory, Arnold studied the fast motions of a tri–axial body, in a force field which is invariant under rotations about the axis e_z (passing through the body fixed point), as is the case of gravity. In such a case, $\mathcal{V}$ does not depend on the precession angle φ, v does not depend on φ_3 and $I_3 = J_z$ is an integral of motion. Consequently, the system can be reduced to a system with *two* degrees of freedom, described by the Hamilton function

$$h_{I_3}(I_1,I_2,\varphi_1,\varphi_2) = k(I_1,I_2) + v_{I_3}(I_1,I_2,\varphi_1,\varphi_2) , \tag{4.12}$$

---

[22] Possibly, quasi–convex. Property *iii)* is not explicitly stated in the quoted references; nevertheless, it is an obvious consequence of what said in [51].

with $v_{I_3}(I_1, I_2, \varphi_1, \varphi_2) = v(I_1, I_2, I_3, \varphi_1, \varphi_2)$. Since, as Arnold showed, k is isoenergetically nondegenerate, one can apply proposition 1.5 to h_{I_3}. Taking into considerations (4.11), one gets the following

Proposition 1.8 (Arnold, Neishtadt[23]) *Under the stated hypotheses, there exists a constant c such that, for all motions which at the initial time have a sufficiently large angular velocity, and are not too near to the separatrices of the unperturbed motion, one has the uniform estimates*

$$|G(t) - G(0)| \leq c \,, \qquad |I_2(t) - I_2(0)| \leq c \tag{4.13a, b}$$

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

Remark 1.3: One should complete the statement of this proposition by specifying the size of the excluded region near the separatrices (this would require a careful study of the analyticity properties of the function $k(I_1, I_2)$ near the separatrices). Results about the splitting of the separatrices can be found in [52, 76, 43, 54, 9]. A special consideration concerns the motions near the rotations about the major and minor inertia axes, and those with m parallel to the 'vertical' axis e_z. Since the action–angle variables are there singular, the estimates (4.13) for such motions cannot be obtained by invoking proposition 1.5. No mention to this problem is made in [3,62]. We think that one should be able to prove the validity of the estimates (4.13) for such motions by combining an obvious consistency argument (the motion cannot escape too far from the singularities, otherwise it enters the 'regular' region, where (4.13) holds) with conservation of energy (this method will be successfully used in a quite similar case in section 5.3).

We mention that the estimates (4.13) are optimal, as one can argue observing that (4.7) implies

$$\dot{I}_j = \frac{\partial v}{\partial \varphi_j}(I, \varphi) = \mathcal{O}(1) \,.$$

Indeed, in [22] some examples are given in which the variations of G are $\mathcal{O}(1)$, in times $\mathcal{O}(1)$.[24]

D. Our approach. The main purpose of the present thesis is to study the fast motion of a rigid body in an arbitrary external force field, which is not assumed to possess the above rotational invariance. In such a case, the potential energy v depends on all the three angle variables φ_1, φ_2, φ_3, the system has three degrees of freedom, and the unperturbed Hamiltonian is degenerate. Consequently, as observed in remark 1.2, KAM theory cannot be used in general; moreover, even if the system has the right, very special properties which make applicable KAM theory, then this does not provide stability results for *all* the fast motions. For these reasons, we shall resort to Nekhoroshev's theory.

---

23) Neishtadt contribution [62] concerns the estimates (4.13): in Arnold's paper [3], although this fact is there not clearly stated, one had $\|I(t) - I(0)\| \leq \mathcal{O}(\omega^a)$ for some $a > 0$.

24) If $\mathcal{V}$ depends only on one of the two angles ψ, θ, the system reduces to a system with one degree of freedom. Conservation of energy then implies that the variations of G and L are smaller than in (4.13), precisely $\mathcal{O}(1/\omega)$. An example is the fast motion of a Lagrangian top.

At first sight, this program could appear to be a trivial application of Nekhoroshev's theorem. Indeed, the Hamilton function $k(I_1, I_2)$ of the free rigid body is steep (proposition 1.7), so that Nekhoroshev's theorem can be applied to the system described by the Hamiltonian $k + v$, if the potential v is analytic.

There are however two problems. The first one is obviously that, because of the degeneracy of the Euler–Poinsot system, and the presence of singularities of the action–angle variables, the conclusions drawn from Nekhoroshev's theorem can be of no utility. As already mentioned, this problem requires a careful study, which is the object of chapter 2.

The second problem is the following. If one adapts the general estimates (3.11), (3.12) to the case at hand, using the relation (4.11) (with $\lambda \sim \omega \sim 1/\sqrt{\varepsilon}$) and the values (3.13) for a and b (as if k were convex), then one finds

$$|I_j(t) - I_j(0)| \le \mathcal{O}(\omega^{7/8}) , \qquad j = 1, 2 \tag{4.14}$$

and $T_0 \sim \exp(\omega^{1/48})$. Estimates (4.14) are very poor, if not even useless.

Better estimates on the two exponents a and b can be obtained by reconsidering the proof of the theorem, adapting it to the case at hand. The proof of Nekhoroshev's theorem for systems with two frequencies is in principle almost trivial, because of the simplicity of the geometry of resonances in the plane. However, our aim is that of getting the optimal estimates (4.13) which is found within KAM theory, namely $a = 1/2$. To obtain such an estimate within Nekhoroshev approach is not completely trivial, and requires a great care (and suitable techniques) in perturbation theory.

Just for the sake of simplicity, we shall restrict the analysis to the case of a symmetric rigid body. Such a case is simpler than the tri–axial one, since the action I_2 coincides with the gyroscopic component L of the angular momentum, and the function $k(G, L)$ is explicitly written. Moreover, the separatrices are absent in the symmetric case. For such a system, we shall prove that, for all initial conditions, one has

$$|G(t) - G(0)| \le \mathcal{O}(\omega^0) , \quad |L(t) - L(0)| \le \mathcal{O}(\omega^0) \tag{4.15}$$

for $|t| \sim \exp\sqrt{\omega}$. In addition, a detailed description of the motions will be given.

1.A Appendix: on the proof of the Arnold–Liouville theorem

We prove here statement *2* of proposition 1.1. Since both I and I' are invertible functions of $f_1, \ldots, f_m$, there exists a diffeomorphism $I' = \mathcal{I}(I)$. Let $\varphi' = \Phi(I, \varphi)$. Hence, the equality $\sum dI \wedge d\varphi = \sum dI' \wedge d\varphi'$ implies

$$\left[\frac{\partial \mathcal{I}}{\partial I}(I)\right]^t = \left[\frac{\partial \Phi}{\partial \varphi}(I, \varphi)\right]^{-1} . \qquad (A.1)$$

This shows that $\frac{\partial \Phi}{\partial \varphi}$ is independent of φ. Thus, there exist a matrix $C(I)$ and a vector $\mathcal{F}(I)$, which depend smoothly on I, and are such that

$$\varphi' = C(I)\varphi + \mathcal{F}(I) .$$

Since the mapping $\varphi \mapsto C(I)\varphi + \mathcal{F}(I)$ is a linear automorphism of the torus $I = \text{const}$, the matrix $C(I)$ has necessarily integer entries, and determinant ± 1. Consequently, such a matrix must also be independent of I, in each connected component of $B \cap B'$. Finally, equation (A.1) implies $I' = (C^t)^{-1} I + \text{const}$. In the proposition, we have written $A = C^t$.

1.B Appendix: proof of proposition 1.2

Let us first recall how a tangent vector to $SO(3)$ is identified to a vector of $\mathbb{R}^3$. Let $t \mapsto \mathcal{R}_t$ be a curve in $SO(3)$, let $\dot{\mathcal{R}}$ be its derivative at $t = 0$, and denote $\mathcal{R} = \mathcal{R}_0$. Then, to the point $(\mathcal{R}, \dot{\mathcal{R}}) \in T_{\mathcal{R}} SO(3)$ one associates the point $(\mathcal{R}, \rho) \in SO(3) \times \mathbb{R}^3$, where $\rho \in \mathbb{R}^3$ is uniquely defined by the condition $\dot{\mathcal{R}} u = \rho \times \mathcal{R} u$ for any vector $u \in \mathbb{R}^3$ (this definition is correct, since the matrix $\dot{\mathcal{R}} \mathcal{R}^t$ is antisymmetric, see [7]).

In this way, as in section 1.2, we identify both $TSO(3)$ and $T^*SO(3)$ to $SO(3) \times \mathbb{R}^3$. Moreover, we identify in the same manner $T^*(SO(3) \times \mathbb{R}^3) = T^*SO(3) \times T^*\mathbb{R}^3$ with $SO(3) \times \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3$. Here, the order of the components is defined by the identification of the spaces $\{\mathcal{R}\} \times \mathbb{R}^3 \times \{m\} \times \mathbb{R}^3$ with $T^*_{\mathcal{R}} SO(3) \times T^*_m \mathbb{R}^3$. In other words, writing $(\mathcal{R}, \rho; m, \mu)$ we interpret $\mathcal{R} \in SO(3)$, $\rho \in \mathbb{R}^3 \equiv T^*_{\mathcal{R}} SO(3)$, $m \in \mathbb{R}^3$ (in fact, the cotangent vector to $SO(3)$ in $\mathcal{R}$), $\mu \in T_m \mathbb{R}^3$.

We now compute the differentials of the three real functions defined on $SO(3) \times \mathbb{R}^3$ by

$$\begin{aligned} k(\mathcal{R}, m) &= \frac{1}{2} \langle m, A^{-1} m \rangle \\ G^2(\mathcal{R}, m) &= \langle m, m \rangle \\ J_\zeta(\mathcal{R}, m) &= \langle m, e_\zeta \rangle \end{aligned}$$

One finds

$$\begin{aligned} dk\,(\mathcal{R}, m) &= (\mathcal{R}, 0; m, A^{-1} m) \\ dG^2\,(\mathcal{R}, m) &= (\mathcal{R}, 0; m, 2m) \\ dJ_\zeta\,(\mathcal{R}, m) &= (\mathcal{R}, -m \times \mathcal{R} e_\zeta; m, \mathcal{R} e_\zeta) \end{aligned} \qquad (B.1)$$

The last expression is obtained in the following way: if ρ denotes the vector of $\mathbb{R}^3$ corresponding to the cotangent vector $d\mathcal{R}$, one has

$$\begin{aligned} dJ_z(\mathcal{R},m) &= \langle dm,\, \mathcal{R}e_\zeta\rangle \;+\; \langle m,\, (d\mathcal{R})e_\zeta\rangle \\ &= \langle dm,\, \mathcal{R}e_\zeta\rangle \;+\; \langle m,\, \rho\times\mathcal{R}e_\zeta\rangle \\ &= \langle dm,\, \mathcal{R}e_\zeta\rangle \;-\; \langle m\times\mathcal{R}e_\zeta,\, \rho\rangle \end{aligned}$$

From (B.1) it follows that dk, dG^2 and dJ_ζ are everywhere linearly independent except at the points $(\mathcal{R}, m)$ such that the equations

$$\begin{aligned} c_1\, A^{-1}m \;+\; c_2\, m \;+\; c_3\mathcal{R}e_\zeta \;&=\; 0 \\ c_3\, m\times\mathcal{R}e_\zeta \;&=\; 0 \end{aligned}$$

have a non trivial solution (c_1, c_2, c_3). Now:

i) if $c_3 = 0$, then the first equation has nontrivial solutions iff either $m = 0$ or m is an eigenvector of A^{-1}.

ii) let $c_3 \neq 0$. Then, the second equation implies $m = \lambda\mathcal{R}e_\zeta$ for some real number λ (which can assumed different from zero), and the first equation is equivalent to

$$\lambda c_1\, A^{-1}\mathcal{R}e_\zeta \;+\; (\lambda c_2 + c_3)\mathcal{R}e_\zeta = 0$$

which has the solution $c_1 = 0$, $c_2 = -c_3/\lambda$.

In this way, we conclude that the gradients of k, G^2 and J_ζ are linearly dependent only on the equilibria, the steady rotations about the three inertia axes, and all the states with m parallel to e_ζ. The same also holds for k, G and J_ζ. To obtain the properness set Σ_ζ we have moreover to exclude the whole level set $2a_2G^2 = k$ since, after the removal of the steady rotations about the middle inertia, it is not compact. All the other level sets are easily seen to be compact, and this completes the proof of the first statement of proposition 1.2.

We now prove that $\Sigma_\zeta \subset SO(3)\times\mathbb{R}^3$ is the union of four connected components. To this purpose, we describe the complement C of such set. One has $C = C_0 \cup C_\zeta$, where

$$\begin{aligned} C_0 &= SO(3)\,\times\, C^* \\ C^* &= \{m\in\mathbb{R}^3 \,:\, G^2(m) = 2a_ik(m) \text{ for some } i = 1,2,3\} \end{aligned}$$

and

$$\begin{aligned} C_\zeta &= \bigcup_{m\in\mathbb{R}^3\setminus C^*} \{m\}\,\times\, C_{\zeta,m} \\ C_{\zeta,m} &= \{\mathcal{R}\in SO(3) \,:\, m\times\mathcal{R}e_\zeta = 0\} \end{aligned}$$

Now, C^* divides $\mathbb{R}^3$ into four disjoint connected components (refer to figure 1.2a). On the contrary, C_ζ does not disconnect $SO(3)\times\mathbb{R}^3$ since, as we now show, for no vector $m\in\mathbb{R}^3\setminus C^*$ does $C_{\zeta,m}$ disconnect $SO(3)$. Observe preliminarily that for any m such that $m\times e_\zeta \neq 0$, $C_{\zeta,m}$ is diffeomorphic to $\mathbb{T}^2$ (S^1, if $m\times e_\zeta = 0$): each $\mathcal{R}\in C_{\zeta,m}$ can be represented as the composition of a rotation about m, a rotation about e_ζ, and a rotation which take e_ζ to coincide with m. Let us now introduce coordinates (referring to the case $m \times e_\zeta \neq 0$). Consider two orthogonal frames $\{e_j, e_l, e_m\}$ and $\{e_\delta, e_\varepsilon, e_\zeta\}$, e_m being the unit vector in the direction of m, and define the Euler angles θ, φ, ψ as, respectively, the angles between e_m and e_ζ, e_j and $e_m \times e_\zeta$, $e_m \times e_z$ and e_δ. In this way, $C_{\zeta,m}$ is identified with $\{\theta_0\} \times S^1 \times S^1$, for a certain angle θ_0, $0 < \theta_0 < \pi$. It is now evident that $C_{\zeta,m}$ does not disconnect $SO(3)$: for any two points $(\theta_1, \varphi_1, \psi_1)$, $(\theta_2, \varphi_2, \psi_2)$ not belonging to $C_{\zeta,m}$ (i.e. $\theta_1 \neq \theta_0$, $\theta_2 \neq \theta_0$), there exists a curve $[0,1] \ni t \mapsto (\theta(t), \varphi(t), \psi(t))$ with $\theta(t) \neq \theta_0$ for all $t \in [0,1]$ which joins them. The degenerate case $e_\zeta \times m = 0$ is also obvious. ∎

Chapter 2

Perturbation theory of degenerate systems

This chapter is devoted to the study of perturbation theory for systems which do not possess global action–angle variables, and are also properly degenerate. The appropriate setting for this study is a theory by Nekhoroshev [64] of degenerate integrable systems, which is shortly reviewed in section 1. In section 2 we illustrate the difficulties that arise with perturbation theory in these cases. The overcoming of such difficulties is then obtained, in section 3, by a 'global' formulation of perturbation theory.

2.1 The geometry of integrable systems

A. Generalized action–angle variables. The appropriate notion for the description of the geometry of the phase space of integrable (possibly degenerate) systems is that of *angular fibering* (this name is not of general use; it is introduced in [64]). Following [64], we define such a notions in a constructive way, which employs coordinates:

Definition 2.1 *A connected symplectic manifold M of dimension $2m$ is called an* angular fibering of order n, $0 \le n \le m$, *if the following holds: M has a covering by sets M_λ, $\lambda \in \Lambda$ (some index set), each of which is provided by coordinates $(I^i_\lambda,\ \varphi^i_\lambda \bmod 2\pi,\ p^j_\lambda,\ q^j_\lambda)$ $(i = 1, ..., n,\ j = 1, ..., m-n)$ with the following properties:*

i) the symplectic two-form of M, restricted to M_λ, is $\sum_{i=1}^n dI^i_\lambda \wedge d\varphi^i_\lambda + \sum_{j=1}^{m-n} dp^j_\lambda \wedge dq^j_\lambda$;

ii) for any pair $(\lambda, \mu) \in \Lambda \times \Lambda$ one has in every connected component of $M_\lambda \cap M_\mu$:

$$\begin{aligned} I_\mu &= (A_{\lambda\mu})^{-1} I_\lambda \;+\; a_{\lambda\mu} \\ \varphi_\mu &= (A_{\lambda\mu})^t \varphi_\lambda \;+\; \mathcal{F}_{\lambda\mu}(I_\lambda, p_\lambda, q_\lambda) \\ p_\mu &= \mathcal{P}_{\lambda\mu}(I_\lambda, p_\lambda, q_\lambda) \\ q_\mu &= \mathcal{Q}_{\lambda\mu}(I_\lambda, p_\lambda, q_\lambda) \end{aligned} \qquad (1.1)$$

where $A_{\lambda\mu} \in O(\mathbb{Z}, n)$, $a_{\lambda\mu} \in \mathbb{R}^n$, *and* $\mathcal{F}_{\lambda\mu}$, $\mathcal{P}_{\lambda\mu}$, $\mathcal{Q}_{\lambda\mu}$ *are vector valued functions (if* $M_\lambda \cap M_\mu$ *is not connected, then* $A_{\lambda\mu} \ldots, \mathcal{Q}_{\lambda\mu}$ *may differ from one component to the other).*

The local coordinates $(I_\lambda, p_\lambda, \varphi_\lambda \bmod 2\pi, q_\lambda)$ will be called (local) *generalized action–angle variables*; the I's will be called *actions* and the φ's *angles*; sometimes, we shall refer to the variables p and q as to the *degenerate* variables.

An example of angular fibering is the (subset foliated into invariant tori of the) phase space of a completely integrable system (Arnold–Liouville theorem).

Geometrically, an angular fibering of order n is a principal bundle with fiber $\mathbb{T}^n$, and structure group the 'affine' group of the n–dimensional torus. The latter is the group of the toral automorphisms $\varphi \bmod 2\pi \to (A\varphi + \alpha) \bmod 2\pi$, with $\alpha \in \mathbb{R}^n$ and $A \in O(\mathbb{Z}, n)$. The base B of the bundle is defined locally by the projection $(b_\lambda, \varphi_\lambda) \mapsto b_\lambda$, where $b_\lambda = (I_\lambda, p_\lambda, q_\lambda)$.

We are interested in the geometry of the base B. The transformation equations (1.1) imply that B is foliated into $2m$–dimensional leaves, defined locally by the equations $I_\lambda =$ const. Consider the quotient

$$\mathcal{I} = B/\sim \tag{1.2}$$

where $\sim$ is the equivalence relation defined by the property of belonging to the same leaf.

One important property of the angular fibering, made manifest by the equation (1.1), is that each fiber has an affine structure which is inherited, locally, by $\mathcal{I}$ (i.e., so to say, by the actions). There is an important case in which such an affine structure is defined globally in $\mathcal{I}$, which is then an affine manifold.

Assume that, within the notations of definition 2.1, the sets M_λ have pairwise connected intersections, and that there exist matrices $A_\lambda \in O(\mathbb{Z}, n)$, $\lambda \in \Lambda$, such that

$$A_{\lambda\mu} = (A_\lambda)^{-1} A_\mu \qquad \forall \mu, \lambda \in \Lambda\,. \tag{1.3}$$

The cocycle identity (1.3) is manifestly equivalent to the existence of a mapping $I = (I^1, \ldots, I^n) : M \to \mathbb{R}^n$ such that

$$I_\lambda = (A_\lambda)^{-1} I + a_\lambda\,, \qquad a_\lambda \in \mathbb{R}\,, \tag{1.4}$$

for any $\lambda \in \Lambda$. Hence, $\mathcal{I}$ has an atlas $\{\hat{I}_\lambda, \mathcal{I}_\lambda\}_{\lambda\in\Lambda}$, where $\mathcal{I}_\lambda = \mathcal{I} \cap M_\lambda$, whose transition mappings are the identity: $\hat{I}_\lambda = \hat{I}_\mu = I$.

Definition 2.2 *Let the angular fibering* M *be such that its chart domains have connected intersections, and (1.3) holds. Then we say that* M *possesses* global actions $I : M \to \mathbb{R}^n$. *We also say that* $\mathcal{I}$ *is the* action space.

It can be proven [64] that a sufficient condition for M to possess global actions is that its base B be simply connected.

The existence of global actions does not imply the existence of global generalized action–angle variables, i.e. the existence of an atlas for M made of a single chart. Clearly, when global actions exist, one can choose the local coordinates with $A_{\lambda\mu} = \mathbb{1}$ and $a_{\lambda\mu} = 0$

for all $\lambda, \mu \in \Lambda$, so that $I_\lambda = I$ for all $\lambda \in \Lambda$. Nevertheless, obstructions to globality still remain. First, the mapping $I = (I^1, \ldots, I^n) : M \to \mathbb{R}^n$ might be not injective. Moreover, the other coordinates p, q, φ might be defined only locally on M; in particular, the functions $\mathcal{F}_{\lambda\mu}$ might not vanish.

Finally, we notice that when global actions do exist, then each leaf of the foliation of B is globally defined by the equation $I = \text{const}$; furthermore, it possesses an atlas made of local coordinates (p_λ, q_λ), $\lambda \in \Lambda$.

Remark 2.1: There exists an alternative point of view: an angular fibering of order n is a foliation of a symplectic manifold, with isotropic leaves diffeomorphic to tori of dimension n; its base is a Poisson manifold, which has a natural decomposition into its symplectic leaves (our $I_\lambda = \text{const}$). Such a point of view is pursued in [33].

B. Nekhoroshev's approach to degeneracy. As a matter of fact, for our study of perturbation theory on an angular fibering, the transformation equations (1.1) will be all that one needs. Nevertheless, we go here a little further in this theory, so to provide some motivations for the subsequent developments. In this subsection, we describe Nekhoroshev's characterization of degeneracy.

In Nekhoroshev's approach, a 'degenerate' integrable system is a Hamiltonian system on a symplectic manifold of dimension $2m$, which possesses $2m - n \geq m$ independent integrals of motion $f = \{f_1, \ldots, f_{2m-n}\}$ (with some $1 \leq n \leq m$), with the property that each of the first n of them is pairwise in involution with all of them:

$$\{f_i, f_j\} = 0 , \qquad i = 1, \ldots, n , \quad j = 1, \ldots, 2m - n . \tag{1.5}$$

This generalizes the notion of completely integrable system (which is recovered when $n = m$). As in section 1.1, we denote by $\mathcal{L}_f(z)$ the connected components of the level sets (which have now dimension n) and by M_f the properness set of f. The basic results of Nekhoroshev's approach is the following generalization of the Arnold–Liouville theorem:

Proposition 2.1 (Nekhoroshev) *Let M be a symplectic manifold of dimension $2m$, and let $f = \{f_1, \ldots, f_{2m-n}\}$ be a set of real functions defined on M which satisfy (1.5). Then, the properness set M_f is an angular fibering of order n. Locally, the n actions $I_\lambda^1, \ldots, I_\lambda^n$ and, respectively, the $2m - n$ coordinates $(I_\lambda, p_\lambda, q_\lambda)$ on the basis are invertible functions of $f_1, \ldots f_n$ and, respectively, of $f_1, \ldots, f_{2m-n}$.*

For the proof of this proposition we demand to [64] (it is an immediate consequence of theorem 1 and of proposition 1 of such paper). We only remark that the leaves $\mathcal{L}_f(z)$, $z \in M_f$, are tori of dimension n: in fact, they are compact, connected n–dimensional manifolds which have n pairwise commuting and everywhere linearly independent tangent vectors fields (the Hamiltonian vector fields of the functions $f_1, \ldots, f_n$). Furthermore, the n (local) action variables $I_\lambda^1, \ldots, I_\lambda^n$ are constructed, as in the non–degenerate case, by integrating the symplectic one–form of M over n independent cycles of the tori $\mathcal{L}_f(z)$.

If the functions $f_1, \ldots, f_{2m-n}$ are integrals of motion of a Hamiltonian system on M, then the fibers $\mathbb{T}^n$ of M_f are invariant under its flow. Precisely, one has the following

Proposition 2.2 (Nekhoroshev) *Under the same conditions of proposition 2.1, assume that the functions $f_1, \ldots, f_{2m-n}$ are integrals of motion of a Hamiltonian system defined by the Hamilton function $k : M \to \mathbb{R}$. Then k is, locally, a function of the action variables only.*

For the easy proof, see [64]. Proposition 2.2 means that one has $k\big|_{M_\lambda} = k_\lambda \circ I_\lambda$, for some functions $k_\lambda : \mathbb{R}^n \to \mathbb{R}$ ($\lambda \in \Lambda$). The equations of motion are then, in local generalized action–angle coordinates:

$$\begin{aligned} \dot{I}_\lambda &= 0 \\ \dot{p}_\lambda &= 0 \\ \dot{\varphi}_\lambda &= \omega_\lambda(I) \\ \dot{q}_\lambda &= 0 \end{aligned}$$

where $\omega_\lambda = \partial k_\lambda / \partial I_\lambda$: all the motions are quasi periodic with (at most) n frequencies.

Remark 2.2: Nekhoroshev's definition of a 'degenerate integrable system' does not require that $m = \frac{1}{2}\dim M$ integrals of motion are pairwise in involution. However, m integrals of motion in involution always exist *locally*: the coordinate functions I_λ, p_λ. We cannot say at present whether (1.5) implies the global existence of m commuting integrals. Furthermore, it should be of interest to investigate the relation of Nekhoroshev's theory with the more recent 'noncommutative integrability' of [37].

C. Proper degeneracy. Completely integrable systems which are properly degenerate can be naturally described within the previous setting: the role of the $2m - n$ integrals of motion of Nekhoroshev's theory is played, at least locally, by the m action variables and the $m - n$ angles corresponding to the null frequencies. However, proposition 2.2 is too restrictive, since it requires the simultaneous linear independence of all the differentials of the $2m - n$ integrals of motion. Miming the cases of the Kepler system, and of the rigid body, we shall now weaken this condition.

Let us refer to the simple but interesting case, already discussed in section 1.1D, of a completely integrable system which has just one integral of motion in excess. The following considerations can be adapted easily to more general situations. Thus, we consider a Hamiltonian system with m degrees of freedom and $m+1$ integrals of motion $f_1, \ldots, f_{m+1}$, and assume that $f' = \{f_1, \ldots, f_{m-1}, f_m\}$ and $f'' = \{f_1, \ldots, f_{m-1}, f_{m+1}\}$ are two sets of functions pairwise in involution. Let $M_{f'}$ and $M_{f''}$ be the correspondent properness sets, and assume that both of them are proper subsets of $M_{f'} \cup M_{f''}$.

In such a situation, Nekhoroshev's theory properly applies to the system in the set $M_{f'} \cap M_{f''}$: it states that this set is an angular fibering of order $n = m - 1$, and the motions are quasiperiodic with n frequencies (moreover, the coordinates on the base are functions of all the integrals $f_1, \ldots, f_{m+1}$, while the actions are functions only of the first $m - 1$ of them). By itself, this result is quite obvious: by the Arnold–Liouville theorem, one knows that not only $M_{f'} \cap M_{f''}$, but also $M_{f'}$ and $M_{f''}$ are angular fiberings of order m (and then $m - 1$, too). Nevertheless, one can easily extend Nekhoroshev's results to the larger set $M_{f'} \cup M_{f''}$:

Proposition 2.3 *In the stated hypotheses, assume that $M_{f'} \cap M_{f''}$ is dense in $M_{f'} \cup M_{f''}$. Then $M_{f'} \cup M_{f''}$ is an angular fibering of order $n = m-1$. Moreover, if both $M_{f'}$ and $M_{f''}$ possesses global actions, then $M_{f'} \cup M_{f''}$ possesses global actions, too.*

Proof. Let us consider a system of local action–angle coordinates (I', φ') of the angular fibering $M_{f'}$, defined in a domain $M'_\lambda \subset M_{f'}$. The functions $\varphi'_1, \dots, \varphi'_m$ are angular coordinates on the m–dimensional tori $\mathcal{L}_{f_1,\dots,f_m}$.

Let us now restrict ourselves to $M'_\lambda \cap M_{f''}$. Such a domain is foliated by the $(m-1)$–dimensional tori $\mathcal{L}_{f_1,\dots,f_m,f_{m+1}}$, each of which is the intersection of a torus $\mathcal{L}_{f_1,\dots,f_m}$ with a level surface of the function f_{m+1}. Thus, in $M'_\lambda \cap M_{f''}$ the tori $\mathcal{L}_{f_1,\dots,f_m}$ are foliated into the $(m-1)$–dimensional tori $\mathcal{L}_{f_1,\dots,f_m,f_{m+1}}$. This implies that there exist a matrix $B'(f_1,...,f_m) \in O(\mathbb{Z}, m)$ and a vector $b'(f_1,...,f_m) \in \mathbb{R}^m$ such that, if one defines

$$\tilde{\varphi}' = \left[B'(f_1,...,f_m)\right]^t \varphi' + b'(f_1,...,f_m)$$

then $(\tilde{\varphi}'_1, \dots, \tilde{\varphi}'_{m-1})$ are angular coordinates on the tori $\mathcal{L}_{f_1,\dots,f_m,f_{m+1}}$; moreover, the angular coordinate $\tilde{\varphi}_m$ is, on every torus $\mathcal{L}_{f_1,\dots,f_m}$, a function of f_{m+1} alone, so that one has $\tilde{\varphi}'_m = \tilde{\varphi}'_m(f_1,...,f_m,f_{m+1})$. By an obvious argument of continuity, B' is a constant matrix. Let us define

$$\tilde{I}' = (B')^{-1} I'$$

and define correspondingly

$$\begin{aligned} \psi' &= (\tilde{\varphi}'_1, \dots, \tilde{\varphi}'_{m-1}), \qquad & q' &= \tilde{\varphi}'_m \\ J' &= (\tilde{I}'_1, \dots, \tilde{I}'_{m-1}), \qquad & p' &= \tilde{I}'_m . \end{aligned}$$

In this way, we have constructed canonical coordinates ψ', q', J', p' on $M'_\lambda \cap M_{f''}$, such that ψ' are angular coordinates on $\mathcal{L}_{f_1,\dots,f_m,f_{m+1}}$ and J', p', q' are (locally invertible) functions of $f_1,...,f_m,f_{m+1}$. Notice that such coordinates can be extended to the whole domain M'_λ: indeed, $M'_\lambda \cap M_{f''}$ is certainly dense in M'_λ, B' is constant and b', being independent of f_{m+1}, can be extended to all of M'_λ.

Consider now any system of local action–angle coordinates (I'', φ'') of the angular fibering $M_{f''}$, whose domain M''_μ has non–empty intersection with M'_λ. Here $\varphi''_1 \dots, \varphi''_m$ are coordinates on the tori $\mathcal{L}_{f_1,\dots,f_{m-1},f_{m+1}}$. As before, we construct in M''_μ a system of canonical coordinates $(J'', \varphi'', q'', \psi'')$ such that J'', p'', q'' are (locally invertible) functions of $f_1, \dots, f_{m+1}$ and, in $M''_\mu \cap M_{f'}$, $\psi''_1, \dots, \psi''_{m-1}$ are coordinates on the tori $\mathcal{L}_{f_1,\dots,f_{m+1}}$.

So, in the set $M'_\lambda \cap M''_\mu$, ψ' and ψ'' are coordinates on the same tori. This implies

$$\psi'' = A^t \psi' + a(f_1, \dots, f_m + 1)$$

for some $A \in O(\mathbb{Z}, n)$ and $a \in \mathbb{R}^{m-1}$. Since $dJ'' \wedge d\psi'' + dp'' \wedge dq'' = dJ' \wedge d\psi' + dp' \wedge dq'$ and J'', p'', q'' do not depend on the angles ψ' (in fact, they are functions of $f_1, \dots, f_{m+1}$), one obtains

$$J'' = A^{-1} J' .$$

Moreover, let us divide $M'_\lambda \cap M''_\mu$ in subsets, in each of which both J', p', q' and J'', p'', q'' are invertible functions of $f_1, \ldots, f_{m+1}$. In each of such subsets, one has

$$p'' = \mathcal{P}(J', p', q'), \qquad q'' = \mathcal{Q}(J', p', q').$$

Since $M_{f'} \cap M_{f''}$ is dense in $M_{f'} \cup M_{f''}$, we can construct in this way an atlas of generalized action–angle coordinates for $M' \cup M''$.

Finally, observe that the statement about the existence of global actions is obvious, by virtue of the very construction of the actions J', J''. ∎

The above result has a certain interest in itself, since it characterizes the global structure of a subset of the phase space which is larger than those to which the Arnold–Liouville theorem applies. Furthermore, this result provides general motivations for the study of the following sections.

2.2 The need for a chart–independent perturbation theory

In this section we indicate, by a simple argument, the difficulties that arise in the perturbation study of a degenerate system, when it does not possess global (generalized) action–angle variables.

Let us refer to the simple case of an angular fibering M of order n and dimension m, which possesses an atlas of just two charts, with generalized action–angle coordinates $\mathcal{C}_\lambda = (b_\lambda, \varphi_\lambda) : M_\lambda \to B_\lambda \times \mathbb{T}^n$, $\lambda = 1, 2$. Here, M_1 and M_2 are the two chart domains (thus $M = M_1 \cup M_2$), B_1 and B_2 two open subsets of $\mathbb{R}^{2m-n}$, and we have written $b_\lambda = (I_\lambda, p_\lambda, q_\lambda)$.

Let $k : M \to \mathbb{R}$ be a function which 'depends only on the actions': $k|_{M_\lambda} = k_\lambda \circ I_\lambda$ for some functions $k_\lambda : B_\lambda \to \mathbb{R}$, $\lambda = 1, 2$. Let $v : M \to \mathbb{R}$ be any function, and consider the Hamiltonian system defined on M by the Hamilton function

$$h = k + \varepsilon v . \tag{2.1}$$

The local representatives of h have the form

$$h_\lambda(I_\lambda, p_\lambda, \varphi_\lambda, q_\lambda) = k_\lambda(I_\lambda) + \varepsilon v_\lambda(I_\lambda, p_\lambda, \varphi_\lambda, q_\lambda) \tag{2.2}$$

and define two Hamiltonian systems on the phase spaces $B_\lambda \times \mathbb{T}^n$, $\lambda = 1, 2$.

By the standard methods of perturbation theory, one can construct suitable normal forms for h_1 and h_2, separately. For definiteness, assume that in some subset of B_λ one has a non–resonant normal form, up to some order $r > 1$, for the system described by h_λ, $\lambda = 1, 2$ (notice that we take r to be the same for the two systems). This means that, for each $\lambda = 1, 2$, there exists a canonical transformation Φ_λ, $(b_\lambda, \varphi_\lambda) = \Phi_\lambda(b'_\lambda, \varphi'_\lambda)$, defined in some subset of $B_\lambda \times \mathbb{T}^n$, such that $h'_\lambda = h_\lambda \circ \Phi_\lambda$ has the form

$$h'_\lambda(b'_\lambda, \varphi'_\lambda) = k_\lambda(b'_\lambda) + \varepsilon g_\lambda(b'_\lambda) + \varepsilon^r f_\lambda(b'_\lambda, \varphi'_\lambda) \tag{2.3}$$

with certain functions g_λ and f_λ. Let

$$\Delta_\lambda = \sup_{(b_\lambda,\varphi_\lambda)} \|(b'_\lambda,\varphi'_\lambda)-(b_\lambda,\varphi_\lambda)\|$$

($\|\cdot\|$ being some norm) be the 'deformation' of the canonical transformation Φ_λ. Typically, the two deformations will be small, say of order ε^β for some $\beta > 0$.

The use of the normal form (2.3) to control the variation of the actions is well known and elementary, if the motion remain within the domain in which h'_λ is defined. Indeed, from (2.3) one gets $\dot I_\lambda = \varepsilon^r \frac{\partial f_\lambda}{\partial \varphi_\lambda}$. Hence

$$\|I'_\lambda(t) - I'_\lambda(0)\| \leq |t|\,\varepsilon^r F_\lambda$$

for a certain constant F_λ. One then writes

$$I_\lambda(t) - I_\lambda(0) = \left[I_\lambda(t) - I'_\lambda(t)\right] + \left[I'_\lambda(t) - I'_\lambda(0)\right] + \left[I'_\lambda(0) - I_\lambda(0)\right] . \qquad (2.4)$$

Estimating the first and the last term by the deformation of the canonical transformation, which we assume to be of order ε^β, one finally gets

$$\|I_\lambda(t) - I_\lambda(0)\| \leq \text{const}\,\varepsilon^\beta \qquad \text{for all } 0 \leq t \leq \frac{\text{const}}{\varepsilon^{r-\beta}} . \qquad (2.5)$$

How to adapt this procedure to the case in which the motion does not remain within a single chart domain for all the time interval of interest? Once more, we observe that no problem arises for a non–degenerate system, namely if $n = m$: the estimate (2.5) implies that the motion does not leave one of the two chart domains, at least if their intersection is not too thin. The same conclusion could be reached for the perturbation theory near a resonance, if a geometric construction of Nekhoroshev's type is used.

In degenerate cases, instead, the normal forms (2.4) do not allow one to control the motion of the additional 'degenerate' variables p_λ, q_λ, which can then escape the chart domains. Naively, one can use consecutively the two normal forms: the one for h_1 when the phase point is in M_1, the one for h_2 when it is in M_2. To do this, one switches from one coordinate system to the other each time the phase point goes through some given hypersurface[25] $N \subset M_1 \cap M_2$ (see figure 2.1).

In this way, if $t_1 < t_2 < \ldots < t_s$ are the instants at which the motion intersects (transversally) N in the time interval $(0,t)$, one can write (assuming that $x(0)$ and $x(t)$ both belong to M_1):

$$\begin{aligned} I(t) - I(0) &= \left[I_1(t) - I'_1(t)\right] + \left[I'_1(t) - I'_1(t_s)\right] + \left[I'_1(t_s) - I_1(t_s)\right] \\ &+ \left[I_2(t_s) - I'_2(t_s)\right] + \left[I'_2(t_s) - I'_2(t_{s-1})\right] + \left[I'_2(t_{s-1}) - I_1(t_{s-1})\right] \\ &+ \qquad \cdots\cdots\cdots\cdots \\ &+ \left[I_1(t_1) - I'_1(t_1)\right] + \left[I'_1(t_1) - I'_1(0)\right] + \left[I'_1(0) - I_1(0)\right] . \end{aligned} \qquad (2.6)$$

[25] Codimension one submanifold

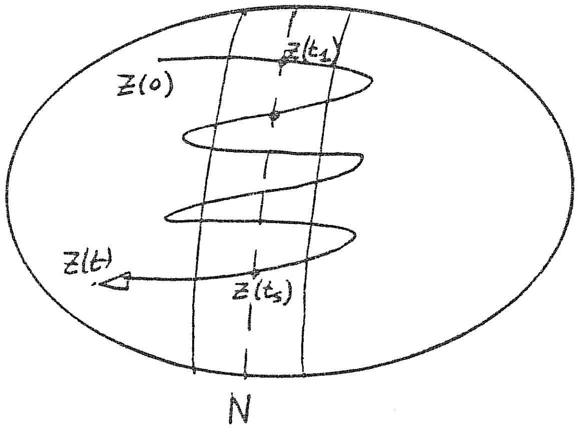

Figure 2.1

Using the normal forms (2.3), the second term of each line is estimated by $|t_j - t_{j-1}|\varepsilon^r F_\lambda$. The other terms can only be estimated by the deformation Δ_λ of the canonical transformation. This leads to

$$\|I_1(t) - I_1(0)\| \;\leq\; |t|\,\varepsilon^r F \;+\; 2(s+1)\Delta$$

with $F = \max_\lambda F_\lambda$ and $\Delta = \max_\lambda \Delta_\lambda$. This result is completely unsatisfactory, because in general one does not have any control on the number s.

The origin of the problem is that, in equation (2.5), one does not cancel terms in diagonal, precisely the last term of each line with the first one of the successive line. Let us even assume that M possesses global actions, so that $I_1(x) = I_2(x) \equiv I(x)$ whenever $x \in M_1 \cap M_2$. Since I'_λ depends not only on $I_\lambda \equiv I$, but also on $p_\lambda, q_\lambda, \varphi_\lambda$, one may not conclude $I'_1(t) = I'_2(t)$ although $I_1(t) = I_2(t)$. In order to be able to draw this conclusion, one should know that the two canonical transformations Φ_1 and Φ_2 are related in the right way in $M_1 \cap M_2$, precisely that they are the local representatives of a single canonical transformation defined (in some subset of) the manifold.

2.3 Perturbation theory on an angular fibering

This section deals with perturbation theory for degenerate systems which do not possess a global coordinate system of generalized action–angle variables. Our aim is that of working out a 'global' canonical transformation (and then a 'global' normal form) from the 'local' canonical transformations constructed within each coordinate system. Since these 'local' canonical transformations are constructed by means of Fourier series, whose harmonics depend on the local resonance properties, one expects that the following two questions play an important role:

i) are resonances intrinsically defined on the manifold?

ii) may one reconstruct a function from the series expansions of its local representatives? As will be seen, these questions are quite simple, at least if 'global actions' exist, in the sense of definition 2.2. The key point is that the affine structure of the 'action space' makes the resonance properties intrinsically defined.

In all of this section, M is an angular fibering of order n and dimension $2m$, which has an atlas $\{\mathcal{C}_\lambda, M_\lambda;\ \lambda \in \Lambda\}$, $\mathcal{C}_\lambda : M_\lambda \to B_\lambda \times \mathbb{T}^n$, with generalized action–angle coordinates: $\mathcal{C}_\lambda(z) \mapsto (b_\lambda(z), \varphi_\lambda(z))$, with $b_\lambda = (I_\lambda, p_\lambda, q_\lambda)$ (see figure 2.2).

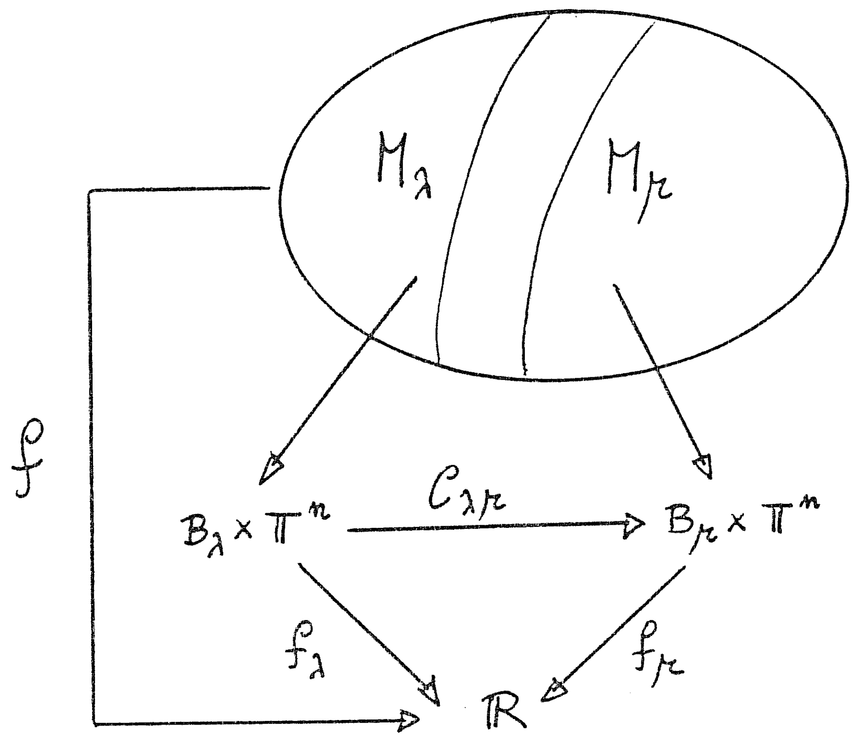

Figure 2.2

We assume that the intersections of the chart domains $M_\lambda \cap M_\mu$ are connected. The transition mappings $\mathcal{C}_{\lambda\mu} = \mathcal{C}_\mu \circ \mathcal{C}_\lambda^{-1}$ of the atlas have then the form (1.1). Furthermore, we denote $\mathcal{F}(\Sigma)$ the space of the real smooth functions defined on a manifold Σ.

A. Resonant manifolds. Let $k : M \to \mathbb{R}$ be a function which 'depends only on the action variables': it satisfies $k\big|_{M_\lambda} = k_\lambda \circ I_\lambda$ for all $\lambda \in \Lambda$, where the functions $k_\lambda : B_\lambda \to \mathbb{R}$ can be regarded as its local representatives. Let $\omega_\lambda = \frac{\partial k_\lambda}{\partial I_\lambda}$ be the frequency vector of k_λ.

A vector $\nu \in \mathbb{Z}^n$ is called a *resonance* of k_λ at the point $I_\lambda \in B_\lambda$ if $\langle \omega_\lambda(I_\lambda), \nu \rangle = 0$; the point I_λ and the vector $\omega_\lambda(I_\lambda)$ are said to be resonant with ν; $|\nu| = \sum |\nu_j|$ is the *order* of the resonance. The set of all the resonances of k_λ at I_λ is a 'sublattice' of $\mathbb{Z}^n$, which we call the *resonant lattice* of k_λ at I_λ. Conversely, given a subset $\mathcal{L}$ of $\mathbb{Z}^n$, the $\mathcal{L}$*–resonant* *manifold* of k_λ is the subset of all the points of B_λ which are resonant with some vector of $\mathcal{L}$; we shall denote it by $r_{\mathcal{L}}$.

The obvious globalization of these notions to the function $k : M \to \mathbb{R}$ is based on the following observation: let $z \in M_\lambda \cap M_\mu$ for some λ, μ in Λ. Then $\nu_\lambda \in \mathbb{Z}^n$ is a resonance of k_λ at $I_\lambda(z)$ iff $(A_{\lambda\mu})^{-1}\nu_\lambda$ is a resonance of k_μ at $I_\mu(z)$ (this is easily checked using the definition of the frequency vectors, $k_\mu \circ I_\mu = k_\lambda \circ I_\lambda$ and (1.1)). It follows that the resonant lattices $\mathcal{L}_\mu$ of k_μ at $I_\mu(z)$ and $\mathcal{L}_\lambda$ of k_λ at $I_\lambda(z)$ are related by $\mathcal{L}_\mu = (A_{\lambda\mu})^{-1}\mathcal{L}_\lambda$.

Consider now the globalization of the resonant manifolds. The question is whether a resonant manifold defined within a chart domain can be continued outside it. Let us first consider the case in which M has 'global actions'. Consider a collection $\{\mathcal{L}_\lambda;\, \lambda \in \Lambda\}$ of subsets $\mathcal{L}_\lambda$ of $\mathbb{Z}^n$, which satisfy the relations

$$\mathcal{L}_\mu \;=\; (A_{\lambda\mu})^{-1}\, \mathcal{L}_\lambda \quad \text{for all } \mu\,,\ \lambda \in \Lambda\,. \tag{3.1}$$

Then the $\mathcal{L}_\lambda$–resonant submanifolds $r_{\mathcal{L}_\lambda} \subset B_\lambda$ are easily seen to be the local images in the charts $\mathcal{C}_\lambda$ of a submanifold of the 'action space' $\mathcal{I}$, which can be considered as the resonant submanifold corresponding to the family $\{\mathcal{L}_\lambda;\, \lambda \in \Lambda\}$.

However, if M does not have global actions, then the above characterization may be meaningless, since it is not assured that, starting with some resonance lattice $\mathcal{L}_\lambda$, there exist a family of lattices $\mathcal{L}_\mu$, $\mu \in \Lambda$, which satisfies (3.1) (i.e. the matrices $A_{\lambda\mu}$ could be not 'reconciliable'). As discussed later in subsection G, this fact seems not to have very serious consequences. Nevertheless, we shall restrict our treatment to the good case in which global actions do exist. In fact, we do not know of any Hamiltonian system in which this condition is not satisfied.

B. Fourier series. Let B be an open set in $\mathbb{R}^{2m-n}$. Any function $f \in \mathcal{F}(B \times \mathbb{T}^n)$ can be expanded in Fourier series

$$f \;=\; \sum_{\nu \in \mathbb{Z}^n} \langle f \rangle_\nu E_\nu$$

where the functions $E_\nu : \mathbb{T}^n \to \mathbb{R}$ and $\langle f \rangle_\nu : B \to \mathbb{R}$ are defined, respectively, by $E_\nu(\varphi) = \exp i\,\langle \nu,\, \varphi\rangle$ and

$$\langle f \rangle_\nu(b) \;=\; (2\pi)^{-n} \int_{\mathbb{T}^n} f(b,\varphi)\, \exp(-i\,\langle \nu,\, \varphi\rangle)\, d\varphi\,. \tag{3.2}$$

If $\mathcal{L}$ is a subset of $\mathbb{Z}^n$, let $\mathcal{F}_{\mathcal{L}}(B \times \mathbb{T}^n)$ be the subspace of $\mathcal{F}(B \times \mathbb{T}^n)$ of the functions whose Fourier spectrum has support contained in $\mathcal{L}$. Correspondingly, we define the linear projector $\Pi_{\mathcal{L}}$ onto $\mathcal{F}_{\mathcal{L}}(B \times \mathbb{T}^n)$ as

$$\Pi_{\mathcal{L}} f(b,\varphi) \;=\; \sum_{\nu \in \mathcal{L}} \langle f \rangle_\nu(b) \exp i\,\langle \nu,\, \varphi\rangle \qquad \forall f \in \mathcal{F}(B \times \mathbb{T}^n)\,. \tag{3.3}$$

Lemma 2.4 *Let the angular fibering M have the property stated at the beginning of the section. Consider a family of functions $\{f_\lambda;\, \lambda \in \Lambda\}$, with $f_\lambda \in \mathcal{F}(B_\lambda \times \mathbb{T}^n)$ for each*

$\lambda \in \Lambda$. *These functions are the local representatives of a function* $f \in \mathcal{F}(M)$ *iff their Fourier components satisfy the equations*

$$\langle f_\mu \rangle_\nu \big(\mathcal{B}_{\lambda\mu}(b_\lambda)\big) \;=\; e^{-i\langle \nu,\, a_{\lambda\mu}(b_\lambda)\rangle}\, \langle f_\lambda \rangle_{A_{\lambda\mu}\nu}(b_\lambda) \tag{3.4}$$

for all $\nu \in \mathbb{Z}^n$, *all* $b_\lambda \in B_\lambda \cap \mathcal{B}_{\lambda\mu}^{-1}(B_\mu)$ *and all* $\mu, \lambda \in \Lambda$ *such that* $M_\mu \cap M_\lambda$ *is not empty.*

Proof. The functions $\{f_\lambda\}$ are the representatives of a function f iff they satisfy $f_\mu = f_\lambda \circ \mathcal{C}_{\lambda\mu}$ for all μ and λ. In turn, these relations are seen to be equivalent to (3.4) by virtue of (3.2) and of the transformation equations (1.1). ∎

The functions on an angular fibering M cannot be expanded in Fourier series, unless M is a trivial bundle (the existence of global actions is not sufficient, since the angular coordinates might not be defined globally). Nevertheless, if M possesses global actions, then the notion of 'having Fourier components only on a subset of $\mathbb{Z}^n$'can be readily generalized. Indeed, let $\{\mathcal{L}_\lambda;\, \lambda \in \Lambda\}$ be a collection of subsets of $\mathbb{Z}^n$ which satisfy (3.1). Define the subspace $\mathcal{F}_{\{\mathcal{L}_\lambda;\, \lambda\in\Lambda\}}$ in the following way: $f \in \mathcal{F}_{\{\mathcal{L}_\lambda;\, \lambda\in\Lambda\}}$ iff its local representative $f_\lambda = f \circ \mathcal{C}_\lambda^{-1}\big|_{M_\lambda}$ belongs to $\mathcal{F}_{\mathcal{L}_\lambda}(B_\lambda \times \mathbb{T}^n)$, for any $\lambda \in \Lambda$ (such a definition makes sense, since (3.4) shows that $f_\lambda \in \mathcal{F}_{\mathcal{L}_\lambda}(B_\lambda \times \mathbb{T}^n)$ iff $f_\mu \in \mathcal{F}_{\mathcal{L}_\mu}(B_\mu \times \mathbb{T}^n)$). Correspondingly, we define in an obvious way a projector $\Pi_{\{\mathcal{L}_\lambda;\, \lambda\in\Lambda\}} : \mathcal{F}(M) \to \mathcal{F}_{\{\mathcal{L}_\lambda;\, \lambda\in\Lambda\}}(M)$.

Let us remark that, if global actions $I : M \to \mathbb{R}^n$ do exist, then one may take all the local actions I_λ coincident with the I (i.e. $A_{\lambda\mu} = \mathbb{1}$ and $a_{\lambda\mu} = 0$ in (1.1)). In such a case, we simply write $\mathcal{F}_\mathcal{L}(M)$ and $\Pi_\mathcal{L}$. In this case, one may also define in an obvious way the cut–off decomposition $f = f^{\le N} + f^{>N}$ for the functions $f \in \mathcal{F}(M)$.

C. The Lie method. We give now an introduction to the so called 'Lie method' for the construction of the canonical transformations. Such a method is fully geometric, then useful for our purposes. A more complete account of the method is given in the appendix at chapter 4.

Within the Lie method, one constructs canonical transformations near the identity by means of Hamiltonian flows. Let M be a symplectic manifold and $\chi : M \to \mathbb{R}$ a smooth function. Let X be the Hamiltonian vector field of χ †. We shall denote $(\tau, z) \to \Phi_\tau^\chi(z)$ the flow of X, and we shall refer to it as to the Hamiltonian flow of χ.

Let Φ_τ^χ be the map at time τ of the flow of X. Standard theorems on ordinary differential equations assure that, if $|\tau|$ is sufficiently small, then Φ_τ^χ is defined in an open non–empty subset M_τ of M; more precisely, Φ_τ^χ is a diffeomorphism of M_τ onto its image, which is contained in M. Furthermore, the mapping Φ_τ^χ depends smoothly on τ, and is canonical. Since $\Phi_0^\chi = \mathbb{1}$, such a mapping is in some sense 'near the identity': for instance, in any reasonable norm one has $\|\Phi_\tau^\chi(z) - z\| = \mathcal{O}(\tau)$.

The *Lie transform* is the family of transformations of functions $f \mapsto f \circ \Phi_\tau^\chi$, depending on the small parameter τ. Sometimes in applications it is useful to use only the map at time one of the flow, and to define the Lie transform as $f \mapsto f \circ \Phi_1^\chi$. The two definitions are

† X is defined globally in M by $d\chi = \omega(X, \ )$, where ω is the symplectic two–form on M, or locally, in canonical coordinates (p, q), by $X = -\frac{\partial\chi}{\partial p}\frac{\partial}{\partial q} + \frac{\partial\chi}{\partial q}\frac{\partial}{\partial p}$.

obviously equivalent, because of the identity $\Phi_\tau^\chi = \Phi_1^{\tau\chi}$, which holds for every fixed τ; the smallness conditions on τ are correspondingly turned into equivalent smallness conditions on χ. We shall employ the time–one map.

The Lie method allows a simple and efficient transformation theory of functions (or vector fields), which is essentially based on the well known identity

$$\frac{d}{d\tau}(f \circ \Phi_\tau^\chi) \;=\; \{\chi, f\} \circ \Phi_\tau^\chi \tag{3.5}$$

relating time derivatives along the flow to Poisson brackets. It follows, for instance,

$$\begin{aligned} f \circ \Phi_1^\chi \;&=\; f \;+\; R_1^\chi(f) \\ &=\; f \;+\; \{\chi, f\} \;+\; R_2^\chi(f) \end{aligned} \tag{3.6}$$

with

$$\begin{aligned} R_1^\chi(f) \;&=\; \int_0^1 \{\chi, f\} \circ \Phi_\tau^\chi d\tau \\ R_2^\chi(f) \;&=\; \int_0^1 d\tau \int_0^\tau d\tau' \{\chi, \{\chi, f\} \circ \Phi_\tau^\chi\} \circ \Phi_{\tau'}^\chi \end{aligned} \tag{3.7}$$

Expressions (3.6) furnish approximations of orders zero and one to $f \circ \Phi_\tau^\chi$, up to remainders R_s^χ which are, in some sense, of order $|\chi|^s$ ($s = 1, 2$).

D. Construction of the 'global' normal form. We give now an outline of the the construction of a normal form on an angular fibering M, which we assume to possess 'global actions' $I : M \to \mathbb{R}^n$. For simplicity, we refer to an atlas in which the local action coordinates coincide with the global actions I. We consider a Hamilton function

$$h \;=\; k \;+\; \varepsilon\, v \tag{3.8}$$

defined on M, with local representatives

$$h_\lambda(b_\lambda, \varphi_\lambda) \;=\; k(I) \;+\; \varepsilon\, v_\lambda(b_\lambda, \varphi_\lambda)\;, \qquad \lambda \in \Lambda\,. \tag{3.9}$$

Clearly, the integrable Hamiltonian k has the same representatives in all charts, as well as the frequency vector $\omega(I) = \frac{\partial k}{\partial I}(I)$.

The aim of perturbation theory is that of constructing a suitable 'normal form' for the system. As is well known, in typical situations, this has to be done differently in different regions of phase space, depending on the local resonance properties. For instance, one can work in a subset of phase space defined by $I \in U$, $U = U(\mathcal{L}, N)$ being a subset of the action space $\mathcal{I}$ in which the only resonances (exact, or approximate within some quantity) of order less or equal to N are those contained in the lattice $\mathcal{L}$.[26] In such a case, one aims to construct a canonical transformation which conjugates h to a normal form 'adapted' to the lattice $\mathcal{L}$ and to the cut–off N.

---

[26] Thus, perturbation theory is in a sense 'local' in the actions, but 'global' in the other variables.

To this purpose, let us observe that, if χ is any function, then by (3.6) one may write

$$h \circ \Phi_1^\chi \;=\; k \;+\; \{\chi, k\} \;+\; \varepsilon\, v^{\le N} \;+\; R_1^\chi(\varepsilon v) \;+\; R_2^\chi(k) \;+\; \varepsilon\, v^{>N} \;. \tag{3.10}$$

One would like to choose χ so that $\{\chi, k\} + \varepsilon v^{\le N}$ is a function $\in \mathcal{F}_{\mathcal{L}}$. Precisely, one takes

$$\{k, \chi\} \;=\; \varepsilon \left(v^{\le N} \;-\; \Pi_{\mathcal{L}} v^{\le N} \right) . \tag{3.11}$$

This equation may be appropriately called ('linearized', and truncated) Hamilton–Jacobi equation[40] ('homological' equation is also used[8]). To solve it, we resort to coordinates. In terms of the local representatives of the various functions, (3.11) reads

$$\{k_\lambda, \chi_\lambda\} \;=\; \varepsilon \left(v_\lambda^{\le N} \;-\; \Pi_{\mathcal{L}} v_\lambda^{\le N} \right) . \tag{3.12}$$

Since $\{\chi_\lambda, k\} = \left\langle \omega,\, \frac{\partial \chi_\lambda}{\partial I} \right\rangle$, (3.12) can be easily solved by Fourier series techniques: one finds

$$\chi_\lambda \;=\; \varepsilon\, \mathcal{S}_{\omega,\mathcal{L}}^{\le N}\, v_\lambda \tag{3.13}$$

where the (formal) operator $\mathcal{S}_{\omega,\mathcal{L}}^{\le N}$ (whose introduction will somehow shorten some statements below) is defined as follows:

$$\mathcal{S}_{\omega,\mathcal{L}}^{\le N} f \;=\; \sum_{\nu \notin \mathcal{L}\,,\, |\nu| \le N} \frac{\langle f \rangle_\nu}{i\, \langle \omega,\, \nu \rangle}\, E_\nu \qquad\qquad \forall\; f \in \mathcal{F}(B \times \mathbb{T}^n) \tag{3.14}$$

As far as $N < \infty$ and the lattice $\mathcal{L}$ contains all the resonances of ω of order $\le N$ (namely, all vectors $\nu \in \mathbb{Z}^n$ such that $|\nu| \le N$ and $\langle \omega,\, \nu \rangle = 0$), the sum in (3.14) is well defined, and no convergence problem arises. One could then expect to have $\chi_\lambda = \mathcal{O}(\varepsilon)$, if $v_\lambda = \mathcal{O}(1)$. However, in typical situations one works in a region of phase space in which some of the denominators in (3.14) are small: $|\langle \omega,\, \nu \rangle| \sim \varepsilon^{1-a}$ for some $1 > a > 0$. Consequently, one has only $\chi = \mathcal{O}(\varepsilon^a)$.

We shall not discuss here further such a kind of problems. Indeed, our attitude in the present is simply that of assuming that all the necessary estimates have been performed within each coordinate system, in terms of the local representatives of functions. The key point is that, as one easily verifies, the Fourier components of the functions χ_λ constructed in this way do satisfy (3.4) (with $A_{\lambda\mu} = \mathbb{1}$). Hence, by lemma 2.5, we may conclude that the functions χ_λ are the local representatives of a function χ, which is manifestly a solution of the Hamilton–Jacobi equation (3.11). The Hamiltonian flow of χ is then globally defined on the manifold, and the transformed Hamiltonian (3.10) takes the form

$$h \circ \Phi_1^\chi \;=\; k \;+\; \varepsilon\, g \;+\; \varepsilon\, f \;, \tag{3.15}$$

where

$$\begin{aligned} g \;&=\; \Pi_{\mathcal{L}} v^{\le N} \\ f \;&=\; R_1^\chi(v) \;+\; \varepsilon^{-1}\, R_2^\chi(k) \;+\; v^{>N} \;. \end{aligned} \tag{3.16}$$

Explicit estimates for the remainder f can be worked out within each chart, by means of standard techniques.

We shall now formalize this argument, by paying the necessary attention to questions of domains.

E. Formalization of the results. In order to be able to reconstruct a global normal form from the local normal forms in the charts, we shall need that the intersection of the domains on which the latter ones are defined are not too small.

In the formulation of proposition 2.5 below we make reference to domains $B''_\lambda \subset B'_\lambda \subset B_\lambda$ $(\lambda \in \Lambda)$ with the following properties: for any $\lambda, \mu \in \Lambda$ such that the intersection $M_\lambda \cap M_\mu$ of the chart domains is not empty, we assume

$$\begin{aligned} \mathcal{C}_{\lambda\mu}(B'_\lambda \times \mathbb{T}^n) \cap (B'_\mu \times \mathbb{T}^n) &= B'_{\lambda\mu} \times \mathbb{T}^n \\ \mathcal{C}_{\lambda\mu}(B''_\lambda \times \mathbb{T}^n) \cap (B''_\mu \times \mathbb{T}^n) &= B''_{\lambda\mu} \times \mathbb{T}^n \end{aligned}$$

where $B'_{\lambda\mu}$ and $B''_{\lambda\mu}$ are open, not empty, connected subsets of $\mathbb{R}^{2m-n}$. Let us also define

$$\begin{aligned} M' &= \bigcup_{\lambda\in\Lambda} \mathcal{C}_\lambda^{-1}(B'_\lambda \times \mathbb{T}^n) \\ M'' &= \bigcup_{\lambda\in\Lambda} \mathcal{C}_\lambda^{-1}(B''_\lambda \times \mathbb{T}^n) \,. \end{aligned}$$

Proposition 2.6 *Within the above assumptions and notations:*

i) assume that the Fourier series $\chi_\lambda = \mathcal{S}^{\le N}_{\omega,\mathcal{L}} v_\lambda$ *define analytic functions* $\chi_\lambda \in \mathcal{F}(B'_\lambda \times \mathbb{T}^n)$ *for any* $\lambda \in \Lambda$*. Then there exists a unique function* $\chi \in \mathcal{F}(M')$ *of which the* χ_λ *are the local representatives;*

ii) assume, moreover, that, for each $\lambda \in \Lambda$*, the time–one map* $\Phi_1^{\chi_\lambda}$ *of the Hamiltonian flow of* χ_λ *is a diffeomorphism of* $B''_\lambda \times \mathbb{T}^n$ *onto its image. Then,* Φ_1^χ *is a canonical diffeomorphism of* M'' *onto its image, has local representatives* $\Phi_1^{\chi_\lambda}$ $(\lambda \in \Lambda)$ *and conjugates* $h|_{\Phi_1^\chi(M'')}$ *to*

$$h \circ \Phi_1^\chi \;=\; k \;+\; \varepsilon\, \Pi_{\mathcal{L}}\, v^{\le N} \;+\; \varepsilon\, f\,, \tag{3.17}$$

with

$$f \;=\; R_1^\chi(v) \;+\; \varepsilon^{-1}\, R_2^\chi(k) \;+\; v^{>N}\,. \tag{3.18}$$

Proof. The statement in part *i)* follows from the fact that the Fourier components of the functions $\chi_\lambda = \mathcal{S}^{\le N}_{\omega,\mathcal{L}} v_\lambda$ satisfy (3.4). Consider now part *ii)*. Since M is a symplectic manifold, the Hamiltonian vector field X of χ is defined globally on M'. The map at time 1 of its flow is canonical, and has local representatives $\Phi_1^{\chi_\lambda}$ (indeed, one has $\mathcal{C}_{\lambda\mu} \circ \Phi_1^{\chi_\mu} =$

$\Phi_1^{\chi_\lambda} \circ \mathcal{C}_{\lambda\mu}^{-1}$ on the intersection of the domains) and is defined on M''. The statements on the normal form are proven as in the previous subsection. ∎

One can iterate the construction of the normal form. The following result is obvious:

Proposition 2.7 *Consider a lattice $\mathcal{L} \subset \mathbb{Z}^n$. Assume that, for each $\lambda \in \Lambda$, the iterative scheme ($s = 0, 1, 2, \ldots$)*

$$
\begin{aligned}
h_\lambda^{(s)} &= k + \varepsilon\, g_\lambda^{(s)} + \varepsilon\, f_\lambda^{(s)} \\
g_\lambda^{(0)} &= 0 \\
f_\lambda^{(0)} &= v_\lambda \\
g_\lambda^{(s+1)} &= g_\lambda^{(s)} + [\Pi_{\mathcal{L}} f_\lambda^{(s)}]^{\leq N} \\
f^{(s+1)} &= R_1^{\chi_\lambda^{(s)}}\left(g_\lambda^{(s)} + f_\lambda^{(s)}\right) + \varepsilon^{-1} R_2^{\chi_\lambda^{(s)}}(k) + [f_\lambda^{(s)}]^{>N} \\
\chi_\lambda^{(s)} &= \mathcal{S}_{\omega,\mathcal{L}}^{\leq N} f_\lambda^{(s)}
\end{aligned}
$$

defines smooth functions $\chi_\lambda^{(0)}, \ldots, \chi_\lambda^{(r-1)}$, for some positive integer r, on some domains $B'_\lambda \times \mathbb{T}^n$. Assume also that, for each $\lambda \in \Lambda$, the composition of the time-one maps

$$
\Phi_\lambda^{(r)} = \Phi_\lambda^{\chi_\lambda^{(0)}} \circ \Phi_\lambda^{\chi_\lambda^{(1)}} \circ \cdots \circ \Phi_\lambda^{\chi_\lambda^{(r-1)}}
$$

is a diffeomorphism of a domain $B''_\lambda \times \mathbb{T}^n$ (we assume that the domains B'_λ and B''_λ satisfy the conditions specified above). Then, $\Phi_\lambda^{(r)}$, $\lambda \in \Lambda$, are the local representatives of a canonical diffeomorphism $\Phi^{(r)}$ of M'' onto its image, which conjugates h, restricted to $\Phi^{(r)}(M'')$, to

$$
h^{(r)} = k + \varepsilon\, g^{(r)} + \varepsilon\, f^{(r)} \tag{3.19}
$$

where $g^{(r)} \in \mathcal{F}_{\mathcal{L}}$ and $f^{(r)}$ have local representatives $g_\lambda^{(r)}$ and, respectively, $f_\lambda^{(r)}$.

Typically, after r steps, one arrives at a remainder $f^{(r)}$ which is of order ε^{rb}, for some positive number b (usually greater then the number a introduced in connection with the estimate of the small denominators of (3.14)). However, such a condition is here supposed to be tested on each chart.

F. Estimate of the variation of the actions. We now indicate how the 'global' construction of proposition 2.7 solves the problem raised in section 2.1. To this purpose, one could refer to the expansion (2.6). However, it is easier to proceed somehow differently. Let $J : M \to \mathbb{R}$ be a function, and let $J_\lambda : B_\lambda \times \mathbb{T}^n \to \mathbb{R}$ ($\lambda \in \Lambda$) be its local representatives. Let $t \mapsto z_t$ be a motion of the Hamiltonian system defined by h on M, and assume that $z_t \in M''$ for all t in a considered interval. Define $z'_t = \Phi^{(r)}(z_t)$. Then, z'_t is a motion of the Hamiltonian system described by $h^{(r)}$ on $\Phi^{(r)}(M'')$. Thus, we may write

$$
J(z_t) - J(z_0) = [J(z_t) - J(z'_t)] + [J(z'_t) - J(z'_0)] + [J(z'_0) - J(z_0)] \tag{3.20}
$$

and compute the second term at the r.h.s. as

$$J(z'_t) - J(z'_0) = \int_0^t \{h', J\} \circ \Phi_\tau^{h'}(z'_0) d\tau$$

where $\Phi^{h'}$ is the Hamiltonian flow of h'. Thus, if J is a first integral of the unperturbed system, we get

$$J(z'_t) - J(z'_0) = \varepsilon \int_0^t \{g, J\} \circ \Phi_\tau^{h'}(z'_0) d\tau + \varepsilon^{1+ra} \int_0^t \{f, J\} \circ \Phi_\tau^{h'}(z'_0) d\tau .$$

The integrals can be evaluated, or estimated, in the charts. In the simple case considered in section 1.2 ($J = I$, $\mathcal{L} = 0$) one has $\{g, J\} = 0$ and then

$$|J(z'_t) - J(z'_0)| \leq |t| \varepsilon C , \tag{3.21}$$

with

$$C = \sup_{\lambda \in \Lambda} \sup_{(b_\lambda, \varphi_\lambda) \in B_\lambda \times \mathbb{T}^n} \left| \{f_\lambda^{(r)}, J_\lambda\}(b_\lambda, \varphi_\lambda) \right| \sim \varepsilon^{rb} .$$

(for some $b > 0$). Together with (3.20), (3.21) furnishes an estimate for the variation of J in which the effect of the deformation of the canonical transformation is evaluated only two times.

G. On cases with no global actions. Besides the obvious changes due to the fact that the transition matrices $A_{\lambda\mu}$ cannot be taken all equal to the identity, the basic difficulty in this case is that the resonant manifolds could be not defined globally on the manifold.

It is possible that this fact does not prevent, in practice, the possibility of a perturbation theory. The point is that the normal forms constructed in each chart confine the motions of the local action variables. Thus, one expects that the actions can visit only few chart domains, for which no global obstruction to the reconciliability of the matrices $A_{\lambda\mu}$ exist.

Anyway, we do not investigate further such a case, since we are not aware of its occurrence in any Hamiltonian system.

Chapter 3

The action–angle variables of the symmetric rigid body

In this chapter we study the action–angle variables of the symmetric Euler–Poinsot problem, namely the 'Poinsot variables'. In section 1 we define such variables, and state a number of their properties. Section 2 deals more specifically with the use of these coordinates in the rigid body dynamics. Some proofs are collected in section 3.

3.1 The Poinsot variables

As in section 1.2A we identify $T^*SO(3)$ with $SO(3) \times \mathbb{R}^3$, by choosing two frames $\mathcal{B}_b = \{e_1, e_2, e_3\}$ and $\mathcal{B}_s = \{e_x, e_y, e_z\}$ which are, respectively, attached to the body and fixed in space. We think the basis vectors of $\mathcal{B}_s$ as being functions of $\mathcal{R} \in SO(3)$: $e_x = \mathcal{R}e_1$, $e_y = \mathcal{R}e_2$, $e_z = \mathcal{R}e_3$. The 'Poinsot variables' have been already introduced in section 1.2C (see figure 1.3). Here, we formalize their definition, and consider with some care the singularities of such coordinates. With reference to the two frames $\mathcal{B}_b$, $\mathcal{B}_s$, we define the following subset of $SO(3) \times \mathbb{R}^3$:

$$\Sigma(\mathcal{B}_b, \mathcal{B}_s) \;=\; \{(\mathcal{R}, m) \in SO(3) \times \mathbb{R}^3 \;:\; m \times e_z \neq 0\,,\; m \times e_3 \neq 0\}. \tag{1.1}$$

Definition 3.1 *The* Poinsot angles *relative to the two bases $\mathcal{B}_b$, $\mathcal{B}_s$ are the three real functions $j(mod\,2\pi)$, $g(mod\,2\pi)$, $l(mod\,2\pi)$ defined on $\Sigma(\mathcal{B}_b, \mathcal{B}_s)$ as follows:*

- *j is the angle (in the 'fixed' plane $e_x \oplus e_y$ †) from e_x to $e_z \times m$, anticlockwise about e_z;*
- *g is the angle (in the plane orthogonal to m through the origin) from $e_z \times m$ to $m \times e_3$, anticlockwise about m;*

† We denote by $u \oplus v$ the linear space spanned by the two vectors u and v.

- *l is the angle (in the 'moving' plane $e_1 \oplus e_2$) from $m \times e_3$ to e_1, anticlockwise about e_3.*

The Poinsot variables *relative to the frames $\mathcal{B}_b$, $\mathcal{B}_s$ are constituted by the three Poinsot angles and by the three* Poinsot momenta $J = \langle m, e_z \rangle$, $G = \langle m, m \rangle^{1/2}$, $L = \langle m, e_3 \rangle$.

Analytic expressions of the Poinsot angles as functions of $(\mathcal{R}, m)$ are easily obtained. First, one deduces from the very definition of the three angles the expressions

$$\begin{aligned} \cos j &= \frac{\langle e_x, e_z \times m \rangle}{\|e_z \times m\|_e} = -\frac{m_y}{\sqrt{m_x^2 + m_y^2}} \\ \cos g &= \frac{\langle e_z \times m, m \times e_3 \rangle}{\|e_z \times m\|_e \, \|m \times e_3\|_e} \\ \cos l &= \frac{\langle m \times e_3, e_1 \rangle}{\|m \times e_3\|_e} = \frac{m_2}{\sqrt{m_1^2 + m_2^2}} \end{aligned} \tag{1.2}$$

(here and in the following $\| \, \|_e$ denotes the euclidean norm). The second expressions of the first and the third equations reflect the fact that j is the angle between the axis e_y and the orthogonal projection of m into the 'fixed' plane $e_x \oplus e_y$, while l is the angle between the orthogonal projection of m into the 'moving' plane $e_1 \oplus e_2$, and the axis e_2.

We shall see in a moment that the Poinsot angles relative to two frames $\mathcal{B}_b$, $\mathcal{B}_s$ are properly defined in all of the set $\Sigma(\mathcal{B}_b, \mathcal{B}_s)$, which is clearly an open, dense, connected subset of $SO(3) \times \mathbb{R}^3$. On the other hand, they cannot be extended to all of $SO(3) \times \mathbb{R}^3$, since they cannot be defined on the boundary of $\Sigma(\mathcal{B}_b, \mathcal{B}_s)$: indeed, the 'nodal lines' used for their constructions are no more defined if $m = 0$ (i.e. $G = 0$) or else if m is parallel either to e_3 ($|L| = G \neq 0$) or to e_z ($|J| = G \neq 0$). The last two singularities, but not of course the one corresponding to $m = 0$, can in principle be removed by changing the frames $\mathcal{B}_b$ and $\mathcal{B}_s$, i.e. by using other 'Poinsot charts'. Such a possibility is made precise in the following proposition.[27)]

Proposition 3.1

i) With reference to any choice of the frames $\mathcal{B}_b$ and $\mathcal{B}_s$, the Poinsot variables define an analytic diffeomorphism $(\mathcal{R}, m) \mapsto (J, G, L, j, g, l)$ of $\Sigma(\mathcal{B}_b, \mathcal{B}_s)$ onto $P \times \mathbb{T}^3$, where P is the 'piramid'

$$P = \{(J, G, L) \in \mathbb{R}^3 \, : \, G > 0 \, , \, |J| < G \, , \, |L| < G\} \, ; \tag{1.3}$$

ii) such a diffeomorphism is symplectic, relatively to the natural symplectic structure of cotangent bundle of $\Sigma(\mathcal{B}_b, \mathcal{B}_s)$, and to the symplectic two-form $dJ \wedge dj + dG \wedge dg + dL \wedge dl$ of $P \times \mathbb{T}^3$;

---

27) In the application to the rigid body, it is usually appropriate to choose the axes e_1, e_2, e_3 of the reference $\mathcal{B}_b$ to be the inertia axes of the body. Specifically, if the body is symmetric, one takes e_3 to be the inertia symmetry axis; instead, for a tri-axial body, e_3 is chosen differently in the two regions in which the separatrices divide the phase space, precisely as the minimal or the maximal inertia axis. Consequently, in the study of the rigid body, it will be easy to change $\mathcal{B}_s$, and correspondingly remove the singularities $G = |J|$, but it will be not possible to change $\mathcal{B}_b$. Correspondingly, the singularities $G = |L|$ and $G = |J|$ will have a completely different role in perturbation theory, in accordance with the discussion of chapter 1.

iii) let $\mathcal{B}_s^{(1)}$, $\mathcal{B}_s^{(2)}$ be any two 'fixed' frames with non-parallel z-axes and $\mathcal{B}_b^{(1)}$, $\mathcal{B}_b^{(2)}$ two 'moving' frames with non-parallel e_3-axes. Then, the Poinsot variables relative to the four sets $\Sigma(\mathcal{B}_b^{(\mu)}, \mathcal{B}_s^{(\nu)})$ $(\mu, \nu = 1, 2)$ constitute an analytic symplectic atlas for $SO(3) \times [\mathbb{R}^3 \setminus \{0\}]$.

The proof of this proposition is deferred to section 3.3A.

We consider now the relation between Poinsot variables and Euler's canonical coordinates φ, ψ, θ, p_φ, p_ψ, p_θ, obviously relative to the same couple of frames $\mathcal{B}_b$, $\mathcal{B}_s$. Since the Euler angles have a singularity for $\theta = 0$, we restrict ourselves to the set

$$\Sigma'(\mathcal{B}_b, \mathcal{B}_s) \;=\; \Sigma(\mathcal{B}_b, \mathcal{B}_s) \setminus \{(\mathcal{R}, m) : \langle e_z, e_3 \rangle = 0\} . \tag{1.4}$$

Proposition 3.2 *(Deprit[34], Boigey[23,24]) In all of $\Sigma'(\mathcal{B}_b, \mathcal{B}_s)$, the Poinsot variables are analytic invertible functions of the Euler canonical coordinates, and satisfy*

$$J dj + G dg + L dl \;=\; p_\varphi d\varphi + p_\psi d\psi + p_\theta d\theta . \tag{1.5}$$

The transformation equations are

$$\cos\theta \;=\; \Theta\Big(\frac{J}{G}, \frac{L}{G}, g\Big) \tag{1.6a}$$

$$\cos(\varphi - j) \;=\; \Phi\Big(\frac{J}{G}, \frac{L}{G}, g\Big) \tag{1.6b}$$

$$\cos(\psi - l) \;=\; \Psi\Big(\frac{J}{G}, \frac{L}{G}, g\Big) \tag{1.6c}$$

$$p_\theta \;=\; G\sqrt{1 - \Big(\frac{L^2}{G^2}\Big)}\Big[1 - \Psi\Big(\frac{J}{G}, \frac{L}{G}, g\Big)^2\Big] \tag{1.6d}$$

$$p_\varphi \;=\; J \tag{1.6e}$$

$$p_\psi \;=\; L \tag{1.6f}$$

where the functions Θ, Φ, Ψ are defined by

$$\begin{aligned}\Theta(x, y, g) \;&=\; xy \;-\; \sqrt{1 - x^2}\,\sqrt{1 - y^2}\,\cos g \\ \Phi(x, y, g) \;&=\; \Psi(y, x, g) \;=\; \frac{y - x\Theta(x, y, g)}{\sqrt{1 - \Theta(x, y, g)^2}\,\sqrt{1 - x^2}}\end{aligned} \tag{1.7}$$

Moreover, one has

$$G^2 \;=\; p_\psi^2 + p_\theta^2 + (p_\varphi - p_\psi \cos\theta)^2 \,\sin^{-2}\theta . \tag{1.8}$$

The proof of this proposition is demanded to section 3.3B.

Notice that the Euler angles θ, φ, ψ are functions of j, g, l, J/G, L/G. Thus, they are homogeneous functions of degree zero of the momenta G, J, L.[28] This fact has the following obvious, but important, consequence (already stated in section 1.4). Consider an analytic ('purely positional') function $\mathcal{V} : SO(3) \to \mathbb{R}$, and extend it to $T^*SO(3)$ in the trivial way: $\mathcal{V}(\mathcal{R}, m) \equiv \mathcal{V}(\mathcal{R})$. Then, in each set $\Sigma(\mathcal{B}_b, \mathcal{B}_s)$, $\mathcal{V}$ has a local representative v which is an analytic function of j, g, l, J/G, L/G. In particular, v is a homogeneous function of degree zero of the three momenta J, G, L:

$$v(j,g,l,\lambda J,\lambda G,\lambda L) \;=\; v(j,g,l,J,G,L) \qquad \forall\lambda > 0\,. \tag{1.9}$$

We consider now in detail the subset of $SO(3) \times \mathbb{R}^3$ given by

$$\Sigma(\mathcal{B}_b) \;=\; \{(\mathcal{R}, m) \in SO(3) \times \mathbb{R}^3 \;:\; m \times e_3 \neq 0\}\,. \tag{1.10}$$

Notice that, in the case of a symmetric body, if e_3 is the inertia symmetry axis of the body, such a set coincides with the set Σ_* introduced in section 1.3C. Consequently, it is assured by proposition 2.3, that $\Sigma(\mathcal{B}_b)$ is an angular fibering of order two and dimension three, which possesses global actions. The generalized local action–angle variables for $\Sigma(\mathcal{B}_b)$ are the Poinsot variables, and clearly two charts are sufficient for an atlas. We now consider in some detail the transition functions:

Proposition 3.3 *For any choice of the frame $\mathcal{B}_b$, the set $\Sigma(\mathcal{B}_b)$ has an atlas constituted by two Poinsot charts*

$$(j_\mu, g_\mu, l_\mu, J_\mu, G_\mu, L_\mu) \;:\; \Sigma(\mathcal{B}_b, \mathcal{B}_s^{(\mu)}) \to P \times \mathbb{T}^3\,, \qquad \mu = 1, 2\,,$$

where the two 'fixed' frames $\mathcal{B}_s^{(1)}$, $\mathcal{B}_s^{(2)}$ have non-parallel z-axes, with the following transition functions:

$$\begin{aligned} G_2 &= G_1\,, & g_2 &= g_1 + \hat{g}_{12}\Big(\frac{J_1}{G_1}, j_1\Big) \\ L_2 &= L_1\,, & l_2 &= l_1 \\ J_2 &= G_1\,\hat{J}_{12}\Big(\frac{J_1}{G_1}, j_1\Big)\,, & j_2 &= \hat{j}_{12}\Big(\frac{J_1}{G_1}, j_1\Big) \end{aligned} \tag{1.11}$$

where the functions $\hat{g}_{12}$, $\hat{J}_{12}$, $\hat{j}_{12}$ (which depend only on the mutual orientation of the two frames $\mathcal{B}_s^{(1)}$, $\mathcal{B}_s^{(2)}$) are analytic for $-1 < J_1/G_1 < 1$ and $j_1 \in S^1$. Moreover, the change of coordinate is analytically invertible and symplectic.

Thus, G, L, g, l are the action–angle variables of the angular fibering $\Sigma(\mathcal{B}_b)$, while J and j are its 'degenerate' variables. The two actions G and L are globally defined, and map $\Sigma(\mathcal{B}_b)$ onto the open domain of the plane (in fact, an open angle)

$$\mathcal{I} \;=\; \{(G, L) \in \mathbb{R}^2 \;:\; G > 0,\, G > |L|\} \tag{1.12}$$

which can be identified with the 'action space' of definition 2.2.

[28] A canonical transformation which preserves the symplectic one–form, as in (1.5), is called 'homogeneous' (or of the 'Mathieu type') since the coordinates and the momenta of one group of coordinates are homogeneous functions of degrees, respectively, zero and one of the momenta of the other group (see [75]).

3.2 The Poinsot variables and the rigid body

Let now the vectors e_1, e_2, e_3 of the frame $\mathcal{B}_b$ be inertia axes of the body, corresponding, respectively, to the inertia moments a_1, a_2, a_3.

Proposition 3.4 (Deprit) *The Hamilton function of the free rigid body, in local Poinsot coordinates in $\Sigma(\mathcal{B}_b, \mathcal{B}_s)$, is*

$$k(l, G, L) \;=\; \left(\frac{\sin^2 l}{2a_1} + \frac{\cos^2 l}{2a_2}\right)(G^2 - L^2) \;+\; \frac{1}{2a_3} L^2 \,. \tag{2.1}$$

Proof. As noticed before, l is the angle between the axis e_2 and the orthogonal projection m_{12} of m into the plane $e_1 \oplus e_2$. Thus, $m_1 = m_{12} \sin l$ and $m_2 = m_{12} \cos l$. On the other hand, one has $m_{12}^2 = \|m\|_e^2 - m_3^2 = G^2 - L^2$, so that the kinetic energy $\frac{1}{2}\left(\frac{m_1^2}{a_1} + \frac{m_2^2}{a_2} + \frac{m_3^2}{a_3}\right)$ takes the form (2.1). ∎

Local action–angle variables for the 'free' rigid body with tri–axial inertia ellipsoid can be constructed out from the Poinsot variables, by employing the usual method of Liouville (see especially [51,53], and also [3, 64, 69, 9]).

The Poinsot variables are local (generalized) action–angle variables for the symmetric rigid body in every chart domain $\Sigma(\mathcal{B}_b, \mathcal{B}_s)$, if of course the body frame $\mathcal{B}_b$ is chosen appropriately, namely with the axis e_3 coincident to the symmetry inertia axis. In such a case one has $a_1 = a_2$ and the Hamilton function (2.1) reduces to the form (2.6) :

$$k(G, L) \;=\; \frac{1}{2a_1} G^2 \;+\; \frac{a_1 - a_3}{2a_1 a_3} L^2 \,. \tag{2.2}$$

The equations of the motion of the regular precession are

$$\begin{aligned} \dot J &= 0\,, & \dot j &= 0 \\ \dot G &= 0\,, & \dot g &= \omega_1(G) \\ \dot L &= 0\,, & \dot l &= \omega_2(L) \end{aligned} \tag{2.3}$$

with $\omega_1 = G/a_1$ and $\omega_2 = \frac{a_1 - a_2}{a_1 a_3} L$ (see (2.9) of chapter 1).

Consider now the problem of the motion in an external force field, which we assume described by a 'purely positional' and analytic potential energy $\mathcal{V} : SO(3) \to \mathbb{R}$. We shall be interested in studying the motion in the whole subset $\Sigma(\mathcal{B}_b) = \Sigma_*$ of phase space. Such a set is covered by two Poinsot charts, corresponding to two spatial frames with non parallel z–axes.

Let us consider one of these two charts. Let $v : P \times \mathbb{T}^3$ be the representative of $\mathcal{V}$. As remarked, v is an analytic function of the three angles g, j, l and of the two ratios J/G, L/G. Of course, v is bounded in all of $P \times \mathbb{T}^3$:

$$\sup_{z \in P \times \mathbb{T}^3} |v(z)| \;\le\; \sup_{\mathcal{R} \in SO(3)} |\mathcal{V}(\mathcal{R})| \,.$$

However, the *derivatives* of v can (and usually do, as a glance at equations (1.6)–(1.7) indicates) diverge on the boundary $(\partial P) \times \mathbb{T}^3$ of the chart domain (precisely, for $|J| \to G$ and $|L| \to G$). Thus, perturbation theory (in action–angle variables) has to be performed in a subset of $P \times \mathbb{T}^3$, from which a neighbourhood of the boundary has been removed.

Such a subset can be obtained, for instance, by removing from P a layer of uniform width near its boundary.Indeed, let

$$P - u \;=\; \{(J,G,L) \in \mathbb{R}^3 \,:\, G > u\,,\; |J| < G - u\,,\;,\; |L| < G - u\} \tag{2.4}$$

for some positive number u. By Cauchy estimate, v being analytic in $P \times \mathbb{T}^3$, all first derivatives of v are uniformly bounded in $P - u$, for any $u > 0$.[29)]

Remark 3.1 A careful analysis shows that the derivatives of the local representatives of (purely positional) functions are uniformly bounded up to distances $\mathcal{O}(G^{-1})$ from the singularity $G = |L|$. However, we shall not make use of such a fact.

Let us conclude by noticing that the pre–images under the Poinsot variables $\mathcal{C}^{(\mu)}$ ($\mu = 1,2$) of the sets $P - u \times \mathbb{T}^3$ cover[30)] the subset of $SO(3) \times \mathbb{R}^3$

$$\Sigma - u \;=\; \{(\mathcal{R}, m) \,:\, \|m\|_e > u\,,\; |\langle m,\, e_3\rangle| < \|m\|_e - u\}\,, \tag{2.5}$$

which is again a fiber bundle, with global action variables G, L and action space

$$\mathcal{I} - u \;=\; \{(G,L) \in \mathbb{R}^2 \,:\, G > u\,,\; |L| < G - u\}\,. \tag{2.6}$$

3.3 Proofs of propositions 3.1, 3.2 and 3.3

A. Proof of Proposition 3.1. First of all, we show that the mapping $\mathcal{C} : (\mathcal{R}, m) \mapsto (j,g,l,J,G,L)$ is a bijection of $\Sigma(\mathcal{B}_b, \mathcal{B}_s)$ onto $P \times \mathbb{T}^3$. This would not be strictly necessary, since we shall re–obtain such a result as a byproduct of the proof of the analyticity of such mapping. Nevertheless, it may be useful to give a simple and descriptive geometric argument. The mapping $\mathcal{C} : \Sigma(\mathcal{B}_b, \mathcal{B}_s) \to P \times \mathbb{T}^3$ is manifestly well defined. We thus show that its inverse is properly defined from $P \times \mathbb{T}^3$ into $\Sigma(\mathcal{B}_b, \mathcal{B}_s)$. To this purpose, we consider any point $(j,g,l,J,G,L,) \in P \times \mathbb{T}^3$ and proceed as follows:

- first (fig. 3.1), the angle j in the plane $e_x \oplus e_y$, measured from e_x toward e_y, determines direction and orientation of a (non–zero) vector, which we identify with $e_z \times m$; this

29) Of course, one could exclude a larger neighbourhood of $|J| = G$. For instance, if ζ is the angle between the z–axes of the two spatial frames, one could restrict P according to $|J| < G\cos(\zeta/2)$, and correspondingly to switch chart as soon as $|J| = G\cos(\zeta/2)$.

30) Obviously, we assume that the angle between the z–axes of the two spatial frames is not too small.

fixes the plane containing e_z and m. Together with $J = \langle m,\, e_z \rangle$ and $G = \langle m,\, m \rangle$, this completely determines the vector m; since $G \neq 0$ and $|J| < G$, such a vector is different from zero and non parallel to e_z;

- next, consider the plane orthogonal to m through the origin, and in such a plane the angle g, taken anticlockwisely (with respect to m) from $e_z \times m$. This angle determines direction and orientation of a (non–zero) vector, which we interpret as $m \times e_3$. Together with $L = \langle m,\, e_3 \rangle$ this uniquely determines the vector e_3 having unit euclidean norm, which is non parallel to m (since $|L| < G$);
- finally, the angle l in the plane orthogonal to e_3, anticlockwise (with respect to e_3) from $m \times e_3$, fixes e_1. The construction is completed by taking $e_2 = e_3 \times e_1$, and defining $\mathcal{R}$ as the operator which transforms the vectors e_1, e_2, e_3 into, respectively, e_x, e_y, e_z.

In this way we have constructed a mapping $P \times \mathbb{T}^3 \to \Sigma(\mathcal{B}_b, \mathcal{B}_s)$, which is easily seen to be the inverse of $\mathcal{C}$.

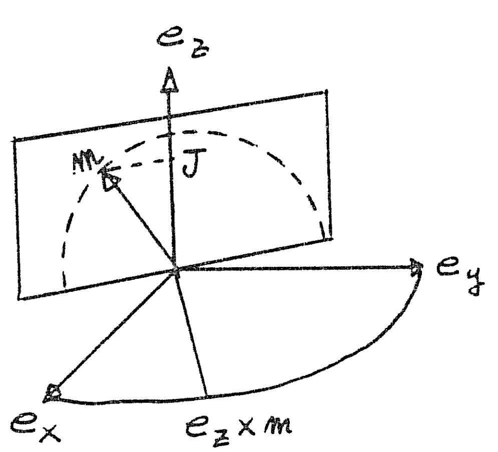

Figure 3.1

The analyticity of $\mathcal{C}$ follows immediately from the expressions (1.2) of the Poinsot variables in terms of $(\mathcal{R}, m)$. Concerning $\mathcal{C}^{-1}$, let us first notice that, as one easily verifies starting with (1.2), the representatives of the vector m in the basis $\mathcal{B}_b$ and $\mathcal{B}_s$ are, respectively,

$$m^{(b)} = \begin{pmatrix} \sqrt{G^2 - L^2}\sin l \\ \sqrt{G^2 - L^2}\cos l \\ L \end{pmatrix} , \qquad m^{(s)} = \begin{pmatrix} \sqrt{G^2 - J^2}\sin j \\ -\sqrt{G^2 - J^2}\cos j \\ J \end{pmatrix} . \tag{3.1}$$

This shows that m is an analytic function of the the Poinsot variables. To show that $\mathcal{R}$ too is analytic, one can construct its matrix relative to the two bases $\mathcal{B}_s$, $\mathcal{B}_b$; indeed, it is easy to see that $\mathcal{R}$ can be expressed as the product of five suitable rotations, whose matrices are explicitly constructed and are manifestly analytic in the Poinsot variables.

In order to prove the other statements *i)* and *ii)* we use some results stated in propositions 3.2 and 3.3 (which are proven below).

First, it is stated in proposition 3.2 that the Poinsot variables are related by a canonical transformation to the Euler coordinates. Precisely, in the set $\Sigma'(\mathcal{B}_b, \mathcal{B}_s)$ defined by (1.4) one has

$$Jdj + Gdg + Ldl = p_\varphi d\varphi + p_\psi d\psi + p_\theta d\theta \,. \tag{3.2}$$

Now, the cotangent bundle $T^*SO(3)$ possesses a symplectic one–form σ which is globally defined and coincides with $p_\varphi d\varphi + p_\psi d\psi + p_\theta d\theta$ whenever the latter is defined (this is assured by Darboux theorem, since the Euler angles are local coordinates on $SO(3)$). Thus, (3.2) implies that one has $Jdj + Gdg + Ldl = \sigma$ in $\Sigma'(\mathcal{B}_b, \mathcal{B}_s)$; by continuity, this relation holds in all of Σ (actually, $\Sigma'(\mathcal{B}_b, \mathcal{B}_s)$ is dense in Σ and the Poinsot variables are continuous in all of Σ). This proves that the Poinsot variables are symplectic coordinates on Σ.

Consider now part *iii)*. The four sets $\Sigma(\mathcal{B}_b^{(\mu)}, \mathcal{B}_s^{(\nu)})$ clearly cover $SO(3) \times [\mathbb{R}^3 \setminus \{0\}]$. Thus, to prove *iii)* we have to show that the different coordinate systems patch in the right way in the intersections of the chart domains. The transition mappings (1.11) of proposition 3.3 indicate that this is the case for the Poinsot variables $(j_1, g_1, l_1, J_1, G_1, L_1)$ in $\Sigma(\mathcal{B}_b, \mathcal{B}_s^{(1)})$ and the Poinsot variables $(j_2, g_2, l_2, J_2, G_2, L_2)$ in $\Sigma(\mathcal{B}_b, \mathcal{B}_s^{(2)})$, where $\mathcal{B}_s^{(1)}$ and $\mathcal{B}_s^{(2)}$ have non–parallel z–axes. It is easy to see that the transition mappings between the other couples of charts have the same good properties as (1.11). Indeed, the transition mapping corresponding to the change of the ('moving') frame is obtained form (1.11) by exchanging the roles of the couples of variables (j, J) and (l, L), while those corresponding to changes of both the frames can be written as compositions of two changes of coordinates of the types above. The proof of proposition 3.1 is complete.

B. Proof of proposition 3.2. We do not repeat here the argument (based on elementary considerations of spherical trigonometry) which proves equation (1.5) (see [34], and also [23,24]).

The analitycity properties of the change of coordinates follow from the transformation equations (1.6), which we now prove. Equation (1.6a) is proven simply by observing that the second equation (1.2) can be written in the form

$$\cos g = \frac{JL - G^2 \cos\theta}{\sqrt{G^2 - L^2}\sqrt{G^2 - J^2}} \,.$$

From figure 1.3 one draws the expressions

$$\cos(\varphi - j) = \frac{\langle e_z \times m,\, e_z \times e_3 \rangle}{\|e_z \times m\|_e \, \|e_z \times e_3\|_e} = \frac{L - J\cos\theta}{|\sin\theta|\,\sqrt{G^2 - J^2}} \,,$$

$$\cos(\psi - l) = \frac{\langle e_z \times m,\, m \times e_3 \rangle}{\|e_z \times m\|_e \, \|m \times e_3\|_e} = \frac{J - L\cos\theta}{|\sin\theta|\,\sqrt{G^2 - L^2}} \,,$$

which give (1.6b,c). The equality (1.6e,f) are obvious. We now prove the expression (1.6f) for p_θ, which is the component of m along the nodal line $e_n = (\sin\theta)^{-1} e_z \times e_3$.

Let γ be the angle between m and e_3. The projection of m along the axis directed and oriented like $e_3 \times (m \times e_3)$ is $G|\sin\gamma|$. The angle between the (oriented) axes $e_3 \times (m \times e_3)$ and e_n is either $l - \psi - \pi/2$ or its complement to 2π; in both the cases, one has $p_\theta = G|\sin\theta|\sin(\psi - l)$. Equation (1.6f) immediately follows, since $\cos\gamma = L/G$. Finally, (1.8) is proven by observing that $G^2 = \langle m, m\rangle = m_1^2 + m_2^2 + m_3^2$ and using the standard expressions

$$
\begin{aligned}
m_1 &= \frac{(p_\varphi - p_\psi \cos\theta)\sin\psi + p_\theta \sin\theta\cos\psi}{\sin\theta} \\
m_2 &= \frac{(p_\varphi - p_\psi \cos\theta)\cos\psi - p_\theta \sin\theta\sin\psi}{\sin\theta}
\end{aligned}
$$

(which can be found, for instance, in [43]). The transformation equations (1.6) are manifestly analytic in $\Sigma'(\mathcal{B}_b, \mathcal{B}_s)$. Their (analytic) invertibility is also easily proven.

C. Proof of proposition 3.3. The equalities $G_2 = G_1$, $L_2 = L_1$ and $l_2 = l_1$ are obvious, since no one of these functions depend on the fixed frame. The basis vectors $e_x^{(2)}$, $e_y^{(2)}$, $e_z^{(2)}$ of the frame $\mathcal{B}_s^{(2)}$ can be written as linear combinations of those of $\mathcal{B}_s^{(1)}$, with coefficients depending only on the mutual orientation of the two frames: $e_a^{(2)} = \sum_b c_{ab} e_b^{(1)}$, with $a, b = x, y, z$. One then easily computes, using (1.2):

$$
\begin{aligned}
J_2 &= J_1 c_{zz} + \left(c_{zx}\sin j_1 - c_{zy}\cos j_1\right)\left(G^2 - J_1^2\right)^{1/2} \\
\cos j_2 &= J_1 c_{zz}\left(G^2 - J_1^2\right)^{-1/2} + \left(c_{yx}\sin j_1 - c_{yy}\cos j_1\right)
\end{aligned}
$$

The angles g_1 and g_2 lie on the same plane, and have one side in common; therefore, they satisfy

$$
\cos(g_1 - g_2) = \frac{\left\langle e_z^{(1)} \times m,\, e_z^{(2)} \times m\right\rangle}{\|e_z^{(1)} \times m\|_e \|e_z^{(2)} \times m\|_e} ,
$$

which implies

$$
g_2 = g_1 - \frac{c_{zz}G_1^2 - J_1 J_2}{\sqrt{G_1^2 - J_1^2}\sqrt{G_2^2 - J_2^2}} .
$$

The analyticity of the transformation equations in the stated domain is manifest, as well as their invertibility. The symplectic character of the transformation is obvious, since by proposition 3.2 each of the two coordinate systems is canonically related to a set of Euler coordinates.

Chapter 4

Perturbation theory: the normal forms

In this chapter we construct, with the techniques of perturbation theory and Nekhoroshev's theorem, the 'normal forms' for the fast rotating symmetric rigid body. Section 1 contains some preliminary materials: the definitions of domains and norms and the (very simple, in the present case) 'geography of resonances'. In section 2 we state the proposition 4.2 on the normal forms, which is then proved in sections 3 and 4. An appendix is devoted to the Lie methods for vector fields. The use of the normal forms for the description of the fast motions of the rigid body is deferred to the next chapter.

4.1 Preliminaries

A. The system. The material in the present chapter covers both the cases of a rigid body with a fixed point and the case with no ixed point. Actually, we are mainly interested in the former case, but the consideration of the latter one introduces no additional complications. The inertia ellipsoid of the body relative to its center of mass (or to its fixed point, depending on the case) is assumed to be symmetric.

The system is defined on the phase space $T^*SO(3) \times Q$, where Q is some open domain in $\mathbb{R}^6$. Let $(p,q) \in \mathbb{R}^6$ be canonical coordinates on Q, $q = (q_1, q_2, q_3)$ being the coordinates of the center of mass and $p = M\dot{q}$ the total linear momentum of the body; M is the mass of the body. Then, the Hamilton function has the form

$$h(\mathcal{R}, m, q, p) \;=\; k(m) \;+\; u(p) \;+\; \mathcal{V}(\mathcal{R}, q) \tag{1.1}$$

where k is given by (2.3) of chapter 1,

$$u(p) \;=\; \frac{\langle p, p\rangle}{2M} \tag{1.2}$$

and $\mathcal{V}$ is the potential energy, which is assumed to depend only on the configuration of the body[31]. The case of a body with a fixed point is recovered simply by taking $Q = \emptyset$, $u = 0$, $\mathcal{V} = \mathcal{V}(\mathcal{R})$, i.e. ignoring the coordinates p, q.

We consider the system (1.1) on the restricted domain $\Sigma_* \times Q$, with Σ_* defined as in section 1.2. The Poinsot coordinates, together with the coordinates (p,q), supply $\Sigma_* \times Q$ with an atlas of two 'Poinsot' charts, which we denote (changing a little the notation)

$$(A_1, A_2, J^{(\lambda)}, \alpha_1^{(\lambda)}, \alpha_2, j^{(\lambda)}, p, q) : \Sigma(\mathcal{B}_b, \mathcal{B}_s) \times Q \to P \times \mathbb{T}^3 \times Q \qquad (\lambda = 1, 2) \tag{1.3}$$

where $A_1 = G$, $A_2 = L$, $\alpha_1^{(\lambda)} = g^{(\lambda)}$ and $\alpha_2 = l$. Of course, we take the coordinates A_1, A_2 and α_2 as coincident in the two charts. It is understood that the axis e_3 of $\mathcal{B}_b$ is the inertia symmetry axis of the body, that the z–axes of the two frames $\mathcal{B}_s^{(1)}$ and $\mathcal{B}_s^{(2)}$ are (for definiteness) orthogonal, and obviously that the origin of these frames is the center of mass or the fixed point, depending on the case. We shall always write $A = (A_1, A_2)$ and $\alpha^{(\lambda)} = (\alpha_1^{(\lambda)}, \alpha_2)$.

Let

$$h^{(\lambda)}(A, J^{(\lambda)}, \alpha^{(\lambda)}, j^{(\lambda)}, p, q) = k(A) + u(p) + v^{(\lambda)}(A, J^{(\lambda)}, \alpha^{(\lambda)}, j^{(\lambda)}, q), \quad \lambda = 1, 2$$

be the local representatives of the Hamilton function (1.1). The function $v^{(\lambda)}$ depends on the variables A, $J^{(\lambda)}$ only through the ratios A_2/A_1 and $J^{(\lambda)}/A_1$. Furthermore, one has

$$k(A) = \frac{1}{2a_1}\left(A_1^2 + \eta A_2^2\right), \qquad \eta = \frac{a_1 - a_3}{a_3}. \tag{1.4}$$

The parameter η is in principle subjected only to the condition $\eta \geq -1/2$. The body is oblate if $\eta < 0$, spherical if $\eta = 0$, prolate if $\eta > 0$. The limiting value $\eta = -1/2$ is attained by plane bodies. We shall assume that the body is not spherical.[32]

B. Properties of $k(A)$. The 'unperturbed Hamiltonian' $k(A)$ is convex if $\eta > 0$, and quasi–convex in the region

$$\mathcal{I} = \{A \in \mathbb{R}^2 \,:\, A_1 > 0,\, A_1 > |A_2|\} \tag{1.5}$$

if $\eta < 0$ (as one readily verifies, referring to definition 1.2). We are here interested in some very elementary properties of $k(A)$, and of the frequency mapping $\omega = \frac{\partial k}{\partial I} : \mathcal{I} \to \mathbb{R}^2$, which is explicitly given by

$$\omega(A) = \frac{1}{a_1}\left(A_1, \eta A_2\right). \tag{1.6}$$

---

[31] In fact, the only important thing is that $\mathcal{V}$ and its first derivatives remain bounded as $\|m\|_e \to \infty$.

[32] All our results become meaningless for $\eta \to 0$ and for $\eta \to \infty$. Nearly spherical bodies could be treated as perturbations of spherical ones.

Let us notice that we shall use, for (real or complex) vectors $z = (z_1, \ldots, z_n)$, the norm

$$\|z\| \;=\; \max_i |z_i| \tag{1.7}$$

(the euclidean norm $\|\cdot\|_e$ will also be used). The norm of $\omega(A)$, namely

$$\|\omega(A)\| \;=\; \frac{1}{a_1}\max\left(|A_1|,\, |\eta|\,|A_2|\right) \tag{1.8}$$

will be the large parameter of perturbation theory. For any $A \in \mathcal{I}$, $\|\omega(A)\|$ is the larger of the two frequencies $|\omega_1(G)|$ and $|\omega_2(L)|$ of the Poinsot precession (see (2.9) of chapter 1).

Lemma 4.1 *Define*

$$\mu(\eta) \;=\; \max(1, |\eta|)\,. \tag{1.9}$$

Then, for any $A \in \mathcal{I}$ one has

$$\frac{1}{a_1}\|A\| \;\le\; \|\omega(A)\| \;\le\; \frac{\mu(\eta)}{a_1}\|A\|. \tag{1.10}$$

Moreover, let $\gamma \neq 0$ be the angle between two vectors $A, A' \in \mathcal{I}$, and γ_ω the angle between $\omega(A)$ and $\omega(A')$. Then

$$\min\left(|\eta|,\, |\eta|^{-1}\right) \;\le\; \left|\frac{\sin\gamma_\omega}{\sin\gamma}\right| \;\le\; \max\left(|\eta|,\, |\eta|^{-1}\right) \tag{1.11}$$

Proof.[33] The inequalities (1.10) are immediately checked (for the left one, the condition $|A_2| < A_1$ is essential). Since $\omega(A)\times\omega(A') = a_1^{-2}\eta A\times A'$, one has $\|\omega(A)\|_e\|\omega(A')\|_e|\sin\gamma_\omega| = a_1^{-2}|\eta|\|A\|_e\|A'\|_e|\sin\gamma|$. This equality is easily seen to imply (1.11). ∎

C. Geography of resonances. The resonant manifolds of the frequency ω in the action space $\mathcal{I}$ are all the straight lines

$$r_\nu \;=\; \{A \in \mathcal{I}\,:\, \nu_1 A_1 + \eta\nu_2 A_2 = 0\}\,, \qquad \nu = (\nu_1, \nu_2) \in \mathbb{Z}^2\,, \tag{1.12}$$

(the point $A = 0$ does not belong to $\mathcal{I}$). Obviously, the correspondence between resonant lines and integer vectors is not one–to–one. Thus, we shall (tacitly) parametrize the resonant lines by vectors of $\mathbb{Z}^2$ whose two components are relative prime numbers. Moreover, we consider the 'one–dimensional lattices' of $\mathbb{Z}^2$, defined as follows: for any $\nu \in \mathbb{Z}^2 \setminus \{0\}$

$$\mathcal{L}_\nu \;=\; \{\nu' \in \mathbb{Z}^2 \setminus \{0\}\,:\, \nu\times\nu' = 0\}$$

---

33) Here and in the following, the cross product $u \times v$ of two vectors u, v of $\mathbb{R}^2$, $\mathbb{Z}^2$ or $\mathbb{C}^2$ denotes obviously the number $u_1 v_2 - u_2 v_1$.

In order to dispose of a uniform notations, we shall formally treat the 'non-resonance' as the resonance with the null vector. Thus, we define

$$\mathcal{L}_0 = \{0\} \subset \mathbb{Z}^2 \tag{1.13}$$

Let us introduce a cut-off N, which at the moment takes the role of a parameter, and consider only the resonant lines of order $|\nu| = |\nu_1|+|\nu_2| \le N$. Around each of them, we set apart a certain 'resonant zone' $\mathcal{R}(\nu, N) \subset \mathcal{I}$ (which will be fixed in a moment), and define correspondingly the zone 'nonresonant within N' (which we denote as the zone resonant with $\nu = 0$):

$$\mathcal{R}(0,N) = \mathcal{I} \setminus \bigcup_{\substack{\nu\in\mathbb{Z}^2 \\ 0<|\nu|\le N}} \mathcal{R}(\nu,N). \tag{1.14}$$

The core of Nekhoroshev's theorem is the construction of a nonresonant normal form in each connected component of $\mathcal{R}(0,N)$, and of a resonant normal form, adapted to the resonant lattice $\mathcal{L}_\nu$, in the resonant zone $\mathcal{R}(\nu,N)$, for each $\nu \in \mathbb{Z}^2$, $1 \le |\nu| \le N$.

The choice of the resonant zones is of crucial importance. The two basic criterions are that the resonant zones must not overlap (otherwise one cannot construct a normal form adapted to a one-dimensional sublattice of $\mathbb{Z}^2$) and, at the same time, that they should not to be too small (the smallest they are, the smallest are the 'small denominators' in the nonresonant zone).

In the case at hand, we find convenient to take the resonant regions the largest as possible, compatibly with the non-overlapping requirement. This choice is possibly not yet optimal, but will allows us to to obtain accurate results.

For a fixed N, we take the resonant zones $\mathcal{R}(\nu,N)$ as follows. Consider an open angle $\mathcal{A}(\nu,N) \subset \mathbb{R}^2$, centered on the straight line orthogonal to ν and with the vertex in the origin of $\mathbb{R}^2$. Then, we define $\mathcal{R}(\nu,N)$ as the preimage of $\mathcal{A}(\nu,N)$ under the diffeomorphism ω (see figure 4.1). Such angular shape is rather natural for the estimate of the small denominators, but is not really important (moreover, it is in some sense fictitious –see subsection D).

The size of $\mathcal{A}(\nu,N)$ is easily determined by the non–overlapping requirement. One easily verifies that the (acute) angle γ_ω between any two (non parallel) vectors ν, ν' of $\mathbb{Z}^2$ is bounded by $|\sin\gamma_\omega| \ge (|\nu|\,|\nu'|)^{-1}$. Thus, taking $\mathcal{A}(\nu,N)$ to be the acute angle whose semi–amplitude satisfies

$$\sin\frac{1}{2}\mathcal{A}(\nu,N) = \frac{1}{2\,N\,|\nu|}, \tag{1.15}$$

one has that the resonant zones

$$\mathcal{R}(\nu,N) = \omega^{-1}\big[\mathcal{A}(\nu,N)\big] \tag{1.16}$$

do not overlap: for all the vectors $\nu, \nu' \in \mathbb{Z}^2$ of 'order' $|\nu|, |\nu'| \le N$, and such that $\nu' \notin \mathcal{L}_\nu$, one has

$$\mathcal{R}(\nu,N) \bigcap \mathcal{R}(\nu',N) = \emptyset \tag{1.17}$$

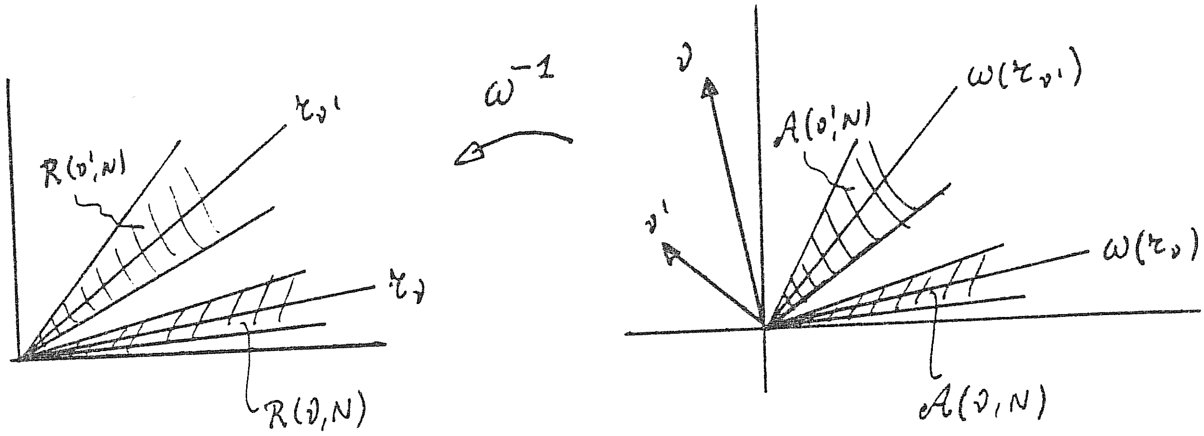

Figure 4.1

The angular size of the zones can be estimated by means of the inequalities (1.11) of lemma 4.1:

$$\frac{1}{2N|\nu|}\max\left(|\eta|,\,|\eta|^{-1}\right)\ \geq\ \sin\frac{1}{2}\mathcal{R}(\nu,N)\ \geq\ \frac{1}{2N|\nu|}\min\left(|\eta|,\,|\eta|^{-1}\right). \tag{1.18}$$

D. A remark on the cut–off. Our aim is that of giving results for large values of $\|\omega(A)\|$. Thus, we introduce a positive parameter Ω and, for any (sufficiently large) value of Ω, say $\Omega \geq \Omega_0$ for a certain threshold Ω_0, we consider the subset of the action space $\mathcal{I}$ in which $\|\omega(A)\| \geq \Omega$.

Correspondingly, in the perturbation theory, the cut–off N will be taken to be a function of Ω. Precisely, we shall take $N(\Omega) = \mathcal{O}(\sqrt{\Omega/\Omega_*})$. Consequently, the angular amplitude of the resonant zones will depend on Ω: $\mathcal{R}(\nu, N(\Omega)) \sim (|\nu|\sqrt{\Omega})^{-1}$. Nevertheless, for the moment it is convenient to treat N and Ω as independent parameters.

E. Domains. Let $\rho = (\rho_1, \ldots, \rho_n)$ be a vector with positive entries. For any point $z \in \mathbb{R}^n$ (or $\mathbb{C}^n$) and any domain $B \subset \mathbb{R}^n$ (or $\mathbb{C}^n$) we define

$$\begin{aligned}
\Delta_{\mathbb{R}}(z,\rho) &= \{y \in \mathbb{R}^n \,:\, |y_i - z_i| \leq \rho_i,\ i = 1,\ldots,n\}\\
\Delta(z,\rho) &= \{y \in \mathbb{C}^n \,:\, |y_i - z_i| \leq \rho_i,\ i = 1,\ldots,n\}\\
B + \rho &= \bigcup_{b\in B} \Delta_{\mathbb{R}}(b,\rho)\\
B_\rho &= \bigcup_{b\in B} \Delta(b,\rho)
\end{aligned} \tag{1.19}$$

Furthermore, we define the set $B - \rho = \{z \in B \,:\, \Delta_{\mathbb{R}}(z,\rho) \subset B\}$. Let us also introduce the complex neighbourhood $\mathcal{S}^1_r$, $r \geq 0$, of the circle S^1 as

$$\mathcal{S}^1_r\ =\ \{y(\mathrm{mod}2\pi) \in \mathbb{C}\,:\, |\mathrm{Im}\, y_i| \leq r\,,\ i = 1,\ldots,n\}. \tag{1.20}$$

Correspondingly, we put $\mathcal{S}_r^2 = \mathcal{S}_r^1 \times \mathcal{S}_r^1$, $\mathcal{S}_r^3 = \mathcal{S}_r^2 \times \mathcal{S}_r^1$.

Let us now consider the domain $P \times \mathbb{T}^3 \times Q$ of the two Poinsot charts (1.2). We introduce an 'extension vector'

$$\rho = (\rho_A, \rho_p, \rho_\alpha, \rho_q) \tag{1.21}$$

whose entries are positive, have the physical dimensions of the corresponding coordinates, and play the role of parameters. For practical convenience, we assume $\rho_\alpha \leq 1$. In order to apply the notations above, we identify ρ with the vector $(\rho_{A_1}, \rho_{A_2}, \rho_J, \rho_{\alpha_1}, \ldots, \rho_{q_3})$, with $\rho_{A_1} = \rho_{A_2} = \rho_J = \rho_A$, $\rho_{\alpha_1} = \rho_{\alpha_2} = \rho_j = \rho_\alpha$ and $\rho_{p_i} = \rho_p$, $\rho_{q_i} = \rho_q$ for all $i = 1, 2, 3$.

Preliminarily, we restrict P so to avoid the singularities on its boundary. Precisely, as in section 3.2, we consider the sets

$$\begin{aligned} \mathcal{I} - 2\rho &= \{A \in \mathcal{I} \,:\, A_1 > 2\rho_A,\, |A_2| < A_1 - 2\rho_A\} \\ P - 2\rho &= \{(A, J) \in P \,:\, A_1 > 2\rho_A,\, |A_2| < A_1 - 2\rho_A,\, |J| < A_1 - 2\rho_A\}. \end{aligned} \tag{1.22}$$

In the following, we shall consider complex sets of the form $D_\rho = (P - 2\rho)_\rho \times \mathcal{S}_\rho^3 \times Q_\rho$, where $\mathcal{S}_\rho^3 = \mathcal{S}_{\rho_\alpha}^3$.

Let us now consider the resonant zones $\mathcal{R}(\nu, N)$. First, we consider their intersections with the 'high frequency' part of the phase space (and with $\mathcal{I} - 2\rho$). With reference to the extension vector ρ (the dependence on which we do not indicate for simplicity), we define for any $\nu \in \mathbb{Z}^2$ with $0 \leq |\nu| \leq N$:

$$\begin{aligned} \mathcal{I}(\nu, N, \Omega) &= \mathcal{R}(\nu, N) \bigcap \left\{ A \in \mathbb{R}^2 \,:\, \|\omega(A)\| \geq \Omega + 2\rho_A \frac{\mu(\eta)}{a_1} \right\} \bigcap (\mathcal{I} - 2\rho) \\ P(\nu, N, \Omega) &= \{(A, J) \in \mathbb{R}^3 \,:\, A \in \mathcal{I}(\nu, N, \Omega)\} \bigcap (P - 2\rho)\,. \end{aligned} \tag{1.23}$$

We shall use real and complex neighbourhoods of these sets. Precisely, for any extension vector σ such that [34] $\sigma \leq \rho$, we shall consider the sets $\mathcal{I}(\nu, N, \Omega) + \sigma$, $\mathcal{I}_\sigma(\nu, N, \Omega)$, $P(\nu, N, \Omega) + \sigma$, $P_\sigma(\nu, N, \Omega)$ defined as in (1.19).

Finally, we observe that the preimages under the two Poinsot charts (1.3) of the set $P(\nu, N, \Omega) \times \mathbb{T}^3 \times Q$ define an angular fibering $\Sigma(\nu, N, \Omega) \times Q$ which is a subbundle of $\Sigma_* \times Q$, and has action space $\mathcal{I}(\nu, N, \Omega)$. Similarly, from the sets $\big(P(\nu, N, \Omega) + \sigma\big) \times \mathbb{T}^3 \times (Q + \sigma)$ one constructs an angular fibering $\big(\Sigma(\nu, N, \Omega) + \sigma\big) \times (Q + \sigma)$. These angular fiberings will be the natural domains for the perturbation theory.

F. Norms. As a last prerequisite, we introduce the norms to be used later on. The supremum norm of a function f in a real or complex domain B will be denoted

$$|f|_B^{(\infty)} = \sup_{b \in B} |f(b)|\,. \tag{1.24}$$

[34] Inequalities of the form $\sigma < \rho$ between vectors are intended to work separately on each entry.

Consider now a domain of the form $B_\rho \times S^2_\rho$, where B is some subset of $P \times S^1 \times Q$. Let $\langle f\rangle_\nu : B_\rho \to \mathbb{C}$, $\nu \in \mathbb{Z}^2$, be the Fourier components (relative to the angles $(\alpha_1, \alpha_2) \in S^2_\rho$) of a function $f : D_\rho \to \mathbb{C}$, so that

$$f(b,\alpha) = \sum_{\nu\in\mathbb{Z}^2} \langle f\rangle_\nu(b)\, e^{i\langle\nu,\alpha\rangle} . \tag{1.25}$$

Then, for any $\sigma \le \rho$ we define the norm

$$|f|_\sigma = \sum_{\nu\in\mathbb{Z}^2} |\langle f\rangle_\nu|^{(\infty)}_{B_\sigma}\; e^{|\nu|\sigma_\alpha} . \tag{1.26}$$

If the function f does not depend on the angles α, one has $f = \langle f\rangle_0$ and (1.26) reduces to the supremum norm in the domain B_σ.

The use of this norm in perturbation theory presents some advantages over the use of the supremum norm. Clearly, it allows a very simple and efficient estimate of the solution of the Hamilton–Jacobi equation and, moreover, of the ultraviolet part of the functions. In our opinion, the latter is the the more important one: in the present study of the rigid body, the use of this norm has been crucial to find estimates for the variation of the actions over times $\exp(\Omega^{1/2})$.

We shall also use a 'local' version of the norm (1.26): for any point $b \in B$ we define

$$|f|_{b,\sigma} = \sum_{\nu\in\mathbb{Z}^2} |\langle f\rangle_\nu|^{(\infty)}_{\Delta(b,\sigma)}\; e^{|\nu|\sigma_\alpha} . \tag{1.27}$$

Actually, such a norm is introduced here in view of an application of secondary importance (it will be used only for the study of the case of scattering, in section 5.5). However, its use does not introduce here any complications at all; on the contrary, its use in perturbation theory is, in our opinion, quite natural. Notice that (1.27) is meaningful for $\sigma = 0$, too:

$$|f|_{b,0} = \sum_{\nu\in\mathbb{Z}^2} |\langle f\rangle_\nu(b)| . \tag{1.28}$$

Moreover, we shall use a 'vector field norm' of functions: for any vector $\sigma \le \rho$ and any function $f : B_\rho \times S^2_\rho \to \mathbb{C}$ we define

$$\|f\|_\sigma = \max_i \left(\frac{1}{\rho_{\bar\imath}}\left|\frac{\partial f}{\partial i}\right|_\sigma\right) \tag{1.29}$$

where the maximum is taken over all the coordinates, $\bar\imath$ is the variable canonically conjugate to i, and $|\cdot|_\sigma$ is the norm (1.26). Clearly, such a norm is nothing else than a norm of the hamiltonian vector field of the function f.

Finally, notice that the norm (1.26) is not equivalent to the supremum norm. However, one can show that, in the analytic case, for any $\delta_\alpha < \rho_\alpha$, and writing $\delta = (0,0,\delta_\alpha,0)$, it results

$$|\langle f\rangle_\nu|^{(\infty)}_{B_\rho\times S^2_{\rho-\delta}} \le |f|_{\rho-\delta} \le \left(\frac{2}{\delta_\alpha}\right)^2 |\langle f\rangle_\nu|^{(\infty)}_{B_\rho\times S^2_\rho} . \tag{1.30}$$

4.2 The normal forms

In this section we state a proposition on the normal forms, which is the central result of the chapter. First, we collect the hypotheses, which essentially reduce to the analyticity of the potential energy $\mathcal{V}$.

We consider the system of Hamiltonian $h = k + u + \mathcal{V}$ as in (1.1), on the phase space $\Sigma_* \times Q$. We assume that each one of the two representatives $v^{(\lambda)} : P \times \mathbb{T}^3 \times Q \to \mathbb{R}$, $\lambda = 1, 2$, of $\mathcal{V}$ in the Poinsot charts (1.3) can be extended analytically to the complex set $(P - 2\rho)_\rho \times S^3_\rho \times Q_\rho$, for some extension vector ρ, the extension being bounded in the vector field norm (1.29). Let us denote (with some abuse of notations)

$$\|\mathcal{V}\|_\rho = \max\left(\|v^{(1)}\|_\rho,\, \|v^{(2)}\|_\rho\right),$$

where the vector field norm (1.29) is evaluated in the set $(P - 2\rho)_\rho \times S^3_\rho \times Q_\rho$. We also assume that $\|u\|_\rho = |u|^{(\infty)}_{Q_\rho}$ is finite (this is obviously a condition on the set Q). Finally, we assume $\rho_\varphi \leq 1$ and $\eta \neq 0$.

Let

$$\begin{aligned} \Omega_* &= \frac{2^{11}}{\rho_\alpha^2}\left(\|u\|_\rho + \|\mathcal{V}\|_\rho\right) \\ \Omega_0 &= \max\left(\Omega_*,\, \frac{2^7\,\mu(\eta)\,\rho_A}{a_1\,\rho_\alpha^3}\right). \end{aligned} \tag{2.1}$$

For any $\Omega \geq \Omega_0$, define the cut-off $N(\Omega)$ by

$$N(\Omega) = \frac{4}{\rho_\alpha}\sqrt{\frac{\Omega}{\Omega_*}} \tag{2.2}$$

and consider the partition of $\mathcal{I}$ into the resonant and non-resonant zones $\mathcal{R}(\nu, N(\Omega))$. We freely use the notations for Fourier series introduced in section 2.3B.

Proposition 4.2 *Within the above hypotheses and notations, for any $\Omega \geq \Omega_0$ and any $\nu \in \mathbb{Z}^2$, $|\nu| \leq N(\Omega)$, there exist a canonical transformation*

$$\Phi_{\nu,\Omega} : \Sigma(\nu, N(\Omega), \Omega) \times Q \to \left(\Sigma(\nu, N(\Omega), \Omega) + \rho\right) \times (Q + \rho) \tag{2.3}$$

which conjugates $h = k + u + \mathcal{V}$ to

$$h' = k + u + \Pi_{\mathcal{L}_\nu}\mathcal{V} + \sqrt{\frac{\Omega_*}{\Omega}}\,\mathcal{G} + \sqrt{\frac{\Omega_*}{\Omega}}\,e^{-\left[\sqrt{\Omega/\Omega_*}\right]}\,\mathcal{F} \tag{2.4}$$

where $[\cdot]$ denotes the integer part and the function $\mathcal{G}$ satisfies $\mathcal{G} = \Pi_{\Lambda_\nu}\mathcal{G}$. The local representatives g, f and v of $\mathcal{G}$, $\mathcal{F}$ and $\mathcal{V}$ in each of the two Poinsot charts satisfy, at any point

$b = (A, J, j, p, q) \in P(\nu, N(\Omega), \Omega) \times S^1 \times Q$:

$$
\begin{aligned}
|g|_{b,0} &\le \frac{2}{3}\,|v|_{b,\rho} \\
|f|_{b,0} &\le \frac{1}{3}\,|v|_{b,\rho} \\
\left|\frac{\partial g}{\partial \alpha_n}\right|_{b,0} &\le \frac{2}{3}\left|\frac{\partial v}{\partial \alpha_n}\right|_{b,\rho} \qquad (n = 1,2) \\
\left|\frac{\partial f}{\partial \alpha_n}\right|_{b,0} &\le \frac{1}{3}\left|\frac{\partial v}{\partial \alpha_n}\right|_{b,\rho} \qquad (n = 1,2)
\end{aligned}
\tag{2.5}
$$

Moreover, $\Phi_{\nu,\Omega}$ is a real analytic diffeomorphism of $\Sigma(\nu, N(\Omega), \Omega)$ onto its image. In each chart, its local representative $\Phi : (b, \alpha) \mapsto (b', \alpha')$ satisfies, at any point $(b, \alpha) \in P(\nu, N(\Omega), \Omega) \times \mathbb{T}^3 \times Q$, for any $i = A, J, j, p, q$ and for any analytic function w:

$$
|b'_i - b_i| \le 2^4 \frac{\delta_\nu}{\Omega}\left(\left|\frac{\partial u}{\partial \bar{\imath}}\right|_{b,\rho} + \left|\frac{\partial v}{\partial \bar{\imath}}\right|_{b,\rho}\right) \le 2^{-5}\sqrt{\frac{\Omega_*}{\Omega}}\,\rho_i\,\frac{\delta_\nu}{N(\Omega)} \tag{2.6a}
$$

$$
|w(b', \alpha') - w(b, \alpha)| \le 2^7 \frac{\delta_\nu}{\Omega}\left(\|u\|_{b,\rho} + \|v\|_{b,\rho}\right)|w|_{b,\rho} \le 2^{-2}\sqrt{\frac{\Omega_*}{\Omega}}\,\frac{\delta_\nu}{N(\Omega)}\,|w|_{b,\rho} \tag{2.6b}
$$

where $\bar{\imath}$ denotes the variable canonically conjugate to i and

$$
\begin{aligned}
\delta_\nu &= |\nu| \qquad \text{if } \nu \ne 0 \\
\delta_0 &= N(\Omega)\,.
\end{aligned}
\tag{2.7}
$$

This proposition is proven in the following two sections.

4.3 Proof: the iterative lemma

A. Generalities. Making reference to the theory of section 2.3, we prove proposition 4.2 by constructing the representatives of the canonical transformation $\Phi_{\nu,\Omega}$ and of the Hamiltonian h' within each chart domain. Thus, we consider one of the two charts, and the Hamilton function $h = k + u + v$ in the phase space $P_\rho(\nu, N(\Omega), \Omega) \times S^3_\rho \times Q_\rho$. Let us remark that in the sequel the 'extension' vector ρ is thought of to be fixed, as in proposition 4.2.

For any $\sigma \le \rho$ we denote

$$
B_\sigma = P_\sigma(\nu, N(\Omega), \Omega) \times S^1_\sigma \times Q_\sigma \tag{3.1}
$$

where S^1_σ is the domain of the angle j, and then write

$$
D_\sigma = B_\sigma \times S^2_\sigma\,. \tag{3.2}
$$

Notice that $B_0 = P(\nu, N(\Omega), \Omega) \times S^1 \times Q$ and $D_0 = B_0 \times \mathbb{T}^2$.

The proof is divided in two parts. In this section we perform just a single 'perturbative step'; in doing this, we shall consider the cut–off N as a free parameter. This procedure will then be iterated in the next section, where the choice (2.2) of the cut–off will also be made.

We shall work with the hamiltonian vector field of the system, rather than with its Hamilton function alone. As explained in [36], such a procedure presents some advantages.

Let D_σ be a domain as in (3.2). The hamiltonian vector field of a ('Hamilton') function $f : D_\sigma \to \mathbb{C}$ is the vector field $F = \sum_i F_i \frac{\partial}{\partial i}$, where

$$\begin{aligned} F_{A_n} &= -\frac{\partial f}{\partial \alpha_n}, & F_J &= -\frac{\partial f}{\partial j}, & F_{p_m} &= -\frac{\partial f}{\partial q_m} \\ F_{\alpha_n} &= \frac{\partial f}{\partial A_n}, & F_j &= \frac{\partial f}{\partial J}, & F_{q_m} &= \frac{\partial f}{\partial p_m} \end{aligned} \tag{3.3}$$

(with $n = 1, 2$ and $m = 1, 2, 3$). As a rule, we denote functions by small letters, and their hamiltonian vector fields by the corresponding capital letters. In particular, the hamiltonian vector field of the 'unperturbed' Hamiltonian $k(A)$ is

$$K = \omega_1(A) \frac{\partial}{\partial \alpha_1} + |\eta| \, \omega_2(A) \frac{\partial}{\partial \alpha_2} \,. \tag{3.4}$$

Furthermore, we shall denote by U and V the hamiltonian vector field of the functions, respectively, u and v, the latter being the local representative of $\mathcal{V}$ in the considered chart.

Let us observe that, if g is a function and G is its hamiltonian vector field, then one has (for any $\nu \in \mathbb{Z}^2$)

$$g = \Pi_{\mathcal{L}_\nu} g \;\Leftrightarrow\; G = \Pi_{\mathcal{L}_\nu} G \,.$$

In all of this section, $|\cdot|_\sigma$ denotes the norm (1.26) relative to the set D_σ. We use the following norms for vector fields. For any $\sigma \le \rho$ and any point $b \in B_\sigma$:

$$\begin{aligned} \|F\|_\sigma &= \max_i \left(\rho_i^{-1} \, |F_i|_\sigma\right) \\ \|F\|_{b,\sigma} &= \max_i \left(\rho_i^{-1} \, |F_i|_{b,\sigma}\right) \end{aligned} \tag{3.5}$$

where the maximum is taken over all the components of the vector field (this is nothing else than the 'vector field norm' of the corresponding Hamilton function, as defined in (1.29)).

We shall denote the cut–off decomposition of a function as $f = \tilde{f} + f^{>N}$. Here $\tilde{f}$, previously denoted $f \le N$ is defined by

$$\tilde{f} = \sum_{\nu' \in \mathbb{Z}^2, |\nu'| \le N} \langle f \rangle_{\nu'} \, E_{\nu'}$$

where $E_{\nu'}$ is defined as in section 2.3B. The cut–off decomposition of a vector field $F = \tilde{F} + F^{>N}$ is defined through its components.

B. The iterative lemma. We consider here, as parameters: a positive number $N \geq 4$, an 'extension vector' σ such that $\sigma \leq \rho$, a positive number x such that $\sigma \geq x\rho$, a vector $\nu \in \mathbb{Z}^2$ such that $|\nu| \leq N$, and a positive number Ω with the dimension of a frequency.

Lemma 4.3 *Assume that the hamiltonian vector fields G and F are analytic and bounded in the norm (3.5) in the domain D_σ, that $G = \Pi_{\mathcal{L}_\nu} G$ and that*

$$\Omega \geq \max\left(\frac{2^6\,\delta_\nu}{x} \|F\|_\sigma \,,\ \frac{8\,\mu(\eta)\,\rho_A}{a_1\,\rho_\alpha} N\,\delta_\nu \right) . \tag{3.6}$$

Then, there exists a canonical transformation $\Phi : D_{\sigma - x\rho} \to D_{\sigma-(x/2)\rho}$ which conjugates $H = K + G + F$ to $H' = K + G' + F'$, where

$$G' \;=\; G \,+\, \Pi_{\mathcal{L}_\nu} \tilde{F}\,; \tag{3.7}$$

F' and its Hamilton function f' satisfy, at any point $b \in B_0$ (for any $i \neq \alpha_1, \alpha_2$ and any $n = 1, 2$):

$$\begin{aligned} |F_i'|_{b,\sigma-x\rho} \leq \frac{2^4\,\delta_\nu}{x\,\Omega} &\Big[|\tilde{F}_i|_{b,\sigma} (\|G\|_{b,\sigma} + 2\|F\|_{b,\sigma}) + \|\tilde{F}\|_{b,\sigma} (|G_i|_{b,\sigma} + 2\,|F_i|_{b,\sigma}) \Big] \\ &+ e^{-Nx\rho_\alpha}\, |F_i^{>N}|_{b,\sigma} \end{aligned} \tag{3.8a}$$

$$|F'_{\alpha_n}|_{b,\sigma-x\rho} \leq \frac{2^5\,\delta_\nu}{x\,\Omega} \|\tilde{F}\|_{b,\sigma} (\|G\|_{b,\sigma} + 2\,\|F\|_{b,\sigma})\, \rho_\alpha \;+\; e^{-Nx\rho_\alpha} |F_{\alpha_n}^{>N}|_{b,\sigma}\, \rho_\alpha \tag{3.8b}$$

$$|f'|_{b,\sigma-x\rho} \leq \frac{2^5\,\delta_\nu}{x\,\Omega} (\|G\|_{b,\sigma} + 2\,\|F\|_{b,\sigma}) |\tilde{f}|_{b,\sigma} \;+\; e^{-Nx\rho_\alpha} |f^{>N}|_{b,\sigma} \,. \tag{3.8c}$$

Furthermore, $\Phi : (b, \alpha) \mapsto (b', \alpha')$ is an analytic diffeomorphism of $D_{\sigma-x\rho}$ onto its image, it is real on real sets, and satisfies for any $b \in B_0$ and any analytic function $w : D_\sigma \to \mathbb{C}$

$$|i' - i|_{b,\sigma-x\rho} \;\leq\; \frac{8\,\delta_\nu}{\Omega} |F_i|_{b,\sigma} \qquad \text{for all } i = A, J, j, p, q \tag{3.9a}$$

$$|w \circ \Phi - w|_{b,\sigma-x\rho} \;\leq\; \frac{32\,\delta_\nu}{x\,\Omega} \|F\|_{b,\sigma}\, |w|_{b,\sigma} \,. \tag{3.9a}$$

The remaining of this section is devoted to the proof of this lemma.

C. Proof: the Hamilton–Jacobi equation. We construct Φ as the time–one map Φ_1^X of a suitable Hamiltonian vector field X. This is the Lie method, which was presented in section 2.3C. Its use for transforming vector fields is briefly accounted for in the appendix at the chapter, where one also finds all estimates which will be used in the sequel. Let us here only recall that for any function w, and any vector field W, one has

$$\begin{aligned} (\Phi_1^X)^* w \;&=\; w + R_1^X(w) \;=\; w + L_X w + R_2^X(w)\,, \\ (\Phi_1^X)^* W \;&=\; W + R_1^X(W) \;=\; W + L_X W + R_2^X(W)\,, \end{aligned}$$

where L_X is the Lie derivative associated to the vector field X, while $\Phi^* w = w \circ \Phi$ and $\Phi^* W = \big((D\Phi^{-1})W\big) \circ \Phi$ are the pull–back of w and W respectively, and $R_k^X = \sum_{s=k}^{\infty}(s!)^{-1} L_x^s$ denotes the k-th remainder of the Lie series.

It is easy to verify that, if the vector field X satisfies the 'Hamilton–Jacobi' equation

$$L_K X = \tilde{F} - \Pi_{\mathcal{L}_\nu} \tilde{F}, \tag{3.10}$$

then $H' = \big(\Phi_1^X\big)^* H$ has the form $K + G' + F'$, with $G' = G + \Pi_{\mathcal{L}_\nu} F$ and

$$F' = R_1^X(G+F) + R_2^X(K) + F^{>N}. \tag{3.11}$$

Since $(L_K X)_i = \sum_j \big(K_j \frac{\partial X_i}{\partial j} - X_j \frac{\partial K_i}{\partial j}\big)$, recalling the expression (3.4) of K, one sees that equation (3.10) reads in components

$$\sum_{m=1}^{2} \omega_m \frac{\partial X_i}{\partial \alpha_m} = \tilde{F}_i - \Pi_{\mathcal{L}_\nu} \tilde{F}_i \qquad (i = A, J, j, p, q)$$

$$\sum_{m=1}^{2} \omega_m \frac{\partial X_{\alpha_n}}{\partial \alpha_m} = \tilde{F}_{\alpha_n} - \Pi_{\mathcal{L}_\nu} \tilde{F}_{\alpha_n} + \sum_{m=1}^{2} X_{A_m} \frac{\partial \omega_n}{\partial A_m} \qquad (m = 1, 2)$$

Thus, it is readily solved by

$$X_i = \sum_{\nu' \notin \mathcal{L}_\nu, |\nu'| \le N} \frac{\langle F_i \rangle_{\nu'}}{\sqrt{-1}\,\langle \omega, \nu' \rangle} E_{\nu'} \qquad \text{for all } i \ne \alpha_1, \alpha_2 \tag{3.12a}$$

$$X_{\alpha_n} = \sum_{\nu' \notin \mathcal{L}_\nu, |\nu'| \le N} \left(\frac{\langle F_{\alpha_n} \rangle_{\nu'}}{\sqrt{-1}\,\langle \omega, \nu' \rangle} - \frac{\langle F_{A_n} \rangle_{\nu'}}{\hat{a}_1\,\langle \omega, \nu' \rangle^2} \right) E_{\nu'} \qquad (n = 1, 2), \tag{3.12b}$$

where $\hat{a}_1 = a_1$ and $\hat{a}_2 = a_1/\eta$. The remaining of this subsection is devoted to the (straightforward) proof of the fact that the vector field X defined by (3.12) satisfies, at any point $b \in B_0$:

$$|X_i|_{b,\sigma} \le \frac{4\sqrt{2}\,\delta_\nu}{\Omega} |\tilde{F}_i|_{b,\sigma} \qquad (i \ne \alpha_1, \alpha_2) \tag{3.13a}$$

$$|X_{\alpha_n}|_{b,\sigma} \le \frac{5\sqrt{2}\,\delta_\nu}{\Omega} \|\tilde{F}\|_{b,\sigma}\,\rho_\alpha \qquad (n = 1, 2). \tag{3.13b}$$

Lemma 4.4 *Consider any point* $A^* \in \mathcal{I}(\nu, N, \Omega)$ *and any vector* $\nu' \in \mathbb{Z}^2 \setminus \mathcal{L}_\nu$ *such that* $0 < |\nu'| \le N$. *Then, for any point* $A \in \Delta(A^*, \sigma_A)$ *one has*

$$|\langle \omega(A), \nu' \rangle| \ge \frac{\Omega}{4\sqrt{2}\,\delta_\nu}. \tag{3.14}$$

Proof of lemma 4.4 First, we show that one has

$$|\langle\omega(A^*),\,\nu'\rangle| \,\geq\, \frac{\Omega}{2\sqrt{2}\,\delta_\nu} \tag{3.15}$$

Let $\nu^\perp \in \mathbb{R}^2$ be any vector orthogonal to ν', and γ the angle between $\omega(A^*)$ and $\nu^\perp$. Then, since $\|\omega(A^*)\| \geq \Omega_*$ (using also $\|\cdot\|_e \geq \|\cdot\|$ and $|\cdot| \leq \sqrt{2}\|\cdot\|_e$) one has

$$|\langle\omega(A^*),\,\nu'\rangle| \,\geq\, \frac{1}{\sqrt{2}}\,\Omega\,|\nu'|\,|\sin\gamma|$$

Since $A^* \notin \mathcal{R}(\nu', N)$, by the very definition of the resonant zones one has $|\sin\gamma| \geq (2N|\nu'|)^{-1}$, so that

$$|\langle\omega(A^*),\,\nu'\rangle| \,\geq\, \frac{\Omega}{2\sqrt{2}\,N}\,.$$

This proves (3.14) if $\nu = 0$. If $\nu \neq 0$, one can do better: since $A^* \in \mathcal{R}(\nu, N)$, one has $\omega(A^*) \in \mathcal{A}(\nu', N)$ (refer to figure 4.1). Thus, the angle γ is at least one half of the angle γ^* between ν' and ν, so that $|\sin\gamma| \geq \frac{1}{2}|\sin\gamma^*|$. As already noticed, $|\sin\gamma^*| \geq (|\nu'|\,|\nu|)^{-1}$, and (3.14) follows.

Consider now a point $A \in \Delta(A^*, \sigma_A)$. One can compute, using also (1.9), (3.14) and $|\nu'| \leq N$:

$$|\langle\omega(A),\,\nu'\rangle| \,\geq\, |\langle\omega(A^*),\,\nu'\rangle| \,-\, |\langle\omega(A^*-A),\,\nu'\rangle| \,\geq\, \frac{\Omega}{2\sqrt{2}\,\delta_\nu} \,-\, \frac{\mu(\eta)\,\sigma_A}{a_1}\,N\,.$$

Inequality (3.14) follows from here, using (3.6), $\sigma_A \leq \rho_A$ and $\rho_\alpha \leq 1$. ∎

We now come back to the inequalities (3.13). The first of them is an immediate consequence of (3.12a), (3.14) and of $\|\omega(A)\| \geq \Omega$. From (3.12b) one gets

$$|X_{\alpha_n}|_{b,\sigma} \,\leq\, \frac{4\sqrt{2}\,\delta_\nu}{\Omega}\,|\tilde F_{\alpha_n}|_{b,\sigma} \,+\, \frac{\mu(\eta)}{a_1}\left(\frac{4\sqrt{2}\delta_\nu}{\Omega}\right)^2 |\tilde F_{A_n}|_{b,\sigma} \qquad (n=1,2)$$

and then, recalling the definition (3.5) of the norm:

$$|X_{\alpha_n}|_{b,\sigma-x\rho} \,\leq\, \frac{4\sqrt{2}\,\delta_\nu}{\Omega}\,\|\tilde F\|_{b,\sigma}\,\rho_\alpha\left[1 \,+\, \frac{\mu(\eta)}{a_1}\,\frac{4\sqrt{2}\,\delta_\nu}{\Omega}\,\frac{\rho_A}{\rho_\alpha}\right] \qquad (n=1,2)$$

By (3.6) the expression inside the brackets is $\leq 1 + N^{-1} \leq 5/4$.

D. The canonical transformation. We now make use of proposition 4.6 of the appendix. First of all, from this proposition one knows that $\Phi_1^X : D_{\sigma-x\rho} \to D_{\sigma-(x/2)\rho}$ is well defined, provided $|X_i|_\sigma \leq x\rho_i/8$ for all the components $i = A, J, \alpha, j, p, q$. Since $|\tilde F_i|_\sigma \leq |F_i|_\sigma \leq \|F\|_\sigma\rho_i$, by (3.13) and (3.6) this condition is fulfilled. The estimates (3.9a,b) of lemma 4.3 follow, respectively, from (A.14) and (A.16a) of proposition 4.6, using again (3.13).

The mapping $\Phi = \Phi_1^X$ is canonical, since X is a hamiltonian vector field: indeed, a Hamilton function χ for X is defined by the Fourier series

$$\chi = \sum_{\nu' \notin \mathcal{L}_\nu,\, |\nu'| \le N} \frac{\langle f \rangle_{\nu'}}{\sqrt{-1}\, \langle \omega, \nu \rangle'} E_{\nu'} \tag{3.16}$$

which is obviously uniformly convergent: as in the case of (3.13a) one has

$$|\chi|_{b,\sigma} \le \frac{4\sqrt{2}\,\delta_\nu}{\Omega} |\tilde{f}|_{b,\sigma} \tag{3.17}$$

E. The remainder. We prove now the estimates (3.8a,b) on the remainder F' as in (3.11). First, we consider the components F_i' with $i \ne \alpha_1, \alpha_2$. From (3.13) and from (A.15a) of proposition 4.6, one gets

$$\left| \left[R_1^X (G+F) \right]_i \right|_{b,\sigma - x\rho} \le \frac{2^4\,\delta_\nu}{x\,\Omega} \left[|\tilde{F}_i|_{b,\sigma} \|G+F\|_{b,\sigma} + \|\tilde{F}\|_{b,\sigma} |G_i + F_i|_{b,\sigma} \right] .$$

From the Hamilton–Jacobi equation (3.10) one has $L_X K = -(\tilde{F} - \Pi_{\mathcal{L}_\nu} \tilde{F})$. Thus, from the inequality (A.15b) of proposition 4.6, taking into account the obvious $\|\tilde{F}_i - \Pi_{\mathcal{L}_\nu} \tilde{F}_i\| \le 2\|\tilde{F}_i\|$, one gets

$$\left| \left[R_2^X (K) \right]_i \right|_{b,\sigma - x\rho} \le \frac{2^5\,\delta_\nu}{x\,\Omega} \|\tilde{F}\|_{b,\sigma} |\tilde{F}_i|_{b,\sigma} .$$

Observe now that, for any function $w : D_\sigma \to \mathbb{C}$ one has

$$|w^{>N}|_{b,\sigma - x\rho} = \sum_{|\nu'| > N} |\langle w \rangle_{\nu'}|_{\Delta(b,\sigma)}^{(\infty)} e^{|\nu'|(\sigma_\alpha - x\rho_\alpha)} \le e^{-Nx\rho_\alpha} |w^{>N}|_{b,\sigma} \tag{3.18}$$

This proves (3.8a). Inequality (3.8b) is proven similarly.

F. The Hamilton function. Let $h = k + g + f$ be the Hamilton function of H. Since Φ_1^X is canonical, the function $h' = h \circ \Phi_1^X$ is the Hamilton function of H', and it has the form $k + g' + f'$, with $g' = g + \Pi_{\mathcal{L}_\nu} f$ and $f' = R_1^X(g+f) + R_2^X(k) + f^{>N}$. Let us prove (3.8c). First, using (A.16b) of proposition 4.6, (A.5a) of proposition 4.5, and observing that, if X and W are hamiltonian vector fields with Hamilton functions, respectively, χ and w, one has $L_X w = -L_W \chi$, one computes

$$\left| R_1^X (g+f) \right|_{b,\sigma - x\rho} \le 2 \left| L_X (g+f) \right|_{b,\sigma - (x/2)\rho} \le \frac{4}{x} \left\| G + F \right\|_{b,\sigma} |\chi|_{b,\sigma} .$$

Next, using (A.16c) and observing that χ satisfies $L_K \chi = \tilde{f} - \Pi_{\mathcal{L}_\nu} \tilde{f}$, one gets

$$\left| R_2^X (k) \right|_{b,\sigma - x\rho} \le \left| L_X (\tilde{f} - \Pi_{\mathcal{L}_\nu} \tilde{f}) \right|_{b,\sigma - (x/2)\rho} \le \frac{4}{x} \left\| \tilde{F} \right\|_{b,\sigma} |\chi|_{b,\sigma} .$$

Inequality (3.8c) follows using (3.17) and (3.18).

The proof of lemma 4.3 is complete.

4.4 Proof: the iteration

A. The strategy. It can be useful to explain briefly the general lines of the proof, and to motivate in particular the choice (2.2) of the cut–off.

Basically, we apply a (large) number r of times the normalization procedure of lemma 4.3, each time with a restriction of the domain $x \sim 1/r$. In such a way, starting with the vector field

$$H^{(0)} = K + G^{(0)} + F^{(0)}, \qquad G^{(0)} = \Pi_{\mathcal{L}_\nu} G^{(0)},$$

we construct iteratively the vector fields $H^{(1)}, H^{(2)}, \ldots, H^{(r)}$, each of which has the form

$$H^{(s)} = K + G^{(s)} + F^{(s)}, \qquad G^{(s)} = \Pi_{\mathcal{L}_\nu} G^{(s)}. \tag{4.1}$$

The remainders $F^{(s)}$ satisfy, by (3.8):

$$\|F^{(s+1)}\| \leq \left[\text{const}\, \frac{r\,\delta_\nu}{\Omega} \left(\|G^{(s)}\| + 2\,\|F^{(s)}\|\right) + e^{-\frac{N\rho_\alpha}{2r}}\right] \|F^{(s)}\|. \tag{4.2}$$

The heart of the analytic part of Nekhoroshev's theorem consists in reducing the remainder at each step of a constant factor, for instance a factor e:

$$\|F^{(s+1)}\| \sim e^{-1}\,\|F^{(s)}\|,$$

and iterate such a procedure a number of times which increases with some inverse power of the small parameter of the problem. In our case, if $r \sim \Omega^\beta$ ($\beta > 0$), then

$$\|F^{(r)}\| \sim e^{-r}\,\|F^{(0)}\| \sim e^{-\Omega^\beta}\,\|F^{(0)}\| \tag{4.3}$$

In order to obtain such a geometric decreasing, both of the two terms inside the square brackets in (4.2) have to be $\sim e^{-1}$. To have $\exp(-N\rho_\alpha/2r) \sim e^{-1}$, we simply take $r \sim N$. After this choice, the condition on the first addend in (4.2) becomes (in the case $\nu_* = 0$, which is the worse)

$$\text{const}\, \frac{N^2}{\Omega} \left(\|G^{(s)}\| + 2\,\|F^{(s)}\|\right) \sim \frac{1}{e}. \tag{4.4}$$

Since $G^{(s)} = G^{(0)} + \sum_{j=0}^{s-1} \tilde{F}^{(j)}$, an elementary induction shows that this condition is satisfied for all $s = 1, \ldots, r$, provided it is satisfied for $s = 0$ (possibly, with some additional factor at the r.h.s.): thus

$$N^2 \leq \frac{\Omega}{\|G^{(0)}\| + \|F^{(0)}\|}\, \text{const}$$

Taking, as is obviously convenient, the largest N compatible with this condition, one arrives at a dependence of the cut–off on Ω of the form (2.2). Since $r \sim N$, this gives $\beta = 1/2$ in (4.3)

There is now, however, a minor point. Starting the iterative construction with $H^{(0)} = H$ (that is, $G^{(0)} = U$, $F^{(0)} = V$), we eventually arrive at $H^{(r)}$ as in (4.1), with $G^{(r)} = G^{(0)} + \sum_{j=1}^{r-1} \tilde{F}^{(j)}$. Unfortunately, we can provide only the poor estimate

$$G^{(r)} - U - \Pi_{\mathcal{L}_\nu} V = \mathcal{O}\Big(\frac{1}{e}\Big)$$

In order to have (as is useful) a more accurate information about $G^{(r)}$, we perform a preliminary perturbative step with a reduction of the domain of order 1, precisely $x = 1/2$ in lemma 4.3. In such a way, we eventually obtain

$$G^{(r)} = U + \Pi_{\mathcal{L}_\nu} V + \mathcal{O}\Big(\sqrt{\frac{\Omega_*}{\Omega}}\Big) .$$

B. Proof of proposition 4.2. We take now the cut–off N as in (2.2), and assume $\Omega \geq \Omega_0$ with Ω_0 as in (2.1). Correspondingly, we define (with $[\cdot]$ denoting the integer part)

$$r = \Big[\frac{N\rho_\alpha}{4}\Big] = \Big[\sqrt{\frac{\Omega}{\Omega_*}}\Big] .$$

Denote $G^{(-1)} = U$, $F^{(-1)} = V$, $g^{(-1)} = u$, $f^{(-1)} = v$, and apply a first time lemma 4.3, with $\sigma = \rho$ and $x = 1/2$. This is possible since, with the above choice of N, (3.6) is implied by $\Omega \geq \Omega_*$. In such a way, we construct a first canonical transformation $\Phi^{(-1)}$, which conjugates $H^{(-1)}$ to

$$H^{(0)} = K + G^{(0)} + F^{(0)}$$

with $G^{(0)} = U + \Pi_{\mathcal{L}_\nu} V$. Estimates on $F^{(0)}$ and its Hamilton function $f^{(0)}$ are worked out from (3.8). Since

$$\exp\left(-\frac{N\rho_\alpha}{2}\right) = \exp\left(-2\sqrt{\frac{\Omega}{\Omega_*}}\right) \leq \frac{1}{2e}\sqrt{\frac{\Omega_*}{\Omega}} ,$$

observing that $|\tilde{V}_i|_{b,\rho} + |V_i^{>N}|_{b,\rho} = |V_i|_{b,\rho}$, one gets from (3.8), for any $b \in B_0$:

$$\begin{aligned}
|F_i^{(0)}|_{b,\rho/2} &\leq \frac{1}{3}\sqrt{\frac{\Omega_*}{\Omega}}\left(|U_i|_{b,\rho} + |V_i|_{b,\rho}\right) && (i \neq \alpha_1, \alpha_2) \\
|F_{\alpha_n}^{(0)}|_{b,\rho/2} &\leq \frac{1}{3}\sqrt{\frac{\Omega_*}{\Omega}}\left(\|U\|_{b,\rho} + \|V\|_{b,\rho}\right)\rho_\alpha && (n = 1,2) \\
|f^{(0)}|_{b,\rho/2} &\leq \frac{1}{3}\sqrt{\frac{\Omega_*}{\Omega}}\left(|u|_{b,\rho} + |v|_{b,\rho}\right)
\end{aligned}$$

We now apply r times lemma 4.3, each time with $x = \frac{1}{2r}$. Let $\rho_s = (\frac{1}{2} - \frac{s}{2r})$. As we now show, after s steps $(s = 1, \ldots, r)$ one arrives at the system described in D_{ρ_s} by

$$H^{(s)} = K + G^{(s)} + F^{(s)} \tag{4.5}$$

with

$$G^{(s)} = G^{(0)} + \sum_{j=1}^{s-1} \Pi_{\mathcal{L}_\nu} \tilde{F}^{(j)} \tag{4.6}$$

and ($f^{(r)}$ being the Hamilton function of $F^{(r)}$):

$$\begin{aligned}
|F_i^{(s)}|_{b,\rho_s} &\le \frac{1}{3}\frac{\Omega_*}{\Omega}\left(|U_i|_{b,\rho} + |V_i|_{b,\rho}\right) e^{-s} && (i \neq \alpha_1, \alpha_2) \\
|F_{\alpha_n}^{(s)}|_{b,\rho_s} &\le \frac{1}{3}\frac{\Omega_*}{\Omega}\left(\|U\|_{b,\rho} + \|V\|_{b,\rho}\right) \rho_\alpha\, e^{-s} && (n = 1, 2) \\
|f^{(s)}|_{b,\rho_s} &\le \frac{1}{3}\frac{\Omega_*}{\Omega}\left(|u|_{b,\rho} + |v|_{b,\rho}\right) e^{-s}
\end{aligned} \tag{4.7}$$

Indeed, assume that the first l steps have been performed. One can then apply once more lemma 4.3, since (4.7) imply that (3.6) is fulfilled. The vector field $H^{(l+1)}$ has again the form (4.5), (4.6), and it is easy to verify that $F^{(l+1)}$ and $f^{(l+1)}$ satisfy (4.7), with $s = l+1$. To this purpose, one simply applies (3.8), taking care of the following two facts. First:

$$\exp\left(-\frac{N\rho_\alpha}{2r}\right) \le e^{-2}\,.$$

Second, $G^{(l)} + 2F^{(l)} = G^{(0)} + \sum_{j=0}^{l-1} \Pi_{\mathcal{L}_\nu} F^{(j)} + 2F^{(l)} = U + \Pi_{\mathcal{L}_\nu} V + F^{(l)} + \sum_{j=0}^{l} \Pi_{\mathcal{L}_\nu} F^{(j)}$. By the induction hypothesis, this gives

$$|G_i^{(l)}| + 2\,|F^{(l)}| \le \left(1 + \frac{1}{3}\frac{2e-1}{e-1}\right)\left(|U_i| + |V_i|\right) < 2\left(|U_i| + |V_i|\right).$$

In such a way, one constructs the final hamiltonian vector field $H^{(r)} = K + G^{(r)} + F^{(r)}$. Its Hamilton function $h^{(r)} = k + g^{(r)} + f^{(r)}$ coincides with the local representative of the function h' given by (2.4), if the functions $g^{(r)}$ and $f^{(r)}$ are related to the local representatives of the two functions $\mathcal{G}$ and $\mathcal{F}$ entering (2.4) via $g^{(r)} = u + \Pi_{\mathcal{L}_\nu} v + g\sqrt{\Omega_*/\Omega}$, $f^{(r)} = e^{-r} f \sqrt{\Omega_*/\Omega}$. Let G and F be the hamiltonian vector fields of f and g; they are related to $G^{(r)}$ and $F^{(r)}$ by completely analogous relations. Then, from (4.7) and (4.6) one works out the estimates

$$\begin{aligned}
\max\left(3|F_i|_{b,0}\,,\ \frac{3}{2}|G_i|_{b,0}\right) &\le |U_i|_{b,\rho} + |V_i|_{b,\rho} && (i \neq \alpha_1, \alpha_2) \\
\max\left(3|F_{\alpha_n}|_{b,0}\,,\ \frac{3}{2}|G_{\alpha_n}|_{b,0}\right) &\le \left(\|U\|_{b,\rho} + \|V\|_{b,\rho}\right)\rho_\alpha && (n = 1, 2) \\
\max\left(3|f|_{b,0}\,,\ \frac{3}{2}|g|_{b,0}\right) &\le |u|_{b,\rho} + |v|_{b,\rho}\,,
\end{aligned}$$

This proves the estimates (2.5).

Finally, the estimates (2.6) on the overall canonical transformation $\Phi = \Phi^{(-1)} \circ \Phi^{(0)} \circ \cdots \circ \Phi^{(r-1)}$ are easy consequences of (3.9a,b) and (4.7). Indeed, one has (for $i \neq \alpha_1, \alpha_2$)

$$|b'_i - b_i| \leq \frac{4\sqrt{2}\,\delta_\nu}{\Omega} \sum_{j=-1}^{r-1} |F_i^{(j)}|_{b,\rho_j} \leq \frac{4\sqrt{2}\,\delta_\nu}{\Omega} \left[|V_i|_{b,\rho} + \frac{1}{3}\left(|U_i|_{b,\rho} + |V_i|_{b,\rho}\right) \sum_{j=o}^{\infty} e^{-j} \right],$$

which implies (2.6a). Concerning (2.6b), one computes

$$\begin{aligned} |w(b',\alpha') - w(b,\alpha)| &\leq \left(\frac{2^6\,\delta_\nu}{\Omega} \|V\|_{b,\rho} + \frac{2^6\, r\, \delta_\nu}{\Omega} \sum_{j=0}^{r-1} \|F^{(j)}\|_{b,\rho_j} \right) |w|_{b,\rho} \\ &\leq \frac{2^7\,\delta_\nu}{\Omega} \left(\|U\|_{b,\rho} + \|V\|_{b,\rho} \right) |w|_{b,\rho} \,. \end{aligned}$$

The proof of proposition 4.2 is complete.

4.A Appendix: the Lie method

This appendix is intended to give a short treatment of the Lie method for functions and vector fields; in particular, we state and prove all the results which have been used in sections 4.3 and 4.4. The estimates presented below are the analog, in the norm (1.26), of the estimates given in [36] for the supremum norm. Let us notice that we shall use freely some of the notations already introduced in the chapter.

A. Estimates on Lie derivatives. If $X = \sum_i X_i \frac{\partial}{\partial z_i}$ is a vector fields on $\mathbb{R}^p$ (or $\mathbb{C}^p$), then the associated Lie derivative L_X acts on functions according to

$$L_X w = \sum_i X_i \frac{\partial w}{\partial z_i} \,. \tag{A.1}$$

If $W = \sum_i W_i \frac{\partial}{\partial z_i}$ is a vector field, then $L_X W$ is the vector field having components

$$\left(L_X W\right)_i = L_X W_i - L_W X_i \,. \tag{A.2}$$

Our first aim is that of giving estimates on the Lie derivative L_X, and its iterations L_X^s, $s \geq 1$, which are necessary for the subsequent treatment of the Lie method.

Let $B \subset \mathbb{R}^m$ ($m \geq 1$) be any domain, and consider the set $D = B \times \mathbb{T}^n$ ($n \geq 1$). The points of D are denoted (b, α), where $b = (b_1, \ldots, b_m) \in B$ and $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{T}^n$.

Consider an 'extension vector' $\rho = (\rho_b, \rho_\alpha)$, with $\rho_b = (\rho_{b_1}, \ldots, \rho_{b_m})$ and $\rho_\alpha = (\rho_\alpha, \ldots, \rho_\alpha)$ (all the n entries are equal). Then, for any extension vector $\sigma \leq \rho$ we define the sets

$$\begin{aligned} B_\sigma &= \bigcup_{b\in B} \Delta(b, \sigma_b) \\ D_\sigma &= B_\sigma \times \mathcal{S}^n_\sigma \end{aligned} \tag{A.3}$$

where $\Delta(b, \sigma_b)$ is the complex polydisk $\{b' \in \mathbb{C}^m : |b'_j - b_j| \leq \sigma_{b_j}, j = 1, ..., m\}$ and $\mathcal{S}^n_\sigma$ is the complex neighbourhood of $\mathbb{T}^n$ obtained as product of n copies of $\mathcal{S}^1_{\sigma_\alpha}$, see (1.20).

We use the same local norms as in sections 4.2 and 4.3: if w and W are, respectively, a function and a vector field defined in D_ρ, then we define, for any point $b \in B$ and any vector σ, $0 \leq \sigma \leq \rho$:

$$\begin{aligned} |w|_{b,\sigma} &= \sum_{\nu\in\mathbb{Z}^n} |\langle w\rangle_\nu|^{(\infty)}_{B_\sigma} e^{|\nu|\sigma_\alpha} \\ \|W\|_{b,\sigma} &= \max_{i=b_1,\ldots,b_m,\alpha_1,\ldots,\alpha_n} \frac{|W_i|_{b,\sigma}}{\rho_i} \end{aligned} \tag{A.4}$$

where $|\nu| = |\nu_1| + \ldots + |\nu_n|$. We shall also use the global norm $|w|_\sigma = \sup_{b\in B} |w|_{b,\sigma}$.

Proposition 4.5 (On Lie derivatives). *Let the vector fields X, W and the function w be defined and analytic in the domain D_ρ. Consider an 'extension vector' $\sigma \leq \rho$ and a positive number x such that $\sigma \geq x\rho$. Then, for any integer $s \geq 1$ and any point $b \in B$ one has*

$$|L_X w|_{b,\sigma-x\rho} \leq \frac{1}{x}\|X\|_{b,\sigma-x\rho}|w|_{b,\sigma} \tag{A.5a}$$

$$\frac{1}{s!}|L^s_X w|_{b,\sigma-x\rho} \leq \left(\frac{e}{x}\|X\|_{b,\sigma}\right)^s |w|_{b,\sigma} \tag{A.5b}$$

$$\left|(L_X W)_i\right|_{b,\sigma-x\rho} \leq \frac{1}{x}\left[\|X\|_{b,\sigma-x\rho}|W_i|_{b,\sigma} + \|W\|_{b,\sigma-x\rho}|X_i|_{b,\sigma}\right] \tag{A.5c}$$

$$\frac{1}{s!}\left|(L^s_X W)_i\right|_{b,\sigma-x\rho} \leq \frac{1}{x}\left[\frac{4}{x}\|X\|_{b,\sigma}\right]^{s-1}\left[|X_i|_{b,\sigma}\|W\|_{b,\sigma} + |W_i|_{b,\sigma}\|W\|_{b,\sigma}\right] \tag{A.5d}$$

where the index i runs over all the components $b_1, \ldots, b_m, \alpha_1, \ldots, \alpha_m$.

Proof of proposition 4.5 The key estimate is (A.5a). To prove it, we need a preliminary estimate of the Lie derivative in the supremum norm. This is well known. In full generality, consider the polydisk $\Delta(z, \delta) = \{z' \in \mathbb{C}^p : |z'_i - z_i| \leq \delta_i\}$ of radii $\delta_i, \ldots, \delta_p$, centered on a point $z = (z_1, \ldots, z_p) \in \mathbb{C}^p$. Let Y and g be, respectively, a vector field and a function which are defined and analytic in $\Delta(z, \delta)$. Then one has

$$\left|(L_Y g)(z)\right| \leq \left(\max_{j=1,\ldots,p} \frac{|Y_j(z)|}{\delta_j}\right) |g|^{(\infty)}_{\Delta(z,\delta)} . \tag{A.6}$$

The proof of (A.6) is straightforward. Since $z + \tau Y(z) \in \Delta(z, \delta)$ for all the complex τ such that $|\tau| < T = \min_j \delta_j |Y_j(z)|^{-1}$, the function $\tau \mapsto g\big(z + \tau Y(z)\big)$ is defined and analytic in the disk $|\tau| \leq T$. Thus, the use of Cauchy inequality gives

$$\left|\frac{d}{d\tau} g\Big(z+\tau Y(z)\Big)\right|_{\tau=0} \leq \frac{1}{T} \sup_{|\tau|<T} \Big|g(z+\tau Y(z)\Big| \leq \left(\max_{j=1,\ldots,p} \frac{|Y_j(z)|}{\delta_j}\right) |g|^{(\infty)}_{\Delta(z,\delta)} .$$

This proves (A.6), since $\frac{d}{d\tau} f\big(z+\tau Y(z)\big)_{\tau=0} = L_Y g(z)$.

We now come to prove (A.5a). Since

$$\big|L_X w\big|_{b,\sigma-x\rho} = \sum_{\nu\in\mathbb{Z}^n} |\langle L_X W\rangle_\nu|^{(\infty)}_{\Delta(b,\sigma-x\rho)} \, e^{|\nu|(\sigma_\alpha - x\rho_\alpha)} \tag{A.7}$$

we have to work out an estimate, in the supremum norm, for the ν–th Fourier component of $L_X w$. It is easy to verify that one has, for any $\nu \in \mathbb{Z}^n$:

$$\langle L_X w\rangle_\nu = \sum_{\substack{\nu',\nu''\in\mathbb{Z}^n \\ \nu'+\nu''=\nu}} \left[\sum_{i=1}^{m} \langle X_{b_i}\rangle_{\nu'} \frac{\partial\langle w\rangle_{\nu''}}{\partial b_i} + \sqrt{-1}\sum_{i=1}^{n} \langle X_{\alpha_i}\rangle_{\nu'}\, \nu''_i \, \langle w\rangle_{\nu''}\right] \tag{A.8}$$

Observe now that, if we introduce (for any $\nu \in \mathbb{Z}^n$) the vector field

$$\langle X\rangle_\nu = \sum_{i=1}^{m} \langle X_{b_i}\rangle_\nu \frac{\partial}{\partial b_i} + \sum_{i=1}^{n} \langle X_{\alpha_i}\rangle_\nu \frac{\partial}{\partial \alpha_i}$$

($\langle X\rangle_\nu$ is nothing else than the ν–th Fourier component of X), then we can rewrite equation (A.8) in the form

$$\langle L_X w\rangle_\nu = \sum_{\nu'+\nu''=\nu} \left[L_{\langle X\rangle_{\nu'}} \left(\langle w\rangle_{\nu''}\, e^{\sqrt{-1}\langle \nu'',\alpha\rangle}\right)\right]_{\alpha=0} \tag{A.9}$$

We can now use (A.6) to estimate each addend of this sum. Denoting by E_ν the base functions of the Fourier series (as in section 2.3B) we find

$$\big|L_{\langle X\rangle_{\nu'}} (\langle w\rangle_{\nu''} E_{\nu''})(b,\alpha)\big| \leq \left(\max_{i=b_1,\ldots,\alpha_n} \frac{|\langle X_i(b)\rangle_{\nu'}|}{x\rho_i}\right) |\langle w\rangle_{\nu''}|^{(\infty)}_{\Delta(b,x\rho_b)} \, |E_{\nu''}|^{(\infty)}_{\Delta(\alpha,x\rho_\alpha)}$$

so that, by (A.9):

$$\big|\langle L_X w\rangle_\nu(b)\big| \leq \sum_{\nu'+\nu''=\nu} \left(\max_i \frac{|\langle X_i(b)\rangle_{\nu'}|}{x\rho_i}\right) |\langle w\rangle_{\nu''}|^{(\infty)}_{\Delta(b,x\rho_b)} \, e^{|\nu''|x\rho_\alpha} .$$

Taking the supremum over $b \in B$ this gives

$$|\langle L_X w\rangle_\nu|^{(\infty)}_{b,\sigma-x\rho} \leq \sum_{\nu'+\nu''=\nu} \left(\max_i \frac{|\langle X_i\rangle_{\nu'}|^{(\infty)}_{b,\sigma-x\rho}}{x\rho_i}\right) |\langle w\rangle_{\nu''}|^{(\infty)}_{\Delta(b,x\rho_b)} \, e^{|\nu''|x\rho_\alpha} .$$

It is now sufficient to insert the last inequality in (A.7), and use there $|\nu| \leq |\nu'| + |\nu''|$ to obtain (A.5a).

We now consider the other estimates. Since $(L_X W)_i = L_X W_i - L_W X_i$, (A.5c) follows from (A.5a). Inequality (A.5d) can be proven by induction. Let us assume it be satisfied for some $s \geq 1$. With the decomposition $\left(L_X^{s+1} W\right)_i = L_X \left(L_X^s W\right)_i - L_{(L_X^s W)} X_i$, using (A.5c) to estimate the two addends (the first one with x replaced by $x/(s+1)$), one gets

$$\frac{\left|\left(L_X^{s+1} W\right)_i\right|_{b,\sigma - x\rho}}{(s+1)!} \leq \frac{1}{x} \|X\|_{b,\sigma} \left|\frac{\left(L_X^s W\right)_i}{s!}\right|_{b,\sigma-(1-\frac{1}{s+1})x\rho} + \frac{1}{(s+1)x} \frac{\|L_X^s W\|_{b,\sigma - x\rho}}{s!} |X_i|_{b,\sigma} .$$

Using (A.5d), the r.h.s. of this inequality is easily seen to be majorized by

$$\frac{1}{4x}\left(\frac{4}{x}\|X\|_{b,\sigma}\right)^s \left[\left(1+\frac{1}{s}\right)^s \left(|X_i|_{b,\sigma}\|W\|_{b,\sigma} + |W_i|_{b,\sigma}\|X\|_{b,\sigma}\right) + \frac{2}{s+1}|X_i|_{b,\sigma}\|W\|_{b,\sigma}\right]$$

Thus, since $\left(1+\frac{1}{s}\right)^s \leq e$ and $s \geq 1$, one gets

$$\frac{\left|\left(L_X^{s+1} W\right)_i\right|_{b,\sigma - x\rho}}{(s+1)!} \leq \frac{1}{4x}\left(\frac{4}{x}\|X\|_{b,\sigma}\right)^s \left[(e+1)\,|X_i|_{b,\sigma}\|W\|_{b,\sigma} + e|W_i|_{b,\sigma}\|X\|_{b,\sigma}\right] ,$$

from which (A.5d) follows with s replaced by $s+1$. Finally, (A.5b) can be proven by a very similar, and somehow simpler, induction. ∎

B. The Lie transform. We give now a sketch of the Lie method for vector fields, which is formally very similar to the Lie method for functions presented in section 2.3C. At the basis of the method is the identity

$$\frac{d}{d\tau}\left(\Phi_\tau^X\right)^* W = \left(\Phi_\tau^X\right)^* L_X W \qquad (A.10)$$

which relates the time–derivative along the flow to the Lie derivative. Here, $\Phi^* W = \left((D\Phi^{-1})W\right) \circ \Phi$ is the pull-back of the vector field W under the mapping Φ. An iterated use of (A.10) leads in a trivial way to the "Lie series" representation of the "Lie transform" $W \mapsto \left(\Phi_1^X\right)^* W$:

$$\left(\Phi_1^X\right)^* W = \sum_{s=0}^{\infty} \frac{1}{s!} L_X^s W . \qquad (A.11)$$

In the analytic case, the convergence of such a series expansion is easily established (see proposition 4.6 below). For any $k = 1, 2, \ldots$, the k–th remainder $R_k^X(W)$ of the Lie series (A.11) is defined by

$$R_k^X W = \sum_{s=k}^{\infty} \frac{1}{s!} L_X^s W . \qquad (A.12)$$

Completely analogous expansions are obtained for functions: the k-th remainder of the Lie series is given by

$$R_k^X w = \sum_{s=k}^{\infty} \frac{1}{s!} L_X^s w . \qquad (A.13)$$

Proposition 4.6 (On Lie transform) *Let the vector fields X, W and the function w be defined and analytic in the domain $D_\rho = B_\rho \times S_\rho^n$ (defined as in (A.3)). Consider a vector $\sigma \le \rho$, a positive number x such that $\sigma \ge x\rho$, and assume $|X_i|_\sigma \le \frac{x\rho_i}{8}$ for all $i = b_1, ..., b_m, \alpha_1, ..., \alpha_n$. Then*

i) the mapping Φ_1^X is an analytic diffeomorphism of $D_{\sigma - x\rho}$ onto $\Phi_1^X(D_{\sigma-x\rho}) \subset D_{\sigma-(x/2)\rho}$, and it is real on real domains. For each point $z \in D_{\sigma-x\rho}$ one has

$$\left|[\Phi_1^X(z)]_i - z_i\right| \le |X_i|_{\Delta(z,x\rho)}^{(\infty)} \qquad \text{for all } i = b_1, \ldots, b_m, \alpha_1, \ldots, \alpha_n . \qquad (A.14)$$

ii) the vector field $(\Phi_1^X)^* W$ is analytic in $D_{\sigma-x\rho}$, and one has, for any $b \in B$:

$$\left|\left(R_1^X W\right)_i\right|_{b,\sigma-x\rho} \le \frac{2}{x} \left[|X_i|_{b,\sigma} \|W\|_{b,\sigma} + |W_i|_{b,\sigma} \|X\|_{b,\sigma}\right] \qquad (A.15a)$$

$$\left|\left(R_2^X W\right)_i\right|_{b,\sigma-x\rho} \le \frac{1}{x} \left[|X_i|_{b,\sigma} \|L_X W\|_{b,\sigma} + |(L_X W)_i|_{b,\sigma} \|X\|_{b,\sigma}\right] \qquad (A.15b)$$

iii) the function $w \circ \Phi_1^X w$ is analytic in $D_{\sigma-x\rho}$, and one has, for any $b \in B$:

$$\left|R_1^X w\right|_{b,\sigma-x\rho} \le \frac{2.5}{x} \|X\|_{b,\sigma} |w|_{b,\sigma} \qquad (A.16a)$$

$$\left|R_1^X w\right|_{b,\sigma-x\rho} \le 2\, |L_X w|_{b,\sigma-(x/2)\rho} \qquad (A.16b)$$

$$\left|R_2^X w\right|_{b,\sigma-x\rho} \le |L_X^2 w|_{b,\sigma-(x/2)\rho} \qquad (A.16c)$$

Remark 4.1: The factor 1/8 in the assumption for $|X_i||_\sigma$ is introduced in order to obtain the estimates on the remainders of the Lie series.

Proof of proposition 4.6 The statements in *i)* are elementary consequences of general properties of ordinary differential equations, and of the fact that the norm (A.4) dominates the supremum norm; in particular, the *a priori* estimate (A.14) is obvious, and implies the inclusion property for $\Phi_1^X(D_{\sigma-x\rho})$. The inequality (A.15a) is very easy to prove: one estimates term by term the series expansion (A.12) by means of (A.5d), and reduces to a geometric series. Concerning (A.15b), one first writes

$$R_2^X W = \sum_{s=1}^{\infty} \frac{1}{s} \frac{1}{(s-1)!} L_X^{s-1}(L_X W)$$

and then proceeds in the same way. The estimates relative to the function w are proven similarly, referring to (A.13), using (A.5b), and taking care of estimating separately the first terms of the series by means of (A.5a). ∎

Chapter 5

Fast motions of the symmetric rigid body

5.1 Introduction

A. Plane of the chapter. In this chapter we use the normal forms of proposition 4.2 to study the fast motion of a symmetric rigid body. Our main object is the motion about a fixed point. Sections 2 and 3 are devoted, respectively, to nonresonant and resonant motions, while in section 4 we extend the results to cover the excluded neighbourhood of the gyroscopic rotations.

In section 5 we give some indications about the fast motion of a rigid body with no a fixed point. The results for this case are much less complete than the previous ones: some important facts about the motion near the resonance are ultimately not understood.

As a rule, we shall not compute the constants entering the various statements. The computation could be done without difficulty. However, in our opinion, the (extremely pessimistic) estimates that one can obtain within such a general approach are essentially meaningless (on the contrary, it would be of great interest to study carefully some specific cases). Although not explicitly stated, such constants depend on the analyticity properties of the potential energy $\mathcal{V}$, and on the parameter η; in particular, most of them vanish or diverge (depending on the case) as $\eta \to 0$ or $\eta \to \infty$, with the consequence that the results become meaningless in these limits.

We shall refer tacitly to the hypotheses and the notations of chapter 4. Everywhere we denote by $A = (G, L)$, $J^{(\lambda)}$, $\alpha^{(\lambda)} = (g^{(\lambda)}, l)$, $j^{(\lambda)}$, p, q $(\lambda = 1, 2)$ the Poinsot variables of the system, and by $(A', J'^{(\lambda)}, \alpha'^{(\lambda)}, j'^{(\lambda)}, p', q') = \Phi_{\nu,\Omega}^{(\lambda)}(A, J^{(\lambda)}, \alpha^{(\lambda)}, j^{(\lambda)}, p, q)$ their image under the local representative of the canonical transformation $\Phi_{\nu,\Omega}$ constructed in proposition 4.2.

We shall be concerned with individual motions $t \mapsto G_t, L_t, J_t^{(\lambda)}, g_t^{(\lambda)}, l_t, j_t^{(\lambda)}, p_t, q_t$. The parameter Ω will be related to the initial conditions of the motion by

$$\Omega = \frac{1}{2} \|\omega(G_0, L_0)\| .$$

Correspondingly, we shall consider only 'fast motions', with $\|\omega(G_0, L_0)\| \geq 2\Omega_*$.

B. Gyroscopic phenomena. Preliminarily, it may be useful to recall some elementary facts about 'gyroscopic' phenomena. The fast motion of a rigid body about a fixed point presents some peculiar qualitative features which are well known (see for instance [56,73]); however, a complete, rigorous theory is still lacking.

A first basic property of the fast motion is the 'gyroscopic stiffness' of the angular momentum vector (see [73]):

Proposition 5.1 *In the motion of a rigid body about a fixed point one has*

$$\dot{\mu} = \frac{\mu \times (\Gamma \times \mu)}{G} \tag{1.1}$$

where $\mu = m/G$ is the unit vector in the direction of the angular momentum and Γ is the external torque (relative to the fixed point).

Proof. Equation (1.1) follows from the balance equation $\dot{m} = \Gamma$, using $\langle \mu, \dot{\mu} \rangle = 0$. ∎

According to equation (1.1), the direction of the angular momentum vector changes slowly in space, with speed $\mathcal{O}(G^{-1})$. Regarding μ as the intersection point of the vector m with the surface of the unit sphere fixed in space (centered in the body fixed point), one can say that such a point moves with velocity $\mathcal{O}(G^{-1})$.

This fact is important, since it indicates that, on time scales shorter than $O(G)$, the direction of the angular momentum vector can be considered as (approximately) fixed in space. Unfortunately, besides this fact, very few informations can be obtained by elementary methods. For instance, from conservation of energy ($k + \mathcal{V} = \text{const}$) one can easily derive the apriori estimate $|G_t - G_0| \leq c(\eta) G_0$ for all $t \in \mathbb{R}$, with a certain constant $c(\eta) < 1$. This assures that a fast motion will perpetually remain fast, but nothing more. No information can be obtained about the motion of the angular momentum relative to the body (thus, on the body motion in space for time scales shorter than $\mathcal{O}(\Omega)$), nor about the motion of the angular momentum vector in space (for time scales larger than $\mathcal{O}(G)$).

In this chapter, we shall give some answers to these questions.

A second feature of the fast motion is the 'tendency to parallelism' with the applied torque. Indeed, from the balance equation $\dot{m} = \Gamma$ one gets, for a motion $t \mapsto \mathcal{R}(t)$ of the body,

$$m(t) - m(0) = \int_0^t \Gamma(\mathcal{R}(\tau))\, d\tau\,.$$

Then, one expects m to move in the direction of an 'averaged' torque (see [56]). This is meaningful, since the body undergoes a fast (two–frequencies) motion. Indeed, one expects one can replace $\mathcal{R}(\tau)$ in the integral by the free motion $\mathcal{R}_0(\tau)$, with a small error. But again, precise statements are lacking.

By Hamiltonian perturbation theory, we shall obtain some results about this fact.[35)]

---

35) In this regard, the fundamental advantage of Hamiltonian perturbation theory (with respect to perturbation theory working directly on the solutions of the equation of motions) is in the fact that one replaces time averages along the motion with the corresponding phase averages (see chapter 4 of [8]).

Frequently, the term 'gyroscopic' is reserved to the case of a symmetric body with a fixed point, which undergoes a fast motion near a rotation about its symmetry inertia axis e_3, i.e. with angular velocity w nearly parallel to e_3. For instance, one can require that, at the initial time, $|w_1|$ and $|w_2|$ are of order $|w_3|^{-\beta}$ for some positive number β. In such a case, one claims that the gyroscopic axis e_3 manifests stiffness, and tendency to parallelism, but it is definitely not clear over which time scale one can assure that the angular velocity remains near the gyroscopic axis.

Just to give an idea of an existing approach to this problem, let us briefly mention here the so-called 'principle of the gyroscopic effect', which has been widely studied in Italy in the last forty years (see for instance [70, 71, 72, 12]). In its simplest form, the principle consists in describing the motion of the gyroscopic axis e_3 by means of the approximated equation

$$a_3\, w_3(0)\, \dot{e}_3 \;=\; \Gamma\,. \tag{1.2}$$

The true motion of the axis e_3 is proven to remain near the solution of (1.2), up to quantities $\mathcal{O}(|w_3|^{-1})$, if at the initial time $|w_1|$ and $|w_2|$ are $\mathcal{O}(|w_3|^{-1})$, and the applied torque Γ is orthogonal to e_3. [36)] However, no estimate of the time interval over which such an approximation holds is given in the papers on the argument. Certainly, such a time scale (at least as could be obtained by the methods of the quoted papers) is extremely short.

Unfortunately, we cannot obtain a detailed description of the motion near a gyroscopic rotation, as it will be better discussed in section 5.4.

5.2 Non-resonant motions

A. Statements. We consider here the non-resonant motions of a rigid body with a fixed point. The Hamilton function is $k + \mathcal{V}$ (see (1.1) of chapter 4). We denote by a bar the average over the two angles $g^{(\lambda)}$ and l, namely we define (with the notations of section 2.3B) $\overline{\mathcal{V}} = \Pi_{\{0\}}\mathcal{V}$. Moreover, if $t \mapsto z_t$ is a motion and $f(z)$ is a function, we write $|f|_0^t$ to mean $|f(z_t) - f(z_0)|$.

Proposition 5.2 *There exist positive constants $c_{11}, \dots, c_{15}$ such that, for any motion $t \mapsto G_t, L_t, J_t^{(\lambda)}, g_t^{(\lambda)}, l_t, j_t^{(\lambda)}$ with $\|\omega(G_0, L_0)\| = 2\Omega \geq 2\Omega_*$ and $(G_0, L_0) \in \mathcal{I}(0, N(\Omega), \Omega)$, one has:*

$$|G_t - G_0| \;\leq\; c_{11}\sqrt{\frac{\Omega_*}{\Omega}} \tag{2.1a}$$

$$|L_t - L_0| \;\leq\; c_{11}\sqrt{\frac{\Omega_*}{\Omega}} \tag{2.1b}$$

---

36) Notice that (1.2) is obtained from (1.1) by assuming $m = a_3 e_3$ and $\langle \Gamma, \mu\rangle = 0$.

$$\left|\dot{g}_t^{(\lambda)} - \frac{G_0}{a_1}\right| \leq c_{12}\sqrt{\frac{\Omega_*}{\Omega}} \qquad (\lambda = 1,2) \tag{2.1c}$$

$$\left|\dot{l}_t - \eta\frac{L_0}{a_1}\right| \leq c_{12}\sqrt{\frac{\Omega_*}{\Omega}} \tag{2.1c}$$

for all $t \in \mathbb{R}$ such that

$$|t| \leq c_{13} \exp\left[\sqrt{\frac{\Omega}{\Omega_*}}\right] ; \tag{2.2}$$

moreover, one has

$$\left|\overline{\mathcal{V}}\right|_0^t = \left|k + \mathcal{V} - \overline{\mathcal{V}}\right|_0^t \leq c_{14}\sqrt{\frac{\Omega_*}{\Omega}} \tag{2.3}$$

for all t such that

$$|t| \leq c_{15}\frac{\Omega_*}{\Omega} \exp\left[\sqrt{\frac{\Omega}{\Omega_*}}\right] . \tag{2.4}$$

Proof. The first two inequalities (2.1) are completely standard: from the normal form Hamiltonian (2.4) of proposition 4.2 one gets

$$\|A_t' - A_0'\| \leq \frac{|t|}{3}\sqrt{\frac{\Omega_*}{\Omega}}\, e^{-\left[\sqrt{\Omega/\Omega_*}\right]} \max_{\lambda,n=1,2}\left|\frac{\partial v^{(\lambda)}}{\partial \alpha_n^{(\lambda)}}\right|_\rho ,$$

so that $\|A_t' - A_0'\| \leq \mathrm{const}\sqrt{\Omega_*/\Omega}$ for times as in (2.2). One then uses the estimate (2.5a) of proposition 4.2 to come back to the original variables. Concerning the other two estimates (2.1) notice, for instance, that $\dot{g}^{(\lambda)} = \frac{\partial(k+v^{(\lambda)})}{\partial G}$, and recall that the local representatives $v^{(\lambda)}$ of $\mathcal{V}$ are homogeneous functions of degree zero of (G, L, J), so that $\frac{\partial v^{(\lambda)}}{\partial G} = \mathcal{O}(G^{-1}) = \mathcal{O}(\Omega^{-1})$.

In order to prove (2.3), let us write

$$\left|(G')^2\right|_0^t \leq \left|G_t' - G_0'\right| \left[|G_t' - G_0'| + 2|G_0' - G_0| + 2|G_0|\right] .$$

Thus, observing that $|G_t' - G_0'| \leq \mathrm{const}(\Omega_*/\Omega)^{3/2}$ on the time scale (2.4), using the estimate for the deformation of the canonical transformation, and the obvious inequality $|G(0)| \leq 2a_1\Omega$, one gets

$$|(G')^2|_0^t \leq \mathrm{const}\sqrt{\frac{\Omega_*}{\Omega}} .$$

Similarly, one finds $|(L')^2|_0^t \leq \mathrm{const}\sqrt{\Omega_*/\Omega}$, so that

$$|k(G_t', L_t') - k(G_0', L_0')| \leq \mathrm{const}\sqrt{\frac{\Omega_*}{\Omega}} .$$

The conservation of energy for the Hamilton function $h' = k + \overline{\mathcal{V}} + \mathcal{O}(\sqrt{\Omega_*/\Omega})$ then gives $|\overline{\mathcal{V}}(z'_t) - \overline{\mathcal{V}}(z'_0)| \leq \mathrm{const}\sqrt{\Omega_*/\Omega}$ (where $z' = (G', ..., j')$). Using (2.6b) of proposition 4.2 to come back to the original variables, and then using once more the conservation of energy, one arrives at (2.4). ∎

B. Description of the fast motion. We now use the above results to give an accurate description of the non–resonant motions in the time interval (2.4).

First, we observe that the body undergoes, approximately, a high frequency Poinsot precession about the angular momentum vector m. This follows from the following facts: for all t as in (2.2), the angle β between e_3 and m is almost constant:

$$|\cos\beta_t \; - \; \cos\beta_0| \;\leq\; \mathrm{const}\left(\frac{\Omega_*}{\Omega}\right)^{3/2}$$

(this follows from $\cos\beta \;=\; L/G$ and (2.1a,b)). Moreover, inequalities (2.1c,d) indicate that the two angular frequencies $\dot{g}$, $\dot{l}$ are almost constant, too.[37] Finally, recall from proposition 5.1 that the direction of the angular momentum vector moves in space with velocity $\mathcal{O}(\Omega^{-1})$, so that it can be considered as essentially constant during a period of g and l.

Let us now consider the motion in space of the angular momentum vector m, whose modulus is, by (2.1a), almost constant. Observe that J/G and j are coordinates on the unit sphere of $\mathbb{R}^3$ (they fix the direction of the vector m relatively to the inertial frame). Furthermore, observe that each local representative $\overline{v}$ of $\overline{\mathcal{V}}$ is an analytic function of J/G, L/G and j. Thus, using $G_t \;=\; G_0 + \mathcal{O}(\Omega^{-1})$ and $L_t \;=\; L_0 + \mathcal{O}(\Omega^{-1})$, one sees that the estimate (2.3) of proposition 5.2 implies

$$\overline{v}\left(\frac{L_0}{G_0}, \frac{J_t}{G_t}, j_t\right) \;=\; \overline{v}\left(\frac{L_0}{G_0}, \frac{J_0}{G_0}, j_0\right) \;+\; \mathcal{O}\left(\sqrt{\frac{\Omega_*}{\Omega}}\right) \;. \tag{2.5}$$

Consider now the intersection curves of the (conical) surfaces of equation $\overline{v}(L_0/G_0, J/G, j) =$ const with the surface of the unit sphere. Equation (2.5) means that the tip of the unit vector μ in the direction of m remains near to one of such curves, within a distance $\mathcal{O}(\sqrt{\Omega_*/\Omega})$ (figure 5.1). Since $\dot{\mu} = \mathcal{O}(\Omega^{-1})$, the motion of μ along such a curve, and transversal to it, is slow.

So, we arrive at the conclusion that, approximately, the motion consists of a (fast) Poinsot precession about the angular momentum vector, which in turn undergoes a (slow) precession in space along a conical surface, which is a level surface of an 'average potential energy'.[38]

---

[37] Notice, in particular, that one has at least $|L_0| \geq \mathcal{O}(\Omega^{1/2})$ in the non–resonant zone, since the angular width of the zone corresponding to the resonance $\nu = (0,1)$ is $\mathcal{O}(N^{-1})$.

[38] Such motion can be interpreted, under certain hypotheses, as taking place in the direction of an 'averaged torque'.

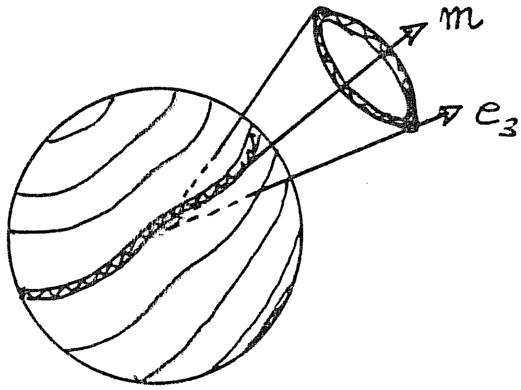

Figure 5.1

In principle, one could obtain precise informations about the precession of the angular momentum vector in space, by means of the normal form Hamiltonian constructed in proposition 4.2. On a time scale on which the exponentially small term can be neglected (with some small error), it provides an effective Hamiltonian for the motion of the variable J' and j', whose local representatives have the form

$$\bar{v}\left(\frac{L'}{G'}, \frac{J'}{G'}, j'\right) + \sqrt{\frac{\Omega_*}{\Omega}} g\left(\frac{L'}{G'}, \frac{J'}{G'}, j'\right) . \tag{2.6}$$

This is the Hamiltonian of a system with one degree of freedom (parametrized by the values of G' and L', which are constant in this approximation). Thus, in this approximation, the system is integrable. If one knows (up to some small error) the function $\mathcal{G}$ in the normal form Hamiltonian (2.4) of proposition 4.2, then one can determine the time evolution of the variables j', J', and consequently also of J'/G'. Informations on j and J/G are then obtained, using the fact that the old variables are near the new ones.

Remarks 5.1: i) Obviously, one can also obtain informations about the motion of the angular velocity vector w. Since the body is symmetric, m, w and e_3 always lie in a plane. Moreover, the angles between such vectors are approximately constant, as well as the modulus $\|w\|_e$ and the projection $w_3 = \langle w, e_3 \rangle$ of the angular velocity.

ii) For the treatment of the non–resonant motions the apparatus of the geometric theory of chapter 2 is not necessary: indeed the system remains certainly within the domain of a single Poinsot chart (the direction of m is confined to the neighbourhood of a curve on the unit sphere of $\mathbb{R}^3$). However, it will be clear in the next section that such a confinement cannot be assured in the case of resonant motions.

5.3 Resonant motions

A. Estimate on the variation of G and L. We consider now a motion starting within a resonant zone: $(G_0, L_0) \in \mathcal{I}(\nu, N(\Omega), \Omega)$ for some $|\nu| \leq N(\Omega)$, where $\|\omega(G_0, L_0)\| = 2\Omega$.

The normal form Hamiltonian h' given by (2.4) of proposition 4.2 indicates that the motion of the (new) actions G', L' is flattened on the 'fast drift' line, namely the straight line through A'_0 parallel to ν. The control of the motion of A' parallel to the fast drift direction is one of the basic ingredients of Nekhoroshev's theorem. It can be achieved by two methods.

On one hand, one can use Nekhoroshev's original method. In the present case, this method reduces essentially to the observation that, if the system escapes a resonant zone, then it enters the non–resonant one, where the non–resonant normal form can be used to stop it. Thus, the variation of the actions is essentially bounded by the width of the resonant zone; more precisely, the length of the segment, contained within the resonant zone, of the line of fast drift through the (image A'_0 of the) initial point A_0. In the present case, the resonant zone has angular width $(|\nu|N)^{-1}$, so that this criterion leads to the poor estimate $\|A_t - A_0\| \leq \mathrm{const}|\,\nu|^{-1}\sqrt{\Omega}$.

The second method, used in [40,15], is based on considerations of energy conservation: it uses the fact that the unperturbed Hamiltonian, if quasi–convex, has the following property: its restriction to the fast drift plane has a (quadratic) extremum at the intersection with the resonant manifold. In the general case, the two methods give essentially equivalent bounds on the variation of the actions. However, in the present case, the method based on energy conservation leads to significantly better results.[39)]

In order to properly formulate the results, we slightly restrict the resonant zone $\mathcal{I}(\nu, N(\Omega), \Omega)$, in such a way to avoid that a point (G, L), moving in the fast drift direction, can arrive in the region where $\|\omega\| \leq \Omega$. Precisely, we define $\hat{\mathcal{I}}(\nu, N(\Omega), \Omega)$ as consisting of those points $A \in \mathcal{I}(\nu, N(\Omega), \Omega)$ such that $\|\omega(A + x\nu)\| \geq 2\Omega$ for all $x \in \mathbb{R}$ for which $A + x\nu \in \mathcal{I}(\nu, N(\Omega), \Omega)$.

Proposition 5.3 *There exist constants c_{21}, c_{22}, c_{23} such that, for any motion with initial data satisfying $\|\omega(G_0, L_0)\| = 2\Omega \geq 2\Omega_*$ and $(G_0, L_0) \in \hat{\mathcal{I}}(\nu, N(\Omega), \Omega)$, one has*

$$|G_t - G_0| \leq c_{21} \tag{3.1a}$$

$$|L_t - L_0| \leq c_{21} \tag{3.1b}$$

$$\left|\dot{g}_t^{(\lambda)} - \frac{G_0}{a_1}\right| \leq c_{22} \qquad (\lambda = 1, 2) \tag{3.1c}$$

$$\left|\dot{l}_t - \eta\frac{L_0}{a_1}\right| \leq c_{22} \tag{3.1c}$$

---

39) This is due to the fact that our system has two frequencies. The reason of this fact is easily understood by referring to figure 5.3, thinking to the effect of the presence of a third action variable.

for all $t \in \mathbb{R}$ *such that*

$$|t| \leq c_{23} \exp \left[\sqrt{\frac{\Omega}{\Omega_*}} \right] . \tag{3.2}$$

Proof. From proposition 4.2 one deduces

$$\left| A'_t - (A'_0 + x_t \nu) \right| \leq \text{const} \sqrt{\frac{\Omega_*}{\Omega}}$$

for all times t as in (3.2), where $t \mapsto x_t$ is some (analytic) function. An argument based on the conservation of energy and the fact that, if h' is the new Hamiltonian as in (2.4) of chapter 4, then one has $|h' - k| \leq$ const, gives

$$\left| k(A'_t) - k(A'_0) \right| \leq c_{24} \tag{3.3}$$

for a certain constant c_{24}, which does not depend on the initial datum (as far as $\Omega \geq \Omega_*$). Inequality (3.3) means that A'_t is confined inside the narrow strip between the two level curves $k(A') = k(A'_0) \pm c_{24}$. The distance of such two level curves, in the region $\mathcal{I}$, is easily estimated to be $\leq$ constΩ^{-1} (one uses $k(A'_0) \sim \Omega^2$). It only remains to evaluate the length $d(\nu, A'_0)$ of the segment of the fast drift line $A'_0 + \mathbb{R}\nu$ which is contained between such two curves. Taking into account the estimate above on the width of the strip (3.3), and the fact that k is a quadratic polynomial, one easily sees that $d(\nu, A'_0) \leq c_{25}$, a constant independent of the initial data. Indeed, $d(\nu, A'_0)$ reaches such a maximum for A'_0 near the resonant line r_ν, and decreases as the distance of A'_0 from r_ν increases (see figure 5.2). The details of the proof can be easily worked out from these considerations. ∎

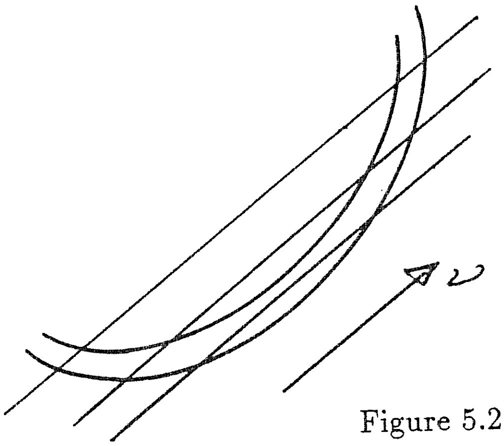

Figure 5.2

Remark 5.2: The proof of proposition 5.3, leading to the estimate $\mathcal{O}(\Omega^0)$ on the variation of G and L, is adapted to the worst case, that is to motions starting sufficiently near the resonant line r_ν, say within an angle $\sim \Omega^{-1}$ around it. One could assure a better confinement of the actions for motions starting at greater distances from r_ν (up to $\mathcal{O}(|\nu|/N)$, on the border of the resonant zone). This can be significant in the case of resonances of low order $|\nu| << N$ (in fact, for any given resonance, when $N \sim \Omega^{1/2} \to \infty$), since the angular amplitude $\sim (|\nu|N)^{-1}$ of the correspondent zone is $\sim \Omega^{-1/2} >> \Omega^{-1}$.

From the estimates (3.1), and the fact that the angular momentum vector moves slowly in space (proposition 5.1), one immediately deduces that the body undergoes approximatively a Poinsot precession about the direction of m (the angle between e_3 and m is now constant up to terms of order Ω_*/Ω). However, on a longer time scale, the motion can be deeply different with respect to the non–resonant case. The reason is that now the angular momentum vector is no more forced to remain close to any given curve on the unit sphere – rather, it can move in a wider region. In fact, the possibility of truly chaotic motions of m, on a time scale of order Ω, cannot be excluded.

This behaviour is in agreement with the general features of the dynamics of nearly integrable Hamiltonian systems near resonances, as depicted in [15].[40] In order to exhibit such a fact, we need to introduce the so–called 'slow' and 'fast' variables (see [65, 15, 9]). This requires some preliminary work, to which the next subsection is devoted.

B. Slow and fast variables. Let $n \in \mathbb{Z}^2$ be a vector such that $\langle n, \nu\rangle = 1$. Since ν_1 and ν_2 are relative prime numbers, the existence of such a vector is a well known fact. Moreover, as is convenient and possible[41], we assume that n satisfies $|n| \le |\nu|$. Consider the matrix

$$C = \begin{pmatrix} n_1 & n_2 \\ -\nu_2 & \nu_1 \end{pmatrix}$$

and perform the canonical change of variables $(G', L', g', l') \to (F, S, \varphi, \sigma)$, where

$$\begin{pmatrix} S \\ F \end{pmatrix} = C \begin{pmatrix} G' \\ L' \end{pmatrix} , \qquad \begin{pmatrix} \sigma \\ \varphi \end{pmatrix} = (C^t)^{-1} \begin{pmatrix} g' \\ l' \end{pmatrix} . \tag{3.4}$$

Notice that one has

$$\begin{aligned} S &= \langle n, A\rangle \\ F &= \nu \times A . \end{aligned} \tag{3.5}$$

Thus, $S(A'+x\nu) = S(A')+x$ and $F(A'+x\nu) = F(A')$ for any real number x. In particular, S is a coordinate in the direction ν of the fast drift.

As a function of the variables $S, F, J', \sigma, \varphi, j'$, each local representative of the normal form Hamiltonian h' takes the form

$$\begin{aligned} &\hat{k}(S,F) + \hat{w}(F,S,J',\sigma,j') + \sqrt{\frac{\Omega_*}{\Omega}}\, \hat{g}(F,S,J',\sigma,j') \\ &+ \sqrt{\frac{\Omega_*}{\Omega}}\, e^{-\left[\sqrt{\frac{\Omega}{\Omega_*}}\right]} \hat{f}(F,S,J',\varphi,\sigma,j') \end{aligned} \tag{3.6}$$

---

40) Because of degeneracy, the dynamics of our system is much similar to that of a three degrees of freedom system in the neighbourhood of a resonance of dimension two: a truly system with two degrees of freedom is essentially integrable, near a resonance.

41) Let $\overline{n} \in \mathbb{Z}^2$ be any vector such that $\langle \overline{n}, \nu\rangle = 1$. Then, for any $t \in \mathbb{Z}$, the vector $n = (\overline{n}_1 - t\nu_2, \overline{n}_2 + t\nu_1)$ satisfies $\langle n, \nu\rangle = 1$. Assume $\nu_1 \ne 0$, $\nu_2 \ne 0$ (the case in which a component of ν vanishes can be considered separately, in an obvious way). One can choose t so to have $-\nu_2 < n_1 < \nu_2$, namely $|n_1| \le |\nu_2| - 1$. From the relation $n_2 = (1 - \nu_1 n_1)/\nu_2$ one then gets $|n_2| \le |\nu_1|$.

where

$$\hat{k}(S,F) = \frac{1}{2\,a_1}\left\langle \begin{pmatrix} S \\ F \end{pmatrix}, \begin{pmatrix} \nu_1^2+\eta\nu_2^2 & -\nu_1 n_2+\eta\nu_2 n_1 \\ -\nu_1 n_2+\eta\nu_2 n_1 & \eta n_1^2+n_2^2 \end{pmatrix}\begin{pmatrix} S \\ F \end{pmatrix}\right\rangle \tag{3.7}$$

while $\hat{w}$, $\hat{g}$ and $\hat{f}$ are obtained from $\Pi_{\mathcal{L}_\nu} v$, g and f, respectively, with the change of variables (3.4), and are easily seen to satisfy (at each point)

$$|\hat{y}| \le |y|\,, \qquad \max_{\psi=\sigma,\varphi}\left|\frac{\partial \hat{y}}{\partial \psi}\right| \le 2\,|\nu|\max_{\xi=g,l}\left|\frac{\partial \hat{y}}{\partial \xi}\right| \qquad (y=v,g,f). \tag{3.8}$$

Let us notice that the resonant line $r_\nu = \{A' : \langle \omega(A'),\, \nu\rangle = 0\}$ has equation

$$\omega_S(S,F) := \frac{\partial \hat{k}}{\partial S}(S,F) = 0\,.$$

This justifies the name 'slow' variables for S and σ. On the other hand, F and φ are called 'fast' variables. Since $\frac{\partial k}{\partial J} \equiv 0$, we consider as 'slow' variables also J' and j'.

According to the Hamilton function (3.6), the 'fast' action F is essentially uncoupled from the 'slow' variables S, σ, J', j' on a time scale which grows exponentially fast with $\sqrt{\Omega/\Omega_*}$. In order to gain some insight into the evolution of the latter variables we introduce, following [15], new coordinates adapted to the initial condition of the motion. Precisely, with reference to a given motion, let A'_* be the intersection point of the fast drift line through the point A'_0, namely $A'_0 + \mathbb{R}\nu$, with the resonant line $r_\nu = \{A' : \langle \omega(A'),\, \nu\rangle = 0\}$ (this is well defined since r_ν and ν are never parallel). Correspondingly, define

$$\begin{pmatrix} S_* \\ F_* \end{pmatrix} = C\,A'_* \tag{3.9}$$

and

$$\begin{aligned} \hat{F} &= F - F_* \\ \hat{S} &= S - S_*\,. \end{aligned} \tag{3.10}$$

Notice that $\hat{F}(0) = 0$, $\hat{S}(0)\nu = A'_0 - A'_*$. Thus, $|\hat{S}(0)|\,|\nu|$ measures the distance (in the fast drift direction) of the point A'_0 from the resonant line. Let us also notice that $\omega_S(S_*,F_*) = 0$.

In the new variables $\hat{S}$, $\hat{F}$, the Hamilton function (3.6) takes the form

$$\begin{aligned} &u_{A'_0}(\hat{F}) + \left[\frac{1}{2}b_\nu \hat{S}^2 + c_\nu \hat{F}\hat{S} + \hat{w}_{A_0}(\hat{F},\hat{S},J',\sigma,j') + \sqrt{\frac{\Omega}{\Omega_*}}\hat{g}_{A'_0}(\hat{F},\hat{S},J',\sigma,j')\right] \\ &\quad + \sqrt{\frac{\Omega_*}{\Omega}}e^{-\left[\sqrt{\frac{\Omega}{\Omega_*}}\right]}\hat{f}_{A'_0}(\hat{F},\hat{S},J',\sigma,\varphi,j') \end{aligned} \tag{3.11}$$

where $b_\nu = \nu_1^2+\eta\nu_2^2$, $c_\nu = -\nu_1 n_2 + \eta\nu_2 n_1$,

$$u_{A'_0}(\hat{F}) = \hat{k}(S_*, F_* + \hat{F}) = \frac{1}{2}(\eta n_1^2 + n_2^2)(F_*+\hat{F})^2 + c_\nu S_* \hat{F}\,,$$

$\hat{w}_{A_0'}(\hat{F},\hat{S},J',\sigma,j') = \hat{w}(F_* + \hat{F}, S_* + \hat{S}, J', \sigma, j')$, and $\hat{g}_{A_0'}$ and $\hat{f}_{A_0'}$ are defined in a similar way. Notice that $|b_\nu| \leq |\nu|^2$, $|c_\nu| \leq |\nu|^2$.

C. Description of the resonant motion. We now come back to the description of the motion. As far as the exponentially small coupling term in (3.11) can be neglected, the time evolution of the 'slow' variables can be thought of as described by an effective Hamiltonian, given by the terms within the square brackets in (3.11). However, the term $c_\nu \hat{F}\hat{S}$ is small: since $\hat{F}(0) = 0$ and $|\hat{S}(t)| = |S_t - S_0| = \mathcal{O}(\sqrt{\Omega/\Omega_*})$, one has

$$\left| c_\nu \, \hat{F}(t)\hat{S}(t) \right| \;=\; \mathcal{O}\left(|t| \exp -\sqrt{\frac{\Omega}{\Omega_*}} \right) \tag{3.12}$$

Thus, as far as such term too can be neglected, one is left with the effective Hamilton function

$$\frac{1}{2} b_\nu \hat{S}^2 \;+\; \hat{w}_{A_0'}(\hat{F},\hat{S},J',\sigma,j') \;+\; \sqrt{\frac{\Omega}{\Omega_*}}\hat{g}_{A_0'}(\hat{F},\hat{S},J',\sigma,j') \,. \tag{3.13}$$

The fundamental difference with respect to the non–resonant case is that now the effective Hamiltonian has two degrees of freedom: the variables J', j' are coupled to the other slow variables S, σ.

Consequently, the motion of the angular momentum vector in space can be deeply different with respect to the non–resonant case (obviously, on times scales $\mathcal{O}(\Omega)$, since in both cases m moves in space with speed $\mathcal{O}(\Omega^{-1})$). If, as one expects to be the typical case, the effective Hamiltonian (3.13) is non–integrable, then the point $(J', j', \hat{S}, \sigma)$ can in principle move widely on each energy shell.

Consider, in particular, a motion starting near the resonant line, say at a distance $|\hat{S}_0|\,|\nu| = \mathcal{O}(1)$ from it, so that $\frac{1}{2} b_\nu \hat{S}^2$ is $\mathcal{O}(1)$. Thus, during the motion, variations of $\hat{w}_{A_0'}$ of order one can in principle take place. It is easy to see that such variations imply that the unit vector μ parallel to m is not confined to the vicinity of a curve on the unit sphere, but rather it will move in a widely extended region of the sphere. A deeper investigation of this situation is certainly needed. For instance, it would be very interesting to exhibit chaotic motions in some specific examples.

It would be also of interest to determine more precisely the time scales on which irregular behaviour appear. For instance, in the case of a low order resonance ($|\nu| \sim 1$), there exist motions starting far away from the resonant line, with $|\hat{S}| \sim \sqrt{\Omega}$. In such a case, the effective Hamiltonian (3.13) is a small perturbation of a free rotator. Thus, it can be expected that the two slow degrees of freedom are essentially uncoupled, on some intermediate time scale. This fact can be also related to remarks 5.2 and 5.4.

D. Some remarks. We conclude this section with some complementary remarks.

Remark 5.3: The transition from the resonant to the non–resonant regime of motion is obviously not so abrupt as presented here. In fact, it is possible to show that motions starting inside the resonant zones, but not too near the resonant lines, manifest, on intermediate time scales, the same characters of non–resonant motions. To this purpose, it is enough to perform the perturbation procedure of chapter 4 with reference to narrower resonant zones. Precisely, one can take the angular size of the resonant zone around r_ν to be of order $(N^\beta|\nu|)^{-1}$ for some $\beta > 1$ (compare with (1.15) of chapter 4). Correspondingly, as one easily sees by repeating the general argument of section 4.4.A, one takes the cut–off as $N(\Omega) \sim \Omega^{1/(1+\beta)}$, while the number r of iterations of the normalization procedure will be taken again to be proportional to N. In this way, one constructs, in the non–resonant zone, a normal form Hamiltonian given by

$$h' = k + u + \overline{\mathcal{V}} + \left(\frac{\Omega_*}{\Omega}\right)^{1/(1+\beta)} \mathcal{G} + \left(\frac{\Omega_*}{\Omega}\right)^{1/(1+\beta)} \mathcal{F} e^{-[(\Omega/\Omega_*)^{1/(1+\beta)}]} .$$

From this, one gets estimates like in proposition 5.2 (with the square root of Ω_*/Ω replaced by its power $1/(1+\beta)$) and a description of the motion similar to that of section 5.2, on a time scale growing with the exponential of $(\Omega_*/\Omega)^{1/(1+\beta)}$. Notice that, for any $\beta > 1$, the relative measure of the union of all resonant zones goes to zero as $\Omega \to \infty$: indeed, if $N(\Omega) \sim (\Omega/\Omega_*)^{1/(1+\beta)}$, then one has

$$\sum_{\substack{\nu\in\mathbb{Z}^2\\ 0<|\nu|\le N(\Omega)}} \frac{1}{N(\Omega)\,|\nu|} \sim \left(\frac{\Omega_*}{\Omega}\right)^{\frac{\beta-1}{\beta+1}} .$$

Remark 5.4: In principle, one could also introduce slow and fast variables with reference to the original action–angle variables:

$$\begin{pmatrix} S(A) \\ F(A) \end{pmatrix} = C \begin{pmatrix} G \\ L \end{pmatrix} , \qquad \begin{pmatrix} \sigma(\alpha) \\ \varphi(\alpha) \end{pmatrix} = (C^t)^{-1} \begin{pmatrix} g \\ l \end{pmatrix} . \tag{3.14}$$

Correspondingly, one can rephrase, in terms of $S(A)$ and $F(A)$, some of the results about the motion. First, let us observe that the distance among the fast and slow variables corresponding to the old and new coordinates satisfies the estimates[42)]

$$\max\left(|F(A') - F(A)|\,,\ |S(A') - S(A)|\right) \le \text{const}\,\frac{|\nu|}{N}\sqrt{\frac{\Omega_*}{\Omega}} . \tag{3.15}$$

This allows one to prove the following estimates, about the motion in the resonant zone:

$$|F(t) - F(0)| \le \text{const}\,\frac{|\nu|}{N}\sqrt{\frac{\Omega_*}{\Omega}}$$
$$|S(t) - S(0)| \le \text{const}\,\frac{1}{|\nu|}$$

---

42) A simple translation of the estimates of proposition 4.2 on the deformation $\|A' - A\|$ gives estimates which are worse than (3.15) for a factor $|\nu|$. The proof of (3.15) requires a careful use of some properties of the canonical transformation constructed in proposition 4.2.

for $|t| \le \text{const}\,\sqrt{\Omega_*/\Omega}\,\exp\left[\sqrt{\Omega/\Omega_*}\right]$.

Remark 5.5: Finally, essentially in view of the study of the case with no fixed point, we consider here the possibility of finding an analog of estimate (2.3). To this purpose, in this resonant case, one can look for the conservation of the energy of the 'fast' motion. Now, using (3.12) and $\frac{d\hat{F}}{dt} = \mathcal{O}(|t|\sqrt{\Omega/\Omega_*}\exp -[\sqrt{\Omega/\Omega_*}])$, one easily obtains

$$\max\left(\left|u_{A_0'}(F)\right|_0^t \, , \, \left|\frac{1}{2}b_\nu\,\hat{S}^2 \, + \, \hat{w}_{A_0'}(F_0, S, J', \sigma, \varphi)\right|_0^t \right) \;\le\; c_{43}\sqrt{\frac{\Omega_*}{\Omega}} \tag{3.16}$$

for times

$$|t| \;\le\; c_{44}\,\frac{\Omega_*}{\Omega}\;\exp\left[\sqrt{\frac{\Omega_*}{\Omega}}\right] \, . \tag{3.17}$$

It would be more interesting to obtain an analogous estimate in term of the original variables (i.e., of the variables $S(A)$, $F(A)$, $\sigma(\alpha)$, $\varphi(\alpha)$, J, j). Unfortunately, we cannot obtain such a separation of the energy $k + \mathcal{V}$ in two parts, separately almost constant, to be interpreted as the energy associated to the fast and, respectively, to the slow motion. The ultimate reason is in the fact that the term $\frac{1}{2}b_\nu\hat{S}^2$ can assume quite large values: up to $\mathcal{O}(\Omega)$, on the border of a resonant zone of low order.

5.4 Near the gyroscopic rotations

So far, we have not considered motions which are too near the gyroscopic rotations. In fact, all the above results are relative to high frequency motions in the subset $\mathcal{I}-2\rho = \{(G,L) \in \mathbb{R}^2 : G > |L| + 2\rho_A\}$ of the action space $\mathcal{I}$. By the way, it is very easy to obtain estimates on the variation of the actions for motions within such an excluded neighbourhood.

To this purpose, let us first observe that, within the excluded neighbourhood, an obstruction to the variation of the actions (G, L) is provided by conservation of energy: the point (G_t, L_t) must be contained within the two level curves $k(G,L) = k(G_0,L_0) \pm |\mathcal{V}|_\rho$. As already remarked, the distance between such curves is $\mathcal{O}(\Omega^{-1})$ if $k(G_0, L_0) \sim \Omega^2$. The fact is now that the equipotential curves of k intersect transversally the lines $G = \pm L$ (see figure 5.3). This implies that a point (G_t, L_t) will move at most of a distance $\mathcal{O}(1)$ before escaping the excluded set $\{G \le |L| - 2\rho_A\}$. Then, an obvious consistency argument based on the results of propositions 5.2 and 5.4 shows that one has $|G_t - G_0| = \mathcal{O}(1)$, $|L_t - L_0| = \mathcal{O}(1)$ for times $|t| = \mathcal{O}(\exp\sqrt{\Omega/\Omega_*})$.

In this way, we can conclude that, for *all* fast motions of the symmetric rigid body with a fixed point, G and L vary at most of quantities of order one, on the above time scale.

However, the present treatment of motions near the gyroscopic rotations is very poor. In particular, we cannot give any accurate geometrical description of such motions.

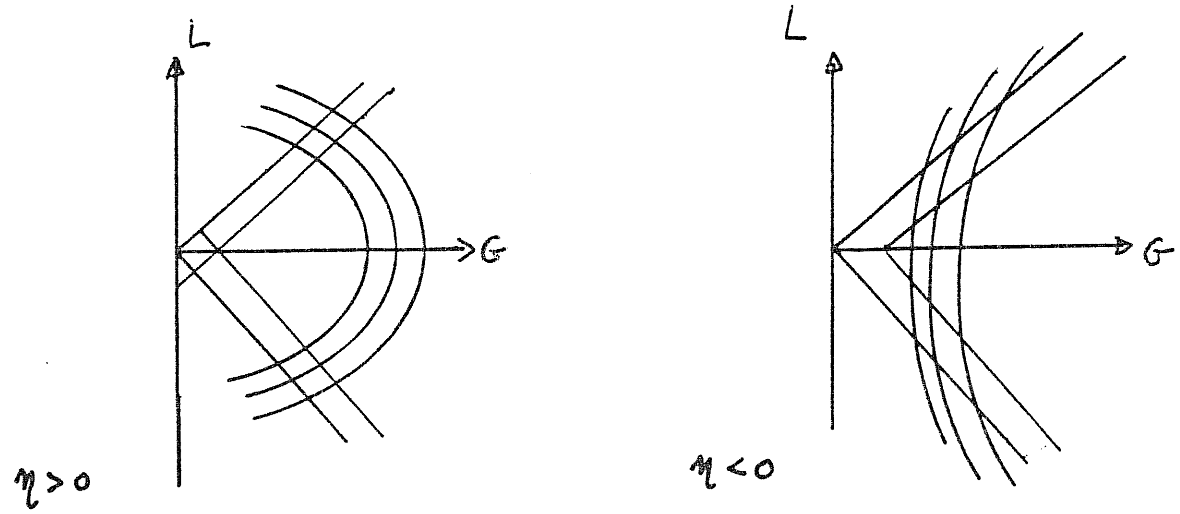

Figure 5.3

As discussed already in the first chapter, this is due to the fact that the action–angle variables are not defined on the gyroscopic rotations, which are isolated periodic orbits. To study the system in the neighbourhood of such orbits one could employ standard techniques, like the study of a Poincaré map.[43)] Another, more promising approach to this problem could be based on the introduction of adapted coordinates, similar to the so-called Poincarè variables for the Kepler problem. These variables, at variance from the action–angle variables ('Delaunay elements'), are defined also on the isolated periodic orbits of the Kepler system, namely the circular orbits (for details, see [1]). Such coordinates are conveniently used in perturbative studies of the three or N body problem (see [4,65]). Because of the analogy between the Kepler system and the rigid body, one can hope that a similar procedure works for the study of a neighbourhood of the gyroscopic rotations of the rigid body.[44)]

5.5 On the motion with no fixed point

In concluding the chapter, we give some indications about the motion of a fast rotating symmetric rigid body with no fixed point. This is an extremely interesting problem ('spin–orbit coupling'). Unfortunately we do not have at present a real understanding of it. The central open question concerns the very possibility of large energy exchanges between the rotational and the translational degrees of freedom, when the frequencies of the rotational motion are near a resonance of the Euler–Poinsot motion.

---

43) To this purpose, one can possibly still use Poinsot coordinates, but with the axis e_3 of the basis $\mathcal{B}_s$ (refer to chapter 2 for notations) coincident with one of the degenerate inertia axes, *not* with the gyroscopic one.

44) This possibility has been suggested to us by A. Gorgilli.

On the contrary, we can give definite results for non–resonant cases. Such results are strictly related to the ones obtained in the study of a simple, planar model in [21] (the body was assumed to move in a plane, with angular velocity orthogonal to the plane), so that we refer to that paper for many details. The basic difference concerns the time scale on which rotational and translational degrees of freedom behave as if they were uncoupled: $\exp\sqrt{\Omega/\Omega_*}$ in the present case, $\exp \Omega/\Omega_*$ in the planar one.

A. Non–resonant motions. We consider the case in which, at the initial time, the energy of the center of mass is $\mathcal{O}(1)$, while that of the motion relative to the center of mass, namely k, is order Ω^2; here, as before, $2\Omega = \|\omega(G_0, L_0)\|$. For definiteness, we assume that the external potential energy $\mathcal{V}$ is analytic and bounded for all $q \in \mathbb{R}^3$ (this is not a crucial hypothesis), and obviously for all $\mathcal{R} \in SO(3)$.

In order to apply proposition 4.2, we need to assume, in particular, $\Omega \geq \Omega_*$. Here, Ω_* depends on the potential $\mathcal{V}$, but also on the maximum value of $\|p\|$ on (a complex neighbourhood of) the domain Q. Now, we take $Q = \mathbb{R}^3 \times \Delta_{\mathbb{R}}(p(0), c\|p(0)\|)$, where c is some positive constant (independent of Ω). Thus, we are faced with a consistency problem: to show that, in the time interval under consideration, $\|p_t\|$ does not increase to much (such a consistency problem is trivial, in the non–resonant case).

Proposition 5.4 *There exist constants $c_{61}, \ldots, c_{67}$ for which the following hold. Consider a positive number $E \geq 2(\|\mathcal{V}\|_\rho + |\mathcal{V}|_\rho)$, and let*

$$\Omega_{**} = \frac{2^{11}}{\rho_\alpha^2} c_{61} \sqrt{\frac{2E}{M}} .$$

Consider a motion of the body with initial data satisfying

$$\begin{aligned} \frac{\langle p(0),\, p(0)\rangle}{2M} &\leq E\,, \\ \|\omega(G_0, L_0)\| &= 2\Omega \geq 2\max\left(\Omega_{**}, \frac{2^7\mu(\eta)\rho_A}{a_1\rho_\alpha^3}\right) \\ (G_0, L_0) &\in \mathcal{I}(0, N(\Omega), \Omega)\,. \end{aligned} \tag{5.1}$$

Then, such a motion is defined in the whole time interval

$$|t| \leq c_{62} \frac{\Omega_*}{\Omega} \exp\left[\sqrt{\frac{\Omega}{\Omega_*}}\right] \tag{5.2}$$

and satisfies

$$\max(|G_t - G_0|, |L_t - L_0|) \leq c_{63} \sqrt{\frac{\Omega_*}{\Omega}} \tag{5.3a}$$

$$\max\left(\left|\dot{g}_t^{(\lambda)} - \frac{G_0}{a_1}\right|, \left|\dot{l}_t - \eta\frac{L_0}{a_1}\right|\right) \leq c_{64} \sqrt{\frac{\Omega_*}{\Omega}} \tag{5.3b}$$

$$\left|\frac{\langle p(0),\, p(0)\rangle}{2M} + \overline{\mathcal{V}}\right|_0^t = \left|k + \mathcal{V} - \overline{\mathcal{V}}\right|_0^t \leq c_{65} \sqrt{\frac{\Omega_*}{\Omega}} . \tag{5.3c}$$

Proof. The proof goes essentially the same way as for proposition 5.2. In addition, one has here to show that $\|p_t\|$ does not increase too much. This a consistency question, which is easily proven by using the energy inequality (5.3c). For details, see the proof of the analogous statement in [21]. ∎

It is clear that, in this non–resonant case, there are no significant differences with respect the case of the body with a fixed point. Obviously, there are now the additional slow variables (p, q), which are coupled with all the other slow variables, but the fast variables G, g, L, l are instead essentially uncoupled; in particular, the energy k of the fast motion is almost constant, up to (high frequency) fluctuations.

B. On a case of scattering. We consider now the case of the scattering of a fast rotating (symmetric) body by a fixed obstacle. We describe the obstacle by a smooth potential, which decays at infinity in an integrable way. The basic question is how much energy can be transferred from the rotational to the translational degree of freedom as a result of a single collision.

Such a problem is of interest, among other things, in connection with a conjecture by Boltzmann[27] and Jeans[46,47] about the 'freezing' of high frequency degrees of freedom in Hamiltonian systems, an argument of interest for the foundations of classical statistical mechanics (see [13, 16, 17, 19, 18, 39, 10, 21]).

We prove that, if far before the collision the body undergoes a non–resonant fast rotation about its center of mass, then the energy exchange between the translational and rotational degrees of freedom is extremely small, precisely it decreases exponentially fast with $\sqrt{\Omega/\Omega_*}$. The proof of this fact uses in an essential way the local estimates on the normal form and the canonical transformations given in proposition 4.2.

In order to proper formulate the results, let us introduce some terminology. We say that we have a "scattering trajectory" $(b(t), \alpha(t))$, with $b(t) = (G_t, L_t, J_t, p_t, j_t, q_t)$, $-\infty < t < +\infty$, if the following conditions are fulfilled:

$$\lim_{t\to\pm\infty} \|q(t)\| = \infty \tag{5.4a}$$

$$\lim_{t\to\pm\infty} p(t) = p^{\pm} \neq 0 \tag{5.4b}$$

$$\beta := \max_{i=1,2} \int_{-\infty}^{+\infty} \left|\frac{\partial v}{\partial \alpha_i}\right|_{z(t),\rho} dt < \infty . \tag{5.4c}$$

In particular, the latter condition assures that $\frac{\partial v}{\partial g}(b(t), \alpha(t))$ and $\frac{\partial v}{\partial l}(b(t), \alpha(t))$ go to zero sufficiently fast for $t \to \pm\infty$, so that the limits $G(\pm\infty)$ and $L(\pm\infty)$ also exist. Of course, one could make assumptions on v (essentially, repulsivity with reference to a fixed scatterer) which ensure the existence of scattering trajectories: however, assuming directly (5.4) is simpler and, in the present framework, more natural.

Proposition 5.5 *For any scattering trajectory, with p^- and $\Omega = \frac{1}{2}\|\omega(G(-\infty), L(-\infty))\|$ satisfying the conditions (5.1), one has*

$$|k|_{-\infty}^{+\infty} \leq \text{const} \exp - \left[\sqrt{\Omega/\Omega^*}\right] . \tag{5.5}$$

Proof. One proceeds as in the proof of proposition 5.4. In addition to the conclusions of proposition 5.4, one has now

$$|G'(+\infty) - G'(-\infty)| \leq \frac{1}{3}\sqrt{\frac{\Omega_*}{\Omega}}\, e^{-\left[\sqrt{\Omega/\Omega^*}\right]} \int_{-\infty}^{+\infty} \left|\frac{\partial v}{\partial g}\right|_{b(t),\rho} dt \leq \frac{1}{3}\beta\sqrt{\frac{\Omega_*}{\Omega}}\, e^{-\left[\sqrt{\Omega/\Omega^*}\right]} . \tag{5.6}$$

On the other hand, by (5.4c), and (2.6a) of chapter 4, at $t = \pm\infty$ one has $G = G'$. Proceeding in the same way for the variable L', one easily proves inequality (5.5). ∎

C. Resonant motions. We give now some indications about the difficulty of extending the previous results to resonant motions. We refer, for definiteness, to the case of subsection A. First of all, we remark that the sharp estimates of proposition 5.3 cannot be assured to hold in the present case. The reason is simply that the conservation of energy now reads

$$k + \frac{\langle p(0), p(0)\rangle}{2M} + \mathcal{V} = \text{const}$$

and one cannot apriori exclude large ($\sim \Omega^2$) energy exchanges between k and $\langle p(0), p(0)\rangle / 2M$. This is the heart of the problem.

A better, though not yet satisfactory approach could be based on the distinction between slow and fast variables. In the present case, the analog of inequality (3.16) of remark 5.5 would read

$$\left| \frac{1}{2} b_\nu \hat{S}^2 + \frac{\langle p(0), p(0)\rangle}{2M} + \hat{w}_{A_0'}(F_0, S, J', p', \sigma, \varphi, q') \right|_0^t \leq \text{const}\sqrt{\frac{\Omega_*}{\Omega}} . \tag{5.7}$$

As already noticed, the term $\frac{1}{2}b_\nu \hat{S}^2$, which could be considered as the energy of the slow motion, is not always small: for motions starting on the border of a resonant zone of low order, such a term is $\mathcal{O}(\Omega)$. This would lead to an a–priori estimate for the variation of p of order $\sqrt{\Omega}$. This is better than the above one, but still not completely satisfactory. We observe that, in order to have smaller values of the 'slow motion' energy $\frac{1}{2}b_\nu \hat{S}^2$, one would significantly restrict the zones corresponding to the low order resonances. Unfortunately, we do not see how to do this (without reducing too much the time scales: see remark 5.3).

Thus, the very question of the possibility of large energy exchanges, in conditions of resonance, is at least at present an open question. It is certainly worth of further analytic and numerical investigations.

Conclusions

In the present thesis we have been concerned with two main topics. On the one hand, there is a 'global' formulation of perturbation theory (chapter 2), which was a necessary prerequisite for the study of the fast rotations of the rigid body. Essentially, we found that certain singularities of the action–angle coordinate systems are not really dangerous for perturbation theory, and the difficulties inherent their presence can be easily overcome. Besides the rigid body, this situation is encountered in the Kepler system. Thus, a further, quite natural application of these ideas could concern the N–body problem: as already remarked in chapter 1, Nekhoroshev's results about the existence of open domains of initial conditions, which give rise to 'planetary motion' (i.e., no close collisions, and not too large escapes) for exponentially long time scales, suffer quite heavily, in our opinion, for the need of excluding the passage through the singularities of the coordinates.

On the other hand, the central argument of the thesis has been the study of the fast rotations of the symmetric rigid body, with a fixed point. In our opinion, Hamiltonian perturbation theory provides a powerful approach to this problem. Our results about the fast motion have been presented in chapter 5. Besides definite bounds on the variations of the action variables G and L, we gained some insight about some features of the body motion.

This study is certainly not yet complete.

First of all, a great lack lies in the restriction to the case of symmetric bodies. This was originally motivated by the will of avoiding, in the first approach to the problem, some technical difficulties. But now, it seems to us that no real difficulty will be met in such an extension.

Of course, a second direction in which this work has to be completed concerns the gyroscopic rotations. To this purpose, one has to study the neighbourhood of a family of periodic orbits, and we are confident that the methods of Hamiltonian perturbation theory will lead to sharp results.

Another question concerns the quality of the estimates. As observed, estimates $\mathcal{O}(1)$ on the variations of G and L are optimal, as far as one looks for estimates uniform in phase space. Clearly, it would be of interest to investigate the optimality of the time scale $\sim \exp(\Omega^{1/2})$ over which such estimates have been proven. Testing the optimality of the exponent $1/2$ would furnish an essential indication about the optimality of perturbation theory, as developed in chapter 4. This is in our opinion an interesting question, since at present we do not see any possibility of improving such a value.

A natural continuation of this work would be the study of some specific example, possibly with the assistance of numerical tools. The possibility of chaotic motions, and the different time scales, certainly deserves a careful consideration.

But foremost, among all open problems, there is the motion with no fixed point. As shortly discussed in chapter 5, motions far from the resonances of the Euler–Poinsot motion are easily treated. But resonant motions are at present basically not understood. We do not think that this will reveal to be a formidable problem, but certainly this problem has to be considered as open. The key question concerns the very possibility of large energy exchanges between the translational and the rotational degrees of freedom. Our feeling is that there exist specific, but interesting situations in which they do not take place. A possible one is the case of scattering, where one can hope of successfully exploiting the fact that, during a collision process, the strong interaction lasts only a very short time.

Moreover, there is another case which is worth of consideration. One should remark that the angular velocity $\omega = a_1^{-1}(G, \eta L)$ of the body can be large either because, as we supposed here, the angular momentum is large (limit of high rotational energy), or else, at fixed finite rotational energy, because the inertia moments are small (limit of point–mass). The latter is the case of a small body of diameter proportional to a small parameter ε. The moments of inertia scale as ε^2, so that for small values of the angular momentum, $G = \mathcal{O}(\varepsilon)$, one has $\omega = \mathcal{O}(\varepsilon^{-1})$ and $k = (G^2 + \eta L^2)/2a_1 = \mathcal{O}(1)$. This case is certainly relevant in connection with the 'Boltzmann–Jeans' conjecture. It was successfully studied in the planar model of reference [21]. In three dimensions an additional difficulty arises, namely the action–angle variables become singular near $G = 0$. Nevertheless, we are confident that these technical difficulties can be overcome.

In conclusion, much work remains to be done in this field, and many problems have to be carefully studied. We hope that the work done for this thesis will provide a concrete basis for it.

References

[1] R. Abraham and J. E. Marsden, *Foundations of Mechanics* (Benjamin, Reading 1978).

[2] H. Andoyer, *Cours de Mécanique Céleste* (Gauthier–Villars, Paris 1923).

[3] V. I. Arnold, *Proof of a theorem of A.N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian.* Usp. Mat. Nauk **18**:5, 13-40 (1963) [Russ. Math. Surv. **18**:5, 9-36 (1963)].

[4] V. I. Arnold, *Small denominators and problems of stability of motion in classical and celestial mechanics.* Usp. Mat. Nauk **18**:6, 91-192 (1963) [Russ. Math. Surv. **18**:6, 85-191 (1963)].

[5] V.I. Arnold, *Instability of dynamical systems with several degrees of freedom.* Dokl. Akad. Nauk. SSSR **156**:1, 9-12 (1964) [Sov. Math. Dokl. **5**, 581-585 (1964)].

[6] V.I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics* (Benjamin, New York 1968).

[7] V. I. Arnold, *Méthodes Mathématiques de la Mécanique Classique* (MIR, Moscow 1976).

[8] V. I. Arnold, *Chapitres supplémentaires de la théorie des équations différentielles ordinaires* (MIR, Moscow 1980).

[9] V. I. Arnold, V.V. Kozlov and A.I. Neishtadt, *Mathematical aspects of classical and celestial mechanics*, in *Dynamical system III*, Volume 3 of *Encyclopedia of Mathematical Sciences*, V.I. Arnold editor (Springer-Verlag, Berlin, 1988).

[10] O. Baldan and G. Benettin, *Classical "freezing" of fast rotations: numerical test of the Boltzmann–Jeans conjecture.* J. Stat. Phys. **62**, 201-219 (1991).

[11] V. Béletski, *Essais sur le mouvement des corps cosmiques* (Mir, Moscow, 1986).

[12] A. Belleni–Morante, *Sul principio dell'effetto giroscopico.* Rendiconti Sem. Mat. Padova **39**, 304-329 (1967).

[13] G. Benettin, L. Galgani and A. Giorgilli, *Boltzmann's ultraviolet cutoff and Nekhoroshev's theorem on Arnold diffusion.* Nature **311**, 444-445 (1984).

[14] G. Benettin, L. Galgani and A. Giorgilli, *Poincaré's Non-Existence Theorem and Classical Perturbation Theory in Nearly-Integrable Hamiltonian Systems*, in *Advances in Nonlinear Dynamics and Stochastic Processes*, R. Livi and A. Politi editors (World Scientific, Singapore 1985).

[15] G. Benettin and G. Gallavotti, *Stability of motions near resonances in quasi–integrable Hamiltonian systems.* J. Stat. Phys. **44**, 293-338 (1986).

[16] G. Benettin, L. Galgani and A. Giorgilli, *Exponential law for the equipartition times among translational and vibrational degrees of freedom.* Phys. Lett.**A 120**, 23 (1987).

[17] G. Benettin, L. Galgani and A. Giorgilli, *Realization of Holonomic Constraints and Freezing of High–Frequency Degrees of Freedom in the Light of Classical Perturbation Theory, part I.* Comm. Math. Phys. **113**, 87-103 (1987)

[18] G. Benettin, *A Completely Classical Mechanism for the "Freezing" of the High-Frequency Degrees of Freedom,* lecture at the workshop "Chaos and Complexity",

R. Livi, S. Ruffo, S. Ciliberto and M. Buiatti editors (World Scientific, Singapore, 1988).

[19] G. Benettin, L. Galgani and A. Giorgilli, *Realization of Holonomic Constraints and Freezing of High–Frequency Degrees of Freedom in the Light of Classical Perturbation Theory. Part II.* Comm. Math. Phys. **121**, 557-601 (1989).

[20] G. Benettin, L. Chierchia and F. Fassò, *Exponential estimates on the one–dimensional Schroedinger equation with bounded analytic potential.* Ann. Inst. Henri Poincaré **51**, 45-66 (1989).

[21] G. Benettin and F. Fassò, *Classical "freezing" of plane rotations: a proof of the Boltzmann–Jeans conjecture*, to appear in J. Stat. Phys.

[22] V.N. Bogaevskii and L.A. Ostrer, *On the fast rotation of a heavy rigid body about a fixed point.* Prikl. Math. Mech. **44**, 999-1009 (1980) [*PMM USSR*, **44**, 713-720 (1981)].

[23] F. Boigey, *Une transformation canonique de Mathieu dans l'espace des phases du mouvement d'un solid mobile autour d'un point fixe.* C. R. Acad. Sc., Ser. A, **272**, 1115-1118 (1971).

[24] F. Boigey, *Une théorie des perturbations en variable angles–actions. Application au mouvement d'un solide autour d'un point fixe. Précession–nutation.* J. de Méc. **11**, 521-543 (1972).

[25] O.I. Bogoyavlenskii, V.S. Vladimirov, I.V. Volovich, A.K. Gushchin, Yu. N. Drozhzhinov, V.V. Zharinov and V.P. Mikhailov, *Trudy Mat. Inst. Steklov* **175** (1986) [Translated in: V.S. Vladimirov, E.F. Mishenko and A.K. Gushchin (editors), *Theoretical and Mathematical Physics* (Proceedings of the Steklov Instute of Mathematics, 1988, Issue 2)]

[26] A.V. Bolsinov, S.V. Matveev and A.T. Fomenko, *Topological classification of integrable hamiltonian systems with two degrees of freedom. List of systems of small complexity.* Usph. Mat. Nauk. **45**:2, 49-77 (1990) [Russian Math. Surv.] **45**:2, 59-94 (1990)]

[27] L. Boltzmann, *Nature* **51**, 413 (1895).

[28] A. D. Bryuno, *Normalization of a Hamiltonian system near an invariant cycle or torus.* Usp. Mat. Nauk **44**, 49-77 (1989) [Russ. Math. Surv. **44**, 53-89 (1989)].

[29] L. Chierchia and G. Gallavotti, *Smooth prime integrals for quasi–integrable hamiltonian systems.* Nuovo Cimento B **67**, 277-295 (1982).

[30] R. Cushman, *Normal Form for Hamiltonian Vectorfields with Periodic Flow*, in *Differential Geometric Methods in Mathematical Physics*, S. Sternberg editor, 125–144 (Reidel, Dordrecht, 1984).

[31] R. Cushman and J. J. Duistermaat, *The quantum mechanical spherical pendulum.* Bull. Am. Math. Soc. **19**, 475-479 (1988).

[32] R. Cushman and H. Knörrer, *The energy–momentum mapping of the Lagrange top*, in *Differential geometric methods in mathematical physics*, H.D. Doebner and J.D. Henning editors, *Lect. Notes Math.* No. 1139 (Springer Verlag, Berlin, 1985).

[33] P. Dazord and T. Delzant, *Le probleme general des variables actions-angles.* J. Diff. Geom. **26**, 223-251 (1987).

[34] A. Deprit, *Free rotation of a rigid body studied in phase plane.* Am. J. Phys. **55**, 424-428 (1967).

[35] J. J. Duistermaat, *On global action–angle coordinates.* Commun. Pure Appl. Math. **33**, 687-706 (1980).

[36] F. Fassò, *Lie series method for vector fields and Hamiltonian perturbation theory.* J. Appl. Math. Phys. (ZAMP) **41**, 843-864 (1990).

[37] A.T. Fomenko and V.V. Trofimov, *Integrable systems on Lie algebras and symmetric spaces*, Advanced studies in contemporary mathematics, volume 2 (Gordon and Breach Science Publishers, New York, 1988).

[38] L. Galgani, A. Giorgilli and J.-M. Strelcyn, *Chaotic motions and transition to stochasticity in the classical problem of the heavy rigid body with a fixed point.* Nuovo Cimento B **61**, 1-19 (1981).

[39] L. Galgani, *Relaxation Times and the Foundations of Classical Statistical Mechanics in the Light of Modern Perturbation Theory,* lectures given at the Noto School "Non-Linear Evolution and Chaotic Phenomena", G. Gallavotti and P.F. Zweifel Editors (Plenum Press, New York, 1988).

[40] G. Gallavotti, *Quasi–Integrable Mechanical Systems,* in *Critical phenomena, Random Systems, Gauge Theories,* K. Osterwalder and R. Stora editors, *Les Houches,* Session XLIII, 1984 (North–Holland, Amsterdam 1986).

[41] A. Giorgilli, A. Delshams, E. Fontich, L. Galgani and C. Simó, *Effective Stability for a Hamiltonian System near an Elliptic Equilibrium Point, with an Application to the Restricted three Body Problem.* J. Diff. Eq. **77**, 167-198 (1989)

[42] H. Goldstein, *Classical Mechanics,* 2^{nd} edition (Addison–Wesley, Reading,86)

[43] P. J. Holmes and J. E. Marsden, *Horseshoes and Arnold diffusion for Hamiltonian systems on Lie groups. Indiana Univ. Math. J.* **32**, 272-309 (1983).

[44] E. Husson, *Acta Math.* **31**, 71 (1908).

[45] Yu. S. Il'yashenko, *A steepness test for analytic functions.* Usp. Mat. Nauk **41**:1, 193-194 (1986).

[46] J.H. Jeans, *On the vibrations set up in molecules by collisions.* Phil. Mag. **6**, 279 (1903).

[47] J.H. Jeans, *On the partition of energy between matter and aether.* Phil. Mag. **10**, 91 (1905).

[48] A. Knauf, *Closed orbits and converse KAM theory.* Nonlinearity **3**, 961-973 (1990).

[49] A.N. Kolmogorov, *On conservation of conditionally periodic motions under small perturbations of the Hamiltonian.* Dokl. Akad. Nauk. SSSR **98**, 527-530 (1954). English translation in G. Casati and G. Ford (Editors), *Lect. Notes in Phys.* No. 93 (Springer Verlag, Berlin, 1979).

[50] A. N. Kolmogorov, *General theory of dynamical systems and classical mechanics* (in Russian), in *Proc. of the 1954 Intern. Congr. Math.* (North Holland, Amsterdam

1957). English translation in: R. Abraham and J. E. Marsden, *Foundations of Mechanics* (Benjamin, Reading 1978).

[51] V. V. Kozlov, *Geometry of "action–angle" variables in Euler–Poinsot problem. Vestn. Moskov. Univ., Mat. Mekh.* **29**, 74-79 (1974).

[52] V. V. Kozlov, *The non–existence of an additional analytic integral in the problem of the motion of an asymmetric heavy solid around a fixed point* Vestn. Moskov. Univ., Mat. Mekh. **1975**, no. 1, 105-110 (1975) [Moskow Univ. Mech. Bull. **30**:1-2, 17-21 (1975)].

[53] V. V. Kozlov, *Methods of qualitative analysis in the dynamics of a rigid body* (in Russian). (Izdatel'stvo Moskovskogo Universiteta, Moskow 1980).

[54] V. V. Kozlov, *Integrability and non–integrability in Hamiltonian mechanics.* Usp. Mat. Nauk **38**:1, 3-67 (1983) [Russ. Math. Surv. **38**:1, 1-76 (1983)].

[55] E. Leimanis, *Some recent advances in the dynamics of rigid bodies and celestial mechanics*, in *Surveys in Appled Mathematics*, vol. **2** (New York, 1958)

[56] T. Levi–Civita and E. Amaldi, *Lezioni di Meccanica Razionale* (Zanichelli, Bologna, 1952)

[57] P. Libermann and C.-M. Marle, *Geometrie Symplectique, bases theoriques de la mecanique* (4 vol.) (Publications mathématiques de l'Université Paris VII, 1986)

[58] L. Markus and K.R. Meyer, *Generic Hamiltonian dynamical systems are neither integrable nor ergodic.* Memoirs of the Am. Math. Soc. **144**, 1-52 (1974)

[59] J. Moser, *Lectures on Hamiltonian Systems. Mem. Am. Math. Soc.* No 81, 1 (1968).

[60] J. Moser, *Regularization of Kepler's problem and the averaging method on a manifold.* Commun. Pure Appl. Math. , **23**, 609-636 (1970).

[61] J. Moser, *Stable and random motions in dynamical systems* (Princeton University Press and Tokio University Press, Princeton 1973).

[62] A.I. Neishtadt, *Estimates in the Kolmogorov theorem on conservation of conditionally periodic motions.* Prikl. Matem. Mekan. **45**:6, 1016-1025 (1981) [J. Appl. Math. Mech. **45**:6, 766-772 (1982)].

[63] N. N. Nekhoroshev, *Behaviour of hamiltonian systems close to integrability.* Funct. Anal. Appl. **5**, 338-339 (1971) [Funk. An. Ego Prilozheniya, **5**, 82-83 (1971)]

[64] N. N. Nekhoroshev, *Action–angle variables and their generalizations.* Trudy Moskov. Mat. Obsc. **26** 181-198 (1972) [*Trans. Moskow Math. Soc.* **26** 180-198 (1972)].

[65] N. N. Nekhoroshev, *An exponential estimate of the time of stability of nearly integrable hamiltonian systems.* Usp. Mat. Nauk **32**:6, 5-66 (1977) [Russ. Math. Surv. **32**:6, 1-65 (1977)].

[66] N. N. Nekhoroshev, *An exponential estimate of the time of stability of nearly integrable hamiltonian systems. II.* Trudy Sem. Petrows. No. 5, 5-50 (1979). Translated in *Topics in Modern Mathematics: Petrovskii Seminar no. 5*, edited by O.A. Oleinik, Consultant Bureau, New York (1985).

[67] H. Poincaré, *Les méthodes nouvelles de la méchanique céleste*, Vols. 1-3 (Gauthier–Villars, Paris, 1892/1893/1899).

[68] J. Pöschel, *Integrability of Hamiltonian systems on Cantor sets*. Commun. Pure Appl. Math. **35**, 653-695 (1982).

[69] Iu. A. Sadov, *The action–angle variables in the Euler–Poinsot problem*. Prikl. Math. Mech. **34**, 962–964 (1970)

[70] A. Signorini, *Complementi alla dinamica dei giroscopi e equazioni del problema completo della balistica esterna*. Atti Accad. Naz. Lincei, Serie 8, vol. 1, Memorie, 1-41 (1946).

[71] F. Stoppelli, *Sul principio dell'effetto giroscopico*. Giornale di Matem. di Battaglini, serie 4, **80** (1950-51).

[72] F. Stoppelli, *Un fenomeno di risonanza nei solidi in rapida rotazione*. Ricerche di Matem., **9**, 213-241 (1960).

[73] J.L. Synge, *Classical Dynamics*. In S. Flügge (editor), *Handbuch der Physik*, vol. III/1 *Principles of classical mechanics and field theory* (Springer, Berlin, 1960).

[74] A. Weinstein, *Symplectic manifolds and their lagrangian submanifolds*. Advances in Math. **6**, 329-346 (1971).

[75] E. T. Whittaker, *A treatise on the analytical dynamics of particles & rigid bodies* (Cambridge Mathematical Library, Cambridge, 1937)

[76] S.L. Ziglin, *Splitting of separatrices, branching of solutions and nonexistence of an integral in the dynamics of a solid body. Trudy Moskov. Mat. Obsc.* **41**, 287-302 (1980) [Trans. Moscow Math. Soc. 1982, No. 1, 283-298 (1982)].