Long term stability of proper rotations of the perturbed Euler rigid body

Giancarlo Benettin[*] Francesco Fassò[†] Massimiliano Guzzo[‡]

Abstract

We study the long term stability of the proper rotations of the perturbed Euler rigid body, in the framework of Nekhoroshev theory. For simplicity we treat here in detail only the kinetically symmetric case (the potential needs not to be symmetric), but we indicate how to extend the results to the triaxial case. We show that the proper rotations around the symmetry axis are Nekhoroshev stable: more precisely, if the initial datum is sufficiently close to a proper rotation, then for a very long time it remains such, and the tip of the unit vector μ parallel to the angular momentum precesses, up to a small noise, along the level curves of a regular function on the unit sphere. If the proper rotations are resonant, chaotic motions with positive Lyapunov exponents are possible, but chaos (unlike the case of ordinary motions, that is motions not close to proper rotations) is always localized, and does not affect in an essential way the motion of the angular momentum in space. Preliminary numerical result indicate that the theory is, in many aspects, optimal, although in some points it can still be improved.

[*]Università di Padova, Dipartimento di Matematica Pura e Applicata, INFM and GNFM (INDAM), Via G. Belzoni 7, 35131 Padova, Italy (E–mail: benettin@math.unipd.it).

[†]Università di Padova, Dipartimento di Matematica Pura e Applicata and GNFM (INDAM), Via G. Belzoni 7, 35131 Padova, Italy (E–mail: fasso@math.unipd.it).

[‡]Università di Padova, Dipartimento di Matematica Pura e Applicata and GNFM (INDAM), Via G. Belzoni 7, 35131 Padova, Italy. (E–mail: guzzo@math.unipd.it).

1 Introduction and statement of the result

1.1 The purpose of the paper

In this paper we are concerned with a classical problem, namely the stability of the proper rotations (rotations about a principal axis of inertia) of the perturbed Euler–Poinsot rigid body (the rigid body with a fixed point in a small positional potential), and prove a result of long term stability in the sense of Nekhoroshev [24]. The Hamiltonian has the form

$$\mathcal{H} = \mathcal{K} + \varepsilon \mathcal{V} \ , \tag{1.1}$$

where $\mathcal{K}$ is the kinetic energy; the potential $\mathcal{V}$ is assumed to depend only on the configuration of the body and to be regular, in fact analytic. The problem is equivalent[1] to the problem of the fast rotating rigid body, with angular velocity $\varepsilon^{-1/2}$, in a (non small) potential $\mathcal{V}$. We consider for simplicity a kinetically symmetric body, with moments of inertia $I_1 = I_2 \neq I_3 \neq 0$ (the potential needs not to be symmetric) and show that, for small ε, the proper rotations around the symmetry axis are stable for a time scale $\mathcal{O}(\exp 1/\varepsilon^a)$, $a > 0$. Stable means here that:

(i) Motions initially close to proper rotations remain close to proper rotations, the modulus of the angular momentum staying almost constant.

(ii) The unit vector μ parallel to the angular momentum precesses in space, up to a small noise, along the level lines of a (known) regular function on the unit sphere.

As indicated in Section 4, such a stability result extends to the triaxial case, too (of course concerning proper rotations about the axes of inertia with minimal or maximal moment of inertia). This paper is based on [4] and [5], where the perturbed Euler rigid body was studied in the framework of Nekhoroshev theory respectively for ordinary motions and for motions close to proper rotations about the symmetry axis; [5], in particular, provides some partial results which are the starting point of this paper, see below Proposition 0. Familiarity with [4] and [5], however, is useful but not necessary.

1.2 The framework

Let (O, e_x, e_y, e_z) be a frame fixed in space, with origin in the fixed point, and (O, e_1, e_2, e_3) be a body frame with the same origin, e_3 being the symmetry axis. The configuration of the body is represented by a matrix $R \in SO(3)$, say the matrix such that $e_1 = Re_x$ and so on. Let $m = (m_1, m_2, m_3)$ and $m^s = (m_x, m_y, m_z) = Rm$ denote the angular momentum of the body respectively in the proper and in the spatial frame. The phase space of the system $T^*SO(3)$ is known to be diffeomorphic to $\mathbb{R}^3 \times SO(3)$, and a dynamical state of the body is identified by a pair (m, R). As is well known, the unperturbed system with Hamiltonian $\mathcal{K}$ is "properly degenerate" (or "superintegrable"), namely it has three degrees of freedom but four constants of motion: the three components of m^s and either m_3 or equivalently $\mathcal{K}$. Correspondingly the phase space is foliated in low dimensional invariant tori $\mathbb{T}^2$, on which motions are quasi–periodic with only two independent frequencies. (See [14] and references therein for comments on the geometric structure of the phase space.) This is evident if one recalls the Euler–Poinsot description of the unperturbed

[1] Precisely, the motions of Hamiltonian (1.1), for given initial data and small ε, are identical, up to a rescaling, to the motions of the same Hamiltonian with $\varepsilon = 1$ and angular velocity rescaled by a factor $\varepsilon^{-1/2}$.

motions, according to which m^s stays fixed in space, while e_3 regularly precesses about it with a frequency Ω_1 depending on $G = \|m\|$ and $L = m_3$, and at the same time the body rotates about e_3, or equivalently m precesses in the body about e_3, with a second independent frequency $\Omega_2(G, L)$. The kinetic energy depends only on G and L,

$$\mathcal{K}(G, L) = \frac{1}{2I_1} G^2 + \frac{\beta}{2I_3} L^2 \ , \qquad \beta = \frac{I_1 - I_3}{I_1} \ ,$$

and $\Omega_1 = \frac{\partial \mathcal{K}}{\partial G}$, $\Omega_2 = \frac{\partial \mathcal{K}}{\partial L}$. Since $0 < I_3 \le 2I_1$, it is $-1 \le \beta < 1$. The value $\beta = -1$ corresponds to a flat body, and will require a special treatment. Besides $\beta = 1$, i.e. $I_3 = 0$, we also exclude the spherically symmetric case $\beta = 0$; quantitatively we assume a condition of the form

$$|\beta|(1 - \beta) \ge \beta_* > 0 \ . \tag{1.2}$$

Understanding the perturbed system means understanding what happens, for small $\varepsilon \ne 0$, to the constants of motion, in particular how m and m^s do move respectively in the body and in space.

For motions which are not close to proper rotations, more precisely for

$$m_\perp \ge \mathrm{const}\, \varepsilon^{1/4} \ , \qquad m_\perp^2 \stackrel{\mathrm{def}}{=} m^2 - m_3^2 \ ,$$

the situation was clarified in [4]. Assuming $\mathcal{V}$ to be analytic, it is there proved that:

(a) *For all initial data,* up to a time scale

$$|t| \le \mathrm{const}\ e^{(1/\varepsilon)^{1/4}} \ , \tag{1.3}$$

in the worst conditions (resonance between Ω_1 and Ω_2) it is

$$|G(t) - G(0)| \ , \ |L(t) - L(0)| \ \le \ \mathrm{const}\, \varepsilon^{1/2}$$

and the body performs an almost regular Euler–Poinsot precession, with almost constant Ω_1 and Ω_2, about the instantaneous direction of the angular momentum m^s, which in turn moves slowly with speed proportional to ε. So, for all initial data the behavior of m in the body is regular up to a small noise.

(b) *If the initial datum is nonresonant,* more precisely if

$$|\nu \cdot \Omega| \ge \mathrm{const}\, \varepsilon^{1/4} \qquad \forall \nu \in \mathbb{Z}^2 \ : \quad 0 < |\nu| \le \mathrm{const}\, \varepsilon^{-1/4} \ ,$$

where $\Omega = (\Omega_1, \Omega_2)$ and as usual

$$|\nu| = |\nu_1| + |\nu_2| \ ,$$

then for the same time scale (1.3) the unit vector μ parallel to m^s precesses slowly, up to a small noise, along the level lines of a regular function $\mathcal{F} : S^2 \to \mathbb{R}$ ($\mathcal{F}$ is constructed and turns out to be close to a suitable average of the potential $\mathcal{V}$). So, in the absence of resonances, m^s moves regularly in space, too, and the only effect of the perturbation is to remove the degeneracy by adding besides Ω_1 and Ω_2 a third small frequency.

The theory provided in [4] does not say anything on the long term behavior of μ for resonant initial conditions, but there is some theoretical evidence (based on the theory of adiabatic invariants [12] [21] [22] [11] [23]), also supported by numerical computations [3], that (for no matter how small ε) chaotic motions, in which μ fills unpredictably a wide region of S^2 in times larger than $1/\varepsilon$ but much shorter than (1.3), do exist.

1.3 The main result

In this paper (going beyond some partial results already proved in [5]) we prove that motions of the perturbed Euler rigid body which are close to proper rotations are substantially more regular than all other motions: even in presence of a resonance, the possibility of large chaotic movements of the angular momentum vector in space, discussed at the end of the previous subsection, is inhibited. In particular, the proper rotations about e_3 are long term stable.

The stability properties turn out to depend on the presence of resonances, and these in turn depend on the arithmetic properties of β. Following [5] we say that β is resonant, for a given ε and with reference to some constant ε_0, if there exists $\nu \in \mathbb{Z}^2$,

$$0 < |\nu| \leq N \ , \qquad N \stackrel{\text{def}}{=} (\varepsilon_0/\varepsilon)^{1/4} \ ,$$

such that

$$\left|\beta - \frac{\nu_1}{\nu_2}\right| < \frac{1}{2N|\nu|} \ ; \tag{1.4}$$

β is instead said to be nonresonant if (1.4) is not satisfied by any $\nu \in \mathbb{Z}^2$, $0 < |\nu| \leq N$. One easily checks that (1.4) is satisfied by at most one independent ν with $|\nu| \leq N$. According to (1.2) it is

$$1 \leq |\nu_1| \leq |\nu_2| \ , \tag{1.5}$$

so the lowest order resonance, characteristic of flat bodies, is $\nu = (-1, 1)$, while for all other resonances it is $|\nu_2| \geq 2$; for all of them, as will be crucial, $\nu_2 \neq 0$.

Proposition 1. *Let $\mathcal{V}$ be analytic, assume $G(0) \geq G_* > 0$ and*

$$m_\perp(0) < M_*\, \varepsilon^{1/4}$$

with some $M_* > 0$; let β satisfy (1.2).

There exist positive constants ε_0, ε_*, A, B, $B_\perp$, C, a, b, $b_\perp$, c and a regular function $\mathcal{F} : S^2 \to \mathbb{R}$ (constructed in the proof, and close to the average of the potential $\mathcal{V}$ along the proper rotations), such that for the time scale

$$|t| \leq A \exp\left(\frac{\varepsilon_*}{\varepsilon}\right)^a \tag{1.6}$$

it is

$$\begin{aligned} \Big|\|m(t)\| - \|m(0)\|\Big| \ &< \ B\,\varepsilon^b && (1.7)\\ |m_\perp(t) - m_\perp(0)| \ &< \ B_\perp\,\varepsilon^{b_\perp} && (1.8)\\ |\mathcal{F}(\mu(t)) - \mathcal{F}(\mu(0))| \ &< \ C\,\varepsilon^c \ , && (1.9) \end{aligned}$$

where as before $\mu = m^s/\|m\|$. Possible choices of the stability exponents are as follows:

(i) In the nonresonant case, $a = \frac{1}{4}$, $b = \frac{3}{4}$, $b_\perp = \frac{1}{2}$, $c = 1$.

(ii) In the resonant case, for all resonances $\nu \neq (-1, 1)$,

$$0 < a \leq \frac{1}{4} \ , \qquad b = \frac{1}{2} \ , \qquad b_\perp = \frac{1}{2}(1 - a) \ , \qquad c = \frac{1}{2} - a \tag{1.10}$$

(c improves up to $c = 1 - a$, if β gets closer to ν_1/ν_2).

(iii) For the resonance $\nu = (-1, 1)$,

$$0 < a < \frac{1}{4}\ , \qquad b = \frac{1}{2}\ , \qquad b_\perp = \frac{1}{3}(1-a)\ , \qquad c = \frac{1}{3}(1-4a)\ .$$

The nonresonant case is included here for completeness, but in fact is already proven in [5], see Section 3.1A. The notion of "average of the potential $\mathcal{V}$ along the proper rotations" is made precise in Section 2.1.

As far as we know, this is the first proof of long term stability of proper rotations of the perturbed Euler rigid body in a non symmetric potential. Such a result can be considered as a rigorous version, extended to a long time scale, of a known heuristic procedure, see for example [17] [28], in which one assumes *a priori* that m and e_3 stay parallel, and replaces the instantaneous torque by its average on a rotation, so deducing an approximate equation for $\dot{m}$; see for example [25] [26] [27]. Clearly, due to the possible presence of chaotic motion, the approximate equation turns out to be valid, at most, for short times. In fact, an essential point of our approach is that we include chaotic resonant motions, where averaging is not allowed (see Section 3.1C,D), and in spite of the presence of chaos it turns out that there exists an approximate constant of motion like $\mathcal{F}$, which persists for exponentially long times.

Proposition 1 applies to any potential $\mathcal{V}$, provided it is analytic as function of the configuration of the body (say analytic as function of the entries of the matrix $R \in SO(3)$ determining the orientation of the body in space); this is the case, in particular, of the gravitational potential. Some comments on the relevance of this paper for Celestial Mechanics are deferred to the Conclusions.

Remark. In presence of resonances, motions near proper rotations can be chaotic (the maximal Lyapunov exponent χ can be positive, $\chi \sim \varepsilon$, and is actually found to be positive in numerical computations [6]), but chaos is localized. In particular, at variance with the case of ordinary motions far from proper rotations, even in presence of resonances μ remains almost confined to a (predictable) curve on S^2. The technical element which is responsible of such extra stability in the vicinity of proper rotations is the presence of a second small quantity, besides ε, in the problem, namely the distance of the initial datum from the proper rotation. Such a smallness was not used in [5], where the stability of resonant motions was missed. Its use in our perturbation construction was suggested in [3].

1.4 Comments

For a better understanding of Proposition 1, it is convenient to comment here on the notion of resonance in the vicinity of proper rotations. For motions close to proper rotations it is natural to introduce, in place of L, the small quantity

$$\Lambda = G - L\ ,$$

related to $m_\perp$ by

$$m_\perp^2 = \Lambda(2G - \Lambda)\ , \tag{1.11}$$

and correspondingly to replace $\Omega = (\Omega_1, \Omega_2)$ by the linear combination $\omega = (\Omega_1 + \Omega_2, -\Omega_2)$. Denoting $K(G,\Lambda) = \mathcal{K}(G, G-\Lambda)$, one easily checks that

$$K(G,\Lambda) = \frac{1}{2I_3}(G^2 + \beta\Lambda^2 - 2\beta G\Lambda) \tag{1.12}$$

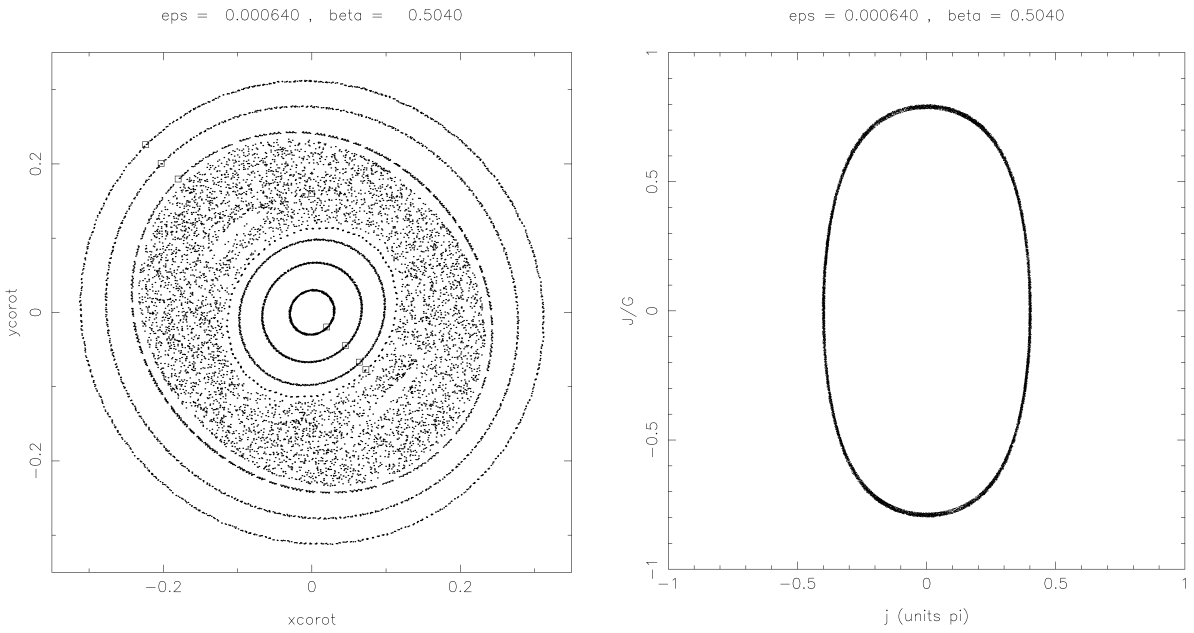


Figure 1: *The motion of m in the body (Poincaré section $m_z = 0$) and of μ on S^2 (projection). Parameters: $\varepsilon = 6.4 \times 10^{-4}$, $\beta = 0.504$, $I_1 = 1$. Initial data: $G = 1$, $\gamma = 0$, $m_z = 0$, $j = 0.4\pi$; $m_1^{(c)}, m_2^{(c)}$ ("xcorot" and "ycorot" in the plot) graphically chosen in the section.*

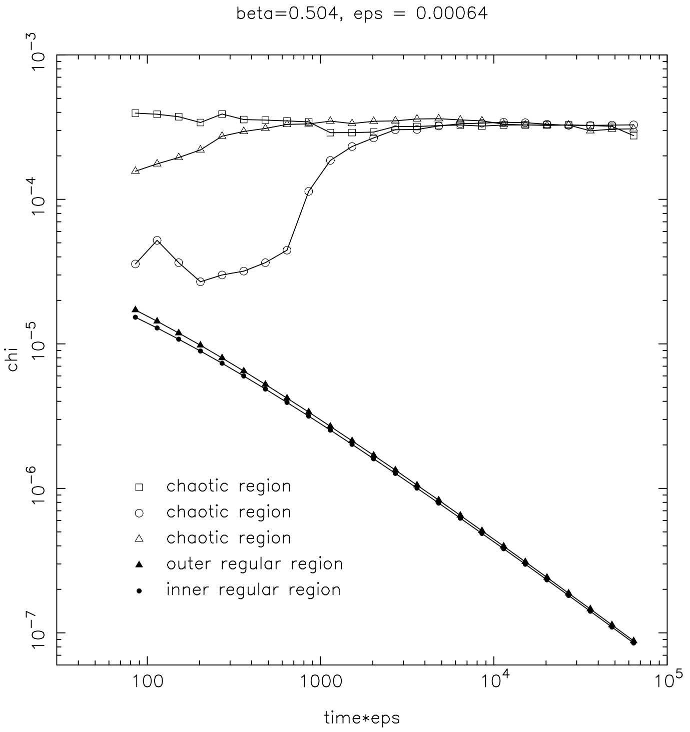


Figure 2: $\chi(x, \xi, t)$ *as function of t, for some tangent vector ξ and three initial data x in the chaotic region, two initial data respectively in the inner and in the outer regular region. Parameters and data as in figure 1.*

and

$$\omega(G,\Lambda)=\Big(\frac{\partial K}{\partial G},\frac{\partial K}{\partial \Lambda}\Big)=\frac{1}{I_3}(G-\beta\Lambda,-\beta(G-\Lambda))\ . \tag{1.13}$$

On the proper rotation it is

$$\omega(G,0)=\frac{G}{I_3}(1,-\beta)\ , \tag{1.14}$$

and the two components of ω are easily identified as the frequency of the proper rotation itself and the frequency of the small oscillations of m about e_3. If (1.4) holds, then $\nu\cdot\omega$ is small in the vicinity of the proper rotation (for $\beta=\nu_1/\nu_2$ it is $\nu\cdot\omega(G,0)=0$, i.e. the proper rotation is exactly resonant). This explains the choice of (1.4) as definition of resonance.

Accurate numerical results suggest that the exponents $b_\perp$ and c as estimated in this theory are optimal, with the exception of the special resonance $\nu=(-1,1)$ for which stability seems to be a little stronger than what predicted here (see Section 3.1.D for a comment on this point). The stability properties can be made evident numerically by following the motion of μ on the unit sphere and of m in the body. To represent μ, we use on S^2 the cylindrical coordinates $(m_z/G, j)$, $j=\arctan m_x/m_y$; concerning m, it is convenient to introduce, for each resonance ν, the "corotating" vector $m^{(c)}=(m_1^{(c)},m_2^{(c)})$ related to (m_1,m_2) by

$$\begin{pmatrix} m_1\\ m_2\end{pmatrix}=\begin{pmatrix}\cos\frac{\nu_1}{\nu_2}\gamma & -\sin\frac{\nu_1}{\nu_2}\gamma\\ \sin\frac{\nu_1}{\nu_2}\gamma & \cos\frac{\nu_1}{\nu_2}\gamma\end{pmatrix}\begin{pmatrix} m_1^{(c)}\\ m_2^{(c)}\end{pmatrix}\ ,$$

where γ is a suitable angle associated to the proper rotation, $\dot\gamma\simeq\omega_1$ (see the next section for the exact definition of γ).

Figure 1, left, shows a typical Poincarè section of $m^{(c)}$, corresponding to the section $m_z=0$; the right part of the figure represents instead the projection of μ on the unit sphere for the big chaotic orbit appearing in the section. The choice of the potential, as a function of the usual Euler angles ϑ,φ,ψ, is a simple one, namely

$$\mathcal{V}=\text{const}\ \sin\vartheta(\sin\vartheta\sin\varphi\sin\psi+\cos\varphi+\sin\vartheta\cos^2\psi)\ .$$

The resonance is $\nu=(1,2)$, with nearly resonant $\beta=0.504$ and $\varepsilon=6.4\times10^{-4}$. Figure 2 shows the behavior in time of

$$\hat\chi(x,\xi,t)=\frac{1}{t}\log\frac{\|D\Phi_x^t\xi\|}{\|\xi\|}\ \longrightarrow\ \chi(x)$$

for different initial data x and some tangent vector ξ (the choice of ξ as usual is irrelevant), Φ^t denoting the Hamiltonian flow. The three upper curves refer to initial data in the chaotic region, the two lower ones to initial data in the inner and in the outer regular regions. By varying ε, it turns out that both images in figure 1 scale exactly as theoretically predicted; in turn χ (in the chaotic region) decreases, as expected, proportionally to ε. Such figures illustrate the main content of Proposition 1: inside resonances, which are met for nearly rational β, motions are indeed chaotic, but chaos is localized and, up to a small noise, μ simply precesses along a line on the sphere. *Such a regular behavior is substantially different from the case of resonant motions far from proper rotations,* where chaotic motions of μ on S^2 are wide for no matter how small ε, and regularity is lost [4] [3]. A more complete numerical analysis of the problem will be published separately [6].

Acknowledgments: we are indebted to A.I. Neishtadt for very helpful discussions and to G. Gallavotti, editor of CMP, as well as to the referees, for valuable suggestions.

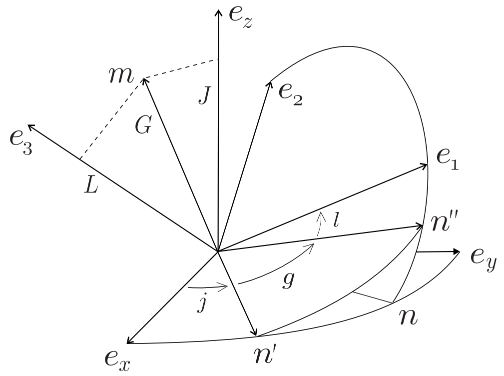


Figure 3: *The action–angle coordinates G, L, J, g, l, j for the symmetric Euler rigid body.*

2 Preliminaries for the proof

2.1 Action–angle and Birkhoff coordinates

The action–angle coordinates (G, L, J, g, l, j) of the symmetric Euler rigid body, known as Andoyer–Deprit coordinates [1] [10], are recalled in figure 3. Three of them are intrinsic (i.e. independent of the choice of the spatial frame), namely the already introduced quantities $G = \|m\|$ and $L = m_3$ entering the expression of K and the angle l, defined as the angle from the nodal line $n'' = m \times e_3$ to e_1 in the plane orthogonal to e_3. The remaining coordinates make instead reference to the spatial frame, and are $J = m_z$, $j =$ angle from e_x to the nodal line $n' = e_z \times m$ in the plane $e_x e_y$ (J/G and j are the already used coordinates of μ on S^2) and $g =$ angle from n' to n'' in the plane orthogonal to the angular momentum. (The angle g depends on the choice of the spatial frame very marginally, for the choice of the origin.) The domain of a chart of such coordinates excludes the intrinsic singularity $e_3 \times m = 0$ and the chart singularity $e_z \times m = 0$; two charts referring to spatial frames with different z–axis cover the subset

$$\mathcal{M}_0 = \{(m, R) \in T^*SO(3) : e_3 \times m \neq 0\} ,$$

that is the whole phase space but the proper rotations, providing an analytic symplectic atlas on it. In fact, the proper rotations are singular leaves of the foliation of $T^*SO(3)$ in tori $\mathbb{T}^2$: a family of invariant tori there shrinks to a circle, see figure 4. Both g and l cease to be defined on the proper rotation, but their sum $\gamma = g + l$ is, there coinciding with the familiar Euler angle $\psi =$ angle from $n = e_z \times e_3$ to e_1. "Less singular" canonical action–angle coordinates, more adapted to work around the proper rotation, are $(G, \gamma, \Lambda, \lambda, J, j)$, where

$$\Lambda = G - L \ , \qquad \lambda = -l$$

(see [5]). Indeed the singularity, now occurring at $\Lambda = 0$, is still present, but it affects only the pair of conjugated variables (Λ, λ), see figure 4. The singularity is removed [18] (see also [5]) by passing from polar–like to Cartesian coordinates

$$x_1 = \sqrt{2\Lambda} \cos \lambda \ , \qquad x_2 = \sqrt{2\Lambda} \sin \lambda \ .$$

In fact, as is proven in [18], both the rotation matrix R and the angular momentum m are analytic functions of $(G, \gamma, x_1, x_2, J, j)$ also for $x_1 = x_2 = 0$; conversely, x_1, x_2 are analytic functions of m, precisely

$$x_1 = \frac{\sqrt{2}m_2}{\sqrt{\|m\| + m_3}} \ , \qquad x_2 = -\frac{\sqrt{2}m_1}{\sqrt{\|m\| + m_3}} \ .$$

Two charts of coordinates $(G, \gamma, x_1, x_2, J, j)$, relative to different spatial frames (for example, frames obtained by a cyclic permutation of the axes), cover the subset

$$\mathcal{M}_+ = \{(m, R) \in T^*SO(3) : m_3 > 0\} \ .$$

In each chart it is possible to assure, for example, $J/G < 3/4$, as we shall later assume.

For technical reasons, in the proof of Proposition 1 it is convenient to use both the adapted action–angle coordinates $(G, \gamma, \Lambda, \lambda, J, j)$, outside a small neighborhood of the proper rotation so as to avoid the singularity, and the Cartesian coordinates $(G, \gamma, x_1, x_2, J, j)$ inside that neighborhood. In the practical work, the real Cartesian coordinates x_1, x_2 are replaced as usual by their complex linear combinations (Birkhoff coordinates)

$$z_1 = \frac{x_1 - ix_2}{i\sqrt{2}} \ , \qquad z_2 = \frac{x_1 + ix_2}{\sqrt{2}}$$

and it is

$$\Lambda(z_1, z_2) = iz_1z_2 \ .$$

Concerning the Hamiltonian, in the adapted action–angle variables it has the form

$$H(G, \gamma, \Lambda, \lambda, J, j) = K(G, \Lambda) + \varepsilon V(G, \gamma, \Lambda, \lambda, J, j) \tag{2.1}$$

with K as in (1.12); in the Birkhoff coordinates it assumes instead the form

$$h(G, \gamma, z_1, z_2, J, j) = K(G, \Lambda(z_1, z_2)) + \varepsilon v(G, \gamma, z_1, z_2, J, j) \ . \tag{2.2}$$

Analytic perturbation theory deals with real analytic functions, namely functions which are analytic and bounded in suitable complex neighborhoods of their real domain and are real for real variables. With reference to the Birkhoff coordinates, we shall call real any domain in which G, γ, J, j are real and x_1, x_2 are real, i.e. z_2, iz_1 are complex conjugate (correspondingly, we shall call real analytic any function which is analytic and bounded in some neighborhood of such a real domain, and is real in the real domain).

As is well known, perturbation theory makes often use of Fourier series expansions. It is worthwhile to recall that in the Birkhoff coordinates the Fourier series of any function $f(G, \gamma, z_1, z_2, J, j)$ is written

$$f(G, \gamma, z_1, z_2, J, j) = \sum_{\nu \in \mathbb{Z}^2} f_\nu(G, \gamma, \Lambda(z_1, z_2), J, j) E_\nu(\gamma, z_1, z_2)$$

with

$$E_\nu(\gamma, z_1, z_2) = \begin{cases} e^{i\nu_1\gamma} \, z_1^{\nu_2} & \text{if } \nu_2 \geq 0 \\ e^{i\nu_1\gamma} \, z_2^{|\nu_2|} & \text{if } \nu_2 < 0 \end{cases}$$

(see [5] for comments). In particular, $v_0(G, \gamma, \Lambda, J, j)$ is the average of the potential v in the Birkhoff variables, and (out of the proper rotation, when both functions are defined)

$$v_0(G, z_1, z_2, J, j) = V_0(G, iz_1z_2, J, j) \ ,$$

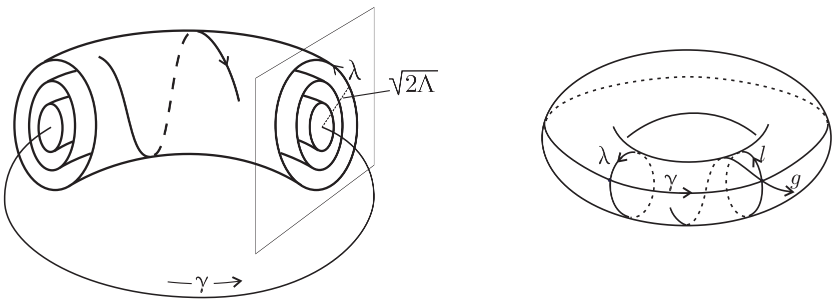


Figure 4: *The change of coordinates from* (G, L, g, l) *to* $(G, \Lambda, \gamma, \lambda)$.

V_0 denoting the usual average of V on γ, λ,

$$V_0(G, \Lambda, J, j) = (2\pi)^{-2} \int_0^{2\pi} \int_0^{2\pi} V(G, \gamma, \Lambda, \lambda, J, j)\, \mathrm{d}\gamma\, \mathrm{d}\lambda\ .$$

The Fourier expansion allows us to define, in any set of variables, the "Fourier norm"

$$\|f\| = \sum_{\nu \in \mathbb{Z}^2} \|f_\nu\|^\infty\ ,$$

which is well known to be useful in perturbation constructions. (The Fourier norm $\|\,.\,\|$ reduces to the sup–norm $\|\,.\,\|^\infty$ for functions independent of the angles). Whenever we say that a function is bounded in some domain, we refer to the Fourier norm. Functions with zero average are necessarily small for small Λ, and this will play a crucial role in the proof of Proposition 1.

Remark. As remarked above, the singularity $m \times e_3 = 0$ is removed by using, near proper rotations, the Cartesian–like variables x_1, x_2 (or z_1, z_2) in place of the polar ones (Λ, λ). However, even restricting to a neighborhood of proper rotations (with m and e_3 almost parallel, but possibly pointing in any direction of the space), we are still confronted with the chart singularities $e_z \times m = 0$, i.e. $J = \pm G$. Our aim is to prove regularity of all motions, so our perturbation treatment must cover the whole sphere $\|m\| = G$, and this requires at least two sets of coordinates. A simple choice [4] [5] is to use two sets of coordinates (J, j) relative to two spatial frames (e_x, e_y, e_z) and (e'_x, e'_y, e'_z), the latter frame being a permutation of the former so as the two z–axes are orthogonal; in this way we can define the domain of each chart by requiring, for example, $|J| < \frac{3}{4}G$, and nevertheless cover the whole phase space we are interested in, actually cover $\mathcal{M}_+$. If, as we assume in Proposition 1, the potential $\mathcal{V}$ is analytic as a function of $R \in SO(3)$, then such charts provide an analytic atlas for $\mathcal{M}_+$. Other choices of coordinates are possible: for example, for potentials for which the z-axis plays a special role, instead of changing the axis it is probably convenient to use, around the poles, Cartesian–like coordinates (defined similarly to x_1, x_2, the components m_x, m_y, m_z of m^s replacing m_1, m_2, m_3; see for example [16]). All choices are in principle equivalent. In general, however, it is not possible to guarantee *a priori* that motions stay inside a single chart,[2] so the problem arises of *how to develop perturbation theory, and discuss about long term stability, in presence of motions which possibly pass from one chart domain to the other.* Such a question is raised and solved

[2]This is evidently the case of the "wide chaotic motions" studied in [4] and [3], where $\mu = m^s/\|m\|$ wanders unpredictably on the unit sphere. But even for regular motions as those found in this paper, requiring that a motion stays inside a single chart domain would require making a choice of the chart depending on the initial datum; this introduces annoying complications in the perturbation work, with no real advantage.

in [13]: provided some elementary care is taken (like keeping the same ultraviolet cut–off for all charts), the perturbation constructions in the different chart domains glue smoothly. In particular the local normal forms, as constructed in the different chart domains, are proven to be the local representatives of a global normal form defined in the union of the chart domains. So, in spite of the presence of chart singularities, local normal forms can be used to provide long term confinement. This idea, after [13], has been succesfully used in connection with the rigid body in [4] and [5]. Conceptually, the prescription is elementary: (i) make any choice of the charts (possibly adapted, if useful, to the potential at hand); write the perturbing potential in each chart, and determine the relevant constants describing it (its norm, its analyticity radii), taking the worst among charts. (ii) Construct the normal form with exponentially small remainder in each chart, and do not pose the question whether motions stay inside the domain of a chart (as might happen with particular potentials, or for a good selection of initial data) or instead wander from one chart to the other: the question is indeed irrelevant; the worst of the local estimates will hold globally. For details and further comments we refer to [13], or to [4], [5]. In the rest of the paper, the problem of chart singularities will be ignored.

2.2 Previous results

Here we recall in a proposition some results proved in [5], which are the starting point of the proof of Proposition 1. We take the liberty to slightly reformulate some statements so as to adapt them to our needs, as well as to include in the proposition some useful estimates which are present only in the proofs, or can be trivially deduced from estimates contained in the proofs. (The main change is the passage from the framework of large angular velocity at $\varepsilon = 1$, which was chosen in [5], to the framework of bounded angular velocity and small ε, that we found convenient to use here.)

The statement below concerns a Hamiltonian of the form (2.2), in a convenient complex neighborhood of a real domain $\mathcal{D}$ of the form

$$G \geq G_* > 0\ , \qquad \Lambda(z_1, z_2) \leq \Lambda_* \varepsilon^{1/2}\ , \qquad |J/G| \leq 3/4\ , \qquad \gamma, j \in S^1\ ;$$

such a neighborhood, denoted $\mathcal{D}_\rho$, depends on an "extension vector" with positive entries $\rho = (\rho_G, \varepsilon^{1/2}\rho_z, \rho_j)$, $\rho_z \leq \Lambda_*^{1/2}$, and is defined by

$$|\Im G|,\ |\Im J| \leq \rho_G\ , \qquad |z_1|,\ |z_2| \leq \rho_z \varepsilon^{1/4}\ , \qquad |\Im \gamma|,\ |\Im j| \leq \rho_j \leq 1\ .$$

The notion of resonance entering the statement is the one introduced in Section 3.1; in the resonant case, to label a resonance, we use (here and in the following) $\nu = (\nu_1, \nu_2)$ such that ν_1, ν_2 have no common divisors and, on account of (1.5), $\nu_2 > 0$.

Proposition 0. *Consider the Hamiltonian (2.2) and assume v is real analytic in $\mathcal{D}_\rho$ defined as above. Let ε be small. There exist $\varepsilon_0, \varepsilon_* > 0$ such that, with reference to the notion (1.4) of resonance (where ε_0 enters):*

(i) In the nonresonant case, there exists an $\varepsilon^{3/4}$–near to the identity canonical transformation

$$(G, \gamma, z_1, z_2, J, j) = \mathcal{C}(G', \gamma', z_1', z_2', J', j')\ , \qquad \mathcal{C} : \mathcal{D}_{\rho/2} \to \mathcal{D}_\rho\ , \tag{2.3}$$

such that the new Hamiltonian $h' = h \circ \mathcal{C}$ is in nonresonant normal form up to an exponentially small remainder, namely

$$\begin{aligned} h'(G', \gamma', z_1', z_2', J', j') \;&=\; K(G', \Lambda') + \varepsilon u(G', \Lambda', J', j') \\ &\quad + \varepsilon\, e^{-(\tilde{\varepsilon}/\varepsilon)^{1/4}} \mathcal{R}(G', \gamma', z_1', z_2', J', j')\ , \end{aligned}$$

where Λ' stands for $\Lambda(z_1', z_2') = i z_1' z_2'$ and u, $\mathcal{R}$ are bounded by quantities independent of ε; u is $\varepsilon^{1/4}$–close to the average v_0 of v.

(ii) In the resonant case, there exists an $\varepsilon^{3/4}$–near to the identity canonical transformation of the form (2.3), which gives the new Hamiltonian h' the ν–resonant normal form

$$\begin{aligned} h'(G',\gamma',z_1',z_2',J',j') &= K(G',\Lambda') + \varepsilon\Big[u(G',\Lambda',J',j') \\ &+ f(G',\gamma',z_1',z_2',J',j')\Big] + \varepsilon\, e^{-(\tilde\varepsilon/\varepsilon)^{1/4}}\mathcal{R}(G',\gamma',z_1',z_2',J',j')\ , \end{aligned} \tag{2.4}$$

where

$$f(G,z_1,z_2,J,j) = \sum_{0\neq n\in\mathbb{Z}} f_n(G,\Lambda,J,j)E_{n\nu}(z_1,z_2,\gamma)\ ; \tag{2.5}$$

u, f and $\mathcal{R}$ are bounded by quantities independent of ε, u is $\varepsilon^{1/4}$–close to v_0, f is as small as $\max(|z_1|^{\nu_2}, |z_2|^{\nu_2})$.

The normalized Hamiltonian (2.4) can be written in action–angle coordinates, giving

$$\begin{aligned} H'(G',\gamma',\Lambda',\lambda',J',j') &= K(G',\Lambda') + \varepsilon\Big[U(G',\gamma',\Lambda',J',j') \\ &+ F(G',\gamma',\Lambda',\lambda',J',j')\Big] + \varepsilon\, e^{-(\tilde\varepsilon/\varepsilon)^{1/4}}\mathcal{R}(G',\gamma',\Lambda',\lambda',J',j')\ , \end{aligned} \tag{2.6}$$

with

$$F(G,\gamma,\Lambda,\lambda,J,j) = \sum_{0\neq n\in\mathbb{Z}} F_n(G,\Lambda,J,j)\, e^{in(\nu_1\gamma+\nu_2\lambda)}\ ,$$

each F_n being divisible by $\Lambda^{\frac12|n\nu_2|}$. All terms U, F and $\mathcal{R}$ are bounded, U is $\varepsilon^{1/4}$–close to the average V_0 of the perturbation and F is divisible by $\Lambda^{\nu_2/2}$. As stressed in Section 1.3, it is always $\nu_2 \neq 0$, so F as well as f in (2.5) are small.

2.3 Corotating coordinates

In the resonant case it is convenient to pass from the coordinates $(G',\gamma',z_1',z_2',J',j')$ to the "corotating coordinates" $(\Gamma,\gamma',w_1,w_2,J',j')$, where

$$w_1 = e^{i\gamma\nu_1/\nu_2} z_1'\ , \qquad w_2 = e^{-i\gamma\nu_1/\nu_2} z_2'\ , \qquad \Gamma = G - \frac{\nu_1}{\nu_2}\Lambda'\ . \tag{2.7}$$

In such coordinates the Hamiltonian (2.4) takes the form (dropping all primes)

$$\begin{aligned} h^{(c)}(\Gamma,\gamma,w_1,w_2,J,j) &= K^{(c)}(\Gamma,\Lambda) + \varepsilon\Big[u^{(c)}(\Gamma,\Lambda,J,j) \\ &+ f^{(c)}(\Gamma,w_1,w_2,J,j)\Big] + e^{-(\tilde\varepsilon/\varepsilon)^{1/4}}\,\mathcal{R}(\Gamma,\gamma,w_1,w_2,J,j)\ , \end{aligned} \tag{2.8}$$

with $\Lambda = i w_1 w_2$ and

$$K^{(c)}(\Gamma,\Lambda) = \frac{1}{2I_3}(\Gamma^2 + M\,\Lambda^2 - 2\delta\Gamma\Lambda)\ ,$$

having denoted

$$\delta = \beta - \frac{\nu_1}{\nu_2}\ , \qquad M = \Big(\frac{\nu_1}{\nu_2} + \delta\Big)\Big(1 - \frac{\nu_1}{\nu_2}\Big)\ ;$$

δ is a small quantity and thanks to (1.2) M is bounded away from zero. The perturbing term $f^{(c)}$ is analytic in a neighborhood of the proper rotation and is small proportionally to $\Lambda^{\nu_2/2}$. The advantage of this change of variables is that γ disappears from the normal part of the Hamiltonian and, as far as the remainder can be neglected, the Hamiltonian reduces to two degrees of freedom, Γ being a parameter. The slow component of the frequency is now

$$\omega^{(c)} = \frac{\partial K^{(c)}}{\partial \Lambda} = M\Lambda - \delta\Gamma$$

and vanishes for

$$\Lambda_{\rm res}(\Gamma) = \frac{\delta}{M}\,\Gamma \ . \tag{2.9}$$

(If $\Lambda_{\rm res}$ is negative, $\omega^{(c)}$ never vanishes, nevertheless if $\Lambda_{\rm res}$ is close to zero $\omega^{(c)}$ is small near $\Lambda = 0$, and this must be taken into account.) Using $\Lambda_{\rm res}$ the corotating kinetic energy assumes the useful expression

$$K^{(c)} = \frac{1}{2I_3}\Big[(1-\delta^2/M)\Gamma^2 + M(\Lambda - \Lambda_{\rm res}(\Gamma))^2\Big] \ . \tag{2.10}$$

The action–angle coordinates corresponding to the corotating coordinates $(\Gamma, \gamma, w_1, w_2, J, j)$ are $(\Gamma, \gamma, \Lambda, \sigma, J, j)$, with

$$\sigma = \frac{\nu_1}{\nu_2}\gamma + \lambda \ ;$$

in such coordinates the Hamiltonian gets the form

$$\begin{aligned} H^{(c)}(\Gamma,\gamma,\Lambda,\sigma,J,j) \ &= \ K^{(c)}(\Gamma,\Lambda) + \varepsilon\Big[U^{(c)}(\Gamma,\Lambda,J,j) \\ &+ \ F^{(c)}(\Gamma,\Lambda,\sigma,J,j)\Big] + e^{-(\tilde\varepsilon/\varepsilon)^{1/4}}\,\mathcal{R}(\Gamma,\gamma,w_1,w_2,J,j) \ . \end{aligned} \tag{2.11}$$

Now $F^{(c)}$ depends only on σ, and is small proportionally to $\Lambda^{\nu_2/2}$.

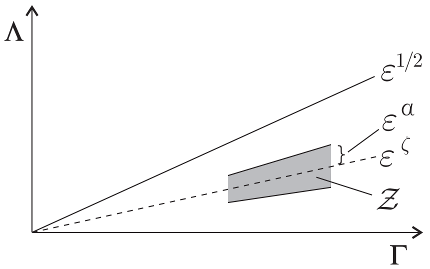
Remark. Let us better explain the change of variables $(\gamma, \lambda) \mapsto (\gamma, \sigma)$, which might surprise since rational rather than integer coefficients are used. (Equivalent considerations hold for the corresponding change in Cartesian variables.) The procedure, in more detail, is as follows:

i) We preliminarily pass, for the angle γ, from the standard circle $S^1 = \mathbb{R}/(2\pi)$ to its covering $S^1_{\nu_2} = \mathbb{R}/(2\nu_2\pi)$. If for clarity we use $\overline{\gamma}$ for the variable on the covering, the ν_2–to–one correspondence is $\gamma = \overline{\gamma}$ mod 2π. The domain of all other variables does not change, the Hamiltonian is the same. Motions in the covering project trivially on motions in the original domain.

ii) In the nonstandard torus $S^1_{\nu_2} \times S^1$ we introduce the smooth one–to–one change of variables $(\overline{\gamma}, \lambda) \mapsto (\overline{\gamma}, \sigma)$ with $\sigma = \frac{\nu_1}{\nu_2}\overline{\gamma} + \lambda$; the extension to the actions is $(G, \Lambda) \mapsto (\Gamma, \Lambda)$ with $\Gamma = G - \frac{\nu_1}{\nu_2}\Lambda$. The resulting Hamiltonian $H^{(c)}$ is periodic of period 2π in σ and of period $2\nu_2\pi$ in $\overline{\gamma}$ (the overline has been there dropped). As is relevant, only σ appears in the normal part of $H^{(c)}$, on which we still need to work, while $\overline{\gamma}$ enters only the already exponentially small remainder.

iii) In the perturbative construction that we are going to start, we average on σ, never on $\overline{\gamma}$, so the uncommon periodicity of $\overline{\gamma}$ is never exploited.[3] As usual, we introduce near–the–identity

[3]More precisely, the canonical transformations we shall construct have the following form: Γ stays constant, the new coordinates $\Lambda', \sigma', J', j'$ are near–the–identity functions of Λ, σ, J, j and Γ (independent of $\overline{\gamma}$), $\overline{\gamma}$ changes a little so as the transformation is canonical.



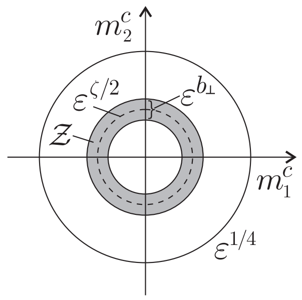


Figure 5: *The resonant zone* $\mathcal{Z}$ *(left) and the corresponding region in the corotating plane* $m_1^{(c)} m_2^{(c)}$ *(right).*

canonical transformations to produce normal forms. Using normal forms we infer confinement of the system in the covering, and confinement goes back to the original domain.

Practically, the presence of noninteger coefficients, and the fact that the periodicity of $H^{(c)}$ on γ is $2\nu_2\pi$, can be safely ignored. (For a similar procedure, see for example [19]; for more general considerations on this subject see [2].)

Remark. It is natural to introduce the real counterpart $x_1^{(c)}, x_2^{(c)}$ of the Birkhoff corotating coordinates,

$$\begin{pmatrix} x_1^{(c)} \\ x_2^{(c)} \end{pmatrix} = \begin{pmatrix} \cos\frac{\nu_1}{\nu_2}\gamma & -\sin\frac{\nu_1}{\nu_2}\gamma \\ \sin\frac{\nu_1}{\nu_2}\gamma & \cos\frac{\nu_1}{\nu_2}\gamma \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \ ;$$

such coordinates are indeed close to the coordinates $(m_1^{(c)}, m_2^{(c)})$ appearing in figure 1, more precisely it is $(m_1^{(c)}, m_2^{(c)}) = (G - \frac{1}{2}\Lambda)^{1/2}(x_2^{(c)}, -x_1^{(c)})$.

3 Proof of Proposition 1

3.1 The proof, up to a technical point.

A. The nonresonant case. In the nonresonant case, part (ii) of Proposition 0 already implies Proposition 1: indeed, G' and Λ' stay obviously almost constant (oscillations are much smaller than ε) for an exponentially long time scale as in (1.6), while G, Λ remain ε–close to them. On the other hand, thanks to the overall energy conservation and to the almost constancy of $K(G', \Lambda')$, the normal term $u(G', \Lambda', J', j')$ also stays almost constant. So, (J', j') stays almost exactly on a level line of

$$\mathcal{F}(J', j') = u(G'(0), \Lambda'(0), J', j') \ ,$$

and in turn (J, j) stays ε–close to (J', j'). The motion is regular, the Lyapunov exponent is as small as $\exp -(\varepsilon_*/\varepsilon)^{(1/4)}$ with $\varepsilon_* = \tilde{\varepsilon}$; the values of the stability exponents appearing in Proposition 1 are immediately recovered.

B. The resonant case $\nu \neq (-1, 1)$*: regular motions.* So, to prove Proposition 1 it suffices to consider the (much less trivial) resonant case. Here we restrict ourselves to $\nu \neq (-1, 1)$; below in paragraph D we consider that special resonance. The proof proceeds as follows: we make any choice of the exponent a in the range specified in (1.6), and consider the resonant value $\Lambda_{\rm res}$ given by (2.9); we take care of a (possibly empty) "resonant zone" $\mathcal{Z}$ surrounding $\Lambda_{\rm res}$, of thickness depending on $\Lambda_{\rm res}$: precisely, we define ζ and α by

$$\Lambda_*\varepsilon^\zeta = \max(\Lambda_*\varepsilon^{1-a}, \Lambda_{\rm res}) \ , \qquad \alpha = \frac{1}{2}(\zeta + 1 - a) \tag{3.1}$$

and then $\mathcal{Z}$ by

$$|\Lambda - \Lambda_{\rm res}| \le \Lambda_*\varepsilon^\alpha \ . \tag{3.2}$$

($\mathcal{Z}$ is empty if $\Lambda_{\rm res} < -\Lambda_*\varepsilon^{1-a}$). Clearly

$$\zeta \le \alpha \le 1 - a \ .$$

Figure 5 provides a sketch of the situation for the case $\Lambda_{\rm res} > \Lambda_*\varepsilon^{1-a}$.

In the complement of $\mathcal{Z}$, more precisely in the slightly larger set $\bar{\mathcal{Z}}$ such that

$$|\Lambda - \Lambda_{\rm res}| \ge \frac{3}{4}\Lambda_*\varepsilon^\alpha \ ,$$

we are able to perform on our Hamiltonians (2.4) or (2.6) a further normalization procedure[4] which eliminates all harmonics $n\nu \ne 0$ entering the perturbation up to an exponentially small remainder. More precisely, we prove that

There exist a constant ε_* and a function Λ' defined in $\bar{\mathcal{Z}}$, which is close to Λ:

$$|\Lambda' - \Lambda| < {\rm const}\,\varepsilon^{(1+\zeta+a)/2} \ ,$$

and stays almost constant:

$$|\dot{\Lambda}'| < {\rm const}\ \varepsilon\, e^{-(\varepsilon_*/\varepsilon)^a} \ . \tag{3.3}$$

This is the only part of the proof requiring perturbation techniques, and is deferred to the next subsection (as is easily imagined, Λ' is the new coordinate Λ of the near to the identity canonical transformation which gives the Hamiltonian the appropriate normal form). The consequence of (3.3) is that for initial data out of $\mathcal{Z}$, and up to the time scale (1.6), the motion is as regular as in the nonresonant case: Λ and Γ stay almost constant, while (J, j) closely follows a level line of $U^{(c)}$; the Lyapunov exponent is negligibly small. In particular, during this time scale, motions starting out of $\mathcal{Z}$ cannot significantly enter it, and conversely, motions starting in $\mathcal{Z}$ cannot significantly escape.

C. The resonant case $\nu \ne (-1, 1)$: bounded chaotic motions. Once such a work is done, the proof of Proposition 1 is practically over: indeed, *to prove the stability of proper rotations as claimed in the proposition, it is not necessary to do any work inside $\mathcal{Z}$.* Motions inside $\mathcal{Z}$ can be chaotic (and are found numerically to be chaotic, see figure 2), but chaos is well localized, and both (J, j) and $m_\perp$ are confined as specified in Proposition 1. Concerning (J, j), the confinement follows from

[4] In principle it would be possible, and simpler, to work only on Hamiltonian (2.4), in Cartesian–like variables. * But as it will be clear in the proof (see in particular the remark at the end of Section 3), in this way the results worsen a little. As a matter of fact, the best values of the stability exponents are found by using Cartesian coordinates in a specially small neighborhood of the proper rotations, and polar coordinates in a larger neighborhood, namely when $\Lambda \ge {\rm const}\,\varepsilon^\alpha$. A similar situation occurred in [15]. **

the energy conservation applied to Hamiltonian (2.11). Indeed, since Γ stays almost constant, the oscillations of $K^{(c)}$, see (2.10), are bounded by $\mathrm{const}\,\varepsilon^{2\alpha}$, while the oscillations of $F^{(c)}$, which is as small as $\Lambda^{\nu_2/2}$, for Λ near $\Lambda_{\rm res} \le \Lambda_*\varepsilon^\zeta$ are bounded at least by $\mathrm{const}\,\varepsilon^\zeta$. As a consequence, the oscillations of $U^{(c)}$ are bounded by

$$\mathrm{const}\,\varepsilon^c \ , \qquad c = \min(\zeta, 2\alpha-1) = 2\alpha - 1 = \zeta - a \ .$$

This gives $c = \frac{1}{2} - a$ for $\Lambda_{\rm res}$ in the bulk of the domain $0 \le \Lambda \le \Lambda_*\varepsilon^{1/2}$, and $c = 1 - 2a$ for $\Lambda_{\rm res}$ very near the origin, i.e. when δ is small and the proper rotation is almost exactly resonant. This is in agreement with Proposition 1, the stability exponent c being as in (1.8).

Finally, to determine the exponent $b_\perp$ appearing in (1.8), we must compute the maximal oscillation $\Delta m_\perp$ of $m_\perp$, which according to (1.11) is bounded by

$$\Delta m_\perp \le \mathrm{const}\left(\sqrt{\Lambda_{\rm res} + \Lambda^*\varepsilon^\alpha} - \sqrt{\Lambda_{\rm res} + \Lambda^*\varepsilon^\alpha}\,\right) \ .$$

If $\Lambda_{\rm res} \le \Lambda_*\varepsilon^{1-a}$, then $\alpha = \zeta = 1 - a$, and so

$$\Delta m_\perp \le \mathrm{const}\,\sqrt{\Lambda_*}\,\varepsilon^{(1-a)/2} \ , \tag{3.4}$$

in agreement with (1.8). In the opposite case it is $\Lambda_{\rm res} = \Lambda_*\varepsilon^\zeta$, and so

$$\Delta m_\perp = \mathrm{const}\,\sqrt{\Lambda^*}\varepsilon^{\zeta/2}\left(\sqrt{1+\varepsilon^{\alpha-\zeta}} - \sqrt{1-\varepsilon^{\alpha-\zeta}}\,\right) \ ;$$

from this expression (3.4) follows as well.

D. The resonance $\nu = (-1, 1)$. As remarked above, the resonance $\nu = (-1, 1)$ is somewhat special. A first reason is that the perturbation is larger, $F^{(c)}$ being small only as $\Lambda^{1/2}$. A second reason is that now, recalling $\beta \ge -1$, it is

$$\delta = \beta + 1 \ge 0 \ , \qquad \Lambda_{\rm res} = -\frac{\delta}{2(1-\delta)}\,\Gamma \le 0 \ ,$$

so the resonance either is exactly on the proper rotation ($\delta = 0$) or is out of the phase space ($\delta > 0$).

We proceed as follows: to take care of the worst situation $\delta = 0$, we exclude a resonant zone $\mathcal{Z}$ defined now by

$$\Lambda \le \Lambda_*\varepsilon^\alpha \ , \qquad \alpha = \frac{2}{3}(1-a) \ , \qquad 0 < a < \frac{1}{4} \ . \tag{3.5}$$

In the complement of $\mathcal{Z}$ we perform as above a normalization procedure, see the next subsection, leading to (3.3); in this way we prove that the motion outside $\mathcal{Z}$ is regular. For motions inside $\mathcal{Z}$, we use as above the energy conservation: since inside $\mathcal{Z}$ the kinetic part oscillates at most by $\mathrm{const}\,\varepsilon^{2\alpha}$, while the oscillations of $F^{(c)}$ are bounded by $\mathrm{const}\,\varepsilon^{\alpha/2}$, the oscillations of U are necessarily bounded by $\mathrm{const}\,\varepsilon^c$, with

$$c = \min\left(\frac{\alpha}{2}, 2\alpha - 1\right) = 2\alpha - 1 = \frac{1}{3}(1-4a) \ .$$

The corresponding exponent $b_\perp$, determining the maximal oscillations of $m_\perp$, is obviously

$$b_\perp = \frac{1}{2}\alpha = \frac{1}{3}(1-a) \ .$$

This concludes the proof of Proposition 1, up to the proof of (3.3).

Remark. The stability of the proper rotation about e_3 in the resonant case $\nu = (-1, 1)$ is in fact slightly better than what we proved here. Indeed, for this special resonance the perturbation has a quite peculiar form which should be taken into consideration. In essence, it turns out that it is possible to reduce the size of the resonant zone by taking α as close as one wishes to $4/5$, in place of $2/3$, the exponent a remaining positive. However, the perturbation scheme leading to (3.3) in the larger domain needs relevant changes (and becomes essentially more complicated). Work is in progress on this point in collaboration with A.I. Neishtadt.

Remark. The necessity of improving $b_\perp$ and c in the resonance $\nu = (-1, 1)$ is evident in the numerical computations mentioned in Section 1.4. On the contrary, for $\nu \neq (-1, 1)$ the exponents seem to be optimal.

3.2 The proof of (3.3)

This is a rather standard perturbation proof, with the only peculiarity that most small quantities — norms of functions, small divisors — are small with Λ: so we must work locally in this variable in a convenient neighborhood of some value $\Lambda_0 \leq \varepsilon^{1/2}\Lambda_*$. We start from the Hamiltonian (2.4) or (2.6), but provisionally forget the small remainder, so as to deal with a system with two degrees of freedom in which Γ is a parameter; at the end of the proof it will be evident how to extend the result to the complete system. We begin with the case of the resonance $\nu = (-1, 1)$, which is easier because the domain has a simple form and moreover, a neighborhood of $\Lambda = 0$ being already excluded, we can safely use the action–angle variables, with no need of the Birkhoff ones.

A. The case $\nu = (-1, 1)$. So, we consider the Hamiltonian (2.6), we cut off the remainder and (with little abuse) denote the resulting function by

$$H(\Lambda, \sigma, J, j) = K(\Lambda) + \varepsilon[U(\Lambda, J, j) + F(\Lambda, \sigma, J, j)] \ , \tag{3.6}$$

with

$$K(\Lambda) = \frac{\Lambda^2}{2I} \ , \qquad I \stackrel{\text{def}}{=} \frac{I_3}{M} = -\frac{I_3}{2(1-\delta)}$$

and

$$F(\Lambda, \sigma, J, j) = \sum_{0 \neq n \in \mathbb{Z}} F_n(\Lambda, J, j)\, e^{in\sigma} \ ; \tag{3.7}$$

F is divisible by $\Lambda^{1/2}$. We consider any Λ_0 such that $\varepsilon^\alpha \Lambda_* \leq \Lambda_0 \leq \varepsilon^{1/2}\Lambda_*$, and introduce the real domain D:

$$\Lambda = \Lambda_0 \ , \qquad \sigma \in S^1 \ , \qquad |J| \leq \frac{3}{4} \ , \qquad j \in S^1 \ .$$

We denote

$$r = \Lambda_0/\Lambda_* \ , \qquad \varepsilon^\alpha \leq r \leq \varepsilon^{1/2} \ ,$$

and consider an extension vector (with positive entries)

$$\rho = (r\rho_\Lambda, \rho_\sigma, \rho_J, \rho j) \ , \qquad \rho_\Lambda = \frac{1}{2}\Lambda_* \ .$$

The extended domain D_ρ is defined (as usual) as the union of the polydisks of radii $r\rho_\Lambda$, ρ_σ, ρ_J, ρ_j, centered in all points of D. We shall assume for simplicity $\rho_\sigma \leq 1$. With no possibility of confusion, we indicate by ρ also the positive number defined by

$$\rho^2 = \max(\rho_\Lambda \rho_\sigma, \rho_J \rho_j) \ ,$$

and denote, as it turns out to be practical, $\Omega = \rho^2/I$. For any extension $\rho' \le \rho$ (the inequality is intended to work on all components), we use the Fourier norm in $D_{\rho'}$, denoted $\|\,.\,\|_{\rho'}$. Norms of derivatives of functions are useful as well, and we find convenient to use the compact notation

$$\|\,.\,\|^{\partial}_{\rho'} = \max\Big(\|\,.\,\|_{\rho'}, r\rho_\Lambda \Big\|\frac{\partial\,.}{\partial\Lambda}\Big\|_{\rho'},\ \rho_\sigma \Big\|\frac{\partial\,.}{\partial\sigma}\Big\|_{\rho'},\ \rho_J \Big\|\frac{\partial\,.}{\partial J}\Big\|_{\rho'},\ \rho_j \Big\|\frac{\partial\,.}{\partial j}\Big\|_{\rho'}\Big)\ . \tag{3.8}$$

Averages on σ will be denoted $\langle\,.\,\rangle$.

We plan to perform a number l of normalization steps, which will later depend on ε, in progressively reduced domains D_{ρ_k}, with

$$\rho_k = \rho(1-2k\eta)\ , \qquad \eta = \frac{1}{4l}\ ,$$

so that $\rho_l = \rho/2$. We assume recursively that after k steps, $0 \le k \le l-1$, the Hamiltonian has the form

$$H^{(k)}(\Lambda,\sigma,J,j) = K(\Lambda) + \varepsilon[U^{(k)}(\Lambda,J,j) + F^{(k)}(\Lambda,\sigma,J,j)]$$

and is real analytic in D_{ρ_k}, $F^{(k)}$ has zero average on σ, and $U^{(k)}$, $F^{(k)}$ satisfy the estimates

$$\|U^{(k)}\|^{\partial}_{\rho_k} \le \mathcal{U}_k\ , \qquad \|F^{(k)}\|^{\partial}_{\rho_k} \le r^{1/2}\,\Psi_k\ , \tag{3.9}$$

with

$$\mathcal{U}_k = \Big(1 + \sum_{i=1}^{k} e^{-k}\Big)\,\mathcal{U} < 2\mathcal{U}\ , \qquad \Psi_k = e^{-k}\,\mathcal{U}\ ; \tag{3.10}$$

the derivative $\frac{\partial U^{(k)}}{\partial\Lambda}$ is however better,[5] namely

$$\Big\|\frac{\partial U^{(k)}}{\partial\Lambda}\Big\|_{\rho_k} \le \frac{\mathcal{U}_k}{r^{1/2}\rho_\Lambda} \tag{3.11}$$

($r^{1/2}$ in place of r in the denominator). As is immediately recognized, Proposition 0 ensures that for $k=0$ there exist an extension vector ρ and a constant $\mathcal{U}$ such that these inequalities do hold; the estimate on $\frac{\partial U}{\partial\Lambda}$ is in fact better, but this is useless.

We perform a near to the identity canonical transformation

$$\mathcal{C} : D_{\rho_{k+1}} \to \mathcal{C}(D_{\rho_{k+1}}) \subset D_{\rho_k}$$

using the Lie method: that is, we introduce an auxiliary (small) Hamiltonian χ, denote by Φ^s_χ its flow, and look for χ such that $\mathcal{C} = \Phi^1_\chi$ behaves as we wish. The new Hamiltonian $H^{(k+1)} = H^{(k)} \circ \mathcal{C}$ has the form

$$H^{(k+1)} = K + \varepsilon U^{(k)} + \varepsilon F^{(k)} + \{K,\chi\} + \varepsilon\mathcal{R}_1[\chi, U^{(k)} + F^{(k)}] + \varepsilon\mathcal{R}_2[\chi, K]\ ,$$

[5] This peculiarity might surprise, but it is easily understood if one makes the decomposition

$$U^{(k)}(\Lambda,J,j) = U^{(k)}_0(J,j) + U^{(k)}_1(\Lambda,J,j)\ , \qquad U^{(k)}_1(0.J,j) = 0\ .$$

Indeed the assumption we are making on $\mathcal{U}_k$ is equivalent to $\|U^{(k)}_0\|^{\partial}_{\rho_k} \le \mathcal{U}_k$, $\|U^{(k)}_1\|^{\partial}_{\rho_k} \le r^{1/2}\mathcal{U}_k$, and this is probably the best way to understand it. In practice, however, using the above decomposition is a little cumbersome.

the remainders $\mathcal{R}_1$, $\mathcal{R}_2$ being defined as

$$\mathcal{R}_1[\chi,\,.\,] = \int_0^1 \{\,.\,,\chi\}\circ\Phi_\chi^s\,\mathrm{d}s\ , \qquad \mathcal{R}_2[\chi,\,.\,] = \int_0^1 \mathrm{d}s\int_0^s \{\{\,.\,,\chi\},\chi\}\circ\Phi_\chi^{s'}\,\mathrm{d}s'\ ;$$

χ is taken to be a solution of $\{K,\chi\}+\varepsilon F^{(k)}=0$, i.e.

$$\frac{\Omega\Lambda}{\rho^2}\frac{\partial\chi}{\partial\sigma} = \varepsilon F^{(k)}\ , \tag{3.12}$$

and thus $U^{(k+1)}$, $F^{(k+1)}$ are given by

$$U^{(k+1)} = U^{(k)} + \langle\mathcal{R}_1+\mathcal{R}_2\rangle\ , \qquad F^{(k+1)} = \mathcal{R}_1+\mathcal{R}_2-\langle\mathcal{R}_1+\mathcal{R}_2\rangle\ . \tag{3.13}$$

We must check that χ is sufficiently small so as $\Phi_\chi^1(D_{\rho_{k+1}})\subset\mathcal{D}_{\rho_k}$, and of course verify that $U^{(k+1)}$, $F^{(k+1)}$ satisfy the recurrent assumptions (3.9), (3.11).

The solution of (3.12) can be written in form of Fourier series, namely

$$\chi = \frac{\varepsilon\rho_\Lambda}{\Omega\Lambda}\sum_{0\neq n\in\mathbb{Z}}\frac{F_n^{(k)}}{in}\,e^{in\sigma}\ .$$

By writing one by one the derivatives of χ, using the recurrent assumption (3.9) on $F^{(k)}$, it is not difficult to deduce that

$$\|\chi\|_{\rho_k}^{\partial} \le \frac{2\varepsilon}{r}\,\frac{\Psi_k}{\Omega\rho_\Lambda}\ .$$

The smallness condition

$$\frac{8\,\varepsilon\,l}{r^{3/2}}\,\frac{\Psi_k}{\Omega\rho^2}\le 1 \tag{3.14}$$

thus ensures that for $(\Lambda,\sigma,J,j)\in D_{\rho_k-\eta\rho}$ it is $\Phi_\chi^s(\Lambda,\sigma,J,j)\in D_{\rho_k}$ for any $|s|\le 1$; this implies, in particular, that the canonical transformation $\mathcal{C}$ from $D_{\rho_{k+1}}\subset D_{\rho_k-\eta\rho}$ to D_{ρ_k} is well defined.

The next estimates, in view of the estimates of the remainders $\mathcal{R}_1$, $\mathcal{R}_2$, concern the Poisson brackets $\{U^{(k)},\chi\}$ and $\{F^{(k)},\chi\}$. By naively using the definition of Poisson bracket it is not difficult to work out the inequalities

$$\|\{U^{(k)},\chi\}\|_{\rho_k}\le\frac{8\varepsilon}{r}\,\frac{\mathcal{U}_k\Psi_k}{\Omega\rho^2}\ , \qquad \|\{F^{(k)},\chi\}\|_{\rho_k}\le\frac{8\varepsilon}{r}\,\frac{\Psi_k^2}{\Omega\rho^2}\ .$$

Good estimates on the remainders are then easily worked out: concerning $\mathcal{R}_1$, from its very definition it follows (also recalling $\mathcal{U}_k<2\mathcal{U}$)

$$\|\mathcal{R}_1\|_{\rho_k-\eta\rho}\le\varepsilon\|\{U^{(k)}+F^{(k)},\chi\}\|_{\rho_k}\le\frac{24\varepsilon}{r}\,\frac{\mathcal{U}}{\Omega\rho^2}\,\Psi_k\ ,$$

while concerning $\mathcal{R}_2$, using $\{K,\chi\}=-\varepsilon F^{(k)}$ and $\Psi_k\le\mathcal{U}_k$, it follows similarly

$$\|\mathcal{R}_2\|_{\rho_k-\eta\rho}\le\frac{1}{2}\varepsilon\|\{F^{(k)},\chi\}\|_{\rho_k}\le\frac{4\varepsilon}{r}\,\frac{\mathcal{U}}{\Omega\rho^2}\,\Psi_k\ .$$

Cauchy estimates, with domain reduction from $D_{\rho_k-\eta\rho}$ to $D_{\rho_{k+1}}$, allow us to estimate the derivatives of $\mathcal{R}_1$, $\mathcal{R}_2$, and the result is

$$\|\mathcal{R}_1+\mathcal{R}_2\|_{\rho_{k+1}}^{\partial}\le\frac{2^6\varepsilon\,l}{r}\,\frac{\mathcal{U}}{\Omega\rho^2}\,\Psi_k\ .$$

From (3.13) it then follows that $U^{(k+1)}$, $F^{(k+1)}$ satisfy the recurrent assumptions (3.9), (3.11) as long as the inequality

$$\frac{2^6\, e\, \varepsilon\, l}{r^{3/2}}\, \frac{\mathcal{U}}{\Omega\rho^2} \le 1 \ , \tag{3.15}$$

stronger than (3.14), is satisfied.

If (3.15) is satisfied, then the non–normal part $\varepsilon F^{(l)}$ of the final Hamiltonian $H^{(l)}$ has norm as small as $\varepsilon\, e^{-l}\mathcal{U}$. Considering the worst case $r = \varepsilon^\alpha$, and choosing l such that

$$l \le \Big(\frac{\varepsilon_*}{\varepsilon}\Big)^a < l+1 \ , \tag{3.16}$$

it turns out that (3.15) is satisfied if

$$\frac{3}{2}\alpha + a = 1 \ , \qquad \varepsilon_*^a = \frac{1}{2^6 e}\, \frac{\Omega\rho^2}{\mathcal{U}} \ .$$

The conclusion is that $F^{(l)}$ is such that

$$\|\varepsilon F^{(l)}\|^{\partial}_{\rho/2} < \varepsilon\, e\, e^{(\varepsilon_*/\varepsilon)^a}\, \mathcal{U} \ , \tag{3.17}$$

and so, if we forget the remainder in (2.6), the inequality (3.3) is satisfied with the appropriate choice (3.5) of α, as claimed in Section 3.1D.

The overall canonical transformation $\mathcal{C}_{\mathrm{tot}}$, produced after l steps, is small: namely denoting $(\Lambda, \sigma, J, j) = \mathcal{C}_{\mathrm{tot}}(\Lambda', \sigma', J', j')$, it is not difficult to deduce (just by summing the maximal possible displacements $\|\frac{\partial\chi}{\partial\sigma}\|_{\rho_k}$ at each step)

$$|\Lambda' - \Lambda| \le \frac{\varepsilon^{1-\alpha/2}}{\Omega} \sum_{k=0}^{l-1} \Psi_k \le 2\varepsilon^{(2+a)/3}\, \frac{\mathcal{U}}{\Omega\rho^2}\, \rho_\Lambda$$

and similarly

$$|\sigma' - \sigma| \le 16\varepsilon^a\, \frac{\mathcal{U}}{\Omega\rho^2}\, \rho_\sigma \ , \qquad |x' - x| \le 8\varepsilon^{(2+a)/3}\, \frac{\mathcal{U}}{\Omega\rho^2}\, \rho_x \quad \text{for} \quad x = J, j \ .$$

The estimate for σ is somehow poor, but for our purposes this is totally irrelevant.

Up to now we considered the truncated Hamiltonian with two degrees of freedom. However, the extension to the complete Hamiltonian is trivial: namely at each perturbation step we maintain the definition $\mathcal{C} = \Phi^1_\chi$ with the same χ, but now we interpret the auxiliary Hamiltonian χ as having three degrees of freedom, the variable Γ being no more a parameter but a canonical coordinate. Since χ does not depend on γ, at each step Γ remains unchanged. The angle γ, entering the remainder of the complete Hamiltonian, of course changes a little, but this is obviously irrelevant, and the remainder itself, after composition with $\mathcal{C}$, remains as small as it was. It should be observed that, since $a < \frac{1}{4}$, the remainder in (2.6) is (much) smaller than the new remainder $\varepsilon F^{(l)}$ estimated in (3.17); the conclusion is that, after a minor correction of the constant in (3.3), the inequality holds for the complete Hamiltonian, too.

B. The case $\nu \ne (-1, 1)$, for Λ not too small. Here we work in action–angle variables and exclude, in addition to the resonant zone $\mathcal{Z}$, a neighborhood of the proper rotation given by $\Lambda < \varepsilon^\alpha \Lambda_*$. The scheme of the proof is very similar to the previous case, with a very minor complication due to the (possible) presence of the resonant zone $\mathcal{Z}$ around $\Lambda_{\rm res} = \varepsilon^\zeta\Lambda_* \geq \varepsilon^\alpha\Lambda_*$. Now the two–degrees of freedom Hamiltonian we start with has the form (3.6), with

$$K(\Lambda) = \frac{1}{2I}(\Lambda - \Lambda_{\rm res})^2 \tag{3.18}$$

and F as in (3.7); F however is smaller, being now divisible (at least) by Λ. For any Λ_0 such that

$$\Lambda_0,\ |\Lambda_0 - \Lambda_{\rm res}|\ \geq \varepsilon^\alpha\Lambda_*\ ,$$

we introduce as above $r = \Lambda_0/\Lambda_*$, but we need a second variable

$$\tilde r = \min(r, |\Lambda_0 - \Lambda_{\rm res}|/\Lambda_*)$$

to be used to bound the denominators and the analyticity domains of functions depending on Λ (previously it was $\tilde r = r$). In particular, $\tilde r$ enters the extension vector, which is now

$$\rho = (\tilde r \rho_\Lambda, \rho_\sigma, J, j)\ ,$$

and also replaces r in the definition (3.8) of the norm $\|\,.\,\|^\partial_{\rho'}$. The recurrent assumption now is

$$\|U^{(k)}\|^\partial_{\rho_k} \leq \mathcal{U}_k\ , \qquad \|F^{(k)}\|^\partial_{\rho_k} \leq r\Psi_k\ ,$$

with $\mathcal{U}_k$, Ψ_k as before, $\frac{\partial U^{(k)}}{\partial\Lambda}$ satisfying however the better inequality

$$\Big\|\frac{\partial U^{(k)}}{\partial\Lambda}\Big\|_{\rho_k} \leq \frac{r}{\tilde r}\,\frac{\mathcal{U}_k}{\rho_\Lambda}\ .$$

The perturbation scheme is then *identical,* and the main estimates found in the way are

$$\|\chi\|^\partial_{\rho_k} \leq 2\frac{r}{\tilde r}\,\frac{\Psi_k}{\Omega}\ , \qquad \|\{U^{(k)},\chi\}\|_{\rho_k} \leq 6\varepsilon\frac{r^2}{\tilde r^2}\,\frac{\mathcal{U}}{\Omega\rho^2}\,\Psi_k\ , \qquad \|\{F^{(k)},\chi\}\|_{\rho_k} \leq 8\varepsilon\frac{r^2}{\tilde r^2}\,\frac{\mathcal{U}}{\Omega\rho^2}\,\Psi_k\ ,$$

leading to

$$\|\mathcal{R}_1 + \mathcal{R}_2\|^\partial_{\rho_{k+1}} \leq 72\varepsilon l\frac{r^2}{\tilde r^2}\,\frac{\mathcal{U}}{\Omega\rho^2}\,\Psi_k\ .$$

The smallness condition replacing (3.15) is

$$72\,\varepsilon\, l\,\frac{r}{\tilde r^2}\,\frac{\mathcal{U}}{\Omega\rho^2} \leq e^{-1}\ , \tag{3.19}$$

and the worst situation is when $\tilde r = \varepsilon^\alpha$, $r = \varepsilon^\zeta + \varepsilon^\alpha < 2\varepsilon^\zeta$; making the same choice of l as in (3.16), condition (3.19) is satisfied if

$$2\alpha - \zeta + a = 1\ , \qquad \varepsilon_* \leq \varepsilon_{**} \stackrel{\rm def}{=} \Big(\frac{1}{18\,e}\,\frac{\Omega\rho^2}{\mathcal{U}}\Big)^{1/a}\ . \tag{3.20}$$

The former equality explains the choice (3.1) of α. The extension of the canonical transformation to the pair of conjugated variables (Γ,γ) is clearly identical as before. However, at variance with the previous case, now the value $a = \frac14$ is possible and correspondingly the remainder in (2.6) can be relevant. To take care of this situation, it is enough to take

$$\varepsilon_* = \min(\tilde\varepsilon, \varepsilon_{**})$$

and to choose appropriately the constant in (3.3). An overall estimate on the canonical transformation is provided by

$$|\Lambda' - \Lambda| \le 4\varepsilon^{(1+\zeta+a)/2}\, \frac{\mathcal{U}}{\Omega\rho^2}\, \rho_\Lambda$$

and

$$|\sigma' - \sigma| \le 4\varepsilon^{a}\, \frac{\mathcal{U}}{\Omega\rho^2}\, \rho_\sigma\ , \qquad |x' - x| \le 4\varepsilon^{(1+\zeta+a)/2}\, \frac{\mathcal{U}}{\Omega\rho^2}\, \rho_x \quad \text{for} \quad x = J,\, j\ .$$

C. The case $\nu \ne (-1,1)$, for small Λ. It remains to consider the case $\nu \ne (-1,1)$ in a neighborhood $\Lambda < \varepsilon^\alpha \Lambda_*$ of the proper rotation, where we cannot enter with the action–angle coordinates due to their singularity at $\Lambda = 0$. This work is useful when such a neighborhood is not included in $\mathcal{Z}$, that is when $\Lambda_{\rm res} > \varepsilon^\alpha\Lambda_*$, to show that the motion is there regular and well confined as regular motions are, with vanishing Lyapunov exponent. (Strictly speaking, for the only sake of proving long term stability of proper rotations, regardless to whether motions are regular or not, such a work is unnecessary.)

We use the Birkhoff (corotating) coordinates and consider the two–degrees of freedom Hamiltonian

$$h(w_1,w_2,J,j) = K(\Lambda) + \varepsilon[u(\Lambda,J,j) + f(w_1,w_2,\sigma,J,j)]\ ,$$

with $\Lambda = iw_1w_2$, $K(\Lambda)$ as in (3.18) and

$$f(w_1,w_2,\sigma,J,j) = \sum_{0\ne n\in\mathbb{Z}} f_{n\nu_2}(w_1,w_2,\sigma,J,j)\ ,$$

having denoted, for $m \in \mathbb{Z}$,

$$f_m(w_1,w_2,J,j) = \hat f_m(\Lambda,\sigma,J,j) w^{|m|}\ , \qquad w = \begin{cases} w_1 & \text{for } m \ge 0 \\ w_2 & \text{for } m < 0 \end{cases}\ .$$

The real domain to be extended is given by

$$w_1 = w_2 = 0\ , \qquad |J| \le \frac{3}{4}\ , \qquad j \in S^1\ ; \tag{3.21}$$

the extension vector, denoting $r = \varepsilon^\alpha$, is now

$$\rho = (r^{\frac12}\rho_w, \rho_J, \rho_j)\ , \qquad \rho_w = \sqrt{\Lambda_*/2}\ ,$$

and the complex extended domain is defined as the union of the polydiscs

$$|w_1|,\, |w_2|\ \le\ r^{1/2}\rho_w\ , \qquad |J - J_0| \le \rho_J\ , \qquad |j - j_0| \le \rho_j$$

for real J_0, j_0 satisfying (3.21). Similarly to the previous case, we denote $\rho^2 = \min(\rho_w^2, \rho_J\rho_j)$ and

$$\|\,.\,\|^\partial_{\rho'} = \max\left(\|\,.\,\|_{\rho'}, r\rho_w \left\|\frac{\partial\,.}{\partial w_1}\right\|_{\rho'},\ r\rho_w\left\|\frac{\partial\,.}{\partial w_2}\right\|_{\rho'},\ \rho_J\left\|\frac{\partial\,.}{\partial J}\right\|_{\rho'},\ \rho_j\left\|\frac{\partial\,.}{\partial j}\right\|_{\rho'}\right)\ .$$

Once more, the proof proceeds along the same scheme. The recurrent hypothesis is that after $k \le l-1$ perturbation steps we deal with

$$h^{(k)}(w_1,w_2,J,j) = K(\Lambda) + \varepsilon[u^{(k)}(\lambda,J,j) + f^{(k)}(w_1,w_2,J,j)]\ ,$$

$f^{(k)}$ containing only the harmonics $n\nu_2$ with $0 \neq n \in \mathbb{Z}$; $u^{(k)}$ and $f^{(k)}$ are real analytic in D_{ρ_k} and satisfy the estimates

$$\|u^{(k)}\|^{\partial}_{\rho_k} \le \mathcal{U}_k \ , \qquad \|f^{(k)}\|^{\partial}_{\rho_k} \le r\Psi_k$$

as well as

$$\Big\|\frac{\partial f}{\partial w_i}\Big\|_{\rho_k} \le r^{1/2}\mathcal{U}_k \ , \quad i = 1,2 \ ,$$

with $\mathcal{U}_k$, Ψ_k as in (3.10). The equation for χ reads now

$$i\frac{\Omega}{\rho^2}(\Lambda - \Lambda_{\mathrm{res}})\Big(w_1\frac{\partial\chi}{\partial w_1} - w_2\frac{\partial\chi}{\partial w_2}\Big) = \varepsilon\, f_k$$

and is solved by

$$\chi = \frac{\varepsilon\rho^2}{i\Omega(\Lambda - \Lambda_{\mathrm{res}})}\sum_{0\neq n\in\mathbb{Z}}\frac{f_{n\nu_2}}{n\nu_2} \ .$$

The procedure then goes on exactly as in the previous cases, the relevant estimates being

$$\|\chi\|^{\partial}_{\rho_k} \le \frac{3}{2}\frac{\varepsilon}{r^{1/2}}\frac{\Psi_k}{\Omega\rho^2} \ , \qquad \|\{u_k,\chi\}\|_{\rho_k} \le 8\,\varepsilon\frac{\mathcal{U}}{\Omega\rho^2}\Psi_k \ , \qquad \|\{u_k,\chi\}\|_{\rho_k} \le 4\,\varepsilon\frac{\mathcal{U}}{\Omega\rho^2}\Psi_k$$

and finally

$$\|\mathcal{R}_1 + \mathcal{R}_2\|^{\partial}_{\rho_{k+1}} \le 48\,\varepsilon\, l\,\frac{\mathcal{U}}{\Omega\rho^2}\Psi_k \ .$$

The appropriate smallness condition which allows to verify the recurrent assumption on u_k, f_k, also ensuring $\mathcal{C}(D_{k+1}) \subset D_k$, is

$$48\,\frac{\varepsilon l}{r}\,\frac{\mathcal{U}}{\Omega\rho^2} \le e^{-1} \ ;$$

using $r = \varepsilon^{\alpha}$ and l as in (3.16) the condition turns into

$$48\, e\, \varepsilon^{1-\alpha-a}\,\frac{\mathcal{U}}{\Omega\rho^2} \le 1$$

and is guaranteed by (3.20). The overall canonical transformation is estimated by

$$|\Lambda' - \Lambda| \le 8\,\varepsilon\,\frac{\mathcal{U}}{\Omega\rho^2}\,\Lambda_* \ .$$

Remark. One might wonder why this perturbation scheme in Birkhoff variables has been used only for $\Lambda < \varepsilon^{\alpha}\Lambda_*$ and not for larger Λ (the obvious advantage would be avoiding the double proof, in two different sets of variables). The answer is that the scheme works, but some results (including the crucial relation (3.20) between α, ζ and a) get worse. The reason is easy to understand: confining Λ in this way requires a correspondingly good confinement of w_1, w_2, but if Λ is large, this produces an unnecessary and "expensive" good confinement of the angle σ, too. The use of the polar–like coordinates Λ, σ allows instead to treat separately the action and the angle, with poorer control on the angle. A similar motivation is at the basis of the analogous choice in [15]. We do not know if it is possible to unify the proofs through a more elaborated construction.

4 The triaxial case

As remarked in the Introduction, we considered the case of a kinetically symmetric body mainly for simplicity: indeed in this case the Andoyer–Deprit variables (G, L, J, g, l, j) are the action–angle variables of the system, and everything can be written explicitly; the same choice was made in [4] and [5]. The results however are more general and naturally extend to the triaxial case, too. In particular the proper rotations about e_1 and e_3, if I_2 is intermediate between I_1 and I_3, are Nekhoroshev stable. Here we produce a sketch of the work to be done to perform such an extension, which includes only minor changes with respect to the symmetric case. It is clearly sufficient to consider motions near the proper rotation about e_3, assuming either $I_1 < I_2 < I_3$ or $I_3 < I_2 < I_1$.

The main question is not the extension of Proposition 1, but the extension of Proposition 0 proved in [5], which is at the basis of Proposition 1. To this purpose the starting point is showing that in the triaxial case, too, there exist Birkhoff coordinates, analytic in a neighborhood of the proper rotation, which integrate the unperturbed system. The existence of such variables is a byproduct of [18]. Indeed, the kinetic energy of the triaxial rigid body in the Andoyer–Deprit coordinates has the expression

$$K(G, L, J, g, l, j) = (G^2 - L^2)\Big(\frac{\sin^2 l}{2I_1} + \frac{\cos^2 l}{2I_2}\Big) + \frac{L^2}{2I_3} \ ;$$

the system is integrable, but the coordinates are not the action–angle ones. We can nevertheless introduce the coordinates $(G, \gamma, \Lambda, \lambda)$ as in Section 2.1, and then the Cartesian coordinates x_1, x_2; the kinetic energy assumes the form

$$k = \frac{1}{2I_3}\Big[G^2 - G(\beta_2 x_1^2 + \beta_1 x_2^2) + \frac{1}{4}(x_1^2 + x_2^2)(\beta_2 x_1^2 + \beta_1 x_2^2)\Big] \ ,$$

with

$$\beta_i = 1 - \frac{I_3}{I_i} \ , \qquad i = 1, 2$$

(both constants are positive for $I_3 < I_2 < I_2$, negative if the opposite inequality holds). The rescaling

$$x_1 = (\beta_2/\beta_1)^{1/4}\tilde{x}_1 \ , \qquad x_2 = (\beta_1/\beta_2)^{1/4}\tilde{x}_2$$

gives then

$$\tilde{k} = \frac{1}{2I_3}\Big[G^2 \mp \sqrt{\beta_1\beta_2}\, G\,(\tilde{x}_1^2 + \tilde{x}_2^2) + \frac{1}{4}(\beta_1\tilde{x}_1^2 + \beta_2\tilde{x}_2^2)(\tilde{x}_1^2 + \tilde{x}_2^2)\Big]$$

(upper sign if β_1, β_2 are positive). Such a Hamiltonian can be put in normal form by means of a near to the identity analytic transformation[6]

$$\tilde{x}_1 = \hat{x}_1 + X_1(\hat{x}_1, \hat{x}_2, G) \ , \qquad \tilde{x}_2 = \hat{x}_2 + X_2(\hat{x}_1, \hat{x}_2, G) \ ,$$

with X_1, X_2 at least quadratic in $\hat{x}_1$, $\hat{x}_2$; denoting $\hat{\Lambda} = \frac{1}{2}(\hat{x}_1^2 + \hat{x}_2^2)$, the Hamiltonian becomes a function of G and $\hat{\Lambda}$, and has the form

$$\hat{k}(G, \hat{\Lambda}) = \frac{1}{2I_3}\Big[G^2 \mp 2\sqrt{\beta_1\beta_2}\, G\hat{\Lambda} + \frac{1}{2}(\beta_1 + \beta_2)\,\hat{\Lambda}^2\Big] + \hat{k}'(G, \hat{\Lambda}) \ ,$$

[6]This is evident, see for example [20], if $\tilde{k}$ is thought of as a system with one degree of freedom (a nonlinear oscillator) parametrized by G; the extension of the transformation to the pair (G, γ) is immediate.

k' being divisible by $\hat{\Lambda}^3$. This expression is the generalization of (1.12) to the triaxial case.

The problem thus becomes essentially identical to the problem studied in [5] and in the present paper. In particular:

(i) Near the proper rotation the frequency has still the form

$$\omega(G,\hat{\Lambda}) = \frac{G}{I_3}(1,-\beta) + \mathcal{O}(\hat{\Lambda}) \ ,$$

with $\beta = \pm\sqrt{\beta_1\beta_2}$, and one easily checks that $-1 \le \beta < 1$. The assumption corresponding to (1.2), ensuring the crucial inequality (1.5), turns out to be

$$|\beta_2|(1-\beta_1) \ge \beta_* > 0 \ .$$

(ii) The Hessian matrix on the proper rotation is

$$\frac{1}{I_3}\begin{pmatrix} 1 & \mp\sqrt{\beta_1\beta_2} \\ \mp\sqrt{\beta_1\beta_2} & \frac{1}{2}(\beta_1+\beta_2) \end{pmatrix} \ ,$$

so $\hat{k}$ is there convex if $\beta_1, \beta_2 > 0$ ($I_3 < I_2 < I_1$) and is immediately seen to be quasi–convex in the opposite case; this is essential to start the perturbation work of [5].

(iii) The perturbation, being analytic in the Birkhoff coordinates, is appropriately small with the distance from the proper rotation exactly as in the symmetric case, the norm of the Fourier component $\nu = (\nu_1, \nu_2)$ being divisible by $\hat{\Lambda}^{|\nu_2|/2}$.

Once these basic properties are satisfied, all of the perturbation work done in [5] and in the present paper proceeds smoothly, with no important changes, and results are the same up to the value of the constants.

5 Conclusions and open problems

A. A final comment.

Let us here shortly summarize what we did in this paper. As remarked in the Introduction, the Euler rigid body has two fast degrees of freedom and a slow degenerate one, which does not appear in the expression of the kinetic energy. In presence of a resonance between the fast frequencies, a new slow degree of freedom is produced, and there is a chance that such degree of freedom interacts with the degenerate one, thus producing chaotic motions. A crucial fact is that degeneracy (complete lack of convexity or steepness, in the language of Nekhoroshev theory) makes useless the standard mechanisms of confinement entering Nekhoroshev theory, and wide chaotic motions become possible. As remarked in the Introduction, chaotic motions in which $\mu = m^s/\|m\|$ wanders unpredictably on wide portions of the unit sphere *for no matter how small* ε have been observed numerically [3], precisely in connection with resonances between the two fast frequencies.

The novelty of this paper is that an important class of motions of the rigid body, namely motions close to proper rotations, are more regular than all the other ones. The key observation is that the distance to the proper rotation can be successfully used as a second small parameter in the system, in addition to ε. Profiting by this fact, we are able to show that for a long time scale *all* motions close to proper rotations either are regular (all Lyapunov exponents vanish) or are possibly chaotic (the maximal Lyapunov exponent is positive), but nevertheless they are confined to a small neighborhood of a predictable regular curve, essentially as in nondegenerate systems (figures 1 and 2). In a sense, the vicinity to a proper rotation removes the degeneracy.

Best results are obtained if the relation between the two small parameters is appropriately tuned, i.e., if the diameter of the neighborhoods of the proper rotation entering the game are appropriately chosen as functions of ε. We made some efforts to be careful on this point, and the result are the powers of ε, i.e. the stability exponents, entering Proposition 1. The main reason of our interest in the value of the exponents is that at least some of them can be computed numerically, and so the theory admits a critical test. The question, important in our opinion, is whether the mechanism producing stability of motions close to proper rotations is now essentially understood, or other elements producing stability are still neglected. Such a test is in progress; preliminary results are encouraging.

B. Applications.

Our Proposition 1 is very general — it holds for any analytic potential — so it is natural to discuss about its possible applications. To this purpose, it is worthwhile to stress that the proposition is interesting only if the potential does depend on the proper rotation (on the angle l, in the Andoyer–Deprit coordinates). Otherwise, if both the kinetic and the potential energy have rotational symmetry about e_3, the system reduces to two degrees of freedom, and everything gets substantially easier. In particular, the mechanism leading to slow chaos, shortly recalled above in paragraph A, is excluded *a priori.* Rigid bodies which are kinetically symmetric, but the potential is not, do exist: for example, restricting to gravitational forces, prismatic bodies, or cylinders with two or more symmetrically disposed legs, as is frequent among artificial satellites.[7] The main interest however, in our opinion, is the case of triaxial bodies, to which Section 4 is devoted. Although in Celestial Mechanics the main objects are symmetric, non symmetric satellites or asteroids are well known to exist (non symmetric artificial satellites exist as well).

As it was suggested to us by the referee, to whom we express our sincere gratitude, there are other problems of Celestial Mechanics which do not fit the assumptions of Proposition 1, but likely can be studied by the same techniques used in this paper and in the same spirit. An important case is the spin–orbit problem for a symmetric rigid body in an elliptic orbit. Indeed the symmetry of the body eliminates the fast angle l, but the mean anomaly of the orbital motion can play the same role. The question then assumes the same form: in presence of a resonance between the spin motion (the angle g) and the orbital motion, a slow degree of freedom is produced, and there is a chance it effectively interacts with (J, j) producing chaotic motions. For models of this kind, see for example [9] and [8]. A rough preliminary analysis makes us confident that the perturbation techniques used in this paper, in principle, extend to those models too. We plan to come back to this problem in a near future.

[7] Adding other forces, for example elastic ones, is an easier way to break the rotational symmetry, but it looks very artificial.

References

[1] H. Andoyer, *Cours de Méchanique Celeste* (Gautier-Villars, Paris 1923).

[2] V.I. Arnol'd, *Geometrical methods in the theory of ordinary differential equations* (Springer–Verlag, Berlin 1982).

[3] G. Benettin, A.M. Cherubini and F. Fassò, *Regular and Chaotic motions of the fast rotating rigid body: a numerical study.* Discr. Cont. Dyn. Sys. 4, 521-540 (2002).

[4] G. Benettin and F. Fassò, *Fast rotations of the rigid body: A study by Hamiltonian perturbation theory. Part I.* Nonlinearity **9**, 137-186 (1996).

[5] G. Benettin, F. Fassò and M. Guzzo, *Fast rotations of the rigid body: A study by Hamiltonian perturbation theory. Part II: Gyroscopic rotations.* Nonlinearity **10**, 1695-1717 (1997).

[6] G. Benettin, F. Fassò and M. Guzzo, paper in preparation.

[7] G. Benettin, P. Hjorth and P. Sempio, *Exponentially long equilibrium times in a one dimensional collisional model of a classical gas.* Journ. Stat. Phys. **94**, 871-892 (1999).

[8] L. Biasco and L. Chierchia, *Effective Hamiltonian for the d'Alembert planetary model near a spin/orbit resonance.* Cel. Mech. & Dyn. Astr **83**, 223–237 (2002).

[9] L. Chierchia and G. Gallavotti, *Drift and diffusion in phase space.* Ann. Inst. H. Poincaré Phys. Théor. **60** (1994).

[10] A. Deprit, *Free rotation of a rigid body studied in phase plane.* Am. J. Phys. **55**, 424 (1967).

[11] Y. Elskens and D.F. Escande, *Slowly pulsating separatrices sweep homoclinic tangles where islands must be small: an extension of classical adiabatic theory.* Nonlinearity **4**, 615–667 (1991).

[12] D.F. Escande, *Change of adiabatic invariant at separatrix crossing; application to slow Hamiltonian Chaos,* in *Advances in Nonlinear Dynamics and Stochastic Processes,* R. Livi and A. Politi Editors, World Scientific, Singapore 1985.

[13] F. Fassò, *Hamiltonian perturbation theory on a manifold.* Cel. Mech. & Dyn. Astr. **62,** 43–69 (1995).

[14] F. Fassò, *The Euler–Poinsot top: a non-commutatively integrable system without global action–angle coordinates.* J. Appl. Math. Phys. **47**, 953-976 (1996).

[15] M. Guzzo, F. Fassò and G. Benettin, *On the stability of elliptic equilibria.* MPEJ - Mathematical Physics Electronic Journal **4**, Paper 1, 1–16 (1998).

[16] J. Henrard, *Virtual singularities in the artificial satellite theory.* Cel. Mech. & Dyn. Astr. **10**, 437–447 (1973).

[17] T. Levi–Civita and U. Amaldi, *Lezioni di Meccanica Razionale,* Vol. 2 – Part 2 (Zanichelli, Bologna 1927).

[18] M.L. Lidov and A.I. Neishtadt, *The method of canonical transformations in problems of the rotation of celestial bodies and Cassini Laws* (in Russian), in: *Determination of the motion of a spacecraft* (in Russian), P.E. Eliasberg Ed., pp. 75–107 (Nauka, Moscow 1975).

[19] A. Morbidelli, *Modern Celestial Mechanics* (Taylor and Francis, London 2002).

[20] J. Moser, *Lectures in Celestial Mechanics,* in *Mem. American Math. Soc.* No 81, 1 (1968).

[21] A.I. Neishtadt, *Change in adiabatic invariant at a separatrix,* Sov. J. Plasma Phys. **12**, 568–573 (1986).

[22] A.I. Neishtadt, *On the change in tha adiabatic invariant on crossing a separatrix in systems with two degrees of freedom,* Prikl. Matem. Mekan. **51**, 750–757 (1987).

[23] A.I. Neishtadt, V.V. Sidorenko and D.V. Treshev, *Stable periodic motions in the problem on passage through a separatrix,* Chaos **3**, 2–11 (1997).

[24] 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)].

[25] 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).

[26] F. Stoppelli, *Sul principio dell'effetto giroscopico.* Giornale di Matem. di Battaglini, serie 4, **80** (1950-51).

[27] F. Stoppelli, *Un fenomeno di risonanza nei solidi in rapida rotazione.* Ricerche di Matem., **9**, 213-241 (1960).

[28] J.L. Synge, *Classical Dynamics.* In: *Handbuch der Physik vol. III/I, Principles of classical mechanics and field theory.* Springer–Verlag (Berlin, 1960).