Nekhoroshev–stability of elliptic equilibria of Hamiltonian systems

Francesco Fassò, Massimiliano Guzzo and Giancarlo Benettin
Università di Padova, Dipartimento di Matematica Pura e Applicata
Via Belzoni 7, 35131 Padova, Italy
e-mail: fasso,guzzo,benettin@math.unipd.it

(First Version: September 1997. Revised Version: March 1998)
(To appear in *Communications in Mathematical Physics)*

Abstract: *We prove a conjecture by N.N. Nekhoroshev about the long–time stability of elliptic equilibria of Hamiltonian systems, without any Diophantine condition on the frequencies. Higher order terms of the Hamiltonian are used to provide convexity. The singularity of the action–angle coordinates at the origin is overcome by working in cartesian coordinates.*

1. Introduction and statement of the result

A. Introduction. In this paper we prove the Nekhoroshev–stability of elliptic equilibria of analytic Hamiltonian systems. The Hamiltonian of the problem is a convergent power series

$$h_2 \;+\; h_3 \;+\; h_4 \;+\; \cdots\,, \tag{1.1}$$

where any h_k is a homogeneous polynomial of degree k in the canonical coordinates $(p,q) \in \mathbb{R}^n$, and

$$h_2 \;=\; \frac{1}{2}\sum_{j=1}^{n}\overline{\omega}_j\left(p_j^2+q_j^2\right)$$

with real $\overline{\omega}_j$. We assume that the frequency vector $\overline{\omega}=(\overline{\omega}_1,\ldots,\overline{\omega}_n)$ does not satisfy any resonance condition up to order four and that the nonlinear term h_3+h_4 has the generic property of 'removing the degeneracy' of h_2, more precisely of providing convexity, in a sense specified below. Under these hypotheses, the Hamiltonian (1.1) can be trasformed with two Birkhoff normalization steps into a Hamiltonian which is integrable and convex (hence anisochronous) up to order five. By studying the latter system with the techniques of Hamiltonian perturbation theory, we prove that any motion starting sufficiently close to the origin remains close to it for times which grow exponentially fast with (a power of) the distance from the origin.

This result was conjectured by Nekhoroshev in his 1977 celebrated paper ([18], section 2.2). It should be noticed that this result does not directly follow from Nekhoroshev's general theorem about the stability of the actions in nearly integrable Hamiltonian systems, because of the singularity at the origin of the action–angle coordinates. Thus, in order to prove it, one should perform perturbation theory using cartesian coordinates, which are regular in a neighbourhood of the equilibrium. This procedure is quite standard for *isochronous* systems, i.e. for constructing the so–called Birkhoff series (see e.g. [20,10,13]), while it seems that the case of *anisochronous* systems had not been understood until very recently [3]. Correspondingly, the problem of the stability of elliptic equilibria over exponentially long times has been so far investigated in the following ways:

On the one hand, the problem has been studied in several papers under the assumption that the frequency $\overline{\omega}$ satisfies a Diophantine condition (see e.g. [10,11,14]). Indeed, as realized by Gallavotti in a closely related problem [8], under this strong nonresonant condition one can regard (1.1) as a small perturbation of the *isochronous* system described by h_2 and avoid any 'geometry of resonances'.[1)]

On the other hand, the full Nekhoroshev's conjecture—without any Diophantine assumption on the frequencies—was investigated by Lochak in [16,17], who however worked in action–angle coordinates. Therefore, in order to consistently avoid the singularities, he had to exclude a small cusp–shaped region around each hyperplane $p_j^2+q_j^2=0$, $j=1,\ldots,n$. Hence, the 'stability region' that he finds, although arriving arbitrarily close to the equilibrium, does not include any neighbourhood of it. (The need of entering in these cusp–shaped regions is typical of Nekhoroshev's approach, while it is not as urgent in KAM theory, see e.g. [5]).

Finally, we mention that a Nekhoroshev–like result for the rather special case of completely resonant frequencies has been very recently obtained in [1], in the framework of a study of the Nonlinear Schrödinger equation.

A way of performing perturbation theory for anisochronous systems using cartesian coordinates was understood in [3] in connection with a problem of rigid body dynamics, namely, the stability of the fast rotations of a rigid body close to 'proper' rotations. The similarity with the problem of the elliptic equilibrium resides in the fact that the proper rotations of a rigid body lie on transversally elliptic circles, on which one pair of action–angle coordinates becomes singular. Here we proceed precisely as there, although we need to generalize and to adapt to the problem at hand the perturbative scheme there introduced. The main difference comes from the fact that the small parameter represents here the distance from the fixed point, while in the rigid body problem it was an independent parameter measuring the strenght of the external forces; moreover, here one needs working with any number of frequencies, and not just with two of them. We shall come back on these points at the end of this section.

---

1) Gallavotti's result, further developed and improved in [2,9,6], concerns small perturbations of systems of uncoupled harmonic oscillators, written in action–angle coordinates; the difference with the problem of elliptic equilibria is that the perturbation is assumed to be analytic in the actions, so that one can work in action–angle coordinates.

B. Statement of the result. We make use of the usual complex canonical coordinates

$$w_j = \frac{p_j - iq_j}{i\sqrt{2}}\,, \qquad z_j = \frac{p_j + iq_j}{\sqrt{2}}\,, \qquad j = 1, \ldots, n\,,$$

and denote by

$$I_j(w_j, z_j) = i w_j z_j \qquad j = 1, \ldots, n\,,$$

the n actions.

As is well known, if the vector $\overline{\omega}$ does not admit any resonance $\nu \cdot \overline{\omega} = 0$ with $\nu \in \mathbb{Z}^n \setminus \{0\}$, $|\nu| \leq 4$, where $|\nu| = \sum_j |\nu_j|$, then one can perform two Birkhoff normalization steps on the Hamiltonian (1.1), thus obtaining a Hamiltonian of the form

$$h(w, z) = k(izw) + f(w, z)\,, \tag{1.2}$$

where

$$k(I) = \overline{\omega} \cdot I + \frac{1}{2} I \cdot AI\,, \tag{1.3}$$

A being a $n \times n$ real symmetric matrix, while f is a series in (w, z) which starts with terms of order five. In any polydisk $|w_j| \leq R$, $|z_j| \leq R$, the canonical transformation leading to (1.2) differs from the identity by terms of order R^2. Both Hamiltonians (1.1) and (1.2) are real whenever $z = -i\overline{w}$ (that is, for real p, q). Throughout the paper, we use the following notation:

$$|I|^\infty = \max_{1 \leq j \leq n} |I_j|\,, \qquad [x] = \text{the integer part of } x \in \mathbb{R}\,.$$

Theorem. *Assume that the Hamiltonian (1.2) is convergent for $|w|^\infty$, $|z|^\infty \leq R_c$, with some $R_c > 0$, and that the quadratic form $I \cdot AI$ is convex, i.e.*

$$\bigl|I \cdot AI\bigr| \geq m\, I \cdot I \qquad \forall I \in \mathbb{R}^n \tag{1.4}$$

with some $m > 0$. Then, there exist positive constants $R \leq R_c$, ε_*, ε_0, c and C such that any motion of the system (1.2), with 'real' initial data $(w(0), z(0)) = (w(0), -i\overline{w}(0))$ with

$$\varepsilon := \frac{|I(0)|^\infty}{R^2} \leq \varepsilon_0\,, \tag{1.5}$$

satisfies

$$|I(t)|^\infty \leq c \left(\frac{\varepsilon}{\varepsilon_*}\right)^{\frac{1}{n}} R^2 \qquad \textit{for} \quad |t| \leq C \exp\left[\frac{1}{2}\left(\frac{\varepsilon_*}{\varepsilon}\right)^{\frac{1}{n}}\right] \tag{1.6}$$

as well as

$$|I(t)|^\infty \leq c \left(\frac{\varepsilon}{\varepsilon_*}\right)^{\frac{1}{2}+\frac{1}{2n}} R^2 \qquad \textit{for} \quad |t| \leq C \exp\left[\frac{1}{2}\left(\frac{\varepsilon_*}{\varepsilon}\right)^{\frac{1}{2n}}\right]. \tag{1.7}$$

Possible values of the constants $R, \ldots, C$ will be provided in the course of the proof (see equations (2.21)–(2.23)).

As an application of this result, we would like to mention here the problem of the Nekhoroshev–stability of the so called Riemann ellipsoids, which will be reported in a forthcoming paper [7]. (See [4,15] and references therein for general informations about the problem). After reduction, the Riemann ellipsoids appear as equilibria of a Hamiltonian system with *four* degrees of freedom, whose frequencies depend on only *two* functions of the momentum. Therefore, the frequencies are certainly not Diophantine for generic values of the momentum, but moreover it is not even clear whether there is any one which is Diophantine at all, so that relaxing such a condition is crucial. In fact, the study of the stability of the Riemann ellipsoids was the occasion for us to become interested in the problem studied in this paper.

Remarks: (i) The following generalizations, not explicitly treated here, are straightforward:

- The case in which k is quasi-convex.
- The case in which the elliptic equilibrium is replaced by a transversally elliptic torus.

In the latter case, in addition to the coordinates (w, z), there are also one or more pairs of nonsingular action-angle coordinates, and everything runs smoothly with no important differences. (This case is in fact close to the situation considered in [3] in connection with the rigid body problem).

(ii) If $\overline{\omega}$ has better arithmetic properties, namely is nonresonant for $|\nu| \leq s-1$ with $s > 5$, then the results can be improved. Specifically, it is possible to obtain a significantly better confinement of the actions, i.e. of order ε, on even longer time scales. However, this result requires a somewhat different treatment (it cannot be obtained just by putting a generic $s > 5$ in place of $s = 5$ in the estimates below) and will be reported in a forthcoming paper [12]. We shall come back on this point at the end of the paper.

(iii) Inequalities (1.6) and (1.7) are in fact unified by

$$|I(t)| \;\leq\; c \left(\frac{\varepsilon}{\varepsilon_*}\right)^{\frac{1+\mu}{n+\mu}} \qquad \text{for} \quad |t| \;\leq\; C \exp\left[\frac{1}{2}\left(\frac{\varepsilon_*}{\varepsilon}\right)^{\frac{1}{n+\mu}}\right] \qquad \forall \mu \;\geq\; 0 \tag{1.8}$$

which is what we shall actually prove.

(iv) The stability time $\exp(\varepsilon^{-1/n})$ of (1.6) is the same which one finds if $\overline{\omega}$ satisfies the Diophantine condition $|\overline{\omega} \cdot \nu| \geq \mathrm{const}\, |\nu|^{-\tau}$, in the limit case $\tau = n-1$ (see [14]). In the Diophantine case, however, there is a much better confinement of the actions, of order ε.

C. On the proof. The proof of the Theorem is in a sense very traditional, being quite similar to the proof of Nekhoroshev theorem. There are however two differences. The first one is the use of the complex coordinates (w, z) instead of the action–angle coordinates. This is achieved with the technique introduced in ref. [3], which resides in the use of suitable 'Fourier series' in these coordinates which, at variance from Taylor series, exactly correspond (out of the origin) to the Fourier series in the angles. Such an idea, which in fact already appears in [20] in connection with an isochronous system, is fully described and exploited in ref. [3], so we do not further comment on it here.

The second difference concerns the meaning of the small parameter, which is important for the confinement of the actions. Proceeding exactly as in the proof of Nekhoroshev theorem, for any value of a cut-off N (which will be eventually chosen to depend on ε as $N \sim \varepsilon^{-1/(n+\mu)}$), we cover the frequency space by suitable 'resonant regions'. On account of the nondegeneracy of the unperturbed Hamiltonian $k(I)$, such a covering has a diffeomorphic image in the action space. We construct a Nekhoroshev normal form only in that particular resonant region to which the equilibrium belongs, and then use this normal form to confine the motions. In this way, one essentially finds that all motions which start within such resonant region remain inside it for times of order $\exp N$.[2] This process is highly nonuniform, since the size of the resonant regions rapidly grows (as a power of $1/\varepsilon$) with the number of resonances.

However, and this is delicate and crucial, although $\overline{\omega}$ is fixed, as a rule one does not know to which resonant region it does belong. The reason is that, unless $\overline{\omega}$ has *very* special arithmetic properties, by increasing the cut-off N (i.e., by considering initial data closer to the equilibrium, and exploring their motions on longer time scales), $\overline{\omega}$ does not remain in one and the same resonant region, but endless jumps from region to region. This fact, which is typical of Nekhoroshev's theory although is not often stressed, is understood by thinking of the effects of increasing the cutoff: on the one hand, new resonances must be taken into account, so that new resonant regions suddendly appear in the frequency space and $\overline{\omega}$ may happen to fall into one of them; on the other hand, the existing resonant regions get thinner, so that $\overline{\omega}$ can exit a resonant region and enter a less resonant one. For a generically given $\overline{\omega}$, this process is very difficult (if not just prohibitive) to describe. For further comments see ref. [3], where the same phenomenon also plays a crucial role.

In this situation, the confinement of the actions is obtained by observing that, if the motion starts at a distance from the equilibrium which is smaller than the radius of the smallest resonant region (so that it certainly starts within the resonant region of $\overline{\omega}$, whichever it is), then, for times of order $\exp N$, it will remain within the radius of the largest resonant region (which certainly contains the resonant region of $\overline{\omega}$). As it turns out, denoting by ε the size of the smallest resonant region (which is the nonresonant one), then the size of the largest one is order $\varepsilon^{\frac{1+\mu}{n+\mu}}$, which accounts for the estimate (1.8).[3]

---

[2] More precisely, the motions must start at a distance from the equilibrium smaller than a certain fraction of the radius of the resonant region. Also, the nonresonant region is by definition the complement of the union of all resonant regions, but here, we mean instead the small polydisc where the nonresonant normal form is constructed.

[3] The largest resonant region will be one of the $(n-1)$–fold resonant regions. This is due to the fact that we shall choose the free parameters of the theory (specifically, the parameter R entering the Theorem) in such a way that the equilibrium never belongs to the completely resonant region. The reason is that, in such a region, in the case $\mu = 0$, one would have a very poor confinement of the actions, order ε^0.

2. Proof of the Theorem

A. Preliminaries. We denote by $\| \ \|$ the Euclidean (or Hermitean) norm and by $| \ |^\infty$ the supremum norm. For $\nu \in \mathbb{Z}^n$ we write $|\nu| = \sum_j |\nu_j|$. We often denote by $\xi = (\xi_1, \ldots, \xi_{2n})$ the coordinates $(w, z) \in \mathbb{C}^{2n}$, and work in polydiscs

$$D_S = \left\{ \xi \in \mathbb{C}^{2n} : |\xi|^\infty \le S \right\}, \qquad S > 0 .$$

For any $S > 0$, let $\mathcal{A}_S$ be the set of all analytic functions $f : D_S \to \mathbb{C}$ which, like the Hamiltonians (1.1) and (1.2), satisfy the reality condition $f(w, -i\overline{w}) \in \mathbb{R}$. For any $\nu \in \mathbb{Z}^n$, let $\mathcal{A}_S^\nu$ be the subset of $\mathcal{A}_S$ consisting of all functions of the form $\hat{f}_\nu(iz_1w_1, \ldots, iz_nw_n)e_\nu(z, w)$, where the complex function $\hat{f}_\nu$ is analytic on the polydisc $\left\{ I \in \mathbb{C}^n : |I|^\infty \le S^2 \right\}$ and

$$e_\nu(w, z) = \prod_{j=1}^n \eta_{\nu_j}^{|\nu_j|}, \qquad \eta_{\nu_j} = \begin{cases} z_j & \text{for } \nu_i > 0 \\ 1 & \text{for } \nu_i = 0 \\ w_j & \text{for } \nu_i < 0 \end{cases} .$$

Correspondingly, for functions $f \in \mathcal{A}_S$ we define the "Fourier series"

$$f = \sum_{\nu \in \mathbb{Z}^n} f_\nu , \qquad f_\nu = \hat{f}_\nu e_\nu \in \mathcal{A}_S^\nu ,$$

and we use the norm

$$|f|_S = \sum_{\nu \in \mathbb{Z}^n} |f_\nu|_S^\infty , \qquad |f_\nu|_S^\infty = \sup_{\xi \in D_S} |f_\nu(\xi)| .$$

The functions f_ν will be called the *harmonics* of f.

We shall also deal with vector fields. If $f = \sum_\nu f_\nu \in \mathcal{A}_S$, we denote by F its Hamiltonian vector field and use the Fourier series

$$F = \sum_{\nu \in \mathbb{Z}^n} F_\nu ,$$

where the 'harmonic' F_ν is by definition the Hamiltonian vector field of f_ν. We also use the norm

$$\|F\|_S = \sum_{\nu \in \mathbb{Z}^n} \|F_\nu\|_S^\infty , \qquad \|F_\nu\|_S^\infty = \frac{1}{R} \max_{j=1,\ldots,2n} \left| \frac{\partial f_\nu}{\partial \xi_j} \right|_S^\infty , \tag{2.1}$$

where R is a positive parameter (the same entering the statement of the Theorem) which will be fixed in the sequel, at the very end of the proof; for the moment, only $R \le R_c$ is important.

Finally, let us consider the Hamiltonian $h = k + f$ as in (1.2), (1.3), and let

$$M = \sup_{I \in \mathbb{C}^n \setminus \{0\}} \frac{|AI|^\infty}{|I|^\infty} ; \tag{2.2}$$

for simplicity, we assume $M \geq m$, where m is the constant entering the convexity condition (1.4). Furthermore, we observe that, since the Taylor series for f in the Hamiltonian (1.2) begins with terms of order $s \geq 5$, for any $0 < S \leq R$ we have

$$|f|_S \leq \left(\frac{S}{R}\right)^s \mathcal{F} R^2 \, , \qquad \|F\|_S \leq \left(\frac{S}{R}\right)^{s-1} \mathcal{F} \, , \tag{2.3}$$

where F is the Hamiltonian vector field of f and

$$\mathcal{F} \; = \; \max\left(\|F\|_R, \, \frac{|f|_R}{R^2} \right) .$$

B. Resonant regions. We begin the proof of the Theorem by recalling the decomposition of the frequency space R^n into resonant regions, which is used in the proof of Nekhoroshev's theorem, with the improved estimates given by Pöschel [19].

For any given $N \geq 1$ we consider all the d-dimensional sublattices Λ of $\mathbb{Z}^n$, $0 \leq d \leq n$, called *N-lattices*, such that:

(i) Λ is generated by d vectors $\nu_1, \ldots \nu_d \in \mathbb{Z}^n$ with $|\nu_j| \leq N$.

(ii) Λ is maximal (it is not properly contained in any sublattice of the same dimension).

We denote by 0 the zero–dimensional lattice, constituted by the null vector alone. The cells of any lattice $\Lambda \neq 0$ have a minimal d–dimensional euclidean volume, which we denote $\|\Lambda\|$; we put $\|0\| = 1$. Moreover, we denote by ω^Λ the orthogonal projection of a vector ω onto a lattice $\Lambda \neq 0$.

The definition of the resonant regions depends on the cutoff N and on two positive parameters b and δ. Specifically, following [19], for any N–lattice Λ of dimension $d \geq 1$ one defines

$$\delta_\Lambda \; = \; (b\,N)^{d-1} \, \frac{\delta}{\|\Lambda\|} \, . \tag{2.4}$$

For any N-lattice Λ of dimension $d = 0, \ldots, n$, one then defines the resonant region B_Λ as the set of all points $\omega \in \mathbb{R}^n$ such that

$$\begin{aligned} \|\omega^\Lambda\| \; &< \; \delta_\Lambda \\ \|\omega^{\Lambda'}\| \; &\geq \; \delta_{\Lambda'} \qquad \text{for any } N\text{-lattice } \Lambda' \text{ of dimension } d+1 \end{aligned} \tag{2.5}$$

(only the former condition for $d = n$, only the latter for $d = 0$). As is clear, for any given $N \geq 1$, the resonant regions cover $\mathbb{R}^n$.

Lemma 1 ([19]): *Consider any $N \geq 1$ and assume $b > \sqrt{2}$. Then, for any N–lattice $\Lambda \neq \mathbb{Z}^n$ and for any $\omega \in B_\Lambda$, one has*

$$|\omega \cdot \nu| \geq \gamma_\Lambda \qquad \forall \nu \in \mathbb{Z}^n \setminus \Lambda \, , \; |\nu| \leq N \, ,$$

with

$$\begin{aligned} \gamma_0 \; &= \; \delta \\ \gamma_\Lambda \; &= \; (b - \sqrt{2}) \, N \, \delta_\Lambda \quad \textit{if} \quad \Lambda \neq 0 \, , \, \mathbb{Z}^n . \end{aligned} \tag{2.6}$$

C. The normal forms. We now construct a normal form for the Hamiltonian (1.2), adapted to the resonance properties of $\overline{\omega}$, up to an exponentially small remainder:

Lemma 2. *Let $k(I) = \overline{\omega} \cdot I + \frac{1}{2} I \cdot AI$ and assume that $f \in \mathcal{A}_R$, $R > 0$, satisfies (2.3). Consider any $N \geq 2$ and let the N–lattice Λ be such that $\overline{\omega} \in B_\Lambda$; let $r = [N/2]$. Consider any positive number $R_\Lambda \leq R$ such that*

$$R_\Lambda^2 \leq \frac{\gamma_\Lambda}{8MN}, \qquad \left(\frac{R_\Lambda}{R}\right)^{s-4} \leq \frac{MR^2}{8\mathcal{F}}. \tag{2.7}$$

Then, there exists a symplectic diffeomorphism $\Phi : D_{R_\Lambda/2} \to \Phi(D_{R_\Lambda/2}) \subset D_{R_\Lambda}$ which satisfies

$$\left|\Phi(\xi) - \xi\right|^\infty \leq \frac{\mathcal{F}}{4MRr}\left(\frac{R_\Lambda}{R}\right)^{s-3} \leq \frac{R_\Lambda}{32r} \tag{2.8}$$

and is such that $(k+f) \circ \Phi$ belongs to $\mathcal{A}_{R_\Lambda/2}$ and has the form

$$k + 2g + e^{-r} f' \tag{2.9}$$

with $g = \sum_{\nu \in \Lambda} g_\nu$ and, denoting by F' the Hamiltonian vector field of f',

$$\max\left(|g|_{R_\Lambda/2}, |f'|_{R_\Lambda/2}\right) \leq \left(\frac{R_\Lambda}{R}\right)^s R^2 \mathcal{F}, \qquad \|F'\|_{R_\Lambda/2} \leq \left(\frac{R_\Lambda}{R}\right)^{s-1} \mathcal{F}. \tag{2.10}$$

Proof of Lemma 2. The proof of Lemma 2 is quite standard, except for the usage of the coordinates (w, z). Hence, we just skecth this proof here. As usual, we construct the diffeomorphism Φ as a composition of r elementary diffeomorphisms. The single step is described in the following

Lemma 3. *Let k, R, N and Λ be as in Lemma 2. Consider two positive numbers x and S such that $xR \leq S \leq R$ and two functions $u, v \in \mathcal{A}_S$ with $u = \sum_{\nu \in \Lambda} u_\nu$. Denote by U and V the Hamiltonian vector fields of u and v, respectively, and assume*

$$S^2 \leq \frac{\gamma_\Lambda}{4MN}, \qquad \frac{16}{x\,\gamma_\Lambda}\|V\|_S \leq 1.$$

Then, there exists a symplectic diffeomorphism $\Psi : D_{S-xR} \to \Psi(D_{S-xR}) \subset D_S$, analytic together with its inverse, which satisfies

$$\left|\Psi(\xi) - \xi\right|^\infty \leq \frac{2R}{\gamma_\Lambda}\|V\|_S \leq \frac{R_\Lambda}{8\,x} \tag{2.11}$$

and is such that $(k+u+v) \circ \Psi$ belongs to $\mathcal{A}_{S-xR}$ and has the form

$$k + \Big(u + \sum_{\nu \in \Lambda, |\nu| \leq N} v_\nu\Big) + v'$$

with v' and its Hamiltonian vector field V' satisfying

$$|v'|_{S-xR} \;\le\; \frac{\beta}{x}\,|v|_S\,, \qquad\qquad \|V'\|_{S-xR} \;\le\; \frac{\beta}{x}\,\|V\|_S \tag{2.12}$$

where $\beta = \max\left(\frac{8}{\gamma_\Lambda}\big(\|U\|_S + 2\|V\|_S\big), \frac{S}{[N]eR}\right)$.

Proof of Lemma 3. We generate Ψ by the Lie method, that is, as the time–one map Φ_1^X of the flow of a Hamiltonian vector field X. (We use the version of the method for vector fields described in [6]). Specifically, if the Hamiltonian χ of X satisfies the so–called homological equation

$$\{k,\chi\} \;=\; \sum_{\nu\notin\Lambda,|\nu|\le N} v_\nu \tag{2.13}$$

then $(k+u+v)\circ\Phi_1^X$ has the required form, with

$$v' \;=\; R_1^X(u+v) \;+\; R_2^X(k) \;+\; v^{>N}\,; \tag{2.14}$$

here $R_p^X \;=\; \sum_{j=p}^{\infty}\frac{1}{j!}L_X^j$ denotes the p–th remainder of the Lie series, L_X being the Lie derivative associated to X, and $v^{>N} = \sum_{|\nu|>N} v_\nu$ is the 'ultraviolet' part of v.

Equation (2.13) is solved by

$$\chi \;=\; \sum_{\nu\notin\Lambda,|\nu|\le N} \frac{v_\nu}{i\omega(I)\cdot\nu}\,,$$

where $\omega(I) = \frac{\partial k}{\partial I}(I) = \overline{\omega} \;+\; AI$. Since

$$|\omega(I)\cdot\nu| \;\ge\; \big|\overline{\omega}\cdot\nu\big| - \big|AI\cdot\nu\big| \;\ge\; \gamma_\Lambda - MS^2N \;\ge\; \frac{3}{4}\gamma_\Lambda\,, \tag{2.15}$$

one has $|\chi|_\sigma \le \frac{4}{3\gamma_\Lambda}\,|v^{\le N}|_\sigma$, with $v^{\le N} \;=\; v - v^{>N}$. We now show that the Hamiltonian vector field X of χ satisfies

$$\|X\|_S \;\le\; \frac{2}{\gamma_\Lambda}\,\|V^{\le N}\|_S\,, \tag{2.16}$$

$V^{\le N}$ being the Hamiltonian vector field of $v^{\le N}$, which in turn implies (2.11). Indeed, using the equality $\nu_j v_\nu \;=\; -w_j V_\nu^{z_j} \;-\; z_j V_\nu^{w_j}$ (where $V_\nu^{\xi_j}$ is the ξ_j–component of V_ν), one verifies that

$$X_\nu^{\xi_j} \;=\; \frac{V_\nu^{\xi_j}}{i\omega\cdot\nu} \pm \frac{\xi_j}{(\omega\cdot\nu)^2}\sum_{l=1}^{n} A_{lj}\left(w_l V_\nu^{z_l} + z_l V_\nu^{w_l}\right)$$

with the minus sign if $j = 1,\ldots,n$ and the plus sign if $j = n+1,\ldots,2n$. Inequality (2.16) is obtained from here, observing that $|w_j|$, $|z_j| \le S$.

The estimates (2.12) are obtained by estimating the various terms of (2.14) and of the analogous expression for V'. For this, one needs the following estimates on the remainders of the Lie series:

$$\begin{aligned}
\|R_1^X(U+V)\|_{S-xR} \ &\le\ \frac{4}{x}\,\|X\|_S\,\|U+V\|_S \\
\|R_2^X(K)\|_{S-xR} \ &\le\ \frac{2}{x}\,\|X\|_S\,\|L_X K\|_S \ \le\ \frac{2}{x}\,\|X\|_S\,\|V\|_S \\
|R_1^X(u+v)|_{S-xR} \ &\le\ \frac{4}{x}\,\|U+V\|_S\,|\chi|_S \\
|R_2^X(k)|_{S-xR} \ &\le\ \frac{4}{x}\,\|X\|_S\,|L_X k|_S \ \le\ \frac{4}{x}\,\|X\|_S\,|v^{\le N}|_S\,.
\end{aligned}$$

The proof of these inequalities is quite standard (some hints are given in [3]). One also needs to estimate the ultraviolet part of functions and vector fields. Recalling that the maximum of an analytic function is attained at the border, one computes

$$\big|f_\nu\big|^\infty_{S-xR} \ =\ \big|\hat f_\nu e_\nu\big|^\infty_{S-xR} \ \le\ \big|\hat f_\nu\big|^\infty_S\, S^{|\nu|}\Big(1-\frac{xR}{S}\Big)^{|\nu|} \ =\ \big|f_\nu\big|^\infty_S\Big(1-\frac{xR}{S}\Big)^{|\nu|}\,,$$

so that

$$\big|f^{>N}\big|_{S-xR} \ \le\ \Big(1-\frac{xR}{S}\Big)^{1+[N]}\big|f^{>N}\big|_S\,.$$

Similarly, observing that, if $f\in\mathcal{A}^\nu_S$, then $\frac{\partial f_\nu}{\partial w_j}\in\mathcal{A}^{\nu+\delta_j}_S$ and $\frac{\partial f_\nu}{\partial z_j}\in\mathcal{A}^{\nu-\delta_j}_S$, where $\delta_j=(\ldots,0,1,0\ldots)$ (1 at the j–th place), one finds $\|F\|_{S-xR}\le(1-xR/S)^{[N]}\|F^{>N}\|_S$. In (2.12), we also used $(1-y)^N\le(eNy)^{-1}$, $y>0$. ■

Lemma 2 is proven by iterating $r=[N/2]$ times the normalization procedure of Lemma 3, each time with $x=R_\Lambda/(2rR)$, beginning with $S=R$ at the first step. One finds that the iteration is possible provided that

$$R_\Lambda^2 \ \le\ \frac{\gamma_\Lambda}{8MN}\,, \qquad\qquad \frac{2^7\,r\,R}{\gamma_\Lambda\,R_\Lambda}\,\|F\|_{R_\Lambda} \ \le\ 1\,, \tag{2.17a,b}$$

and in such a case one has $|\Phi(\xi)-\xi|^\infty\le\frac{4R}{\gamma_\Lambda}\|F\|_{R_\Lambda/2}$, $\max\big(|g|_{R_\Lambda/2},|f'|_{R_\Lambda/2}\big)\le|f|_{R_\Lambda}$, and $\|F'\|_{R_\Lambda}\le\|F\|_{R_\Lambda}$. Using the fact that the Taylor series of f begins with terms of order s, i.e., inequalities (2.3), one verifies (using also (2.17a)) that inequality (2.17b) can be replaced by (2.7b), while the estimates above on Φ and on the normal form reduce to (2.8) and (2.10), respectively. ■

D. Confinement of the actions. The next step in the proof is using the normal form of Lemma 2, together with the convexity of k, so as to produce bounds on the variation of the actions. Here too the use of the coordinates (w,z) introduces only a few changes with respect to the usual proof which uses action–angle coordinates, and we limit ourselves to a sketch.

Lemma 4 *Within the hypotheses and the notation of Lemma 2, assume that* $\overline{\omega} \in B_\Lambda$ *and that the quadratic form* $k(I)$ *satisfies the convexity condition (1.4). Consider any motion* $t \mapsto (w_t, z_t)$ *of the system (1.2) with 'real' initial data* $(w_0, z_0) = (w_0, -i\overline{w}_0)$*.*

(i) If $\Lambda = 0$*, then*

$$|w_0|^\infty \ \le \ \frac{R_0}{8} \quad \Longrightarrow \quad |w_t|^\infty \ \le \ \frac{R_0}{2} \qquad \text{for} \quad |t| \ \le \ \frac{e^{[N/2]}}{4\mathcal{F}} \,. \tag{2.18}$$

(ii) If $\Lambda \ne 0, \mathbb{Z}^n$ *and*

$$R_\Lambda^2 \ \ge \ \frac{32}{m}\, \delta_\Lambda \tag{2.19a}$$

$$\Big(\frac{R_\Lambda}{R}\Big)^{s-4} \ \le \ \frac{mR^2}{2^{12}\mathcal{F}} \, \min\left(1,\ \frac{2^9}{n^{1/4}}\sqrt{\frac{M}{m}}\right)\,, \tag{2.19b}$$

then

$$|w_0|^\infty \ \le \ \frac{R_\Lambda}{8n^{1/4}}\sqrt{\frac{m}{M}} \quad \Longrightarrow \quad |w_t|^\infty \ \le \ \frac{R_\Lambda}{2} \qquad \text{for} \quad |t| \ \le \ \frac{2\, e^{[N/2]}}{\Omega} \tag{2.20}$$

where $\Omega = \sqrt{n}\, \sup_{|I|\le R} \|\overline{\omega} + AI\|$*.*

Proof. Let us prove statement (ii). Let Φ be the symplectic diffeomorphism constructed in Lemma 2, and write $(w,z) = \Phi(w',z')$, $I' = iw'z'$, $\Delta I = I'_t - I'_0$, $\Delta k = k\big(I'_t\big) - k\big(I'_0\big)$, and $\omega_0 = \omega\big(I'_0\big) = \overline{\omega} + AI'_0$. We assume that (w_t, z_t) and (w'_t, z'_t) remain within $D_{R_\Lambda/2}$ for all the times we deal with (the consistency of this assumption can be verified at the end of the proof). The convexity condition (1.4) gives $\frac{m}{2}\, \|\Delta I\|^2 \le \big|\Delta k\big| + \big|\omega_0 \cdot \Delta I\big|$. The conservation of energy for the Hamiltonian (2.9) implies $|\Delta k| \le 6|f|_{R_\Lambda}$, so that, writing also $\omega_0 = \omega_0^\Lambda + (\omega_0 - \omega_0^\Lambda)$, one obtains

$$\frac{m}{2}\, \|\Delta I\|^2 \ \le \ 6\Big(\frac{R_\Lambda}{R}\Big)^s R^2\mathcal{F} \ + \ \|\omega_0^\Lambda\|\, \|\Delta I\| \ + \ |(\omega_0 - \omega_0^\Lambda)\cdot \Delta I|\,.$$

Now, using (2.9) and (2.10c), observing that $\{I'_j, k\} = 0$ for all j, that $\{I'_j, g\} = \sum_{\nu\in\Lambda} i\nu_j g_\nu e_\nu$ are the components of a vector parallel to Λ, and that $\{I'_j, f'\} = iz'_j (F')^{z_j} + iw'_j (F')^{w_j}$, one computes

$$\begin{aligned} |(\omega_0 - \omega_0^\Lambda)\cdot \Delta I| \ &= \ \Big|\sum_{j=1}^n (\omega_0 - \omega_0^\Lambda)_j \int_0^t \{I'_j, e^{-[N/2]} f'\}(s)ds\Big| \\ &\le \ |t|\, \sqrt{n}\, \|\omega_0\|\, e^{-[N/2]} \, \max_j \big|\{I'_j, f'\}\big|^\infty_{R_\Lambda/2} \ \le \ 2\Big(\frac{R_\Lambda}{R}\Big)^s R^2\, \mathcal{F} \end{aligned}$$

for $|t| \le 2\,\Omega^{-1} e^{[N/2]}$. Hence, on this time scale, $\frac{m}{2}\, \|\Delta I\|^2 \le 8\Big(\frac{R_\Lambda}{R}\Big)^s R^2\mathcal{F} + \|\omega_0^\Lambda\|\, \|\Delta I\|$. Taking also into account inequality (2.19b), this inequality is seen to imply

$$\|\Delta I\| \ \le \ \frac{2}{m}\, \|\omega_0^\Lambda\| \ + \ \frac{R_\Lambda^2}{16}\,.$$

Now, since $\overline{\omega} \in B_\Lambda$, one has $\|\omega_0^\Lambda\| \le \|\omega_0 - \overline{\omega}\| + \|\overline{\omega}^\Lambda\| \le \sqrt{n}\, M\, |I_0'|^\infty + \delta_\Lambda$, so that, using also $|I_t'| \le |I_0'| + |\Delta I|$, $m \le M$ and (2.19a),

$$|I_t'|^\infty \ \le\ \frac{3M\sqrt{n}}{m}\, |I_0'|^\infty \ +\ \frac{R_\Lambda^2}{8}\,.$$

This implies $|w_t'|^\infty \le R_\Lambda/2$ and $|w_t|^\infty \le R_\Lambda/2$, provided that $|w_0'|^\infty \le \sqrt{m/2M}R_\Lambda/(4n^{1/4})$. Using (2.8) and (2.19b), one sees that the latter inequality follows from the hypothesis (2.20) on w_0. The proof of statement (i) is similar, and actually easier. ∎

E. Proof of the Theorem. We now show that, if we make the choice

$$R \ =\ \min\Big(R_c,\ \frac{\sqrt{m\,\|\overline{\omega}\|}}{32\,M}\Big) \tag{2.21}$$

and moreover

$$\varepsilon_* \ =\ \min\Big(\frac{1}{64},\ \frac{1}{2\sqrt{n}}\Big)\,, \qquad \varepsilon_0 \ =\ \bigg(\Big(\frac{m}{2^8 M}\Big)^{n-1} a^{\frac{1}{s-4}}\bigg)^{\frac{\mu+n}{\mu+1}}\, \varepsilon_* \tag{2.22}$$

with

$$a \ =\ \min\bigg(1,\ \frac{mR^2}{2^{12}\mathcal{F}}\ \min\Big(1,\ \frac{2^9}{n^{1/4}}\sqrt{\frac{M}{m}}\Big)\bigg)\,,$$

then all the hypotheses for the applicability of Lemma 4 are satisfied whenever the initial data fulfil (1.5), and in such a case the estimates (2.20) and (2.18) of Lemma 4 imply (1.8) with

$$c \ =\ \frac{1}{8}\Big(2^9\frac{M}{m}\Big)^{n-1}\,, \qquad C \ =\ \min\Big(\frac{2}{\Omega},\ \frac{1}{4\mathcal{F}}\Big)\,, \tag{2.23}$$

with Ω as in Lemma 4.

As we shall see below, the restriction $R \le \sqrt{m\|\overline{\omega}\|}/32M$ in (2.21) serves to prevent $\overline{\omega}$ to belong to the completely resonant region $B_{\mathbb{Z}^n}$, i.e., to assure that

$$\|\overline{\omega}\| \ \ge\ \delta_{\mathbb{Z}^n} \ =\ (bN)^{n-1}\,\delta \tag{2.24}$$

(as already mentioned in the Introduction, the reason for this request is the very poor confinement of the actions in the completely resonant region, when $\mu = 0$). If (2.24) is verified, then Lemma 4 can be applied whenever the three inequalities $R_\Lambda \le R$ and (2.7a,b) are satisfied for any $\Lambda \ne \mathbb{Z}^n$, and the two inequalities (2.19a,b) are satisfied for any $\Lambda \ne 0, \mathbb{Z}^n$. The latter four inequalities read

$$R_\Lambda^2 \ \le\ \frac{\gamma_\Lambda}{8MN}\,, \qquad\qquad \Lambda \ne \mathbb{Z}^n \tag{2.25a}$$

$$\Big(\frac{R_\Lambda}{R}\Big)^{s-4} \ \le\ \frac{MR^2}{8\mathcal{F}}\,, \qquad\qquad \Lambda \ne \mathbb{Z}^n \tag{2.25b}$$

$$R_\Lambda^2 \ \ge\ \frac{32}{m}\,\delta_\Lambda\,, \qquad\qquad \Lambda \ne 0,\ \mathbb{Z}^n \tag{2.25c}$$

$$\Big(\frac{R_\Lambda}{R}\Big)^{s-4} \ \le\ \frac{mR^2}{2^{12}\mathcal{F}}\ \min\Big(1,\ \frac{2^9}{n^{1/4}}\sqrt{\frac{M}{m}}\Big)\,, \qquad\qquad \Lambda \ne 0\,, \mathbb{Z}^n\,. \tag{2.25d}$$

In order to fulfil these conditions, we begin by making the (rather natural) choice

$$R_0^2 \;=\; \frac{R^2}{N^{n+\mu}}\,, \qquad \mu \geq 0\,. \tag{2.26}$$

Recalling the definition (2.5) of γ_Λ, we choose δ (for $\Lambda = 0$) and R_Λ (for $\Lambda \neq 0, \mathbb{Z}^n$) so as to fulfil all conditions (2.25a) as equalities:

$$\begin{aligned} \delta \;&=\; 8MNR_0^2 \\ R_\Lambda^2 \;&=\; \frac{b-\sqrt{2}}{b\,\|\Lambda\|}\,(bN)^d\,R_0^2 \qquad\quad \text{if} \quad d \;=\; \dim\Lambda = 1,\ldots,n-1\,. \end{aligned}$$

Similarly, we choose the constant b so as to fulfil (2.25c) as equalities, too:

$$b \;=\; \sqrt{2} \;+\; 2^8\,\frac{M}{m}\,.$$

At this point, as one immediately checks, (2.24) follows from the choice (2.21) of R.

Let us now provisionally introduce the small parameter

$$\eta \;=\; \Big(\frac{R_0}{R}\Big)^2\,;$$

correspondingly, equation (2.26) can be regarded as defining the cutoff N:

$$N \;=\; \Big(\frac{1}{\eta}\Big)^{\frac{1}{n+\mu}}\,.$$

We can now fulfil the two conditions (2.25b,d), for all $\Lambda \neq 0, \mathbb{Z}^n$, just by requiring that R_0, namely η, is small enough, precisely

$$\eta^{\frac{1+\mu}{n+\mu}} \;\leq\; \frac{1}{b^{n-1}}\,\left(\frac{mR^2}{2^{12}\mathcal{F}}\,\min\Big(1, \frac{2^9}{n^{1/4}}\sqrt{\frac{M}{m}}\Big)\right)^{\frac{1}{s-4}} \tag{2.27}$$

(recall that $m \leq M$ and $d \leq n-1$). Finally, for $\Lambda \neq \mathbb{Z}^n$ one has $R_\Lambda^2 \leq b^{n-1}\eta^{(1+\mu)/(n+\mu)}R^2$ and so all inequalities $R_\Lambda \leq R$, $\Lambda \neq \mathbb{Z}^n$, are fulfilled if

$$\eta^{\frac{1+\mu}{n+\mu}} \;\leq\; \frac{1}{b^{n-1}}\,. \tag{2.28}$$

With the above choice of the parameters R, b, δ, and R_Λ, $\Lambda \neq \mathbb{Z}^n$, we may now apply the estimates of Lemma 4 with $\overline{\omega} \in B_\Lambda$ for some $\Lambda \neq \mathbb{Z}^n$. Since $R_0^2 \leq \frac{m}{2^8 M}\,R_\Lambda^2$ for all $\Lambda \neq 0\,, \mathbb{Z}^n$ (as one sees using $\|\Lambda\| \leq N^{n-1}$), it is easy to verify that, if

$$|I(0)|^\infty \;\leq\; \min\Big(\frac{1}{64}, \frac{1}{2\sqrt{m}}\Big)\,R_0^2 \;=\; \varepsilon_*\,\eta\,R^2 \tag{2.29}$$

with ε_* as in (2.22a), then $I(0)$ satisfies either the inequality (2.18) (if $\Lambda = 0$) or the inequality (2.20) (if $\Lambda \neq 0, \mathbb{Z}^n$) of Lemma 4, so that in all cases

$$|I(t)|^\infty \;\le\; \frac{R_\Lambda^2}{4} \;\le\; \frac{b^{n-1}}{4}\,\eta^{\frac{\mu+1}{\mu+n}}\,R^2 \;=\; c\,\eta^{\frac{\mu+1}{\mu+n}}\,R^2$$

with c as in (2.23), for $|t|$ as in (1.8). In the Theorem, we have used the parameter $\varepsilon = \varepsilon_*\eta$ instead than η; the two conditions (2.27) and (2.28) follow from $\varepsilon \le \varepsilon_0$, with ε_0 as in (2.22b). This concludes the proof of the Theorem.

F. Conclusive comment. As a final comment, we would like to discuss why the results do not significantly improve when $s > 5$. We used a generic s in all of the formulas above precisely to keep track of this fact: as the reader can see, the only improvement of a larger s is that the threesold ε_0 gets slightly larger, see (2.22b).

The reason of this fact is easily envisaged: our strategy was to confine the motions in the resonant region of $\overline{\omega}$, and so the confinemement of the action is determined solely by the size of the latter. Hence, in order to improve the confinement of the actions, one should allow the size of the resonant regions to be essentially smaller than the distance from the equilibrium. This requires a finer analysis, in which one further subdvides into different resonant regions what is here regarded as the resonant region of $\overline{\omega}$. This analysis will published elsewhere [12].

Acknowledgments: The idea of this work came during some conversations with Debra Lewis regarding the stability of the Riemann ellipsoids. We are very indebted to her for carrying the problem to our knowledge and for many stimulating discussions. F.F. also wishes to thank Tudor Ratiu for his ospitality at the Mathematics Department of the University of California, Santa Cruz, where part of this work was done, as well as for many useful discussions. This work has been supported by the grant *EC contract ERBCHRXCT940460* for the project *Stability and Universality in Classical Mechanics.* F.F. was also partially supported by DOE contract DEFG03-95ER25245-A000 while visiting UCSC.

Note added in proofs: After this paper had been submitted for publication, we received a preprint by L. Niederman (*Nonlinear stability around an elliptic equilibrium point in an Hamiltonian system*) which treats the same problem and obtains very similar (though not identical) results. From a technical point of view, Niederman uses Lochak's scheme for the proof of Nekhoroshev theorem (the so–called "simultaneous approximation" method), but implemented in cartesian coordinates.

References

[1] D. Bambusi, *Long time stability of some small amplitude solutions in nonlinear Schrdinger equations.* Comm. Math. Phys. **189**, 205–226 (1997).

[2] G. Benettin and G. Gallavotti, *Stability of motions near resonances in quasi–integrable Hamiltonian systems.* J. Stat. Phys. **44,** 293-338 (1986).

[3] 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).

[4] S. Chandrasekhar, *Ellipsoidal Figures of Equilibrium* (Yale University Press, New Haven, 1969).

[5] A. Delshams and P. Gutiérrez *Estimates on invariant tori near an elliptic equilibrium point of a Hamiltonian system.* J. Diff. Eq. **131**, 277–303 (1996).

[6] F. Fassò, *Lie series method for vector fields and Hamiltonian perturbation theory.* J. Appl. Math. Phys. (ZAMP) **41**, 843-864 (1990).

[7] F. Fassò and D. Lewis, paper in preparation.

[8] 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).

[9] A. Giorgilli and L. Galgani, *Rigorous estimates for the series expansions of Hamiltonian perturbation theory.* Celestial Mechanics **37,** 95-112 (1985).

[10] A. Giorgilli, *Rigorous results on the power expansions for the integrals of a Hamiltonian system near an elliptic equilibrium point.* Ann. Inst. Henri Poincaré - Physique Thèorique **48**, 423-439 (1988)

[11] 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)

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

[13] H. Ito, *Convergence of Birkhoff normal forms for integrable systems.* Comment. Math. Helvetici **64,** 412-461 (1989).

[14] À. Jorba and J. Villanueva, *On the normal behaviour of partially elliptic lower dimensional tori of Hamiltonian systems.* Preprint (1996)

[15] D. Lewis, *Bifurcation of liquid drops.* Nonlinearity **6,** 491-522 (1993).

[16] P. Lochak, *Canonical perturbation theory via simultaneous approximation.* Russ. Math. Surv. **47**, 57-133 (1992)

[17] P. Lochak, *Stability of Hamiltonian systems over exponentially long times: the near linear case.* In H. Dumas, K. Meyer, D. Schmidt (eds), *Hamiltonian Dynamical Systems – History, Theory, and Applications*, The IMA Volumes in Mathematics and its Applications **63**, 221-229 (Springer, New York, 1995).

[18] 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)].

[19] J. Pöschel, *Nekhoroshev estimates for quasi–convex Hamiltonian Systems,* Math. Z. **213**, 187–216 (1993).

[20] C. Siegel and J. Moser, *Lectures on Celestial Mechanics* (Springer, Berlin, 1971).