Commun. Math. Phys. 186, 413 – 449 (1997)

Communications in
Mathematical Physics

On the Stability of Realistic Three-Body Problems*

Alessandra Celletti[1], **Luigi Chierchia**[2]

[1] Dipartimento di Matematica, Università dell'Aquila, 67100–Coppito, L'Aquila, Italy.
E-mail: celletti@axscaq.aquila.infn.it
[2] Dipartimento di Matematica, Università di Roma Tre, Largo San Murialdo 1, 00146 Roma, Italy.
E-mail: luigi@matrm3.mat.uniroma3.it

Received: 1 April 1996 / Accepted: 25 October 1996

Abstract: We consider the system Sun–Jupiter–Ceres as an example of a planar, circular, restricted three-body problem and, after substituting the mass ratio of Jupiter/Sun (which is approximately 10^{-3}) with a parameter ε, we prove the existence of stable quasi-periodic motions with frequencies close to the observed (average) frequencies reported in "The Astronomical Almanac" for $|\varepsilon| \leq 10^{-6}$. The proof is "computer-assisted".

Table of Contents

* This work is part of the research program of the European Network on: "Stability and Universality in Classical Mechanics", # ER–BCHRXCT940460.

1. Introduction and Theorem 1.1

1) Since human kind developed a definite mathematical taste (whatever this means) the "stability" of planetary motions might be considered as one of the central questions in mathematics. Nowadays consciousness about new phenomena (pollution, just to name one) has drawn the attention of scientists and non-scientists to other types of stabilities (in other words, extinction of living species will depend more and more upon the chaotic effects of pollution rather than on "the sky falling on our heads"). Nevertheless the stability problem for many-body systems interacting only through gravitation still stems out as one of the more intriguing and rich problems in mathematics. In modern times outstanding contributions came, above all, from H. Poincaré [13], V. I. Arnold [2] and J. Moser [12]. In particular the so-called KAM (Kolmogorov, Arnold, Moser) theory (see [1] and references therein) gave a "positive" answer to the above stability problem in the sense that it proved ([2]) the possibility of the existence of many-body systems ("planetary systems") whose time evolution may be described by a linear flow on a torus ("quasi-periodic motion," here, synonymous of "stable motion"). The drawback of this beautiful result is that, taking the estimates contained in it seriously, it turns out that the mass ratio of the planets of such a hypothetical planetary system with their star should be more or less comparable to the mass ratio of a proton with the Sun.[1]

One then may pose the question of stability of *realistic many-body systems*.

Of course, to give a mathematical content to the word "realistic" is clearly impossible and the best we could do was to get inspiration from our own planetary system. The "simplest non-trivial" three-body problem is the so-called *planar, circular, restricted three-body problem* (see Sect. 2 for definitions): we considered one of the most popular three-body problems of the Solar system, namely Sun, Jupiter and Ceres (one of the major bodies of the so-called asteroid belt). Following Delaunay [6], we then derived a Hamiltonian model. Clearly, in deriving a model one simplifies Nature[2] quite a bit. We then took our model seriously (from the mathematical point of view) and replacing the Jupiter/Sun mass ratio, which is approximately 10^{-3}, with a "perturbative parameter" ε, we asked for how large values of ε one could find quasi-periodic motions with "frequencies" close to the observed frequencies of the Sun–Jupiter–Ceres system (which may be found in the ephemeris [16]).

For such a model we proved stability for $|\varepsilon| \leq 10^{-6}$, being therefore away from "reality" by three orders of magnitude. We leave it to the reader to judge if this is realistic or not. We believe, however, that with some more efforts one should indeed be able to prove stability up to $\varepsilon = 10^{-3}$ and we regard our result as a first step towards a proof of the mathematical stability of realistic many-body problems.

Our result relies basically on two techniques: (i) a (new) KAM scheme presented in Sect. 4 (for the experts: a KAM result in Hamiltonian setting in the style of Moser, Salamon and Zehnder [15] with emphasis on analytical dependence upon parameters); (ii) computer-assisted (rigorous) estimates, which are needed in order to apply "effectively" the KAM scheme to our three-body problem.

It is well known that computers may be used to prove theorems (see, e.g., [7, 11, 10] or think of the famous "four-colour theorem"). We were not really enthusiastic to rely

[1] Mass of proton/mass of Sun= $1.6724 \cdot 10^{-21}/8.4078 \cdot 10^{30} \simeq 1.9891 \cdot 10^{-52}$; compare [8]. For some recent applications of Arnold's result to three-body problems see [14].

[2] For example, to consider Sun–Jupiter–Ceres as a planar, *circular*, restricted three-body problem means, in particular, that one is assuming the Jupiter orbit circular: this is a rather crude approximation and more "realistic" models would include the Jupiter eccentricity, the Saturn "secular" effects on Jupiter, etc. See Sect. 3 for a (partial) justification of our model.

on machines to prove our result but couldn't get away without it. In fact we are pretty sure that with more refined techniques and/or with new ideas one might get better results without the (essentially trivial but) lengthy computations which are the only reason to call in machines.

2) We give now a precise **formulation of the main result**. Let $(\ell, g) \in \mathbb{T}^2 \equiv \mathbb{R}^2/(2\pi\mathbb{Z}^2)$; let[3]

$$L_0 \equiv 0.729305 \,, \quad G_0 \equiv 0.727162 \,, \quad r_0 \equiv 0.001 \,, \tag{1.1}$$

and let

$$B \equiv \{(L, G) \in \mathbb{C}^2 : \; |L - L_0| \le r_0 \,, \; |G - G_0| \le r_0\}, \quad B_0 \equiv B \cap \mathbb{R}^2 \,. \tag{1.2}$$

On $\mathbb{T}^2 \times B_0$, endowed with the standard symplectic form $d\ell \wedge dL + dg \wedge dG$, consider the one-parameter family of Hamiltonian functions given by

$$\begin{aligned} H(\ell, g, L, G; \varepsilon) \; &\equiv \; (\frac{1}{2L^2} - G)^2 \; + \; 2\varepsilon \, (\frac{1}{2L^2} - G) \, R(\ell, g, L, G) \\ &\equiv \; h(L, G) + \varepsilon \, f(\ell, g, L, G) \,, \end{aligned} \tag{1.3}$$

with the "perturbing function" R defined as

$$R(\ell, g, L, G) \; \equiv \sum_{\substack{n \in \mathbb{Z}^2 \\ 0 \le |n_1| + |n_2| \le 10}} R_n(L, G) \cos(n_1 \ell + n_2 g), \tag{1.4}$$

where $R_n \equiv R_{n_1 n_2}$ vanishes unless it belongs to the following list:

$$\begin{aligned} R_{00} \; &\equiv \; \frac{L^4}{4}(1 + \frac{9}{16}L^4 + \frac{3}{2}e^2) \,, \qquad R_{10} \equiv -\frac{L^4 e}{2}(1 + \frac{9}{8}L^4) \,, && (1.5) \\ R_{11} \; &\equiv \; \frac{3}{8}L^6(1 + \frac{5}{8}L^4) \,, \qquad R_{12} \equiv -\frac{L^4 e}{4}(9 + 5L^4) \,, && (1.6) \\ R_{22} \; &\equiv \; \frac{L^4}{4}(3 + \frac{5}{4}L^4) \,, \qquad R_{32} \equiv \frac{3}{4}L^4 e \,, \\ R_{33} \; &\equiv \; \frac{5}{8}L^6(1 + \frac{7}{16}L^4) \,, \qquad R_{44} \equiv \frac{35}{64}L^8 \,, \\ R_{55} \; &\equiv \; \frac{63}{128}L^{10} \,, && (1.7) \end{aligned}$$

and the "eccentricity" e, which is a function of the "action variables" (L, G), is defined as[4]

$$e \equiv e(L, G) \equiv \sqrt{1 - \frac{G^2}{L^2}} \,.$$

Let α be the golden mean ($\alpha \equiv \frac{\sqrt{5}-1}{2}$) and let

$$\Omega_- \equiv \frac{5}{2} + \frac{1}{13 + \alpha} = 2.573432\ldots \,, \qquad \Omega_+ \equiv \frac{5}{2} + \frac{1}{12 + \alpha} = 2.579251\ldots \,.$$

[3] All the following numbers, quantities and functions will be physically motivated in Sects. 2 and 3.

[4] In this paper the letter e will always refer to "eccentricities" and never to the Neper number; the exponential function will be denoted $\exp(x)$.

The "observed average frequency" of Ceres is approximately $-\Omega_c \equiv -2.577107$ (see Sect. 3 below) so that $\Omega_- < \Omega_c < \Omega_+$. Let

$$L_\pm \equiv \Omega_\pm^{-1/3}\,, \qquad G_\pm \;\equiv\; L_\pm\,\sqrt{1-e_0^2}\,,$$

where $e_0 = 0.0766$ is the "observed" eccentricity of Ceres as found in [16]. Finally, we define the h-frequencies

$$\omega^{(\pm)} \;\equiv\; (E_\pm\Omega_\pm,\; E_\pm)\,, \qquad E_\pm \equiv -2(\frac{1}{2L_\pm^2} - G_\pm)\,. \tag{1.8}$$

For later use we point out that $\omega^{(\pm)}$ is a "Diophantine vector" satisfying

$$|\omega^{(\pm)}\cdot n| \geq (\gamma_\pm |n|)^{-1}\,, \quad \forall\, n \in \mathbb{Z}^2\backslash\{0\}\,, \tag{1.9}$$

with $\gamma_\pm$ given by[5]

$$\gamma_\pm = 2|E_\pm|(\sqrt{5} + 24 \pm 1)\,. \tag{1.10}$$

The result discussed in 1) above can now be formulated as follows.

Theorem 1.1. *Let H be as in* (1.3)–(1.5)*; let B_0 be as in* (1.1), (1.2)*; let $\omega^{(\pm)}$ be as in* (1.8)*. Then, for all $0 \leq |\varepsilon| \leq 10^{-6}$ there exist (unique) two-dimensional analytic tori $\mathcal{S}_\varepsilon(\omega^{(\pm)}) \subset B_0 \times \mathbb{T}^2$, depending analytically also on the parameter ε (for $|\varepsilon| \leq 10^{-6}$), on which the H-flow is (analytically) conjugated to the linear flow $\theta \in \mathbb{T}^2 \to \theta - \omega^{(\pm)}t$.*

Remark 1.1. In Sect. 3 we give a physical motivation for having chosen as a perturbating function a trigonometric polynomial. We believe, however, that considering perturbations with an infinite number of non-vanishing Fourier coefficients would lead to essentially the same results making only technically more involved (and more expensive) the proof.

3) (On the proof) The proof of Theorem 1.1 is given in Sect. 5 and, as already mentioned, is a "computer assisted" application of the KAM scheme[6] of Sect. 4.

As it is well known KAM schemes are "Newton algorithms": they are procedures to iteratively construct solutions for certain nonlinear equations (with "loss of regularity"), starting from some initial "approximate solution," with a quadratic rate of convergence. *Our initial approximate solution is a suitable truncation* (actually a "fifth order truncation") *of the so-called "Lindstedt series"* (see [1] for generalities), i.e., of the formal ε-power series solution for the invariant torus equation associated to the looked for quasi-periodic solution. To this initial datum we apply the "KAM algorithm" presented in Sect. 4. The "KAM algorithm" is based on an algebraic scheme which, starting from a given "approximate solution," produces a new function solving the invariant torus equation up to an error which is "quadratically smaller" than the one produced by the starting approximate solution (Sect. 4.1). This algebraic scheme (which, as already mentioned, is new) is equipped with a set of "accurate" estimates (Sect. 4.3). The algebraic scheme plus the set of estimates is what we call "the KAM algorithm." We then work out a criterion (the KAM Theorem of Sect. 4.4) which guarantees the applicability of the KAM

[5] In general, numbers of the form $a = \frac{p}{q} \pm \frac{1}{k+\alpha}$ with p, q, k non negative integers, $q > 0$ and $k \geq 2$, satisfy $|an + m| \geq (\gamma|n|)^{-1}$ for any $n, m \in \mathbb{Z}$, $n \neq 0$ with $\gamma = q^2(k + \alpha)$ (see, e.g., [3]).

[6] See [4] and [5] for general information, references and a different "KAM computer-assisted algorithm".

algorithm an infinite number of times yielding a solution of the invariant torus equation. Such criterion is obtained simplifying the estimates and getting a unique stronger condition (the "KAM condition") ensuring the indefinite applicability of the scheme.

As we already pointed out in previous papers (see [5] and references therein), in concrete applications, it is convenient to iterate a few times the KAM algorithm *before* trying to apply the KAM theorem.

Both the computation of the initial approximate solution and the application of the KAM algorithm are computer-assisted.

We remark that it is quite different to explicitly compute the *initial* approximate solution from the "computation" of the new approximate solutions based on the KAM algorithm: the calculation of the truncated Lindstedt series is completely explicit (we compute numbers!) while the construction of the sequence of quadratically better and better approximate solutions is only *implicitly described* by the KAM algorithm and what we actually compute are bounds on norms relative to such approximate solutions.

4) (On the use of computers) Our proof is "computer-assisted" in the sense that certain formulae (derived below) have been implemented on a computer (a VAX) keeping rigorous control, by means of the so-called "interval arithmetic" (see below), of the numerical errors introduced by the machine. We report in Sects. 5.2 and 5.3 all the computer-aided calculations needed to prove Theorem 1.1. Instead we do not include the computer program which anybody can write by her/himself.[7] We are obviously aware of the (phylosofical?) problem of proceeding in such a way: It is clear that writing a program in a slightly different way or using different machines might (better: will) produce slightly different outputs, which in our case are *intervals of rational numbers* (see below). However we regard the computer-implementation, *once it is clearly settled the type of rigorous method used to control the propagation of numerical errors* (here "interval arithmetic") as a detail at the same level, say, of the details needed to work out explicitly the estimates of the KAM algorithm of Sects. 4.3 and 4.4. Of course, we shall be happy to send to interested readers the computer programs contained in this paper.

Let us now briefly discuss *interval arithmetic* which is the technical tool we used to control the numerical errors introduced by the machines. Real numbers are represented by computers as sign-exponent-fraction quantities, with the length of the exponent and of the fraction depending on the machine. Any result among elementary operations (sum, subtraction, multiplication and division) is rounded by the computer up to a certain decimal digit. To rigorously implement on a computer a certain sequence of formulae, one first reduces such formulae to a sequence of elementary operations.[8] The idea of the "interval arithmetic" is then to construct an interval (exactly representable on the computer) containing the *exact* result of an elementary operation and to replace (in the obvious way) algebra on numbers with algebra on intervals. In our FORTRAN 77 programs we define quadruple precision (H-floating) variables, which are allowed to vary in a range between $0.84 \cdot 10^{-4392}$ and $0.59 \cdot 10^{4392}$. The binary structure of a quadruple precision datum is composed by 128 bits, with 1 sign bit, 15 bits for the exponent and the remaining bits for the fraction. Two extra hidden *guard* bits are used to guarantee the result of an elementary operation "up to 1/2 of the last significant bit" ([17]). The interval containing the result of an elementary operation is therefore obtained increasing or decreasing by one bit the last bit of the mantissa, eventually taking care of

[7] The (11189–line) computer program is "just" a translation in computer language (FORTRAN 77) of the formulae of Sects. 5.1 and 4.3 after the standard arithmetic (basic operations) is replaced by "interval arithmetic" (the "arithmetic routines" may be found, e.g., at pages 153–158 of [4]).

[8] Elementary functions (such as roots, exponentials, trigonometric functions, etc.), will be approximated by a finite sequence of elementary operations using Taylor polynomials keeping track of errors; see [5] §8.3.

the propagation of the carry. For further information and for the necessary routines we refer to [4 and 5].

We finally mention that in the Appendix we report a few computer-assisted data with the following doublefold aim. From one side the reader reproducing our estimates might check her/his results with ours; from the other side the reader who is not going to waste time performing the computations will have an idea of the type of outputs one needs in this paper.

2. Restricted, Circular, Planar Three-Body Problem

Here we recall the Hamiltonian formulation of the "restricted, circular, planar three-body problem" (for general information see [6 or 1]).

Consider first a Keplerian two-body problem made up of two material points ("bodies") P_1 and P_2 with masses m_1 and m_2 and let P_2 revolve on a circular orbit around P_1. Consider now a third body A moving on the orbital plane of P_1 and P_2 and subject to the gravitational attraction of P_1 and P_2. Let the mass m_A of A be much smaller than m_1, m_2 and assume that the motion of P_1 and P_2 is not affected by A. The study of the dynamics associated to such a model is known in the literature as *the circular, planar, restricted three-body problem*. In particular, we shall be interested in phase space regions for which the resulting motion of A is a nearly circular orbit "around" P_1.

A convenient Hamiltonian formulation of such a three-body problem is based upon the classical "planar Delaunay variables" [6]. Let $\mathbb{T} \equiv \mathbb{R}/(2\pi\mathbb{Z})$ and consider the phase space[9]

$$\mathcal{P} = \{(\lambda, \gamma, \psi) \in \mathbb{T}^3\} \times \{(\Lambda, \Gamma, E) \in \mathbb{R}^3 : \Lambda \neq 0, |\Gamma| < |\Lambda|\}$$

endowed with the standard symplectic form $d\lambda \wedge d\Lambda + d\gamma \wedge d\Gamma + d\psi \wedge dE$. Then, the dynamics associated with the circular, planar restricted three-body problem is given by the Hamiltonian flow generated by the Hamiltonian[10]

$$H_0(\lambda, \gamma, \psi, \Lambda, \Gamma, E) \equiv \frac{1}{2\Lambda^2} + E + \varepsilon R_0(\lambda, \gamma - \psi, \Lambda, \Gamma) , \qquad (2.1)$$

where $\varepsilon \equiv m_2/m_1$ and the "perturbation function" R_0 is given as follows. Let $\nu \in \mathbb{T}$ (the "eccentric anomaly") be implicitly defined for $|e| < 1$ ($e \in \mathbb{R}$) by the relation ("Kepler's equation")

$$\lambda = \nu - e \sin \nu ;$$

let $\varphi \in \mathbb{T}$ (the "true anomaly") be implicitly defined (again for $|e| < 1$) by the relation

$$\mathrm{tg}\, \frac{\varphi - \gamma}{2} = \left(\frac{1+e}{1-e}\right)^{1/2} \mathrm{tg}\, \frac{\nu}{2} ,$$

and define the "orbital radius" r as

$$r \equiv \frac{a(1-e^2)}{1 + e\cos(\varphi - \gamma)} , \qquad \text{where} \quad a \equiv \Lambda^2 .$$

The function R_0 in (2.1) is then given by

[9] The Delaunay coordinates λ, γ, ψ are often called, respectively, the "mean anomaly," the "argument of the perihelion" and the "longitude" (of the "planet" P_2).

[10] We have chosen the units of measure in such a way that $m_1 + m_A = 1$ and that the period of P_2 is 2π.

$$R_0(\lambda, \gamma - \psi, \Lambda, \Gamma) \equiv -\Big(r\cos(\varphi - \psi) - \frac{1}{\sqrt{1 + r^2 - 2r\cos(\varphi - \psi)}}\Big) ,$$

where e is defined as

$$e \equiv \sqrt{1 - \frac{\Gamma^2}{\Lambda^2}} .$$

We recall that a convenient representation of R_0 is obtained by means of Legendre polynomials[11]: if $r < 1$ (which will be the case for our specific model) one finds

$$R_0 = 1 + \sum_{j=2}^{\infty} r^j P_j(\cos(\varphi - \psi)) .$$

A trivial reduction shows that the dynamics generated by (2.1) may be described by a two–degree–of–freedom Hamiltonian: under the canonical (or "symplectic") transformation

$$(\ell, g, \tau) \equiv (\lambda, \gamma - \psi, \psi) , \qquad (L, G, T) \equiv (\Lambda, \Gamma, \Gamma + E) ,$$

H_0 takes the form

$$H_1(\ell, g, L, G) \;=\; \frac{1}{2L^2} \;-\; G \;+\; \varepsilon\, R_0(\ell, g, L, G) , \tag{2.2}$$

having omitted the dummy variable T; the phase space is now $\mathbb{T}^2 \times \{(L, G) \in \mathbb{R}^2 : L \neq 0, |G| < |L|\}$.

3. A Model from the Solar System

Let us now focus on the case in which P_1 is the Sun, P_2 is Jupiter and A is Ceres (one of the largest bodies in the asteroid belt). Notice (again) that regarding Sun–Jupiter–Ceres as a planar, circular, restricted three-body problem contains a lot of physical approximations, which we shall not discuss here. But even accepting these basic approximations, the reader will have certainly noticed that the Hamiltonian in (1.3) is different from the Hamiltonian in (2.2): besides the factor $(\frac{1}{2L^2} - G)$ (and 2ε in place of ε), the main difference is that R is a trigonometric polynomial of degree 10 while R_0 contains infinite non vanishing Fourier harmonics. The "selection rule" which led us to the choice of the "physically relevant" Fourier modes is based on the following trivial observation. Among other things, the gravitational effects on Ceres of asteroids and planets and most notably the attraction exerted by Saturn (which, after Jupiter, is the largest planet in the Solar system[12]) have been neglected. *Therefore, after having defined a (rough) measure, $\mathcal{G}_{\rm Sa}$, of the Ceres–Saturn attraction, we disregard in the Fourier expansion of R_0 the terms exceeding $\mathcal{G}_{\rm Sa}$ in absolute value*.

In order to define $\mathcal{G}_{\rm Sa}$ we first look up a few astronomical data in the ephemeris (see The Astronomical Almanac [16]). In particular we want to define the "reference values" of L_0 and G_0 for the Sun–Jupiter–Ceres system. Observations of the true motion of Ceres, as found in [16], indicate that Ceres moves on a nearly elliptical orbit of "average eccentricity"

[11] $P_0(x) = 1$; $P_1(x) = x$; $P_{k+1}(x) = \frac{(2k+1)P_k(x)x - kP_{k-1}(x)}{k+1}$, (for $k \geq 1$).

[12] Even though the orbit of Ceres is closer to the orbit of Mars than to the orbit of Saturn, the difference in mass makes the gravitational attraction of Saturn the largest one after that of Jupiter.

$$e_0 \equiv 0.0766 \tag{3.1}$$

and whose average semimajor axis is approximately 0.532. Hence, the corresponding "average frequency" of Ceres, computed by Kepler's third law, yields a value of

$$-\Omega_c \equiv -2.577107 \ .$$

Since $-\Omega_c \simeq \partial_L H_1|_{\varepsilon=0} = -L^{-3}$ we take as "reference L-value" the quantity

$$L_0 \equiv 0.729305 \simeq \Omega_c^{\frac{1}{3}}$$

and, since $G_0 = L_0\sqrt{1-e_0^2}$, we take, as "reference G-value" the quantity

$$G_0 \equiv 0.727162 \ .$$

Such reference values have been taken as center of the *analyticity domain for the "action variables"*; see (1.1) and (1.2). Notice that with our choice of the analyticity radius r_0 one finds that the function $e(L,G)$ satisfies

$$0.019799 < |e(L,G)| < 0.106364 \ , \qquad \forall\, (L,G) \in B \ .$$

Let us turn now to the definition of $\mathcal{G}_{\rm Sa}$. In general, for planets whose orbits have a larger semimajor axis than that of Ceres, the "secular term" of H_1 is given by[13] ε ($\equiv$ mass of the planet/mass of the Sun) times the term $R_{00} \equiv R_{00}(L;e)$ in (1.5). Keeping in mind that, in the integrable limit, L is the ratio of the semimajor axis of Ceres with that of the planet we define (for planets whose orbits have a larger semimajor axis than that of Ceres)

$$\mathcal{G}_{\rm P} \equiv \varepsilon(P) \times R_{00}(L(P);e_0), \tag{3.2}$$

where e_0 is the observed "average eccentricity" of Ceres (3.1), $\varepsilon(P)$ is the mass ratio of the planet P and of the Sun and $L(P)$ is the ratio of the semimajor axis of Ceres with that of the planet P. Looking up the "true" values in [16] one finds

$$\mathcal{G}_{\rm Sa} \equiv \mathcal{G}_{\rm Saturn} = 6.3778 \cdot 10^{-6} \ .$$

For comparison purposes we report also the value for Jupiter, which is

$$\mathcal{G}_{\rm Jupiter} = 7.8850 \cdot 10^{-5} \ .$$

Neglecting in the expansion of εR_0 those terms whose size is smaller than[14] $\mathcal{G}_{\rm Sa}$, one is led to consider a "three-body problem" governed by the Hamiltonian

$$H_2(\ell,g,L,G;\varepsilon) \ \equiv \ \frac{1}{2L^2} \ - \ G \ + \ \varepsilon \ R(\ell,g,L,G) \tag{3.3}$$

with R given in (1.4) and (1.5) of Sect. 1.

The final modification of H_2 which gives the Hamiltonian in (1.3), (1.4) is due to merely technical reasons.

As mentioned above, our results are based on computer-assisted KAM theory, and one of the standard hypotheses of KAM theorems is that the unperturbed Hamiltonian

[13] The "secular term" of H_1 is the average over the angular variables ℓ and γ of the "perturbation" εR_0; the computation is immediately checked using, e.g., the above mentioned expansion in terms of Legendre Polynomials.

[14] More precisely we omit all the terms such that $\varepsilon|R_n(L_0,G_0)| < \mathcal{G}_{\rm Sa}$.

($\varepsilon = 0$) is non-degenerate, *i.e.* has an invertible Hessian matrix on its domain of analyticity. In the case of (3.3), the unperturbed Hamiltonian is given by

$$h_0(L, G) = \frac{1}{2L^2} - G\,,$$

whose Hessian matrix is *not* invertible. There are a few well known methods to overcome this minor problem[15] and it turns out that for our purposes the most convenient one is to follow Poincaré's trick [13], which consists in replacing the Hamiltonian H_2 by its square.[16] Therefore we let $H_3 \equiv (H_2)^2$:

$$\begin{aligned} H_3(\ell, g, L, G; \varepsilon) = \; & \Big(\frac{1}{2L^2} - G\Big)^2 + 2\varepsilon \Big(\frac{1}{2L^2} - G\Big) R(\ell, g, L, G) \\ & + \varepsilon^2 \,[R(\ell, g, L, G)]^2\,. \end{aligned} \tag{3.4}$$

The Hessian of the unperturbed Hamiltonian ($H_3|_{\varepsilon=0}$) is equal to

$$A \equiv A(L, G) \equiv \begin{pmatrix} \frac{5}{L^6} - \frac{6G}{L^4} & \frac{2}{L^3} \\ \frac{2}{L^3} & 2 \end{pmatrix}, \tag{3.5}$$

and, if $(L, G) \in B$, one has

$$\begin{aligned} |\det A| &= \Big| \frac{12}{L^4} \Big(\frac{1}{2L^2} - G\Big) \Big| \\ &\geq \frac{12}{(L_0 + r_0)^4} \Big(\frac{1}{2(L_0 + r_0)^2} - (G_0 + r_0) \Big) \geq 8.830153\,. \end{aligned}$$

To be consistent with the criterion that led us to the Hamiltonian (3.3), we have to omit the term of order ε^2 in (3.4) and this leads us to the Hamiltonian (1.3) introduced in Sects.(1, 2).

4. A KAM Theorem

Here we prove a KAM result, in the style of [5], which will be the basis of the proof of Theorem 1.1.

First we provide a "KAM algorithm" (in the Hamiltonian context), which yields a sequence of quadratically better and better approximations to the conjugacy function of a maximal invariant (Diophantine) torus, and then we formulate a criterion ensuring the applicability of the algorithm an infinite number of times and hence the existence of an invariant torus. Technically, the algorithm, which does not use symplectic transformations (used, instead, in the original works of the masters), may be viewed as a Hamiltonian version of the Lagrangian approach developed in the eighties by Moser, Salamon and Zehnder (see [5] and references therein).

4.1. Algebraic Scheme. Let us consider a smooth (later real-analytic) Hamiltonian $h(x, y)$, where x varies on the standard N-torus $\mathbb{T}^N \equiv \mathbb{R}^N/(2\pi\mathbb{Z}^N)$ and y varies in

[15] For example, one can replace the non-degeneracy hypothesis with a "iso-energetic non-degeneracy" (see, e.g., [1]), which is satisfied by h_0.

[16] Note that the dynamics generated by a Hamiltonian function $h = h(q, p)$ and by h^2 coincide up to a time scale: if $z(t) = (q(t), p(t))$ is an h-motion then $z(2Et)$, with $E = h(z(0))$ is the corresponding h^2-motion.

some open ball $B^N \subset \mathbb{R}^N$; (x, y) are standard symplectic coordinates.[17] The problem is to construct an invariant N-torus $\mathcal{S}$ on which the flow is conjugated to the linear flow $\theta \in \mathbb{T}^N \to \theta + \omega t$ for some "rationally independent" vector[18] $\omega \in \mathbb{R}^N$. The $\mathcal{S}$-embedding function $\theta \in \mathbb{T}^N \to (\theta + u(\theta), v(\theta)) \in \mathbb{T}^N \times B^N$ is immediately seen to satisfy the following quasi-linear, degenerate PDE on $\mathbb{T}^N$:

$$\begin{aligned} \omega + Du - h_y(\theta + u, v) &= 0 \ , \\ Dv + h_x(\theta + u, v) &= 0 \ , \end{aligned} \tag{4.1}$$

where D denotes the derivatives in the ω direction:

$$D \equiv \omega \cdot \partial_\theta \equiv \sum_{i=1}^N \omega_i \frac{\partial}{\partial \theta_i} \ , \tag{4.2}$$

and h_x, h_y denote the gradient of h with respect to x, y.

As usual, we assume that ω is a *Diophantine vector*, *i.e.* there exist $\gamma > 0$ and a positive integer τ such that

$$|\omega \cdot n| \equiv \Big| \sum_{i=1}^N \omega_i n_i \Big| \geq (\gamma |n|^\tau)^{-1} \ , \qquad \forall\, n \in \mathbb{Z}^N \backslash \{0\} \ . \tag{4.3}$$

The starting point of a KAM algorithm is an *approximate* solution (u, v), which solves (4.1) up to some "error." In order to formulate a precise result we need some notations and some assumptions.

Given a function $u : \mathbb{T}^N \to \mathbb{R}^N$ we denote by u_θ or by $\partial_\theta u$ its Jacobian matrix

$$(u_\theta)_{ij} \equiv \frac{\partial u_i}{\partial \theta_j} \ ;$$

h_{xy} denotes the matrix with entries

$$\Big(h_{xy} \Big)_{ij} \equiv \frac{\partial^2 h}{\partial x_i \partial y_j} \ ;$$

thus h_{yx} is the transpose of h_{xy}, *i.e.* $h_{yx} = h_{xy}^T$; if A is a square matrix, we denote by $A^\#$ the antisymmetric part of A times two:

$$A^\# \equiv A - A^T \ ;$$

finally, if $\theta \in \mathbb{T}^N \to f(\theta) \in \mathbb{R}^s$ $(s \geq 1)$ is a smooth function with vanishing mean value, *i.e.*

$$\langle f \rangle \equiv \frac{1}{(2\pi)^N} \int_{\mathbb{T}^N} f(\theta)\, d\theta \ = 0 \ ,$$

(and if D is as in (4.2), (4.3)), we denote by $D^{-1} f$ the unique solution with vanishing mean value of the equation $Dg = f$; such a solution in Fourier expansion has the form

$$D^{-1} f = \sum_{n \in \mathbb{Z}^N \backslash \{0\}} \frac{f_n}{i\omega \cdot n} \exp(in \cdot \theta) \ ,$$

where f_n denote Fourier coefficients and $i = \sqrt{-1}$.

[17] That is, the symplectic structure is given by the standard 2-form $\sum_{i=1}^N dx_i \wedge dy_i$.
[18] *I.e.* if $\omega \cdot n = 0$ for some $n \in \mathbb{Z}^N$, then n must be 0.

Assumption 4.1. *Let $\theta \in \mathbb{T}^N \to (u, v) \in \mathbb{R}^N \times \mathbb{R}^N$ be a smooth function and let $\mathcal{M}$ and h^0_{yy} be the matrices*[19]

$$\mathcal{M} \equiv I + u_\theta \ , \qquad h^0_{yy}(\theta) \equiv h_{yy}\Big(\theta + u(\theta), v(\theta)\Big) \ . \tag{4.4}$$

Denoting[20]

$$\mathcal{T} \equiv \mathcal{M}^{-1} h^0_{yy} \mathcal{M}^{-T} \ , \tag{4.5}$$

we assume that, for any $\theta \in \mathbb{T}^N$, the matrices $\mathcal{M}$, h^0_{yy} and $\langle \mathcal{T} \rangle$ are invertible.

Proposition 4.1. *Let h, u, v and ω satisfy, respectively, Assumption 4.1 and* (4.3) *and define f and g by*

$$\begin{aligned} \omega + Du - h_y(\theta + u, v) &= f \ , \\ Dv + h_x(\theta + u, v) &= g \ . \end{aligned} \tag{4.6}$$

Then, if we define the vector/matrix-valued functions $b(\theta)$ and $\mathcal{B}(\theta)$ by

$$b \equiv v_\theta^T f - \mathcal{M}^T g \ , \qquad \mathcal{B} \equiv \Big(\mathcal{M}^T g_\theta - v_\theta^T f_\theta\Big)^\# \ ,$$

we have

$$\langle b \rangle = 0 \ , \qquad \langle \mathcal{B} \rangle = 0 \ . \tag{4.7}$$

Furthermore, the following equation holds[21]*:*

$$\begin{aligned} \omega + Du' - h_y(\theta + u', v') &= f' \ , \\ Dv' + h_x(\theta + u', v') &= g' \ , \end{aligned} \tag{4.8}$$

where u', v', f', g' are defined at the end of the following list of definitions[22]*:*

$$\begin{aligned}
& b_0 \equiv \mathcal{T}\Big(D^{-1}b + c_1\Big) - \mathcal{M}^{-1} f \ , \qquad c_1 \equiv \langle \mathcal{T} \rangle^{-1} \Big(\langle \mathcal{M}^{-1} f \rangle - \langle \mathcal{T} D^{-1} b \rangle\Big) \ , \\
& z \equiv \mathcal{M}\Big(D^{-1} b_0 + c_2\Big) \ , \qquad c_2 \equiv -\langle \mathcal{M} D^{-1} b_0 \rangle \ , \\
& w \equiv \Big(h^0_{yy}\Big)^{-1} \Big(Dz - h^0_{yx} z + f\Big) \ , \\
& q_1 \equiv h_x(\theta + u + z, v + w) - h^0_x - h^0_{xx} z - h^0_{xy} w \ , \\
& q_2 \equiv h_y(\theta + u + z, v + w) - h^0_y - h^0_{yx} z - h^0_{yy} w \ , \\
& q_3 \equiv f_\theta^T w - g_\theta^T z - \mathcal{M}^T q_1 \ , \\
& f' \equiv -q_2 \ , \\
& g' \equiv \mathcal{M}^{-T} \Big(D\Big[(D^{-1}\mathcal{B})\mathcal{M}^{-1} z + \mathcal{M}^T (h^0_{yy})^{-1} f_\theta \mathcal{M}^{-1} z\Big] - q_3\Big) \ , \\
& u' \equiv u + z \ , \qquad v' \equiv v + w \ .
\end{aligned}$$

[19] I is the identity matrix. To be precise we should replace, in (4.4), u with $p \circ u$, p being the projection of $\mathbb{R}^N$ onto $\mathbb{T}^N$; however we shall omit, here and in other circumstances, such projection.

[20] The superscript $^{-T}$ denotes the transpose of the inverse: $A^{-T} = (A^{-1})^T$.

[21] Here and in what follows, the prime attached to a function will never denote derivates but just new functions.

[22] As above if $h = h(x, y)$, $h^0(\theta)$ denotes the function $h(\theta + u(\theta), v(\theta))$.

Remark 4.1. (i) It is immediate to check that if we replace f and g by εf and εg then $z, w = O(\varepsilon)$ and $f', g' = O(\varepsilon^2)$, *i.e. the errors associated to u' and v' are quadratically smaller than the errors associated to u and v*. We shall call a couple (u, v) as in (4.6) an *approximate solution* for (4.1) and the relative couple (f, g) the *error function*.
(ii) Note that the constants c_1 and c_2 are defined so that the functions b_0 and z have vanishing mean value.

Proof. To check the first of (4.7), observe that

$$\partial_\theta h^0 = \mathcal{M}^T h_x^0 + v_\theta^T h_y^0 \ ;$$

now, multiply the first of (4.6) by v_θ^T, the second by $-\mathcal{M}^T$, add them together and use integration by parts to get rid of the terms containing θ-derivatives.

To check the second of (4.7), take the θ-gradient of (4.6) to obtain

$$\begin{aligned} D\mathcal{M} &= h_{yx}^0 \mathcal{M} + h_{yy}^0 v_\theta + f_\theta \ , \\ Dv_\theta &= -h_{xx}^0 \mathcal{M} - h_{xy}^0 v_\theta + g_\theta \ . \end{aligned} \tag{4.9}$$

Let

$$\mathcal{A} \equiv \Big(\mathcal{M}^T v_\theta\Big)^{\#} \ , \tag{4.10}$$

and notice (integration by parts) that

$$\langle \mathcal{A} \rangle = 0 \ . \tag{4.11}$$

From (4.9), it follows that the matrix $\mathcal{A}$ satisfies the equation

$$D\mathcal{A} = \mathcal{B} \ , \tag{4.12}$$

from which the second of (4.7) follows at once.

The first of (4.8) follows immediately from the definitions:

$$\begin{aligned} \omega + Du' - h_y(\theta + u', v') &= \omega + Du + Dz - h_y(\theta + u + z, v + w) \\ &= h_y(\theta + u, v) - h_y(\theta + u + z, v + w) + Dz + f \\ &= -q_2 - h_{yx}^0 z - h_{yy}^0 w + Dz + f \\ &= f' \ . \end{aligned} \tag{4.13}$$

The check of the second of (4.8) is more tricky. First observe that last identity in (4.13) can be rewritten as

$$Dz = h_{yx}^0 z + h_{yy}^0 w - f \ . \tag{4.14}$$

Next, from the definition of z, it follows that

$$D\Big(\mathcal{M}^T (h_{yy}^0)^{-1} \Big[\mathcal{M} D(\mathcal{M}^{-1} z) + f\Big]\Big) = b \ . \tag{4.15}$$

Solving for h_{yy}^0 in the first of (4.9) and inserting the obtained expression in the definition of w, we get

$$\begin{aligned} w &= (h_{yy}^0)^{-1}\Big(Dz - \Big[(D\mathcal{M})\mathcal{M}^{-1} - h_{yy}^0 v_\theta \mathcal{M}^{-1} - f_\theta \mathcal{M}^{-1}\Big] z + f\Big) \\ &= v_\theta \mathcal{M}^{-1} z + (h_{yy}^0)^{-1}\Big(f_\theta \mathcal{M}^{-1} z + \mathcal{M} D(\mathcal{M}^{-1} f) + f\Big) \ . \end{aligned} \tag{4.16}$$

From the definition of g', (4.15), (4.16), (4.10)÷(4.12), it follows that

$$\mathcal{M}^T g' + q_3 + b = D(\mathcal{M}^T w - v_\theta^T z)\ , \quad i.e. \quad g' = \mathcal{M}^{-T}\Big(D(\mathcal{M}^T w - v_\theta^T z) - b - q_3\Big)\ .$$

From this identity, recalling the definitions of b and q_3 and using (4.9) to eliminate f_θ and g_θ, one obtains

$$g' = \mathcal{M}^{-T} v_\theta^T \Big[h^0_{yx} z + h^0_{yy} w - f\Big] + \Big[g + h^0_{xx} z + h^0_{xy} w + q_1\Big] + Dw - \mathcal{M}^{-T} v_\theta^T D z\ ,$$

which, in view of (4.14), the definition of q_1 and the second of (4.6), yields the second of (4.8). $\square$

4.2. Analytic Tools. From now on we shall work in the *real-analytic category*; in this section we review some basic technical facts.

We shall consider the Banach space of periodic functions f real-analytic on the torus $\mathbb{T}^N$, admitting (for some prefixed $\xi > 0$) analytic extension on the closed strip

$$\Delta_\xi \equiv \{\theta \in \mathbb{C}^N :\ |\,\mathrm{Im}\,\theta_i| \le \xi\ ,\ \forall\, i = 1, ..., N\}\ ,$$

equipped with the "Fourier norm"

$$\|f\|_\xi \equiv \sum_{n\in\mathbb{Z}^N} |f_n| \exp(|n|\xi)\ ; \tag{4.17}$$

in $\mathbb{C}^N$ (and its subsets $\mathbb{R}^N$ and $\mathbb{Z}^N$) we shall use the 1-norm,

$$|y| \equiv |y|_1 \equiv \sum_{i=1}^N |y_i|\ .$$

If $f : \mathbb{T}^N \to \mathbb{C}^s$ is a vector valued, real-analytic function, analytic on Δ_ξ, its norm is defined as $\|f\|_\xi \equiv \sum_i \|f_i\|_\xi$, which coincides with (4.17) if f_n denotes the s-vector whose components are given by the Fourier coefficients of the components of f. These definitions are immediately extended to matrix/tensor-valued functions by making use of the standard "operator norm": *e.g.* if $A(\theta)$ is a matrix-valued periodic functions with analytic extension on Δ_ξ, we set

$$\|A\|_\xi \equiv \sup_{c\in\mathbb{C}^N : |c|=1} \|Ac\|_\xi\ ;$$

or, if $\partial_x^3 f$ is the tensor of order three of the derivatives of a periodic, real-analytic function $f : \Delta_\xi \to \mathbb{C}$, its Fourier norm is given by

$$\|\partial_x^3 f\|_\xi \equiv \sup_{|b|=|c|=1} \sum_{i=1}^N \|\sum_{j,k=1}^N \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k} b_k c_j\|_\xi\ .$$

Finally, we shall also consider functions (possibly vector/matrix/tensor-valued) $h = h(x, y)$ periodic in x and real-analytic on the closed domain $\Delta_\xi \times \widehat{B}_r(y_0)$, where $\widehat{B}_r(y_0)$

$\equiv \widehat{B}_r^N(y_0)$ denotes the closed complex ball of radius r around $y_0 \in \mathbb{C}^N$. For any such function, which admits the expansion[23] (convergent on $\Delta_\xi \times \widehat{B}_r(y_0)$)

$$h(x,y) = \sum_{\substack{n\in\mathbb{Z}^N \\ k\in\mathbb{N}^N}} h_{n,k} \exp(in\cdot x)\,(y-y_0)^k\ , \tag{4.18}$$

we set

$$\|h\|_{\xi,r} \equiv \sum_{\substack{n\in\mathbb{Z}^N \\ k\in\mathbb{N}^N}} |h_{n,k}| \exp(|n|\xi)\, r^{|k|}\ . \tag{4.19}$$

The elementary properties of interest in the present context are collected in the following

Lemma 4.1. (i) *Let $f : \mathbb{T}^N \to \mathbb{R}$ have an analytic extension on Δ_ξ (for some $\xi > 0$) and let $\omega \in \mathbb{R}^N$ be a rationally independent vector. Then, for all $0 < \delta \le \xi$, for any $p \in \mathbb{Z}$ and for any $k \in \mathbb{N}$ or any $k \in \mathbb{N}^N$ one has*[24]

$$\|D^{-p}\partial_x^k f\|_{\xi-\delta} \le \|f\|_\xi\ \sigma_{pk}(\delta),$$

where, if $k = 0$, f is assumed to have vanishing mean value, and

$$\sigma_{pk}(\delta) \equiv \sup_{\{n\in\mathbb{Z}^N\setminus\{0\}:f_n\neq 0\}} \Big(\pi_{nk}|\omega\cdot n|^{-p}\ \exp(-\delta|n|)\Big)\ ,$$

$$\pi_{nk} \equiv \begin{cases} |n|^k\ , & \text{if } k\in\mathbb{N}\ , \\ |n^k|\ , & \text{if } k\in\mathbb{N}^N\ . \end{cases}$$

(ii) *Let $f, g : \mathbb{T}^N \to \mathbb{R}$ have an analytic extension to Δ_ξ, then*

$$\|fg\|_\xi \le \|f\|_\xi \|g\|_\xi\ .$$

(iii) *Let $0 < \xi < \bar{\xi}$; let $h : \mathbb{T}^N \times \{y \in \mathbb{R}^s : |y - y_0| \le r\} \to \mathbb{R}$ have an analytic extension on $\Delta_{\bar{\xi}} \times \widehat{B}_r^s(y_0)$, $f : \mathbb{T}^N \to \mathbb{R}^N$ and $g : \mathbb{T}^N \to \mathbb{R}^s$ have analytic extension on Δ_ξ. Assume that $\|f\|_\xi \le \bar{\xi} - \xi$, $\|g - y_0\|_\xi \le r$. Then, denoting $\phi(\theta) \equiv (\theta + f(\theta), g(\theta))$, one has*

$$\|h\circ\phi\|_\xi \le \|h\|_{\bar{\xi},r}\ .$$

Proof. (i) The claim follows immediately by expanding f in Fourier series.
(ii) In the following sums the indices n, m run over $\mathbb{Z}^N$,

$$\begin{aligned}
\|fg\|_\xi &\equiv \sum_n |(fg)_n| \exp(|n|\xi) = \sum_n \Big|\sum_m f_m g_{n-m}\Big|\ \exp(|n|\xi) \\
&\le \sum_{n,m} \exp|m| \exp|n-m||f_m|\ |g_{n-m}| \\
&= \|f\|_\xi \|g\|_\xi\ .
\end{aligned}$$

(iii) In the following sums the indices n, m run over $\mathbb{Z}^N$, the index k runs over $\mathbb{N}^s$ and j over $\mathbb{N}$. Using (ii), one gets

[23] We use the standard notation: $(y-y_0)^k = \prod_{i=1}^k (y_i - y_{0i})^{k_i}$.

[24] If $k \in \mathbb{N}$, $\partial_x^k f$ denotes the k-tensor of the derivatives of f; if $k \in \mathbb{N}^N$, $\partial_x^k f = \frac{\partial^{|k|} f}{\partial x_1^{k_1}\cdots\partial x_N^{k_N}}$.

$$\begin{aligned}
|h\circ\phi|_\xi &= \sum_n \Big| \sum_{m,k,j} \frac{h_{m,k}}{j!}\, i^j \, \Big((m\cdot f)^j (g-y_0)^k\Big)_{n-m}\Big| \ \exp(|n|\xi)\\
&\le \sum_{n,m,k,j} \exp(|m|\xi)\frac{|h_{m,k}|}{j!}\ \Big|\Big((m\cdot f)^j (g-y_0)^k\Big)_{n-m}\Big| \ \exp(|n-m|\xi)\\
&= \sum_{m,k,j}\frac{|h_{m,k}|}{j!}\ \|(m\cdot f)^j (g-y_0)^k\|_\xi \exp(|m|\xi)\\
&\le \sum_{m,k,j}\frac{|h_{m,k}|}{j!}\ \|m\cdot f\|_\xi^j \ \|g-y_0\|_\xi^{|k|}\ \exp(|m|\xi)\\
&\le \sum_{m,k,j}\frac{|h_{m,k}|}{j!}\ |m|^j \ (\bar\xi-\xi)^j \ r^{|k|} \exp(|m|\xi)\\
&= \|h\|_{\bar\xi,r}\ . \qquad \square
\end{aligned}$$

Remark 4.2. (i) Note that the result might be empty if ω is "too well approximable by rationals vectors." If ω satisfies (4.3) then one checks easily that

$$\sigma_{pk}(\delta) \le \begin{cases} \gamma^p \delta^{-(|k|+p\tau)}(|k|+p\tau)!\ , & \text{if } p\ge 0,\\ \Omega^{|p|}\delta^{-(|k|+|p|)}(|k|+|p|)!\ , & \text{if } p<0\ , \end{cases} \qquad \Omega\equiv\max\ |\omega_i|\ . \tag{4.20}$$

(ii) It is easy to check that (i) and (iii) of Lemma 4.1 holds also if f, respectively, h are vector valued.

4.3. KAM Algorithm. Here we describe the "KAM algorithm" associated to the scheme of Sect. 4.1, *i.e.* we equip the algebraic scheme described in Proposition 4.1 with "accurate" and detailed estimates so as to end up with a map $\mathcal{K}$, which to given bounds on norms of the relevant objects relative to a certain approximate solution (u,v), associates corresponding bounds on the new approximation (u',v'). More precisely, let us start by making quantitative the hypotheses formulated in Sect. 4.1.

Assumption 4.2. *Let* $0<\xi<\bar\xi$, r, $\overline{E}$, $E_{p,q}$ *(*$p,q\in\mathbb{N}$*) be such that* $(x,y)\to h(x,y)$ *is real analytic on* $\Delta_{\bar\xi}\times\overline{B}_r^N(y_0)$ *and*

$$\|(h_{yy})^{-1}\|_{\bar\xi,r}\le\overline{E}\ ,\qquad \|\partial_x^p\partial_y^q h\|_{\bar\xi,r}\le E_{p,q}\ ; \tag{4.21}$$

assume that the approximate solution (u,v) *is real analytic on* Δ_ξ *and let* U, V, M, $\overline{M}$, $\widetilde{V}$, F, G, $\widetilde{T}$ *be positive numbers bounding the following norms:*

$$\begin{aligned}
&\|u\|_\xi\le U\ ,\quad \|v\|_\xi\le V\ ,\quad \|\mathcal{M}\|_\xi\le M\ ,\quad \|\mathcal{M}^{-1}\|_\xi\le\overline{M}\ ,\\
&\|v_\theta\|_\xi\le\widetilde{V}\ ,\quad \|f\|_\xi\le F\ ,\quad \|g\|_\xi\le G\ ,\quad |\langle\mathcal{T}\rangle^{-1}|\le\widetilde{T}\ ,
\end{aligned}\tag{4.22}$$

where $\mathcal{M}$, $\mathcal{T}$, f *and* g *are, respectively, as in* (4.4), (4.5) *and* (4.6). *Finally assume that*

$$U\le\bar\xi-\xi\ ,\qquad \rho\equiv\|v-y_0\|_\xi<r\ . \tag{4.23}$$

Now let $0 < \delta \le \xi$ and define

$$\xi' \equiv \xi - \delta \ . \tag{4.24}$$

In the rest of this section we shall define the *KAM map*, i.e., an explicit map

$$\mathcal{K} : (U, V, M, \overline{M}, \widetilde{V}, F, G, \widetilde{T}) \to (U', V', M', \overline{M}', \widetilde{V}', F', G', \widetilde{T}'), \tag{4.25}$$

where $(U', ..., G', \widetilde{T}')$ are bounds on the norms $\|u'\|_{\xi'}$,..., $\|g'\|_{\xi'}$ and on the number $|\langle \mathcal{T}'\rangle^{-1}|$, with u', v', f', g' defined in Proposition 4.1, while $\mathcal{M}'$ and $\mathcal{T}'$ are (obviously) defined as

$$\mathcal{M}' \equiv I + u'_\theta \equiv \mathcal{M} + z_\theta \ , \qquad \mathcal{T}' \equiv \mathcal{M}'^{-1} h_{yy}(\theta + u', v')\, \mathcal{M}'^{-T} \ .$$

We start now to work out the necessary estimates. By Lemma 4.1 and (4.22), we get

$$\|\mathcal{T}\|_\xi \le T \equiv \overline{M}^2 E_{0,2} \ .$$

Remark 4.3. In principle, the bound on $\|\mathcal{T}\|_\xi$ could be improved replacing $E_{0,2}$ by $\|(h^0_{yy})^{-1}\|_\xi$ (without invoking point (iii) of Lemma 4.1); in practice, however, such a norm is difficult to evaluate accurately and one would not get significantly better estimates.

All other estimates are immediately obtained from Lemma 4.1 (and from the definitions given in Proposition 4.1). Here is the complete list:

$$\begin{aligned}
\|\mathcal{T}^{-1}\|_\xi \ &\le\ \overline{T} \equiv M^2 \overline{E} \ ; \\
\|b\|_\xi \ &\le\ F_* \equiv \widetilde{V} F + M G \ ; \\
\|\mathcal{B}\|_{\xi - \frac{\delta}{2}} \ &\le\ s_1 F_* \ , \qquad s_1 \equiv 2\sigma_{01}\Big(\frac{\delta}{2}\Big) \ ; \\
|c_1| \ &\le\ \widetilde{T}\Big(\overline{M} F + T\sigma_{10}(\xi) F_*\Big) \ ;
\end{aligned}$$

(note that in the last estimate we have used the fact that the supremum norm $\sup_{\mathbb{T}^N} |\cdot|$ is dominated by the 0-Fourier norm $\|\cdot\|_0$);

$$\begin{aligned}
\|\mathcal{A}\|_{\xi'} \ &\equiv\ \|D^{-1}\mathcal{B}\|_{\xi'} \le s_2 F_* \ , \qquad s_2 \equiv \sigma_{10}\Big(\frac{\delta}{2}\Big) s_1 \ ; \\
\|b_0\|_{\xi - \frac{\delta}{2}} \ &\le\ B_* \equiv T\sigma_{10}\Big(\frac{\delta}{2}\Big) F_* + T\widetilde{T}\Big(\overline{M} F + T F_* \sigma_{10}(\xi)\Big) + \overline{M} F \ ; \\
|c_2| \ &\le\ M \|D^{-1} b_0\|_0 \le M\sigma_{10}\Big(\xi - \frac{\delta}{2}\Big) B_* \ ; \\
\|z\|_{\xi'} \ &\le\ B_* s_3 \ , \qquad s_3 \equiv M\sigma_{10}\Big(\frac{\delta}{2}\Big) + M^2 \sigma_{10}\Big(\xi - \frac{\delta}{2}\Big) \ ; \\
\|u'\|_{\xi'} \ &\equiv\ \|u + z\|_{\xi'} \le U + \|z\|_{\xi'} \\
&\le\ U' \equiv U + B_* s_3 \ ; \\
\|Dz\|_{\xi'} \ &=\ \|(D\mathcal{M})(D^{-1} b_0 + c_2) + \mathcal{M} b_0\|_{\xi'} \\
&\le\ B_* s_4 \ , \qquad s_4 \equiv s_3 \sigma_{-10}(\delta) + M \ ; \\
\|z_\theta\|_{\xi'} \ &=\ \|\mathcal{M}_\theta (D^{-1} b_0 + c_2) + \mathcal{M} D^{-1} \partial_\theta b_0\|_{\xi'}
\end{aligned}$$

$$
\begin{aligned}
&\le B_* s_5 \;, \qquad s_5 \equiv \sigma_{01}(\delta) s_3 + M\sigma_{11}\Big(\frac{\delta}{2}\Big) \;; \\
\|Dz_\theta\|_{\xi'} &= \|(D\mathcal{M}_\theta)(D^{-1}b_0 + c_2) + D\mathcal{M}D^{-1}\partial_\theta b_0 + \mathcal{M}_\theta b_0 + \mathcal{M}\partial_\theta b_0\|_{\xi'} \\
&\le B_* s_6 \;, \\
&\quad s_6 \equiv \sigma_{-11}(\delta) s_3 + M\Big[\sigma_{-10}(\delta)\sigma_{11}\Big(\frac{\delta}{2}\Big) + \sigma_{01}(\delta) + \sigma_{01}\Big(\frac{\delta}{2}\Big)\Big] \;; \\
\|w\|_{\xi'} &\le \overline{E}\Big(\|Dz\|_{\xi'} + E_{1,1}\|z\|_{\xi'} + F\Big) \\
&\le \overline{E}(s_7 B_* + F) \;, \qquad s_7 \equiv s_4 + E_{1,1}s_3 \;; \\
\|v'\|_{\xi'} &\equiv \|v + w\|_{\xi'} \le V + \|w\|_{\xi'} \\
&\le V' \equiv V + \overline{E}(s_7 B_* + F) \;; \\
\|w_\theta\|_{\xi'} &= \|\partial_\theta (h^0_{yy})^{-1}(Dz - h^0_{yx}z + f) + \\
&\quad (h^0_{yy})^{-1}(Dz_\theta - \partial_\theta h^0_{yx} z - h^0_{yx} z_\theta + f_\theta)\|_{\xi'} \\
&\le \sigma_{01}(\delta)\overline{E}(\|Dz\|_{\xi'} + E_{1,1}\|z\|_{\xi'} + F) \\
&\quad + \overline{E}\Big[\|Dz_\theta\|_{\xi'} + \sigma_{01}(\delta)E_{1,1}\|z\|_{\xi'} + E_{1,1}\|z_\theta\|_{\xi'} + \sigma_{01}(\delta)F\Big] \\
&\le \overline{E}B_* s_8 + \overline{E}F s_9 \;, \\
&\quad s_8 \equiv \sigma_{01}(\delta)s_7 + s_6 + \sigma_{01}(\delta)E_{1,1}s_3 + E_{1,1}s_5 \;, \\
&\quad s_9 \equiv 2\sigma_{01}(\delta) \;; \\
\|q_1\|_{\xi'} &\le \frac{1}{2}\|h_{xxx}\|_{\bar{\xi},r}\|z\|^2_{\xi'} + \|h_{xxy}\|_{\bar{\xi},r}\|z\|_{\xi'}\|w\|_{\xi'} + \frac{1}{2}\|h_{xyy}\|_{\bar{\xi},r}\|w\|^2_{\xi'} \\
&\le s_{10}B_*^2 + s_{11}B_*F_* + s_{12}F^2 \;, \\
&\quad s_{10} \equiv \frac{1}{2}E_{3,0}s_3^2 + \overline{E}E_{2,1}s_3 s_7 + \frac{1}{2}E_{1,2}\overline{E}^2 s_7^2 \;, \\
&\quad s_{11} \equiv s_3 E_{2,1}\overline{E} + s_7 E_{1,2}\overline{E}^2 \;, \\
&\quad s_{12} \equiv \frac{1}{2}\overline{E}^2 E_{1,2} \;; \\
\|f'\|_{\xi'} &\equiv \|q_2\|_{\xi'} \le \frac{1}{2}\|h_{yxx}\|_{\bar{\xi},r}\|z\|^2_{\xi'} + \|h_{yxy}\|_{\bar{\xi},r}\|z\|_{\xi'}\|w\|_{\xi'} \\
&\quad + \frac{1}{2}\|h_{yyy}\|_{\bar{\xi},r}\|w\|^2_{\xi'} \\
&\le F' \equiv s'_{10}B_*^2 + s'_{11}B_*F + s'_{12}F^2 \;, \\
&\quad s'_{10} \equiv \frac{1}{2}E_{2,1}s_3^2 + \overline{E}E_{1,2}s_3 s_7 + \frac{1}{2}E_{0,3}\overline{E}^2 s_7^2 \;, \\
&\quad s'_{11} \equiv s_3 E_{1,2}\overline{E} + s_7 E_{0,3}\overline{E}^2 \;, \\
&\quad s'_{12} \equiv \frac{1}{2}\overline{E}^2 E_{0,3} \;; \\
\|q_3\|_{\xi'} &\le \sigma_{01}(\delta)F\|w\|_{\xi'} + \sigma_{01}(\delta)G\|z\|_{\xi'} + M\|q_1\|_{\xi'} \\
&\le s_{13}B_*F + s_{14}B_*G + s_{15}F^2 + s_{16}B_*^2 \;, \\
&\quad s_{13} \equiv \sigma_{01}(\delta)\overline{E}s_7 + Ms_{11} \;, \qquad s_{14} \equiv \sigma_{01}(\delta)s_3 \;, \\
&\quad s_{15} \equiv \sigma_{01}(\delta)\overline{E} + Ms_{12} \;, \qquad s_{16} \equiv Ms_{10} \;;
\end{aligned}
$$

$$
\begin{aligned}
\|g'\|_{\xi'} &= \|\mathcal{M}^{-T}\Big[\mathcal{B}\mathcal{M}^{-1}z + \mathcal{A}(D\mathcal{M}^{-1})z + \mathcal{A}\mathcal{M}^{-1}Dz \\
&\quad +(D\mathcal{M}^{T})(h^0_{yy})^{-1}f_\theta\mathcal{M}^{-1}z\Big] \\
&\quad +D(h^0_{yy})^{-1}\ f_\theta\mathcal{M}^{-1}z + (h^0_{yy})^{-1}Df_\theta\ \mathcal{M}^{-1} \\
&\quad +(h^0_{yy})^{-1}f_\theta(D\mathcal{M}^{-1})\ z + (h^0_{yy})^{-1}f_\theta\mathcal{M}^{-1}Dz - \mathcal{M}^{-T}q_3\|_{\xi'} \\
&\le \overline{M}\Big[s_1F_*\overline{M}B_*s_3 + s_2F_*\sigma_{-10}(2\delta)\overline{M}B_*s_3 + s_2F_*\overline{M}B_*s_4 \\
&\quad +\sigma_{-10}(2\delta)\overline{E}\sigma_{01}(2\delta)F\overline{M}B_*s_3\Big] \\
&\quad +\sigma_{-10}(\delta)\overline{E}\sigma_{01}(2\delta)F\overline{M}B_*s_3 + \overline{E}\sigma_{-11}(2\delta)F\overline{M}B_*s_3 \\
&\quad +\overline{E}\sigma_{01}(2\delta)F\sigma_{-10}\overline{M}B_*s_3 + \overline{E}\sigma_{01}(2\delta)F\overline{M}B_*s_4 \\
&\quad +\overline{M}\Big(s_{13}B_*F + s_{14}B_*G + s_{15}F^2 + s_{16}B_*^2\Big) \\
&\le G' \equiv s_{17}F_*B_* + s_{18}\overline{E}FB_* + s_{14}\overline{M}B_*G + s_{15}\overline{M}F^2 + s_{16}\overline{M}B_*^2\ , \\
&\quad s_{17} \equiv \overline{M}^2(s_1s_3 + \sigma_{-10}(2\delta)s_2s_3 + s_2s_4)\ , \\
&\quad s_{18} \equiv \overline{M}s_3\Big[\overline{M}\sigma_{-10}(2\delta)\sigma_{01}(2\delta) + \sigma_{-10}(\delta)\sigma_{01}(\delta) + \sigma_{-11}(2\delta) \\
&\qquad +\sigma_{01}(2\delta)\sigma_{-10}(2\delta)\Big] + \overline{M}\sigma_{01}(2\delta)s_4 + \overline{M}\ \overline{E}^{-1}s_{13}\ ; \\
\|\mathcal{M}'\|_{\xi'} &\equiv \|\mathcal{M} + z_\theta\|_{\xi'} \le M + \|z_\theta\|_{\xi'} \\
&\le M' \equiv M + B_*s_5\ .
\end{aligned}
$$

To define $\overline{M}'$ we have to distinguish two cases according to whether $\overline{M}B_*s_5$ is greater or smaller than 1. In the first case we define $\overline{M}'$ to be infinite:

$$\overline{M}B_*s_5 \ge 1 \qquad \Longrightarrow\ \overline{M}' \equiv \infty\ ,$$

while if

$$\overline{M}B_*s_5 < 1, \tag{4.26}$$

we proceed as follows:

$$
\begin{aligned}
\|\mathcal{M}'^{-1}\|_{\xi'} &= \|(I + \mathcal{M}^{-1}z_\theta)^{-1}\mathcal{M}^{-1}\|_{\xi'} \le \overline{M}(1 - \overline{M}\|z_\theta\|_{\xi'})^{-1} \\
&\le \overline{M}' \equiv \overline{M}(1 - \overline{M}B_*s_5)^{-1}\ .
\end{aligned}
$$

Next,

$$
\begin{aligned}
\|v'_\theta\|_{\xi'} &\equiv \|v_\theta + w_\theta\|_{\xi'} \le \widetilde{V} + \|w_\theta\|_{\xi'} \\
&\le \widetilde{V}' \equiv \widetilde{V} + \overline{E}B_*s_8 + \overline{E}Fs_9\ .
\end{aligned}
$$

Of course, as it is clear from the definition of $\overline{M}'$, the KAM algorithm will be of some use only if (4.26) is satisfied. To complete the computation of $\mathcal{K}$, it remains to bound $|\langle\mathcal{T}'\rangle^{-1}|$, *i.e.* to compute $\widetilde{T}'$. If $\overline{M}' = \infty$, we define also $\widetilde{T}' \equiv \infty$, otherwise we proceed as follows. A bit of algebra shows that if we define

$$\begin{aligned}
\mathcal{C} \;&\equiv (I + \mathcal{M}^{-1} z_\theta)^{-1} - I = \sum_{k=1}^{\infty} (-\mathcal{M}^{-1} z_\theta)^k \ , \\
\mathcal{C}' \;&\equiv \mathcal{C}\mathcal{M}^{-1} \ , \\
\mathcal{C}'' &\equiv h_{yy}(\theta + u', v') - h^0_{yy} \ ,
\end{aligned}$$

then we can write $\mathcal{T}'$ as

$$\mathcal{T}' = \mathcal{T} + \mathcal{C}_* \ ,$$

where

$$\begin{aligned}
\mathcal{C}_* \;\equiv\; & \mathcal{M}^{-1} h^0_{yy} \mathcal{C}'^T + \mathcal{M}^{-1}\mathcal{C}''\mathcal{M}^{-1} + \mathcal{M}^{-1}\mathcal{C}''\mathcal{C}'^T + \mathcal{C}' h^0_{yy}\mathcal{M}^{-1} \\
& + \mathcal{C}' h^0_{yy} \mathcal{C}'^T + \mathcal{C}'\mathcal{C}''\mathcal{M}^{-1} + \mathcal{C}'\mathcal{C}''\mathcal{C}'^T \ .
\end{aligned}$$

Thus, since $\sup_{\mathbb{T}^N} |z_\theta| \le \|z_\theta\|_0$ (and a bound on $\|z_\theta\|_0$ is obtained exactly as above replacing δ with ξ), we obtain

$$\begin{aligned}
\sup_{\mathbb{T}^N} |\mathcal{C}| \;&\le\; \overline{M}\|z_\theta\|_0 (1 - \overline{M}\|z_\theta\|_0)^{-1} \\
&\le\; C \equiv \overline{M} B^0_* s^0_5 (1 - \overline{M} B^0_* s^0_5)^{-1} \ ; \\
\sup_{\mathbb{T}^N} |\mathcal{C}'| \;&\le\; C' \equiv C\overline{M} \ ; \\
\sup_{\mathbb{T}^N} |\mathcal{C}''| \;&\le\; E_{1,2}\|z\|_0 + E_{0,3}\|w\|_0 \\
&\le\; C'' \equiv E_{1,2} B^0_* s^0_3 + E_{0,3}\overline{E}(s^0_7 B^0_* + F) \ ; \\
\sup_{\mathbb{T}^N} |\mathcal{C}_*| \;&\le\; C_* \equiv 2\overline{M}\ E_{0,2} C' + \overline{M}^2 C'' + 2\overline{M} C'' C' \\
&\qquad + C'^2 E_{0,2} + C'^2 C'' \ , \qquad (4.27)
\end{aligned}$$

where $B^0_*, s^0_5, s^0_3, s^0_7$ are defined as above but with δ replaced by ξ. Now, if

$$\widetilde{T} C_* \ge 1 \qquad \Longrightarrow\ \widetilde{T}' = \infty \ ,$$

otherwise, if

$$\widetilde{T} C_* < 1 \ , \qquad (4.28)$$

we have

$$\begin{aligned}
|\langle \mathcal{T}' \rangle^{-1}| &= |(I + \langle \mathcal{T}\rangle^{-1}\langle \mathcal{C}_*\rangle)^{-1} \langle \mathcal{T}\rangle^{-1}| \\
&\le \widetilde{T}' \equiv \widetilde{T}(1 - \widetilde{T} C_*)^{-1} \ . \qquad (4.29)
\end{aligned}$$

The computation of the map $\mathcal{K}$ is completed.

Remark 4.4. (*"KAM algorithm"*). If (4.23) is satisfied when U, v, ξ are replaced by U', v', ξ' then the map $\mathcal{K}$ can be re-applied and, iterating when possible, one obtains a sequence $(u^{(j)}, v^{(j)})$ of approximate solutions [with relative error functions $(f^{(j)}, g^{(j)})$] and corresponding norm-bounds U_j, V_j,...,$\widetilde{T}_j$. More precisely, fix numbers $\delta_j \searrow 0$ such that $\sum \delta_j < \xi$ (where $\delta_0 \equiv \delta$ as above); let, for $j \ge 1$, $\xi_j \equiv \xi_{j-1} - \delta_{j-1}$; call the above "first approximate solution" $(u^{(0)}, v^{(0)})$, $(f^{(0)}, g^{(0)})$ the relative error function and attach to $\mathcal{M}$, $\mathcal{T}$ and to their norm bound an index 0. Then, given, for $j \ge 0$, $(u^{(j)}, v^{(j)})$

and the relative norm bounds $(U_j, V_j, ..., \widetilde{T}_j)$ we let $(u^{(j+1)}, v^{(j+1)})$ be the approximate solution constructed in[25] Proposition 4.1 and, if conditions (4.23), (4.26) and (4.28) are satisfied (see, again, footnote 25), then $(U_{j+1}, ..., \widetilde{T}_{j+1}) = \mathcal{K}(U_j, ..., \widetilde{T}_j)$ are the norm bounds controlling the new approximate solution $(u^{(j+1)}, v^{(j+1)})$.

4.4. KAM Theorem. Here we prove a KAM theorem based on the KAM algorithm described above.[26]

Theorem 4.1. *Let ω satisfy* (4.3), *let Assumption 4.2 hold, let*

$$0 < \hat{\xi} < \xi\ , \qquad \hat{\delta} \equiv \frac{1}{2}\,\frac{\xi - \hat{\xi}}{\xi}\ ,$$

let η be the following norm on (f, g)

$$\eta \equiv \max\{\overline{E}\ E_{0,2}\ ,\ E_{0,2}\widetilde{T}\}\ \max\{\gamma F\ ,\ \gamma^2 E_{0,2}\widetilde{V} F\ ,\ \gamma^2 E_{0,2} G\}\ ,$$

define the following parameters related to ω and to the quantities introduced in Assumption 4.2:

$$\begin{aligned}
\Omega\ &\equiv \max_{1\le i\le N} |\omega_i|\ , \qquad \Omega_1 \equiv \max\{\Omega\ ,\ E_{1,1}\}\ ,\\
\Omega_*\ &\equiv \max\{E_{2,1}\ ,\ \overline{E}E_{1,2}\Omega_1\ ,\ E_{0,3}\overline{E}^2\Omega_1^2\}\ ,\\
H_*\ &\equiv \max\{E_{3,0}\ ,\ \overline{E}E_{2,1}\Omega_1\ ,\ E_{1,2}\overline{E}^2\Omega_1^2\}\ ,\\
H'_*\ &\equiv \max\{H_*\ ,\ \overline{E}\Omega\Omega_1\},\ H''_* \equiv \max\{H'_*\ ,\ \widetilde{V}\Omega_*\}\ ,\\
\alpha_*\ &\equiv \max\{\overline{E}\ E_{0,2}\ ,\ E_{0,2}\widetilde{T}\}\cdot\max\{\gamma^2 E_{0,2} H''_*\ ,\ \gamma\Omega_*\}\ ,\\
\theta\ &\equiv \max\{(\xi - \hat{\xi})^{2\tau}\ ,\ (\xi - \hat{\xi})^{2\tau+2}\}\ ,\\
\alpha_{**}\ &\equiv \theta \max\{1\ ,\ \frac{\overline{E}\Omega_1}{r - \rho}\ ,\ \frac{\overline{E}\Omega_1}{\widetilde{V}}\ ,\ \widetilde{T}E_{0,2}\ ,\ \widetilde{T}E_{1,2}\ ,\ \widetilde{T}E_{0,3}\overline{E}\Omega_1\}\ ,\\
\alpha\ &\equiv \max\{\alpha_*\ ,\ \alpha_{**}\}\ .
\end{aligned}$$

There exists a polynomial ν in $(\xi, \hat{\delta})$ satisfying[27]

$$\frac{5}{4} \le \nu(\xi, \hat{\delta}) \le 21 + 88\max\{\xi\ ,\ \xi^6\}\ , \quad \forall \xi > 0\ ,\ \forall\ 0 < \hat{\delta} < \frac{1}{2}\ ,$$

such that, if

[25] Of course, the unprimed quantities in Proposition 4.1 (besides $\overline{E}$ and $E_{p,q}$ which remain unchanged) correspond here to the index j while the primed ones correspond here to the index $j + 1$. Analogously, one has to attach in the obvious way an index j or $j + 1$ in the formulae defined in this section; for example $\overline{T}_j = M_j^2\overline{E}$, $F_{*j} = \widetilde{V}_j F_j + M_j G_j$, $s_1 = 2\sigma_{01}(\delta_j/2)$, etc.

[26] We recall that in order to apply such a theorem in an "effective way," one should apply it only after a few iterations of the map $\mathcal{K}$: see [5] for more information.

[27] The polynomial ν (of degree 12 and with positive coefficients) is explicitly given in the proof below. For later use we report also the following values:

$$\nu(1, \frac{1}{4}) = 17.4281...\ , \quad \nu(0.0025, \frac{1}{4}) = 6.5190... \tag{4.30}$$

$$\eta\alpha\, M^7\overline{M}^9(\xi-\hat{\xi})^{-2(2\tau+1)}2^{8\tau+13}\,\tau!^4\,\nu(\xi,\hat{\delta})\leq 1\,, \tag{4.31}$$

then there exists a solution $(\tilde{u},\tilde{v})$ *of* (4.1)*, which is real-analytic on* $\Delta_{\hat{\xi}}$ *and is* η*-close to* (u,v)*. Furthermore* $\langle\tilde{u}\rangle = \langle u\rangle$ *and the solution* $(\tilde{u},\tilde{v})$ *is the unique solution in an* η *neighbourhood of* (u,v) *such that* $\langle\tilde{u}\rangle = \langle u\rangle$*.*

For us it will be particularly important to investigate *dependence upon parameters*: assume that the Hamiltonian $h = h(x,y;\mu)$ depends analytically also upon a set of parameters varying in some complex domain

$$\mu\in\mathcal{D}\subset\mathbb{C}^p\,.$$

Then the coefficients $h_{n,k}$ in (4.18) will be analytic functions of $\mu\in\mathcal{D}$ and we shall *redefine the norm in* (4.19) *by*

$$\|h\|_{\xi,r}\equiv\sum_{\substack{n\in\mathbb{Z}^N\\ k\in\mathbb{N}^N}}\Big(\sup_{\mu\in\mathcal{D}}|h_{n,k}|\Big)\,\exp(|n|\xi)\,r^{|k|}\,. \tag{4.32}$$

Analogously, for a function $f(x;\mu)$ analytic in $\Delta_\xi\times\mathcal{D}$ we shall *redefine the Fourier norm* (4.17) *by*

$$\|f\|_\xi\equiv\sum_{n\in\mathbb{Z}^N}\Big(\sup_{\mu\in\mathcal{D}}|f_n|\Big)\exp(|n|\xi)\,.$$

Note that we are not changing the symbols of the norms since *the domain* $\mathcal{D}$ *will remain unchanged in the proofs*.

Finally, Assumptions 4.1 and 4.2 will be modified as follows:

Assumption 4.3. *Let* $(\theta,\mu)\in\mathbb{T}^N\times\mathcal{D}\to(u,v)\in\mathbb{C}^N\times\mathbb{C}^N$ *be a regular function of* θ *and let* $\mathcal{M}$ *and* h^0_{yy} *be the matrices*

$$\mathcal{M}\equiv I+u_\theta\,,\qquad h^0_{yy}\equiv h_{yy}\Big(\theta+u(\theta;\mu),v(\theta;\mu);\mu\Big)\,. \tag{4.33}$$

We assume that, if we denote

$$\mathcal{T}\equiv\mathcal{M}^{-1}h^0_{yy}\mathcal{M}^{-T}\,, \tag{4.34}$$

then for any $(\theta,\mu)\in\mathbb{T}^N\times\mathcal{D}$*, the matrices* $\mathcal{M}$*,* h^0_{yy} *and* $\langle\mathcal{T}\rangle$ *are invertible.*

Assumption 4.4. *Let* $0<\xi<\bar{\xi}$*,* r*,* $\overline{E}$*,* $E_{p,q}$ $(p,q\in\mathbb{N})$ *be such that* $(x,y,\mu)\to h(x,y;\mu)$ *is real analytic on* $\Delta_{\bar{\xi}}\times\overline{B}^N_r(y_0)\times\mathcal{D}$ *and such that* (4.21) *holds; assume that* (u,v) *is real analytic on* $\Delta_\xi\times\mathcal{D}$ *and let* $U,V,M,\overline{M},\widetilde{V},F,G,\widetilde{T}$ *be positive numbers for which* (4.22) *holds*[28] *(with* $\mathcal{M}$*,* $\mathcal{T}$*,* f *and* g *as in* (4.33)*,* (4.34) *and* (4.6)*). Finally, assume that* (4.23) *holds.*

With these modifications one obtains

Theorem 4.2. *Let* ω*,* ξ*,* $\hat{\xi}$*,* $\hat{\delta}$*,* ν *be as in Theorem 4.1; assume that* h*,* (f,g) *satisfy Assumption 4.4. Let* α *and* η *be defined as in Theorem 4.1 and let* (4.31) *hold. Then there exists a solution* $(\tilde{u},\tilde{v})$ *of* (4.1)*, which is real-analytic on* $\Delta_{\hat{\xi}}\times\mathcal{D}$ *and is* η*-close to* (u,v)*. Furthermore* $\langle\tilde{u}\rangle = \langle u\rangle$ *and the solution* $(\tilde{u},\tilde{v})$ *is the unique solution in an* η *neighbourhood of* (u,v) *such that* $\langle\tilde{u}\rangle = \langle u\rangle$*.*

[28] Last inequality must be rewritten as $\sup_{\mu\in\mathcal{D}}|\langle\mathcal{T}\rangle^{-1}|\leq\widetilde{T}$.

Remark 4.5. (i) We shall not prove here the local uniqueness (under the constraint $\langle \tilde{u} \rangle = \langle u \rangle$) of the solution $(\tilde{u}, \tilde{v})$ and we refer the reader to [5] (where a similar statement is proven in full detail).
(ii) Precise estimates concerning the η closeness (which is of course in the sense of the $\| \cdot \|_{\hat{\xi}}$ norm) can be easily deduced from the detailed estimates given in the proof.

Proof. (of Theorem 4.1.) The argument is by induction. We define, for $j \geq 0$,

$$\xi_j \equiv \hat{\xi} + \frac{\xi - \hat{\xi}}{2^j} , \qquad \delta_j \equiv \xi_j - \xi_{j+1} = \frac{\xi - \hat{\xi}}{2^{j+1}} .$$

We assume (by induction) to have constructed (by iterating $j \geq 0$ times the scheme of Proposition 4.1 starting with $(u^{(0)}, v^{(0)}) \equiv (u, v)$) the approximate solution $(u^{(j)}, v^{(j)})$; denote by $(U_j, ..., \widetilde{T}_j)$ the relative norm-bounds[29] and assume that, for $0 \leq i \leq j$,

$$\begin{gathered} \|u^{(i)}\|_{\xi_i} \leq \overline{\xi} - \xi_i , \qquad \|v^{(i)} - y_0\|_{\xi_i} \leq r , \\ \widetilde{V}_i \leq 2\widetilde{V}_0 , \qquad M_i \leq 2M_0 , \qquad \overline{M}_i \leq 2\overline{M}_0 , \qquad \widetilde{T}_i \leq 2\widetilde{T}_0 . \end{gathered} \tag{4.35}$$

We want to show that (4.35) is satisfied also for $i = j + 1$ and that the norms of $z^{(i)}$, $w^{(i)}$, $f^{(i)}$ and $g^{(i)}$ decay exponentially fast as i tends to infinity so that the claim of the theorem follows.

We now list estimates on various quantities entering the definition of the map $\mathcal{K}$ (*i.e.* on the definitions of the parameters, introduced in the previous section, which bound the norms of the relevant objects associated to[30] $u^{(i)}, v^{(i)}$); the estimates are completely elementary and we shall give details only for some of the first estimates (all the other estimates being obtained in a complete analogous way).

In the following estimates we shall make systematic use of (4.35), (4.20), of the fact that $\tau \geq 1$ and of the following simple observations[31]:

$$\gamma\Omega \geq 1 , \quad M_0 \geq 1 , \quad \overline{M}_0 \geq 1 , \quad \overline{E}E_{0,2} \geq 1 , \quad \frac{\delta_j}{\xi_j} = 1 - \frac{\xi_{j+1}}{\xi_j} \leq \hat{\delta} , \tag{4.36}$$

where

$$\hat{\delta} \equiv \frac{\delta_0}{\xi_0} = \frac{1}{2} \frac{\xi - \hat{\xi}}{\xi} . \tag{4.37}$$

Here is the list:

$$\begin{aligned} &\sigma_{01}\Big(\frac{\delta_i}{t}\Big) \leq t\delta_i^{-1} , \ (\forall\ t > 0) ; \quad \sigma_{10}(\xi_i) \leq \gamma\xi_i^{-\tau}\ \tau! ; \\ &\sigma_{10}\Big(\frac{\delta_i}{2}\Big) \leq 2^\tau\gamma\delta_i^{-\tau}\tau! , \quad \sigma_{10}\Big(\xi_i - \frac{\delta_i}{2}\Big) \leq \sigma_{10}(\xi_i) \leq \gamma\xi_i^{-\tau}\tau! ; \end{aligned}$$

---

[29] See the previous section. Note that, with these definitions, $\xi_0 = \xi$, $U_0 = U$,...,$M_0 = M$, etc. The "norm-parameters" are defined in the previous section where the primed quantities correspond to $u^{(i+1)}, v^{(i+1)}$ while the unprimed ones correspond to $u^{(i)}, v^{(i)}$.

[30] Recall that the primed quantities of Sect. 4.3 correspond here to the index $i + 1$ while the unprimed ones to the index i, that the index $i = 0$ corresponds to the "initial" approximate solution, namely to the quantities defined in Assumption 4.2 and, finally, that $E_{p,q}$ and $\overline{E}$ are independent of i.

[31] The first relation in (4.36) follows from (4.3) letting $n = e^{(i)}$, where $\{e^{(i)}\}$ is the standard orthonormal basis of $\mathbb{R}^N$; the second and third relations follow by observing that $e^{(1)}$ is an eigenvector with eigenvalue 1 of $\mathcal{M}^T(\theta_0)$ and of $\mathcal{M}^{-T}(\theta_0)$, if θ_0 is a critical point of $u_1(\theta)$; the last two relations are obvious.

$$\sigma_{-10}\Big(\frac{\delta_i}{t}\Big) \le t\,\Omega\delta_i^{-1}\,,\ (\forall\, t>0);\quad \sigma_{11}\Big(\frac{\delta_i}{2}\Big) \le \gamma 2^{\tau+1}\delta_i^{-(\tau+1)}(\tau+1)!\,;$$
$$\sigma_{-11}\Big(\frac{\delta_i}{t}\Big) \le 2t^2\,\Omega\delta_i^{-2}\,,\ (\forall\, t>0)\,;$$
$$T_i \le 4\overline{M}_0^2 E_{0,2}\,;\quad \overline{T}_i \le 4M_0^2\overline{E}\,,\quad F_{*i} \le 2\widetilde{V}_0 F_i + 2M_0G_i\,;$$
$$s_1 \le 4\delta_i^{-1}\,;\quad s_2 \le 2^{2+\tau}\gamma\delta_i^{-(\tau+1)}\tau!\,;$$
$$s_3 \le M_0 2^{\tau+1}\gamma\delta_i^{-\tau}\tau! + 4M_0^2\gamma\xi_i^{-\tau}\tau! = M_0^2 2^{\tau+1}\gamma\delta_i^{-\tau}\tau!\Big[\frac{1}{M_0} + 2\Big(\frac{\delta_i}{2\xi_i}\Big)^\tau\Big]$$
$$\le M_0^2 2^{\tau+1}\gamma\delta_i^{-\tau}\tau!\Big(1+\hat{\delta}\Big) \equiv M_0^2 2^{\tau+1}\gamma\delta_i^{-\tau}\tau!\,\nu_1\,,\ \nu_1 \equiv 1+\hat{\delta}\,;$$
$$s_4 \le (\gamma\Omega)M_0^2 2^{\tau+1}\delta_i^{-(\tau+1)}\tau!\,\nu_2\,,\ \nu_2 \equiv \nu_1 + \frac{\delta_0^2}{2}\,;$$
$$s_5 \le \gamma M_0^2 2^{\tau+2}\delta_i^{-(\tau+1)}(\tau+1)!\,\nu_3\,,\ \nu_3 \equiv 1 + \frac{\nu_1}{4}\,;$$
$$s_6 \le (\gamma\Omega)M_0^2 2^{\tau+2}\delta_i^{-(\tau+2)}(\tau+1)!\,\nu_4\,,\ \nu_4 \equiv 1 + \frac{\nu_1}{2} + \frac{3}{16}\delta_0^2\,;$$
$$s_7 \le (\gamma\Omega_1)M_0^2 2^{\tau+1}\delta_i^{-(\tau+1)}\tau!\,\nu_5\,,\ \nu_5 \equiv \nu_2 + \delta_0\nu_1\,;$$
$$s_8 \le (\gamma\Omega_1)M_0^2 2^{\tau+2}\delta_i^{-(\tau+2)}(\tau+1)!\,\nu_6\,,\ \nu_6 \equiv \nu_4 + \frac{\nu_5}{4} + \frac{\nu_1\delta_0}{4} + \nu_3\delta_0\,;$$
$$s_9 \le 2\delta_i^{-1}\,;$$
$$s_{10} \le \gamma^2 H_* M_0^4 2^{2\tau+1}\delta_i^{-2(\tau+1)}\tau!^2\,\nu_7\,,\ \nu_7 \equiv (\nu_5+\nu_1\delta_0)^2\,;$$
$$s_{11} \le \gamma^2 H_* M_0^2 2^{\tau+1}\delta_i^{-(\tau+1)}\tau!\,\nu_8\,,\ \nu_8 \equiv \nu_5 + \nu_1\delta_0\,;$$
$$s_{12} \le \frac{\gamma^2 H_*}{2}\,;$$
$$s'_{10} \le \gamma^2\Omega_* M_0^4 2^{2\tau+1}\delta_i^{-2(\tau+1)}\tau!^2\,\nu_7\,;$$
$$s'_{11} \le \gamma^2\Omega_* M_0^2 2^{\tau+1}\delta_i^{-(\tau+1)}\tau!\,\nu_8\,;$$
$$s'_{12} \le \frac{\gamma^2\Omega_*}{2}\,;$$
$$s_{13} \le \gamma^2 H'_* M_0^3 2^{\tau+1}\delta_i^{-(\tau+2)}\tau!\,\nu_9\,,\ \nu_9 \equiv \nu_5 + 2\delta_0\nu_8\,;$$
$$s_{14} \le \gamma M_0^2 2^{\tau+1}\delta_i^{-(\tau+1)}\tau!\,\nu_1\,;$$
$$s_{15} \le \gamma^2 H'_* M_0\delta_i^{-1}\,\nu_{10}\,,\ \nu_{10} \equiv 1+\delta_0\,;$$
$$s_{16} \le \gamma^2 H_* M_0^5 2^{2\tau+2}\delta_i^{-2(\tau+1)}\tau!^2\,\nu_7\,;$$
$$s_{17} \le \gamma(\gamma\Omega)(M_0\overline{M}_0)^2 2^{2\tau+5}\delta_i^{-2(\tau+1)}\tau!^2\,\nu_{11}\,,\ \nu_{11} \equiv \nu_2 + \frac{\nu_1}{2} + \frac{\nu_1\delta_0^2}{2}\,;$$
$$s_{18} \le \gamma^2\overline{E}^{-1}H'_* M_0^3\overline{M}_0^2 M_0 2^{\tau+3}\delta_i^{-(\tau+2)}\tau!\,\nu_{12}\,,\ \nu_{12} \equiv \frac{9}{8}\nu_1 + \frac{\nu_9}{2} + \frac{\nu_2}{4}\,.$$

Let now η_i be as η (defined in the text of Theorem 4.1) but with F_i, G_i in place of F, G. Then,

$$\gamma B_{*i} \le \eta_i M_0\overline{M}_0^4 2^{\tau+4}\delta_i^{-\tau}\tau!\,\nu_{13}\,,\ \nu_{13} \equiv 1 + 4\hat{\delta} + \frac{9}{16}\delta_0;$$
$$\|z^{(i)}\|_{\xi_{i+1}} \le \eta_i M_0^3\overline{M}_0^4 2^{2\tau+5}\delta_i^{-2\tau}\tau!^2\,\nu_{14}\,,\ \nu_{14} \equiv \nu_1\nu_{13}\,;$$

$$
\begin{aligned}
&\|z_\theta^{(i)}\|_{\xi_{i+1}} \leq \eta_i M_0^3 \overline{M}_0^4 2^{2\tau+6} \delta_i^{-(2\tau+1)} \tau!^2 (\tau+1)\, \nu_{15}\ ,\ \nu_{15} \equiv \nu_3 \nu_{13}\ ;\\
&\|w^{(i)}\|_{\xi_{i+1}} \leq \eta_i (\overline{E}\Omega_1) M_0^3 \overline{M}_0^4 2^{2\tau+5} \delta_i^{-(2\tau+1)} \tau!^2\, \nu_{16}\ ,\ \nu_{16} \equiv \nu_5 \nu_{13} + \frac{\delta_0^3}{2^7}\ ;\\
&\|w_\theta^{(i)}\|_{\xi_{i+1}} \leq \eta_i (\overline{E}\Omega_1) M_0^3 \overline{M}_0^4 2^{2\tau+6} \delta_i^{-2(\tau+1)} \tau!^2 (\tau+1)\, \nu_{17}\\
&\qquad\qquad \nu_{17} \equiv \nu_6 \nu_{13} + \frac{\delta_0^3}{2^8}\ ;\\
&\|f^{(i+1)}\|_{\xi_{i+1}} \leq \Omega_*\, \eta_i^2 M_0^6 \overline{M}_0^8 2^{4\tau+9} \delta_i^{-(4\tau+2)} \tau!^4\, \nu_{18}\\
&\qquad\qquad \nu_{18} \equiv \nu_7 \nu_{13}^2 + \frac{\delta_0^3}{2^6} \nu_8 \nu_{13} + \frac{\delta_0^6}{2^{14}}\ ;\\
&\|g^{(i+1)}\|_{\xi_{i+1}} \leq H'_*\, \eta_i^2 M_0^7 \overline{M}_0^9 2^{4\tau+12} \delta_i^{-(4\tau+2)} \tau!^4\, \nu_{19}\ ,\\
&\qquad\qquad \nu_{19} \equiv \nu_7 \nu_{13}^2 + \frac{\delta_0}{2} \nu_{11} \nu_{13} + \frac{\delta_0^2}{2^6} \nu_{12} \nu_{13} + \frac{\delta_0^3}{2^7} \nu_1 \nu_{13} + \frac{\delta_0^5}{2^{14}} \nu_{10}\ ; \qquad (4.38)
\end{aligned}
$$

(in the last inequality we used also the fact that from $\overline{E}E_{0,2} \geq 1$ and $\gamma\Omega_1 \geq 1$ it follows that $\gamma\Omega \max\{\widetilde{V}F, G\} \leq H'_*\eta_i$). We let now

$$\nu \equiv \nu(\xi_0, \hat\delta) \equiv \nu_{19} + \frac{1}{4} \qquad (4.39)$$

and observe that all the $\nu's$ are greater than or equal to 1. Putting together the above definitions of the various ν_i's and recalling that $\delta_0 = \xi\hat\delta$, one finds

$$
\begin{aligned}
\nu \;=\; &\frac{5}{4} + 10\,\hat\delta + 33\,\hat\delta^2 + 40\,\hat\delta^3 + 16\,\hat\delta^4 + \frac{47}{8}\,\hat\delta\xi + \frac{101}{2}\,\hat\delta^2\xi + \frac{1161}{8}\,\hat\delta^3\xi\\
&+ \frac{329}{2}\,\hat\delta^4\xi + 64\,\hat\delta^5\xi + \frac{5257}{512}\,\hat\delta^2\xi^2 + \frac{39527}{512}\,\hat\delta^3\xi^2 + \frac{50415}{256}\,\hat\delta^4\xi^2\\
&+ 194\,\hat\delta^5\xi^2 + 64\,\hat\delta^6\xi^2 + \frac{77319}{8192}\,\hat\delta^3\xi^3 + \frac{455303}{8192}\,\hat\delta^4\xi^3 + \frac{6105}{64}\,\hat\delta^5\xi^3\\
&+ 50\,\hat\delta^6\xi^3 + \frac{1131}{256}\,\hat\delta^4\xi^4 + \frac{8415}{512}\,\hat\delta^5\xi^4 + \frac{921}{64}\,\hat\delta^6\xi^4 + \frac{15447}{16384}\,\hat\delta^5\xi^5\\
&+ \frac{925}{512}\,\hat\delta^6\xi^5 + \frac{1369}{16384}\,\hat\delta^6\xi^6\ .
\end{aligned}
$$

We note, for later use, that[32]

$$\nu \geq \max\{\ \frac{\nu_{18}}{4}\ ,\ \nu_{14}\ ,\ \nu_{15}\ ,\ \nu_{16} + \delta_0 \nu_{14}\ \}\ . \qquad (4.40)$$

Thus, (using the inequalities $\alpha \geq \alpha_*$ and $\nu \geq \max\{\frac{\nu_{18}}{4}, \nu_{19}\}$), we find

$$\eta_{i+1} \leq \kappa \lambda^i \eta_i^2\ ,$$

where[33]

$$\kappa \equiv \alpha M_0^7 \overline{M}_0^9 2^{8\tau+13} (\xi - \hat\xi)^{-(4\tau+2)} \tau!^4\, \nu\ , \qquad \lambda \equiv 2^{4\tau+2}\ .$$

Iterating, for all $0 \leq i \leq j+1$, we get[34]

[32] As it is immediate to check after having written down explicitly the definition of ν (which, recalling that $\delta_0 = \xi_0\hat\delta$, turns out to be a polynomial of degree 12 in ξ_0 and $\hat\delta$ with positive (rational) coefficients) and of the other ν_k's.

[33] The reason for having in this formula α in place of α_* will be plain when we shall check the inductive assumptions (4.35) for $j+1$.

[34] Recall that $\sum_{k=0}^{i-1} 2^k = 2^i - 1$, $\sum_{k=1}^{i-1}(i-k)2^{k-1} = 2^i - i - 1$, (and that $\eta_0 = \eta$).

$$\eta_i \le \frac{(\eta\kappa\lambda)^{2^i}}{\kappa\lambda^{i+1}} \ ,$$

whence, in view of the "KAM condition" (4.31) (which is now recognized as equivalent to require $\eta\kappa\lambda \le 1$),

$$\eta_i \le \frac{1}{\kappa\lambda^{i+1}} \ . \tag{4.41}$$

This bound allows to get simple estimates on the norms of $z^{(i)}$ and $w^{(i)}$ for $0 \le i \le j$. Using (4.41) and the facts that

$$\nu \ge \max\{\nu_{14} \ , \ \nu_{15}\} \quad \text{and} \quad \alpha \ge \max\{(\xi - \hat{\xi})^{2\tau+2} \ , \ (\xi - \hat{\xi})^{2\tau}\} \ ,$$

from (4.38) we get

$$\max\{\|z^{(i)}\|_{\xi_{i+1}} \ , \ \|z_\theta^{(i)}\|_{\xi_{i+1}}\} \le \min\{\delta_i \ , \ \frac{1}{\overline{M}_0} \ \frac{1}{2^{15}} \ \frac{1}{8^i}\} \ . \tag{4.42}$$

Analogously, using

$$\nu \ge \nu_{16} \ , \quad \alpha \ge \overline{E}\Omega_1 \ (\xi - \hat{\xi})^{2\tau+1} \ \max\{\frac{1}{r-\rho} \ , \ \frac{1}{\widetilde{V}}\} \ ,$$

from (4.38) there follows

$$\|w^{(i)}\|_{\xi_{i+1}} \le \min\{r - \rho \ , \ \widetilde{V}\} \ \frac{1}{2^{18}} \ \frac{1}{8^i} \ . \tag{4.43}$$

To check the inductive hypotheses (4.35) we shall also need simple bounds on the constants C_i, C_i', C_i'', C_{*i} [recall (4.27) and (4.29)]. Using

$$\nu \ge \max\{\nu_{15} \ , \ \nu_{16} + \delta_0\nu_{14}\}$$

and the fact that

$$\alpha \ge (\xi - \hat{\xi})^{2\tau+1} \ \max\{1 \ , \ \widetilde{T}E_{0,2} \ , \ \widetilde{T}E_{1,2} \ , \ \widetilde{T}E_{0,3}\overline{E}\Omega_1\} \ ,$$

and that

$$\overline{M}_i B_{*i} s_5 \le \frac{1}{\overline{M}_0^2} \ \frac{1}{2^{14}} \ \frac{1}{8^i} \ \min\{1 \ , \ \frac{1}{\widetilde{T}E_{0,2}}\} \le \frac{1}{2^{14}} \ ,$$

we obtain

$$\begin{aligned}
&C_i \le \min\{1 \ , \ \frac{1}{\widetilde{T}E_{0,2}}\}\frac{1}{\overline{M}_0^2}\frac{1}{2^{13}} \ \frac{1}{8^i} \ , \\
&C_i' \le \min\{1 \ , \ \frac{1}{\widetilde{T}E_{0,2}}\}\frac{1}{\overline{M}_0}\frac{1}{2^{12}} \ \frac{1}{8^i} \ , \\
&C_i'' \le \max\{E_{1,2} \ , \ E_{0,3}\overline{E}\Omega_1\} \ \eta_i \ M_0^3\overline{M}_0^4 \ \delta_i^{-2\tau+1} \ 2^{2\tau+5}\tau!^2(\nu_{16} + \delta_0\nu_{14}) \ , \\
&C_{*i} \le \widetilde{T}_0^{-1}\frac{1}{2^8} \ \frac{1}{8^i} \ .
\end{aligned} \tag{4.44}$$

We are ready to check (4.35) for $i + 1$: From (4.35) and (4.42) we find

$$\|u^{(i+1)}\|_{\xi_{i+1}} \le \|u^{(i)}\|_{\xi_i} + \|z^{(i)}\|_{\xi_{i+1}} \le \overline{\xi} - \xi_i + \delta_i = \overline{\xi} - \xi_{i+1} \ .$$

From (4.23) and (4.43), we get

$$\begin{aligned}\|v^{(i+1)} - y_0\|_{\xi_{i+1}} &\leq \|v^{(0)} - y_0\|_{\xi} + \sum_{j=0}^{i} \|w^{(j)}\|_{\xi_{i+1}} \\ &\leq \rho + \frac{r-\rho}{2^{18}} \sum_{j=0}^{\infty} \frac{1}{8^j} \leq r\ .\end{aligned}$$

Finally, using (4.43), (4.42), (4.44) one easily obtains the remaining inductive hypotheses.

Observe that from (4.41), (4.42) and (4.43) it follows that the error functions $f^{(i)}$ and $g^{(i)}$ go to zero exponentially fast, while $u^{(i)}$ and $v^{(i)}$ converge (exponentially fast) to real-analytic functions $\tilde{u}$ and $\tilde{v}$.

This concludes the proof of Theorem 4.1. □

Proof. (of Theorem 4.2) The reader will have no difficulty in checking that the previous proof goes through word by word so that the claim follows from uniformity in $\mu \in \mathcal{D}$. □

5. Proof of Theorem 1.1

The proof of Theorem 1.1 will be divided in three steps: (1) construction of the starting approximate solution (Remark 4.1) using ε-expansions; (2) bounds (4.22) on the norms relative to starting approximate solution with $\mathcal{D} \equiv \{\varepsilon \in \mathbb{C} : |\varepsilon| \leq \varepsilon_0 \equiv 10^{-6}\}$, see Assumption 4.4; (3) iteration of the map $\mathcal{K}$ (4.25) and application of Theorem 4.2.

5.1. Step 1: Formal ε-expansion and initial approximate solution. The Hamiltonian H in (1.3) contains ε as a parameter: ε corresponds to the parameter μ of Theorem 4.2 and $\mathcal{D}$ corresponds to the complex ball $\{\varepsilon \in \mathbb{C} : |\varepsilon| \leq \varepsilon_0\}$ for some $\varepsilon_0 > 0$ to be determined below.[35] As was well known to Poincaré, Lindstedt & Co., one may *compute formally* the ε-expansion of quasi-periodic *formal* solutions ("Lindstedt series"; see [1]). *Our starting approximate solution will be a suitable truncation of such a formal expansion.*

Here we deduce a few elementary formulae which allow us to explicitly compute recursively the formal solution.

Let $N = 2$, $x \equiv (\ell, g)$, $y \equiv (L, G)$ and let us rewrite explicitly Eq. (4.1) for the Hamiltonian H:

$$\begin{aligned}\omega_1 + Du_1(\theta) &= -2(\frac{1}{2v_1^2} - v_2)\frac{1}{v_1^3} + 2\varepsilon(\frac{1}{2v_1^2} - v_2)R_L(\theta + u, v) - 2\varepsilon\frac{1}{v_1^3}R(\theta + u, v), \\ \omega_2 + Du_2(\theta) &= -2(\frac{1}{2v_1^2} - v_2) + 2\varepsilon(\frac{1}{2v_1^2} - v_2)R_G(\theta + u, v) - 2\varepsilon R(\theta + u, v), \\ Dv_1(\theta) &= -2\varepsilon(\frac{1}{2v_1^2} - v_2)R_\ell(\theta + u, v), \\ Dv_2(\theta) &= -2\varepsilon(\frac{1}{2v_1^2} - v_2)R_g(\theta + u, v)\ ; \end{aligned} \tag{5.1}$$

[35] Clearly H is an entire function of ε and the restriction on $\mathcal{D}$ is needed in order to meet the basic condition (4.31); the choice of the "best value for ε_0" (i.e., the largest one) has been done simply by "trial and error".

here ω and D are short for $\omega^{(\pm)}$ and for $\omega^{(\pm)} \cdot \partial_\theta$, $\theta \in \mathbb{T}^2$ and recall (1.8), (1.9) and (1.10).

It is well known that (5.1) admits a *formal solution*

$$\tilde{u}(\theta) \sim \sum_{j=0}^{\infty} \tilde{u}^{(j)}(\theta)\, \varepsilon^j \ , \qquad \tilde{v}(\theta) \sim \sum_{j=0}^{\infty} \tilde{v}^{(j)}(\theta)\varepsilon^j \ , \tag{5.2}$$

with $\tilde{u}^{(j)}$ and $\tilde{v}^{(j)}$ being vector-valued *real-analytic on* $\mathbb{T}^2$ functions:

$$\tilde{u}^{(j)} \equiv (\tilde{u}_1^{(j)}, \tilde{u}_2^{(j)}) \ , \qquad \tilde{v}^{(j)} \equiv (\tilde{v}_1^{(j)}, \tilde{v}_2^{(j)}) \ .$$

Furthermore, such a formal solution is *uniquely determined* by requiring that the averages of the $\tilde{u}^{(j)}$'s vanish:

$$\langle \tilde{u}^{(j)} \rangle = 0 \ , \qquad (\forall\, j \geq 0) \ .$$

In particular this implies that

$$\tilde{u}^{(0)} \equiv 0 \ .$$

Instead (as one checks immediately by inserting (5.2) into (5.1) and looking at the order zero in ε) $\tilde{v}^{(0)}$ has to be chosen so that

$$\partial_{(L,G)} H \Big|_{\varepsilon=0} (\tilde{v}^{(0)}) = \omega \ , \qquad i.e. \ , \qquad \tilde{v}^{(0)} \equiv (L_\pm, G_\pm) \ .$$

Remark 5.1. The formal solvability of (5.1) implies that the right-hand sides of the last two equations of (5.1) *have vanishing mean value*[36] *over* $\mathbb{T}^2$. The averages of the $\tilde{v}^{(j)}$'s have then to be chosen so that the first two equations in (5.1) are solvable. This leaves free the averages of the $\tilde{u}^{(j)}$'s, which, as already said, will be taken to be zero.

As initial approximate solution we take

$$u^{(0)}(\theta;\varepsilon) \equiv \sum_{j=0}^{5} \tilde{u}^{(j)}(\theta)\, \varepsilon^j \ , \qquad v^{(0)}(\theta;\varepsilon) \equiv \sum_{j=0}^{5} \tilde{v}^{(j)}(\theta)\, \varepsilon^j \tag{5.3}$$

(recall that, with the above conventions, $\tilde{v}^{(0)} \equiv (L_\pm, G_\pm)$, $\tilde{u}^{(0)} \equiv 0$ *and* $\langle \tilde{u}^{(j)} \rangle = 0$*).*

We proceed by writing down the explicit formulae which, implemented on a machine, allow to compute[37] the functions $\tilde{u}^{(j)}$, $\tilde{v}^{(j)}$ or, more precisely, *allow to compute intervals of real numbers containing the Fourier coefficients of the* $(\tilde{u}^{(j)}, \tilde{v}^{(j)})$*'s.*

Recalling the explicit form of the function R [see (1.4), (1.5)], one sees that the right-hand sides of (5.1) have the form

$$\sum_{i=1}^{M} r_i\, \varepsilon^{s_i}\, (v_1^{(0)})^{p_i}\, (v_2^{(0)})^{q_i}\, e^{\sigma_i}(v_1^{(0)}, v_2^{(0)})\, c_{n_i}(\theta + u^{(0)}) \ , \tag{5.4}$$

where: $M < \infty$; r_i are rational numbers; s_i, p_i, q_i, σ_i are integers obeying the constraints

$$0 \leq s_i \leq 1 \ , \quad -5 \leq p_i \leq 10 \ , \quad 0 \leq q_i \leq 1 \ , \quad -1 \leq \sigma_i \leq 1 \ ;$$

[36] For any periodic function $f(\theta)$, $\theta \in \mathbb{T}^N$, the integral over $\mathbb{T}^N$ of Df vanishes.

[37] Notice that since the Hamiltonian H is a trigonometric polynomial in x, the functions $(\tilde{u}^{(j)}, \tilde{v}^{(j)})$ are also trigonometric polynomials.

$n_i \in \mathbb{Z}^2$ with $|n_i| \le 10$; finally $c_n(x)$ is either $\cos n \cdot x$ or $\sin n \cdot x$. We shall denote by $[\cdot]_j$ the operator that acts on a (formal) ε-power series, $\sum \varepsilon^k a^{(k)}$, by associating to it the $j^{\rm th}$ coefficient $a^{(j)}$:

$$\Big[\sum_{k\ge 0} \varepsilon^k a^{(k)}\Big]_j \equiv a^{(j)} \; .$$

Let $p, q, \sigma \in \mathbb{Z}$ with $q > 0$ and $|\sigma| \le 1$ and let us compute the $j^{\rm th}$ ε-coefficient of, respectively, $a^p b^q e^\sigma(a, b)$ and of $c_n(\theta + \varphi)$ where a, b and φ are formal ε-power series (with periodic real-analytic coefficients) given by

$$a \sim \sum_{k\ge 0} a_k(\theta)\varepsilon^k \; , \quad b \sim \sum_{k\ge 0} b_k(\theta)\varepsilon^k \; ,$$
$$\varphi \sim \sum_{k\ge 0} \varphi^{(k)}(\theta)\varepsilon^k \equiv \Big(\sum_{k\ge 0} \varphi_1^{(k)}(\theta)\varepsilon^k, \sum_{k\ge 0} \varphi_2^{(k)}(\theta)\varepsilon^k\Big) \; .$$

As above, we denote $y_0 \equiv \tilde v^{(0)}$ and write the expansions of $e(y)$ and of $e^{-1}(y) \equiv 1/e(y)$ as[38]

$$e(y) \equiv \sum_{h\in\mathbb{N}^2} e_h (y - y_0)^h \; , \quad e^{-1}(y) \equiv \sum_{h\in\mathbb{N}^2} \tilde e_h (y - y_0)^h \; .$$

Then, for $p \ge 0$, one finds

$$[a^{-p} b^q e^\sigma(a, b)]_j = \sum_{(k,h)\in I_{-p,q}} \binom{p}{k_1}\binom{q}{k_2} y_{01}^{k_1-p}\, y_{02}^{q-k_2}\, e_{\sigma,h}\, a_{k_3}^{[h_1]}\, \tilde a_{k_4}^{[k_1]}\, b_{k_5}^{[h_2+k_2]} \; , \tag{5.5}$$

where

$$I_{-p,q} \equiv \{(k, h) \in \mathbb{N}^5 \times \mathbb{N}^2 : \; 0 \le k_1 \le p \; , \; 0 \le k_2 \le q \; , \; k_3 \ge h_1 \; , \; k_4 \ge k_1 \; , \; k_5 \ge k_2 + h_2 \; , \; k_3 + k_4 + k_5 = j \; \} \; ;$$

$$e_{\sigma,h} \equiv \begin{cases} \tilde e_h & \text{if } \sigma = -1 \\ e_h & \text{if } \sigma = 1 \\ \delta_{0|h|} & \text{if } \sigma = 0 \end{cases} \; .$$

$c^{[k]}(\varepsilon)$ denotes the $k^{\rm th}$ power of a formal power series $c \sim \sum_{j\ge 0} c_j \varepsilon^j$ and $c_j^{[k]}$ its $j^{\rm th}$ ε-coefficient; finally $\tilde a$ is the power series defined by

$$\tilde a \sim \frac{1}{a} - \frac{1}{a_0} \; .$$

Analogously, for $p > 0$, one gets

$$[a^p b^q e^\sigma(a, b)]_j = \sum_{(k,h)\in I_{p,q}} \binom{p}{k_1}\binom{q}{k_2} y_{01}^{p-k_1} y_{02}^{q-k_2} e_{\sigma,h}\, a_{k_3}^{[h_1+k_1]}\, b_{k_4}^{[h_2+k_2]} \; , \tag{5.6}$$

[38] We use standard multiindex notation: if $n \in \mathbb{N}^N$, $|n| = \sum_{i=1}^N n_i$; if $z \in \mathbb{R}^N$, $\partial_z^n = \frac{\partial^{|n|}}{\partial z_1^{n_1} \cdots \partial z_n^{n_N}}$; $n! = n_1! \cdots n_N!$; $z^n = z_1^{n_1} \cdots z_N^{n_N}$.

where:

$$I_{p,q} \equiv \{(k,h) \in \mathbb{N}^4 \times \mathbb{N}^2 : \ 0 \le k_1 \le p \ , \ 0 \le k_2 \le q \ , \\ k_3 \ge k_1 + h_1 \ , \ k_4 \ge k_2 + h_2 \ , k_3 + k_4 = j \ \} \ . \tag{5.7}$$

Computing $[c_n(\theta + \varphi)]_j$ is clearly equivalent to evaluate $[\exp(in \cdot \varphi)]_j$ and letting

$$E_n^{(0)} \equiv 1 \ , \quad \forall \, n \in \mathbb{Z}^2 \ ,$$

$$E_n^{(k)}(\theta) \equiv \frac{1}{k} \sum_{\ell=1}^{k} \ell \, E_n^{(k-\ell)} \, (\theta) \, n \cdot \varphi_\ell(\theta) \ ,$$

one gets

$$\Big[\exp(in \cdot \varphi) \Big]_j = E_n^{(j)}(\theta) \ . \tag{5.8}$$

Inserting (5.2) in (5.1) one obtains recursive relations of the type

$$D\tilde{u}^{(k)} = A\tilde{v}^{(k)} + \Phi^{(k)} \ , \quad D\tilde{v}^{(k)} = \Psi^{(k)} \ , \tag{5.9}$$

where A is the (constant) matrix given in (3.5) evaluated at $(L, G) \equiv (L_\pm, G_\pm)$ and the vectors $\Phi^{(k)}$ and $\Psi^{(k)}$ depend on $\tilde{u}^{(0)}$,...,$\tilde{u}^{(k-1)}$, $\tilde{v}^{(0)}$,...,$\tilde{v}^{(k-1)}$ (and on θ) and can be explicitly written down by using the remark leading to (5.4) and the expansions (5.5), (5.6) and (5.8).

Assume now that $(\tilde{u}^{(j)}, \tilde{v}^{(j)})$, for $j = 0, ..., k-1$ are known and let us determine $(\tilde{u}^{(k)}, \tilde{v}^{(k)})$. Inverting the operator D in the second of (5.9) (recalling Remark 5.1) we let

$$\tilde{v}^{(k)} \equiv \overline{v}^{(k)} + D^{-1}\Psi^{(k)} \ ,$$

where $\overline{v}^{(k)}$ denotes the average of $\tilde{v}^{(k)}$ and it has to be determined so that the equations for $D\tilde{u}^{(k)}$ have a right-hand side with vanishing mean value, i.e.,

$$\overline{v}^{(k)} \equiv -A^{-1} \langle \Phi^{(k)} \rangle \ , \qquad \tilde{u}^{(k)} \equiv D^{-1}\Big(A\tilde{v}^{(k)} + \Phi^{(k)} \Big) \ .$$

The formulae for the recursive computation of the functions $(\tilde{u}^{(j)}, \tilde{v}^{(j)})$ (and hence of our choice of the initial approximate solution) are complete. In Appendix A we report the number of Fourier coefficients of the functions $(\tilde{u}^{(j)}, \tilde{v}^{(j)})$ for $j \le 5$ and, as an example, the list of intervals trapping the Fourier coefficients of the first component of $\tilde{u}^{(1)}$.

5.2. Step 2: Norm bounds relative to the initial approximate solution. Having defined the initial approximate solution as the fifth order truncation (5.3) of the ε-expansion of the formal solution of (5.1), we want now to estimate the relative norm parameters as defined in (4.22).

We attach an index 0 to the quantities related to our starting approximate solution (i.e. the fifth order truncation (5.3) of the ε-expansion of the formal solution of (5.1)); thus the symbols ξ, U, V,...,$\widetilde{T}$ of Sect. 4.3 correspond here to ξ_0, U_0, V_0,...,$\widetilde{T}_0$ (see Remark 4.4). Let

$$\xi_0 \equiv 0.2 \ , \qquad \varepsilon_0 \equiv 10^{-6} \ ,$$

and recall that the norm in (4.32) contains also a supremum taken over the complex parameter region $\mathcal{D}$,

$$\mathcal{D} \equiv \{\varepsilon \in \mathbb{C} : \ |\varepsilon| \le \varepsilon_0\} \ .$$

The evaluation of U_0, V_0, M_0 and $\widetilde{V}_0$ are easily obtained having computed the "explicit" (in the sense of interval arithmetic) form of the approximate solution (u, v). Having (u, v) given as

$$u^{(0)} = \sum_{j=1}^{5} \varepsilon^j \sum_{0<|n|\le\nu_j} \tilde{u}_n^{(j)} \exp(in\cdot\theta)\ , \qquad v^{(0)} = \sum_{j=0}^{5} \varepsilon^j \sum_{|n|\le\nu'_j} \tilde{v}_n^{(j)} \exp(in\cdot\theta)$$

(where the ν_j, ν'_j are the trigonometric degrees listed in Appendix A), we let U_0 and V_0 be upper bounds on

$$\sum_{j=1}^{5} \varepsilon_0^j \sum_{0<|n|\le\nu_j} |\tilde{u}_n^{(j)}| \exp(|n|\xi_0)\ , \qquad \sum_{j=0}^{5} \varepsilon_0^j \sum_{|n|\le\nu'_j} |\tilde{v}_n^{(j)}| \exp(|n|\xi_0)\ .$$

The computer-assisted evaluations of such sums yields the values[39]

$$\begin{aligned}
U_0^{(+)} &\equiv 1.3191129131218200558424179887264441 \cdot 10^{-5}\ ,\\
U_0^{(-)} &\equiv 1.3098175930298179879143983045402 15 \cdot 10^{-5}\ ,\\
V_0^{(+)} &\equiv 1.4562203598716627221542610477705 3\ ,\\
V_0^{(-)} &\equiv 1.4573172438104795787786953864709 7\ .
\end{aligned}$$

It is easy to check that these bounds imply

$$\Big(\sup_{\substack{\theta\in\Delta_{\xi_0}\\ \varepsilon\in\mathcal{D}}} |\theta + u^{(0)}|\ ,\ \sup_{\substack{\theta\in\Delta_{\xi_0}\\ \varepsilon\in\mathcal{D}}} |v^{(0)}|\Big) \in B\ .$$

Analogously, since[40]

$$\mathcal{M} \equiv I + \sum_{j=1}^{5} \varepsilon^j \sum_{0<|n|\le\nu_j} i\Big(\tilde{u}_n^{(j)} \otimes n\Big) \exp(in\cdot\theta)\ ,$$

$$v_\theta^{(0)} = \sum_{j=1}^{5} \varepsilon^j \sum_{|n|\le\nu'_j} i\Big(\tilde{v}_n^{(j)} \otimes n\Big) \exp(in\cdot\theta)\ ,$$

one gets the computer-assisted evaluations

[39] Recall that the plus sign corresponds to the case $\tilde{v}^{(0)} = (L_+, G_+)$ while the minus sign corresponds to $\tilde{v}^{(0)} = (L_-, G_-)$. Note also that interval arithmetic yields (as the name says) intervals with rational endpoints trapping the actual quantity one is computing; but since we need (usually) *upper bounds* we shall report only the right endpoints of the computed intervals which shall *define* our norm-bounds.

[40] If $a, b \in \mathbb{C}^N$, we denote by $a \otimes b$ the $N \times N$ matrix with entries $(a \otimes b)_{ij} = a_i b_j$. We also recall that by default we use the 1-norm on vectors so that the ("operator") norm on a 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ may be bounded as follows

$$\|A\| = \sup_{|x|_1=1} |Ax| \le \min\Big\{ \max\{|a|, |b|\} + \max\{|c|, |d|\}\ , \max\{|a|, |c|\} + \max\{|b|, |d|\}\Big\}\ .$$

$$
\begin{aligned}
M_0^{(+)} &\equiv 1.000021266465664701238677170875501\ , \\
M_0^{(-)} &\equiv 1.000021439075256126271432431716151\ , \\
\widetilde{V}_0^{(+)} &\equiv 7.453571858517957815592730443874432 \cdot 10^{-7}\ , \\
\widetilde{V}_0^{(-)} &\equiv 7.531270519848394217447212176311194 \cdot 10^{-7}\ .
\end{aligned} \tag{5.10}
$$

The evaluation of $\overline{M}_0$ is immediately obtained using (5.10):

$$
\begin{aligned}
\|\mathcal{M}^{-1}\|_{\xi_0} &\equiv \|(I + u_\theta)^{-1}\|_{\xi_0} \leq \frac{1}{1 - \|u_\theta\|_{\xi_0}} \leq \frac{1}{1 - \tilde{V}_0^{(\pm)}} \\
&\leq \overline{M}_0^{(\pm)}\ ,
\end{aligned}
$$

which leads to

$$
\begin{aligned}
\overline{M}_0^{(+)} &\equiv 1.000021266917936881337383036975 81\ , \\
\overline{M}_0^{(-)} &\equiv 1.000021439534899928447344044122831\ .
\end{aligned}
$$

To estimate $\widetilde{T}_0$ write the matrix $\mathcal{T}$ as

$$
\mathcal{T} \equiv A_0 + \mathcal{B}\ ,
$$

where $A_0 \equiv A(L_\pm, G_\pm)$ [$A(L, G)$ being the invertible matrix defined in (3.5)]. Hence

$$
\|\langle\mathcal{T}\rangle^{-1}\| = \|(I + A_0^{-1}\langle\mathcal{B}\rangle)^{-1}A_0^{-1}\| \leq \frac{\|A_0^{-1}\|}{1 - \|A_0\|^{-1}\|\mathcal{B}\|_0}\ .
$$

To estimate $\|\mathcal{B}\|_0$ we write explicitly the definition of $\mathcal{B}$ and use the following computer-assisted bounds

$$
\begin{aligned}
\|R\|_0 &\leq 1.437804901538657151668121552901 55\ , \\
\|R_L\|_0 &\leq 68.807159920918346704866573003865 1\ , \\
\|R_G\|_0 &\leq 64.306054561211428014777669344126 3\ , \\
\|R_{LL}\|_0 &\leq 56614.107211642990416531771025631 5\ , \\
\|R_{LG}\|_0 &\leq 56755.023460663906373116854159863 5\ , \\
\|R_{GG}\|_0 &\leq 28205.787242113238403228513377231 5\ .
\end{aligned}
$$

In this way one obtains

$$
\begin{aligned}
\widetilde{T}_0^{(+)} &\equiv 8.567100579602744072852019085722171\ , \\
\widetilde{T}_0^{(-)} &\equiv 8.559621778472845438209667021368791\ .
\end{aligned}
$$

A bit more delicate is the evaluation of F_0 and G_0, *i.e.* of the norms of the error functions f and g. We shall follow the classical *Cauchy's majorant method* (see [9], chapter 5 for generalities).

We recall that if

$$
\alpha(y) \equiv \sum_{k \in \mathbb{N}^p} \alpha_k (y - y_0)^k \qquad \text{and} \qquad \overline{\alpha}(y) \equiv \sum_{k \in \mathbb{N}^p} \overline{\alpha}_k (y - y_0)^k
$$

are smooth functions of y in a neighbourhood of $y_0 \in \mathbb{R}^p$, $\bar\alpha$ is said to *majorize* α and is denoted

$$\alpha \prec \bar\alpha\ ,$$

if

$$|\alpha_k| \le \bar\alpha_k\ , \qquad \forall\ k\ .$$

For example, for any $n \in \mathbb{Z}^p$,

$$\cos n\cdot x \prec \cosh(|n_1|x_1+\cdots+|n_p|x_p)\ , \quad \sin n\cdot x \prec \sinh(|n_1|x_1+\cdots+|n_p|x_p)\ . \tag{5.11}$$

We need a little technical lemma.

Lemma 5.1. *Let $\nu, p \in \mathbb{Z}_+$ and let*

$$\alpha(x,y) \equiv \sum_{n\in\mathbb{Z}^p:|n|\le\nu} c_n(x)\alpha_n(y)\ ,$$

where $c_n(x)$, for each n, is either $\cos n\cdot x$ or $\sin n\cdot x$ ($x \in \mathbb{T}^p$) and the α_n's are analytic functions on a complex ball around $y_0 \in \mathbb{R}^p$. Fix $M \in \mathbb{N}$. Let $b_0 \equiv y_0$ and, for $1 \le j \le M$, let $a^{(j)}(\theta)$ and $b^{(j)}(\theta)$ be given $\mathbb{R}^p$-valued functions of $\theta \in \mathbb{T}^p$ analytic on Δ_ξ for some $\xi > 0$. Let $\bar\alpha_n(y)$ be analytic functions such that

$$\alpha_n \prec \bar\alpha_n\ , \qquad (\forall\ |n| \le \nu)\ .$$

Let $\bar a_0 \equiv (\xi,\ldots,\xi)$, $\bar b_0 \equiv (|y_{01}|,\ldots,|y_{0p}|)$ and, for $1 \le j \le M$ let $\bar a^{(j)}$ and $\bar b^{(j)}$ be p-vectors with nonnegative components such that

$$\|a_i^{(j)}\|_\xi \le \bar a_i^{(j)}\ , \qquad \|b_i^{(j)}\|_\xi \le \bar b_i^{(j)}\ , \qquad (\forall\ 1 \le j \le M\ , \quad \forall\ 1 \le i \le p)\ .$$

Finally, denote

$$\bar c_n(x) \equiv \begin{cases} \cosh(|n_1|x_1+\cdots+|n_p|x_p) & \text{if } c_n = \cos n\cdot x \\ \sinh(|n_1|x_1+\cdots+|n_p|x_p) & \text{if } c_n = \sin n\cdot x \end{cases}$$

and let

$$\tilde\alpha(\varepsilon,\theta) \equiv \alpha\Big(\theta + \sum_{j=1}^M \varepsilon^j a^{(j)}(\theta), \sum_{j=0}^M \varepsilon^j b^{(j)}(\theta)\Big)\ ,$$

$$\beta(\varepsilon,\theta) \equiv \sum_{j\ge M+1} \varepsilon^j \Big[\tilde\alpha(\cdot,\theta)\Big]_j\ ,$$

$$\varphi(\varepsilon) \equiv \sum_{|n|\le\nu} \bar c_n\Big(\sum_{j=0}^M \varepsilon^j \bar a^{(j)}\Big)\ \bar\alpha_n\Big(\sum_{j=0}^M \varepsilon^j \bar b^{(j)}\Big)\ ,$$

where, of course, the operator $[\cdot]_j$ refers to ε-expansions. Then, for all $\bar\varepsilon > 0$,

$$\|[\tilde\alpha(\cdot,\theta)]_j\|_\xi \le [\varphi]_j\ , \qquad \forall\ j \ge 0\ ;$$

$$\sup_{|\varepsilon|\le\bar\varepsilon} \|\beta\|_\xi \le \varphi(\bar\varepsilon) - \sum_{j=0}^M \bar\varepsilon^j [\varphi]_j\ . \tag{5.12}$$

Proof. Observe that, by (5.11), for any fixed $\theta \in \Delta_\xi$, one has

$$c_n\Big(\sum_{j=0}^{M} \varepsilon^j a^{(j)}(\theta)\Big) \prec \overline{c}_n\Big(\sum_{j=0}^{M} \varepsilon^j \overline{a}^{(j)}\Big) ,$$

and also that[41]

$$\alpha_n\Big(\sum_{j=0}^{M} \varepsilon^j b^{(j)}(\theta)\Big) \prec \overline{\alpha}_n\Big(\sum_{j=0}^{M} \varepsilon^j \overline{b}^{(j)}\Big) .$$

These relations imply that for all $\theta \in \Delta_\xi$, $\tilde{\alpha}(\cdot,\theta) \prec \varphi$, which implies immediately (5.12). □

In order to apply the lemma to our situations we need the following obvious majorizations. Let $z \in \mathbb{C}$ with $|z| < 1$; let[42] y_{10}, y_{20} be positive numbers and let $|y_1 - y_{10}| < 1$, $|y_2 - y_{20}| < 1$; finally let s be a positive integer. Then

$$\begin{aligned}
&\sqrt{1-z} \prec 2 - \sqrt{1-z}\,, \qquad (1-z)^{-\frac{1}{2}} \prec (1-z)^{-\frac{1}{2}}\,;\\
&\frac{y_2}{y_1} \prec \frac{y_2}{2y_{10}-y_1}\,, \qquad \Big(\frac{y_2}{y_1}\Big)^2 \prec \Big(\frac{y_2}{2y_{10}-y_1}\Big)^2\,;\\
&e(y_1,y_2) \prec \overline{e}(y_1,y_2) \equiv 2 - \sqrt{1-\Big(\frac{y_2}{2y_{10}-y_1}\Big)^2}\,;\\
&e^{-1}(y_1,y_2) \prec \left(1-\Big(\frac{y_2}{2y_{10}-y_1}\Big)^2\right)^{-1/2}\,;\\
&y_1^{-s} \prec \Big(\frac{1}{2y_{10}-y_1}\Big)^s\,, \qquad -2\Big(\frac{1}{2y_1^2}-y_2\Big) \prec 2y_2 + \Big(\frac{1}{2y_{10}-y_1}\Big)^2\,.
\end{aligned}$$

Using the above observations, one gets the following evaluations of the norm bounds F_0 and G_0:

$$\begin{aligned}
F_0^{(+)} &\equiv 1.8755301827538371921261973556 72179 \cdot 10^{-23}\,,\\
F_0^{(-)} &\equiv 1.881443228644026073096600123652624 \cdot 10^{-23}\,,\\
G_0^{(+)} &\equiv 1.524340973886308626744645639896068 \cdot 10^{-20}\,,\\
G_0^{(-)} &\equiv 1.573241367452682546591639160273634 \cdot 10^{-20}\,.
\end{aligned}$$

5.3. Step 3: Application of the KAM algorithm and of the KAM Theorem. We proceed to apply the KAM algorithm and the KAM Theorem worked out in Sect. 4.3: recall 3) of Sect. 1 and Remark 4.4. Also this step is computer-assisted; however double precision (rather than quadruple precision) will usually be good enough. We discuss in detail the case with initial data (L_+, G_+); the case with initial data (L_-, G_-) is completely analogous[43] and we spare the reader more data which would not shed much more light.

[41] Recall the basic facts of majorization theory (see, e.g., [9]): sum, multiplication and composition are preserved by majorizations.

[42] To avoid confusion with the norm parameter G_0 we denote here the action-variables by (y_1, y_2).

[43] Actually, the (L_+, G_+)-case "converges" after three iterations of $\mathcal{K}$ while the (L_-, G_-) case after four.

Remark 5.2. The quantities with an index 0 refer (as in Step 2 above) to our choice of the *initial* approximate solution. To such initial approximate solution we apply a few times [respectively three times for the case with initial data (L_+, G_+) and four times for the case with initial data (L_-, G_-)] the KAM algorithm obtaining new approximate solutions $(u^{(1)}, v^{(1)})$,...,$(u^{(3)}, v^{(3)})$ [for the case with initial data (L_+, G_+), while for the case with initial data (L_-, G_-) we consider also $(u^{(4)}, v^{(4)})$]. The *input* of the KAM Theorem will be respectively $(u, v) \equiv (u^{(3)}, v^{(3)})$ (with the relative norm bounds given by the KAM algorithm) for the case with initial data (L_+, G_+) and $(u, v) \equiv (u^{(4)}, v^{(4)})$ for the case with initial data (L_-, G_-). *Quantities referring to the input of the KAM Theorem will carry no index* (and beware of the difference between $\tilde{u}^{(j)}$ and $u^{(j)}$).

Since we shall consider only the case with initial data (L_+, G_+) we shall drop from the notation the suffix $^{(+)}$.

To apply the KAM algorithm we have to fix the values of the "angle" analyticity widths, namely the values of $\bar{\xi}$ and ξ_0 [see (4.23)] and of $\xi_1 \equiv \xi'$ [see (4.24)] and[44] $\xi_2 \equiv \xi_1'$, $\xi_3 \equiv \xi_2'$. We choose[45]

$$\bar{\xi} \equiv 0.201\ , \quad \xi_0 \equiv 0.2\ , \quad \xi_1 = 0.1\ ,$$
$$\xi_2 = 0.05\ , \quad \xi_3 \equiv 0.0025\ .$$

We also fix the value r appearing in (4.23) as $r \equiv 0.001$.

After three steps of the KAM algorithm we shall apply the KAM Theorem, we therefore let

$$\xi \equiv \xi_3 \equiv 0.0025\ , \qquad \hat{\xi} \equiv \frac{\xi}{2} \equiv 0.00125\ .$$

Next, we need estimates on the derivatives of the Hamiltonian $h(x, y)$ appearing in the definitions of Ω,...η in Theorem 1.1 [recall (4.21)]. These estimates (which do not change in the iteration) are straightforward and one obtains:

$$\begin{aligned}
\overline{E} &\le 0.83456917806280141684592592214367 6\ ,\\
E_{0,2} &\le 56.548643863316208467789066978800 6\ ,\\
E_{1,1} &\le 1.1866945834052331676669774467302 61 \cdot 10^{-3}\ ,\\
E_{0,3} &\le 24606.142191612735926875935854661 2\ ,\\
E_{1,2} &\le 3.9626581982761491937011045794301 0\ ,\\
E_{2,1} &\le 2581.0532581721429733799270162472 5\ ,\\
E_{3,0} &\le 1.7479497240057940449872322169573 51 \cdot 10^{-4}\ .
\end{aligned}$$

Now we iterate the map $\mathcal{K}$ three times: conditions (4.23), (4.26) and (4.28) are satisfied for $j = 0, 1, 2, 3$ (as it is easy to check using the values $(U_j, ..., \widetilde{T}_j)$ reported here). The (double precision) iterated values of the map $\mathcal{K}$ are the following. After the first iteration:

$$\begin{aligned}
&U_1 \le 1.309817593322689 \cdot 10^{-5}\ , \quad V_1 \le 1.45622036006177\ ,\\
&M_1 \le 1.00002126646582\ , \quad \overline{M}_1 \le 1.00002126691809\ ,\\
&\tilde{V}_1 \le 7.453583421789768 \cdot 10^{-7}\ , \quad F_1 \le 4.006354647833255 \cdot 10^{-24}\ ,\\
&G_1 \le 1.368656716949379 \cdot 10^{-25}\ , \quad \tilde{T}_1 \le 0.834604675233853\ .
\end{aligned}$$

[44] Recall that $\xi_{j+1} = \xi_j - \delta_j$, where the δ_j's are such that $\delta_0 > \delta_1 > \delta_2$... and $\sum \delta_j < \xi$.

[45] These values have been chosen by "optimizing" (trial and error) the KAM algorithm.

After the second iteration:

$$
\begin{aligned}
&U_2 \le 1.309817593322718 \cdot 10^{-5} , \quad V_2 \le 1.45622036006177 , \\
&M_2 \le 1.00002126646582 , \quad \overline{M}_2 \le 1.00002126691809 , \\
&\tilde{V}_2 \le 7.453583426051443 \cdot 10^{-7} , \quad F_2 \le 1.209224397282491 \cdot 10^{-31} , \\
&G_2 \le 2.301233775774239 \cdot 10^{-33} , \quad \tilde{T}_2 \le 0.834604675233862 .
\end{aligned}
$$

After the third iteration:

$$
\begin{aligned}
&U_3 \le 1.309817593322718 \cdot 10^{-5} , \quad V_3 \le 1.45622036006177 , \\
&M_3 \le 1.00002126646582 , \quad \overline{M}_3 \le 1.00002126691809 , \\
&\tilde{V}_3 \le 7.453583426051448 \cdot 10^{-7} , \quad F_3 \le 8.855523608162042 \cdot 10^{-44} , \\
&G_3 \le 3.174732732716713 \cdot 10^{-46} , \quad \tilde{T}_3 \le 0.834604675233862 .
\end{aligned}
$$

The quantities Ω_1,...,η defined in Theorem 4.2 (recalling Theorem 4.1) are immediately computed using the above values and one obtains

$$\alpha \le 88530999255.1887 , \quad \eta \le 1.185668436207269 \cdot 10^{-38} .$$

With such values condition (4.31) is satisfied, in fact we obtained[46]

$$\eta\alpha \, M^7 \overline{M}^9 \xi^{-2(2\tau+1)} 2^{16\tau+23} \, \tau!^4 \le 3.584973875295102 \cdot 10^{-8} < 1 ,$$

so that Theorem 1.1 holds. $\square$

A. Some Computer-Assisted Data

We first report the *trigonometric degrees* ν_j, ν_j' appearing in the Fourier-Taylor expansion of the approximate solution $(u^{(0)}, v^{(0)})$ (obtained as the fifth ε-order truncation of the formal solution)

$$u^{(0)} = \sum_{j=1}^{5} \varepsilon^j \sum_{0<|n|\le\nu_j} \tilde{u}_n^{(j)} \exp in\cdot\theta , \qquad v^{(0)} = \sum_{j=0}^{5} \varepsilon^j \sum_{0<|n|\le\nu_j'} \tilde{v}_n^{(j)} \exp in\cdot\theta .$$

We give the result by components: for $\tilde{u}^{(j)}$ we found:

Order j	Fourier deg. of $\tilde{u}_1^{(j)}$	Fourier deg. of $\tilde{u}_2^{(j)}$
1	5	5
2	15	15
3	25	25
4	35	35
5	45	45

For $\tilde{v}^{(j)}$:

[46] According to the above convention M and $\overline{M}$ are, in fact, the values M_3 and $\overline{M}_3$.

Order j	Fourier deg. of $\tilde{v}_1^{(j)}$	Fourier deg. of $\tilde{v}_2^{(j)}$
0	0	0
1	5	5
2	14	11
3	24	21
4	34	31
5	44	41

The *number of Fourier components* of $u^{(0)}$ and of $v^{(0)}$ is given for each order of the Taylor series expansion in powers of the perturbing parameter ε. We report only those components with Fourier index (n, m) with $n > 0$ or $n = 0$ and $m \geq 0$.

For $u^{(0)}$ we found:

Order j	Fourier coeff. of $\tilde{u}_1^{(j)}$	Fourier coeff. of $\tilde{u}_2^{(j)}$
0	0	0
1	5	5
2	35	35
3	93	93
4	179	179
5	293	293

For $v^{(0)}$ we found:

Order j	Fourier coeff. of $\tilde{v}_1^{(j)}$	Fourier coeff. of $\tilde{v}_2^{(j)}$
0	6	5
1	34	26
2	92	76
3	178	156
4	292	264
5	436	400

Finally, we report now the intervals containing the five components of the function $\tilde{u}_1^{(1)}$. For the initial data (L_+, G_+) we obtained[47]:

$$
\begin{aligned}
&0.673056264392395519996544917278 0 + [58, 87] \cdot 10^{-33} \,, \\
&0.297906244730211719152652671857 1 + [77, 85] \cdot 10^{-33} \,, \\
&1.482946598095398011285563691781 + [19, 25] \cdot 10^{-32} \,, \\
&-0.458934417488553736667977972251 2 + [53, 41] \cdot 10^{-33} \,, \\
&-0.201074523046940409405383663516 4 + [22, 14] \cdot 10^{-33} \,.
\end{aligned}
$$

For the initial data (L_-, G_-) we obtained:

[47] The notation $x + \delta[a, b]$, with $\delta > 0$ and $a < b$, means $[x + \delta a, x + \delta b]$: for example,

$$
\begin{aligned}
&0.6730562643923955199965449172780 + [58, 87] \cdot 10^{-33} = \\
&[0.67305626439239551999654491727805 8, 0.67305626439239551999654491727808 7] \,.
\end{aligned}
$$

$$\begin{aligned}
&0.676604357788015853965640275018\,3 + [17, 48] \cdot 10^{-33}\ ,\\
&0.306184561199587897367472922674\,6 + [45, 54] \cdot 10^{-33}\ ,\\
&1.488798647829940896171271103877 + [83, 90] \cdot 10^{-32}\ ,\\
&-0.465944384342636496164611377214\,3 + [22, 11] \cdot 10^{-33}\ ,\\
&-0.201889937070167172193315337178\,8 + [70, 62] \cdot 10^{-33}\ .
\end{aligned}$$

References

1. Arnold, V. I. (editor): Encyclopaedia of Mathematical Sciences. Dynamical Systems III, Berlin-Heidelberg-New York: Springer-Verlag **3**, 1988
2. Arnold, V. I.: Small divisor problems in classical and Celestial Mechanics. Russ. Math. Surv. **18**, 85–191 (1963)
3. Celletti, A.: Analysis of resonances in the spin-orbit problem in Celestial Mechanics: The synchronous resonance (Part I). J. of App. Math. and Phys. (ZAMP) **41**, 174 (1990)
4. Celletti, A., Chierchia, L.: Construction of Analytic KAM Surfaces and Effective Stability Bounds. Commun. Math. Phys. **118**, 119–161 (1988)
5. Celletti, A., Chierchia, L.: A Constructive Theory of Lagrangian Tori and Computer-assisted Applications. In: Dynamics Reported, C.K.R.T. Jones, U. Kirchgraber, H.O. Walther, Managing Editors, Vol. **4** (New Series), Berlin-Heidelberg-New York: Springer–Verlag, 1995
6. Delaunay, C.: Théorie du Mouvement de la Lune. Mémoires de l'Académie des Sciences **1**, Tome XXVIII, Paris (1860)
7. Eckmann, J.-P., Wittwer, P.: Computer methods and Borel sommability applied to Feigenbaum's equation. Berlin-Heidelberg-New York: Springer Lecture Notes in Physics **227**, 1985
8. Hénon, M.: Explorationes numérique du problème restreint IV: Masses egales, orbites non periodique. Bullettin Astronomique **3**, no. 1, fasc. 2, 49–66 (1966)
9. John, F.: Partial Differential Equations. Berlin-Heidelberg-New York: Springer-Verlag, 1982
10. Koch, H., Schenker, A., Wittwer, P.: Computer-Assisted Proofs in Analysis and Programming in Logic: A Case Study. Preprint Université de Genève (1994)
11. Meyer, K.R., Schmidt, D. (editors): Computer Aided Proofs in Analysis. Berlin-Heidelberg-New York: Springer-Verlag, 1991
12. Moser, J.: Stable and random motions in dynamical systems, with special emphasis on Celestial Mechanics. Ann. Math. Studies; **77**, Princeton, NJ: Princeton University Press, 1973
13. Poincarè, H.: Les Methodes Nouvelles de la Mechanique Celeste. Paris: Gauthier Villars, 1892
14. Robutel, P.: Stability of the planetary three-body problem. II. KAM Theory and existence of quasi-periodic motions. Celestial Mechanics **62**, 219–261 (1995)
15. Salamon, D., Zehnder, E.: KAM theory in configuration space. Comment. Math. Helveticae **64**, 84 (1989)
16. (no author listed) The Astronomical Almanac. Washington, D.C.: U.S. Government Printing Office, 1990
17. (no author listed) Vax Architecture handbook. Digital Equipment Corporation, 1981

Communicated by Ya.G. Sinai