Ergod. Th. & Dynam. Sys. (2009), **29**, 849–873
doi:10.1017/S0143385708000503

Analytic Lagrangian tori for the planetary many-body problem

LUIGI CHIERCHIA† and FABIO PUSATERI‡

† *Dipartimento di Matematica, Università 'Roma Tre', Largo S. L. Murialdo 1, I-00146 Roma, Italy*
(e-mail: luigi@mat.uniroma3.it)

‡ *Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012, USA*
(e-mail: pusateri@cims.nyu.edu)

(*Received* 20 *December* 2007 *and accepted in revised form* 23 *April* 2008)

Abstract. In 2004, Féjoz [Démonstration du 'théoréme d'Arnold' sur la stabilité du système planétaire (d'après M. Herman). *Ergod. Th. & Dynam. Sys.* **24**(5) (2004), 1521–1582], completing investigations of Herman's [Démonstration d'un théoréme de V.I. Arnold. *Séminaire de Systémes Dynamiques et manuscripts*, 1998], gave a complete proof of 'Arnold's Theorem' [V. I. Arnol'd. Small denominators and problems of stability of motion in classical and celestial mechanics. *Uspekhi Mat. Nauk.* **18**(6(114)) (1963), 91–192] on the planetary many-body problem, establishing, in particular, the existence of a positive measure set of smooth (C^∞) Lagrangian invariant tori for the planetary many-body problem. Here, using Rüßmann's 2001 KAM theory [H. Rüßmann. Invariant tori in non-degenerate nearly integrable Hamiltonian systems. *R. & C. Dynamics* **2**(6) (2001), 119–203], we prove the above result in the real-analytic class.

1. *Introduction*

The planetary many-body problem consists of studying the evolution of $(1+n)$ bodies (point masses), subject only to the mutual gravitational attraction, in the case where one of the bodies (the 'Sun') has a mass m_0 considerably larger than the masses m_i of the remaining n bodies (the 'planets'). The Newtonian evolution equations for such problem (in suitable units) are given by

$$m_j \ddot{q}_j = \sum_{k \neq j} m_j m_k \frac{q_k - q_j}{|q_k - q_j|^3}, \qquad j = 0, 1, \ldots, n, \tag{1}$$

where $q_j = q_j(t) \in \mathbb{R}^3$ denotes the position at time t of the jth body, '$|\cdot|$' denotes the Euclidean norm and '$\dot{\ }$' denotes the time derivative.

In [**1**, Ch. III, p. 125], Arnold made the following statement§.

§ The integer n in Arnold's statement corresponds to the above $(1+n)$.

Arnold's statement 1. In the n-body problem there exists a set of initial conditions having a positive Lebesgue measure and such that if the initial positions and velocities belong to this set, the distances of the bodies from each other will remain perpetually bounded.

As is well known, such a statement solves a fundamental problem considered, for several centuries, by astronomers and mathematicians. However, Arnold considered in details only the planar three-body case† and it appears that his indication for extending the result to the general case contains a flaw; cf. the end of [**9**, §1.2, p. 1524].

A complete general proof of Arnold's statement was given in 2004, when Féjoz, completing the work of Herman, proved the following‡.

THEOREM 1. (Arnold, Herman, Féjoz [**9**, §1.2, p. 1523, Théorème 1]) *Si le maximum* $\epsilon = \max\{m_j/m_0\}_{j=1,\dots,n}$ *des masses des planètes rapportées à la masse du soleil est suffisamment petit, les équations* (1) *admettent, dans l'espace des phases au voisinage des mouvements képlériens circulaires et coplanaires, un ensemble de mesure de Lebesgue strictement positive de conditions initiales conduisant à des mouvements quasipériodiques.*

The beautiful proof of this result given in [**9**] (see also [**10**]) relies, on one side, on the elegant C^∞ KAM theory worked out by Herman [**9**, §§2–5], and, on the other side, on the analytical celestial mechanics worked out, in particular, by Poincaré and clarified and further investigated in Paris in the late 1980s by Chenciner and Laskar in the *Bureau des Longitudes*§ and later by Herman himself.

The invariant tori associated with the motions provided by Theorem 1, in view of the just mentioned KAM tools, are C^∞. Now, since the many-body problem is formulated in terms of real-analytic functions, it appears somewhat more natural to seek for *real-analytic invariant manifolds*. This is the problem addressed in this paper. In particular, we shall give a new proof of Arnold's statement, proving the following.

THEOREM 2. *If* $\epsilon = \max\{m_j/m_0\}_{j=1,\dots,n}$ *is small enough, there exists a strictly positive measure set of initial conditions for the* $(1+n)$*-planetary problem (1), whose time evolutions lie on real-analytic Lagrangian tori in the* $6n$*-dimensional phase space*

$$\mathcal{M} := \left\{(q, p) \in \mathbb{R}^{6(1+n)} \mid q_j \neq q_k, \forall j \neq k \text{ and } \sum_{j=0}^{n} p_j = 0 = \sum_{j=0}^{n} m_j q_j\right\}$$

endowed with the restriction of the standard symplectic form $\sum_{j=0}^{n} dq_j \wedge dp_j = \sum_{\substack{0\le j\le n\\ 1\le k\le 3}} dq_{j,k} \wedge dp_{j,k}$.

Remark 1. Let us collect here a few observations concerning the above statements and respective proofs.

† A few lines after the above-reported statement in [**1**, Ch. III, p. 125], Arnold states: '*We shall consider only the plane three-body problem in detail.* [$\cdots$]. *In the final section a brief indication is given of the way in which the fundamental theorem of Chapter IV is applied in the investigation of the planetary motions in the plane and spatial many-body problems.*'.

‡ For a more detailed statement, see the statements in Remark 1(v).

§ Compare, e.g., the *Notes Scientifiques et techniques du Bureau des Longitudes S 026* and *S 028* by, respectively, Chenciner and Laskar, and Chenciner.

(i) The proof of Theorem 2 given below is similar in strategy to that in [**9**] but technically different and it is based on an *analytic* (rather than smooth) KAM theory for properly degenerate Hamiltonian systems (see also point (iv) below). On the other hand, it is conceivable—notwithstanding the presence of strong degeneracies (see points (ii) and (iii) below)—to prove regularity and uniqueness results for the planetary problem so as to deduce that the invariant tori in [**9**] are indeed analytic, a fact which does not follow from our proofs†.

(ii) The evolution equations (1) are Hamiltonian and admit seven integrals, namely, the Hamiltonian (energy)

$$H := \sum_{j=0}^{n} \frac{|p_j|^2}{2m_j} - \sum_{0 \le k < j \le n} \frac{m_j m_k}{|q_k - q_j|},$$

the three components of the total linear momentum $M := \sum_{j=0}^{n} p_j$ and the three components of the total angular momentum $C := \sum_{j=0}^{n} p_j \times q_j$, where '$\times$' denotes the usual skew vector product in $\mathbb{R}^3$. As a reflection of the invariance of Newton's equation (1) under changes of inertial reference frames, the Hamiltonian system associated with the $(1 + n)$-body problem may be studied on the symplectic, invariant $6n$-dimensional manifold $\mathcal{M}$ defined above, where, in addition to the total linear momentum, also the coordinates of the barycenter of the system vanish ('reduction of the total linear momentum'). However, the reduced $(1 + n)$-body Hamiltonian still admits, in addition to the energy, three integrals given by the components of $C = (C_x, C_y, C_z)$. Incidentally, such integrals are not commuting since, if $\{\cdot, \cdot\}$ denotes the natural Poisson bracket on $\mathcal{M}$, one has the cyclical relations $\{C_x, C_y\} = C_z$, $\{C_y, C_z\} = C_x$ and $\{C_z, C_x\} = C_y$; but, for example, $|C|^2$ and C_z are two independent, commuting integrals.

(iii) The reasons why, notwithstanding the development of KAM theory in the early 1960s, it took so long to give a complete proof of Arnold's statement are technical in nature and are related to the strong degeneracies of the planetary problem (degeneracies, which are related to the abundance of integrals mentioned in (ii)). The planetary $(1 + n)$-body problem is perturbative, the unperturbed limit being obtained by considering n decoupled two-body problems formed by the Sun and the jth planet. Now, the two-body problem in space is a three-degrees-of-freedom problem, but, once it is put into (Delaunay) action-angle variables, it depends only on one action (the action L proportional to the square root of the semi-major axis of the Keplerian ellipse on which the two bodies revolve). Systems of this kind are called *properly degenerate* and standard KAM theory does not apply. This difficulty, however, was overcome by Arnold—essentially by refined normal form theory—in the case of the planar three-body case, to which he could apply his 'fundamental theorem' [**1**, Ch. IV]. Indeed, Arnold's approach, in view of Jacobi's reduction of the nodes, could be extended [**17**] to the spatial three-body case ($n = 2$) but not to the general case‡ (spatial, $n > 2$). Furthermore (but not independently), in higher

† Recent interesting progresses in the study of uniqueness of invariant Lagrangian manifolds appeared in [**5**] and, in particular, in [**8**]; however, as far as regularity is concerned, to the best of our knowledge, the only complete proven statement is [**19**, Theorem 4, §4, p. 34], which covers only the C^∞ non-degenerate case.

‡ However, for a remarkable symplectic extension of Jacobi's reduction of the nodes in higher dimension, see [**4, 7**].

dimension, there appear two *secular resonances* (see (52) below), which prevent direct application of any kind of KAM machinery. We mention that the way we overcome this last difficulty here is slightly different from that used in [**9**]: roughly speaking, in [**9**] it is introduced a modified Hamiltonian, which is then considered on the symplectic submanifold of vertical total angular momentum; here, we consider, instead, an extended phase space by adding an extra degree of freedom and consider on it a modified non-degenerate Hamiltonian.

(iv) The main technical tool for us is the analytic KAM theory for weakly non-degenerate systems worked out by Rüßmann in [**18**]; the main results of Rüßmann's theory (in the case of Lagrangian tori) are recalled in §2.1 (see, also, Lemma 8 in §2.3.3). The extension of this theory to properly degenerate systems is explained in §2.2 and proved in §2.3 (which constitutes the longest and most technical part of the paper). In §3, using several results reported in [**9**], the proof of Theorem 2 is given.

(v) Finally, we mention very briefly a few problems related to the context considered here.

- Describe, in detail, the motions that take place on the Lagrangian tori. Let us clarify this point. From the proof of Theorem 2 given below, in view of the indirect argument used [**9**, Lemma 82, p. 1578] we cannot conclude that the 'true' motion is quasi-periodic; on the other hand, using different arguments, Arnold and Féjoz say that the motion, in the general case, is quasi-periodic and takes place on $(3n-1)$-dimensional tori.

  Statement 1. (Arnold [**1**, p. 127]) 'Thus, the Lagrangian motion is conditionally periodic and to the n_0 'rapid' frequencies of the Keplerian motion are added n_0 (in the plane problem) or $2n_0 - 1$ (in the space problem) 'slow' frequencies of the secular motions'.

  Statement 2. (Féjoz [**9**, p. 1566]) *THÉORÈME 60. Pour toute valeur des masses $m_0, m_1, \ldots, m_n > 0$ et des demi grands axes $a_1 > \cdots > a_n > 0$, il existe un réel $\epsilon_0 > 0$ tel que, pour tout ϵ tel que $0 < \epsilon < \epsilon_0$, le flot de l'hamiltonien F (défini en (28)) possède un ensemble de mesure de Lebesgue strictement positive de tores invariants de dimension $3n-1$, de classe C^∞, quasipériodiques et ϵ-proches en topologie C^0 des tores képlériens de demi grands axes $(a_1, \ldots, a_n)$ et d'excentricités et d'inclinaisons relatives nulles; de plus, quand tend vers zéro la densité des tores invariants au voisinage de ces tores képlériens tend vers un.*

  Moreover, in the spatial three-body case ($n = 2$) the Lagrangian tori are actually four-dimensional (not $5 = 3n - 1$) and the number of *independent* frequencies is four (cf. [**17**]).
- Use Boigey–Deprit's symplectic variables [**4, 7**] and try to extend Arnold's approach to the general spatial case.
- Give asymptotic (as $\epsilon \to 0$) estimates on the measure of Lagrangian invariant tori.
- Apply some of the above results to subsystems of the Solar system (for some progress in this direction, see [**6**]).

https://doi.org/10.1017/S0143385708000503 Published online by Cambridge University Press

2. *Analytic Lagrangian tori for properly degenerate systems*

In this section we first recall a result due to Rüßmann concerning analytic perturbations of weakly non-degenerate Hamiltonian systems (§2.1) and then show how such a result may be used to give an analytic version of Herman's C^∞ KAM theorem on properly degenerate systems (i.e. nearly integrable systems, which when the perturbation parameter vanishes depend on less action variables than the number of degrees of freedom). The statement of the analytic theorem for properly degenerate systems is given in §2.2 and is proved in §2.3.

2.1. *Rüßmann's theorem for weakly non-degenerate systems.*

We start by fixing some notation.

- If $a, b \in \mathbb{R}^n$, then $\langle a, b\rangle := \sum_{i=1}^n a_i b_i$ and $|a| := |a|_2 := \langle a, a\rangle^{1/2}$.
- If g is a μ-times continuously differentiable function ($\mu \in \mathbb{N}$) from an open set $B \subset \mathbb{R}^n$ to $\mathbb{R}^m$, the μth (tensor) derivative of g in $b \in B$ is denoted by $(a_1, \ldots, a_\mu) \to \partial^\mu g(b)(a_1, \ldots, a_\mu)$, $a_j \in \mathbb{R}^n$, $j = 1, \ldots, \mu$; if $a_1 = \cdots = a_\mu$, we shall write $\partial^\mu g(b)(a)^\mu$.
- We have $|\partial^\mu g(b)| := \max_{a\in\mathbb{R}^n, |a|=1} |\partial^\mu g(b)(a, \ldots, a)|$ and $|\partial^\mu g|_A := \sup_{b\in A} |\partial^\mu g(b)|$.
- By $C^\mu(B, \mathbb{R}^m)$ we denote the Banach space of all μ-times continuously differentiable functions $g : B \to \mathbb{R}^m$ with bounded derivatives up to order μ, endowed with the norm $|g|_B^\mu = \sup_{0\le\nu\le\mu} |\partial^\nu g|_B < \infty$.

The key notion of non-degeneracy is as follows†.

Definition 1. (Rüßmann non-degeneracy condition) A real-analytic function

$$\omega : y \in B \subset \mathbb{R}^n \longrightarrow \omega(y) = (\omega_1(y), \ldots, \omega_m(y)) \in \mathbb{R}^m$$

is called *R-non-degenerate* if B is a non-empty open connected set in $\mathbb{R}^n$ and if for any $c = (c_1, \ldots, c_m) \in \mathbb{R}^m \setminus \{0\}$ one has

$$y \longrightarrow \langle c, \omega\rangle := \sum_{i=1}^m c_i \omega_i \neq 0$$

or, equivalently, if the range $\omega(B)$ of ω does not lie in any $(m-1)$-dimensional linear subspace of $\mathbb{R}^m$. We call ω *R-degenerate* if it is not R-non-degenerate.

The following lemma is a simple consequence of R-non-degeneracy and analyticity.

LEMMA 1. *Let $\omega : B \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be R-non-degenerate. Then for any non-empty compact set $\mathcal{K} \subset B$ there exist numbers $\mu_0 = \mu_0(\omega, \mathcal{K}) \in \mathbb{Z}_+$ and $\beta = \beta(\omega, \mathcal{K}) > 0$ such that*

† This terminology, nowadays, seems to be standard (see, e.g., [**20**]); however many authors, in addition to Rüßmann, contributed to its formulation, including Arnold, Margulis, Pyartli, Parasyuk, Bakhtin and Sprindzhuk.

https://doi.org/10.1017/S0143385708000503 Published online by Cambridge University Press

$$\max_{0\le\mu\le\mu_0} |\partial_y^{\mu}\langle c, \omega(y)\rangle^2| \ge \beta \quad \textit{for all } c \in \mathcal{S}^{m-1}, \textit{ for all } y \in \mathcal{K} \tag{2}$$

where $\mathcal{S}^{m-1} := \{c \in \mathbb{R}^m : |c|_2 = 1\}$.

For the proof see [**18**, Lemma 18.2, p. 185].

In view of Lemma 1 one can give the following.

Definition 2. Let $\mathcal{K}$ and B be as in the preceding lemma and let $\omega : y \in B \longrightarrow \mathbb{R}^m$ be a real-analytic and R-non-degenerate function. We define $\mu_0(\omega, \mathcal{K}) \in \mathbb{Z}_+$, the *index of non-degeneracy* of ω with respect to $\mathcal{K}$, as the smallest positive integer such that

$$\beta := \min_{y\in\mathcal{K},\, c\in\mathcal{S}^{m-1}} \max_{0\le\mu\le\mu_0} |\partial^{\mu}\langle c, \omega(y)\rangle^2| > 0. \tag{3}$$

The number $\beta = \beta(\omega, \mathcal{K})$ is called the *amount of non-degeneracy* of ω with respect to $\mathcal{K}$.

Remark 2. If a real-analytic function $\omega : B \to \mathbb{R}^m$ admits the existence of μ_0 and β as in (3), for some compact $\mathcal{K}$ containing an open ball, then it is R-non-degenerate.

The following result, which concerns the existence of maximal (Lagrangian) tori only, is a particular case of the main theorem in [**18**], where lower-dimensional tori are also treated†.

THEOREM 3. (Rüßmann [**18**]) *Let $\mathcal{Y}$ be an open connected set of $\mathbb{R}^n$ and $\mathbb{T}^n$ the usual n-dimensional torus $\mathbb{R}^n/2\pi\mathbb{Z}^n$. Consider a real-analytic Hamiltonian*

$$H(x, y) = h(y) + P(x, y)$$

defined for $(x, y) \in \mathbb{T}^n \times \mathcal{Y}$ endowed with the standard symplectic form $dx \wedge dy$. Let $\mathcal{K}$ be any compact subset of $\mathcal{Y}$ with positive n-dimensional Lebesgue measure $\mathrm{meas}_n\, \mathcal{K} > 0$ *and fix $0 < \epsilon^\star <$ $\mathrm{meas}_n\, \mathcal{K}$. Let $\mathcal{A}$ be an open set in $\mathbb{C}^n/2\pi\mathbb{Z}^n \times \mathbb{C}^n$ on which H can be analytically extended and such that $\mathbb{T}^n \times \mathcal{K} \subset \mathcal{A}$. Assume that the frequency application $\omega := \nabla h$ is R-non-degenerate on $\mathcal{Y}$; let μ be any integer greater than or equal to $\mu_0(\omega, \mathcal{K})$ (the index of non-degeneracy of ω with respect to $\mathcal{K}$) and let β be as in (3) with μ_0 replaced by μ.*

Then, for any fixed $\tau > n\mu$, there exist $\epsilon_0 = \epsilon_0(\epsilon^\star, n, \mu, \beta, \tau, \omega, \mathcal{K}) > 0$ and $\gamma = \gamma(\epsilon^\star, n, \tau, \mu, \beta, \omega, \mathcal{K}) > 0$ such that if

$$|P|_{\mathcal{A}} := \sup_{\mathcal{A}} |P| \le \epsilon_0 \tag{4}$$

the following is true. There exist a compact set

$$\mathcal{K}^\star \subset \mathcal{K} \quad \text{with } \mathrm{meas}_n\ \mathcal{K}^\star > \mathrm{meas}_n\ \mathcal{K} - \epsilon^\star \tag{5}$$

and a Lipschitz mapping

$$X : (b, \xi, \eta) \in \mathcal{K}^\star \times \mathbb{T}^n \times \mathcal{U} \longrightarrow \mathbb{T}^n \times \mathcal{Y},$$

where $\mathcal{U}$ is an open neighborhood of the origin in $\mathbb{R}^n$, such that:

† See [**18**, Theorem 1.7, p. 127]. We refer to [**15**] for more details on how to obtain Theorem 3 from the general results in [**18**].

https://doi.org/10.1017/S0143385708000503 Published online by Cambridge University Press

(i) *the mapping*

$$(\xi, \eta) \longmapsto (x, y) = X(b, \xi, \eta)$$

defines, for every $b \in \mathcal{K}^\star$, a real-analytic symplectic transformation† *close to the identity on $\mathbb{T}^n \times \mathcal{U}$;*

(ii) *the map*

$$(b, \xi) \in \mathcal{K}^\star \times \mathbb{T}^n \longrightarrow X_0(b, \xi) := X(b, \xi, 0)$$

is a bi-Lipschitz homeomorphism;

(iii) *the transformed Hamiltonian $H^\star := H \circ X$ is in the form*‡*:*

$$H^\star(b, \xi, \eta) = h^\star(b) + \langle \omega^\star(b), \eta \rangle + O(|\eta|^2)$$

for every $b \in \mathcal{K}^\star$ and $(\xi, \eta) \in \mathbb{T}^n \times \mathcal{U}$;

(iv) *the new frequency vector $\omega^\star$ satisfies, for all b in $\mathcal{K}^\star$, the Diophantine inequality*

$$|\langle k, \omega^\star(b) \rangle| \geq \frac{\gamma}{|k|_2^\tau} \quad \textit{for all } k \in \mathbb{Z}^n \setminus \{0\}. \tag{6}$$

Remark 3. We make the following remarks.

(i) From Theorem 3 we immediately obtain that for any $b \in \mathcal{K}^\star$ the n-dimensional tori

$$\mathcal{T}_b := X_0(b, \mathbb{T}^n)$$

are invariant for H and the H-dynamics is analytically conjugate to $\xi \to \xi + \omega^\star(b)t$. Furthermore, as follows from (5) and point (ii), the measure of $\bigcup_{b \in \mathcal{K}^\star} \mathcal{T}_b$ is proportional to $(\text{meas}_n\, \mathcal{K} - \epsilon^\star)(2\pi)^n$ and, hence, tends to the full measure linearly when $\epsilon^\star$ tends to zero.

(ii) A (technical) difference between Rüßmann's theorem and the formulation given above in Theorem 3 is the choice of μ as any integer greater or equal than the actual index of non-degeneracy of ω, while in [**18**] μ is chosen equal to the index of non-degeneracy of ω. In fact, it is easy to check (see, e.g., [**15**]) that Rüßmann's theorem holds in this slightly more general case, which will however be important in our applications.

Another difference between Theorem 3 and Rüßmann's original formulation concerns the way the small divisors are controlled. Rüßmann uses a very general approach based on 'approximation functions'; however, such an approach is too general for our application and cannot be applied directly. Nevertheless, it is easy to follow a more classical approach (cf., again, [**15**]) based on Diophantine inequalities of the form (6), which will be good enough for the application to properly degenerate systems; cf. also Remark 5(ii) below.

† That is, it preserves the symplectic form $dx \wedge dy$.

‡ Here and in what follows $f(\eta) = O(g(\eta))$ means that there exists a constant C such that $|f(\eta)| \leq C|g(\eta)|$ for small enough η.

2.2. *A KAM theorem for properly degenerate systems.* Let d and p be positive integers; let $\mathcal{B}$ an open set in $\mathbb{R}^d$, $\mathcal{U}$ some open neighborhood of the origin in $\mathbb{R}^{2p}$ and ϵ a 'small' real parameter. Consider a Hamiltonian function H_ϵ of the form

$$H_\epsilon(\varphi, I, u, v) = h(I) + \epsilon f(\varphi, I, u, v), \tag{7}$$

which is real-analytic for

$$(\varphi, I, (u, v)) \in \mathbb{T}^d \times \mathcal{B} \times \mathcal{U} =: \mathcal{M},$$

where $\mathcal{M}$ is endowed with the standard symplectic form

$$d\varphi \wedge dI + du \wedge dv.$$

The 'perturbation' f is assumed to have the form

$$\begin{cases} f(\varphi, I, u, v) = f_0(I, u, v) + f_1(\varphi, I, u, v), \displaystyle\int_{\mathbb{T}^d} f_1(\varphi, I, u, v)\, d\varphi = 0, \\ f_0(I, u, v) = f_{00}(I) + \displaystyle\sum_{j=1}^{p} \Omega_j(I) \frac{u_j^2 + v_j^2}{2} + O(|(u, v)|^3). \end{cases} \tag{8}$$

Observe that for every $\bar{I} \in \mathcal{B}$, the Hamiltonian $h + \epsilon f_0$ possesses the invariant isotropic (non-Lagrangian) torus

$$\mathcal{T}^d_{\bar{I}} := \mathbb{T}^d \times \{\bar{I}\} \times \{0\} \subset \mathcal{M}$$

with corresponding quasi-periodic flow

$$\varphi(t) = (\partial_I h(\bar{I}) + \epsilon \partial_I f_{00}(\bar{I}))t + \varphi_0, \quad I(t) \equiv \bar{I}, \quad (u(t), v(t)) \equiv 0.$$

The purpose is to find Lagrangian invariant tori for H_ϵ close to $(d + p)$-tori of the form

$$\mathcal{T}^{d+p}_{\bar{I},w} = \mathbb{T}^d \times \{\bar{I}\} \times \{(u, v) \in \mathbb{R}^{2p}, |(u_j, v_j)|^2 = 2w_j, \ \forall j = 1, \ldots, p\} \tag{9}$$

for $\bar{I}$ in $\mathcal{B}$ and $w \in (\mathbb{R}_+)^p$ small.

THEOREM 4. *Consider a real-analytic Hamiltonian function H_ϵ as in (7) and (8), and assume that the 'frequency map'*

$$I \in \mathcal{B} \longrightarrow (\omega(I), \Omega(I)) := (\nabla h(I), \Omega_1(I), \ldots, \Omega_p(I)) \in \mathbb{R}^d \times \mathbb{R}^p \tag{10}$$

is R-non-degenerate. Then, if ϵ is sufficiently small, there exists a positive measure set of phase space points belonging to real-analytical, Lagrangian, H_ϵ-invariant tori, which are close to $\mathcal{T}^{d+p}_{\bar{I},w}$ as in (9) with $w_j = O(\epsilon)$; furthermore, the H_ϵ-flow on such tori is quasi-periodic with Diophantine frequencies.

Remark 4. We make the following remarks.

(i) The above theorem may be viewed as the real-analytic version for Lagrangian tori of the C^∞ KAM theorem by Herman (see, in particular, [**9**, Theorem 57, p. 1559]). Under stronger non-degeneracy assumptions the above theorem corresponds to the 'fundamental theorem' in [**1**].

https://doi.org/10.1017/S0143385708000503 Published online by Cambridge University Press

(ii) The term 'properly degenerate' refers to the fact that for $\epsilon = 0$ the Hamiltonian H_0 depends on d action variables, while the number of degrees of freedom is $d + p > d$. In particular, the tori constructed in Theorem 4, as $\epsilon \to 0$, degenerate into lower-dimensional (non-Lagrangian) tori $\mathcal{T}^d_{\bar{I}}$.

(iii) The natural symplectic variables for the KAM theory of the Hamiltonian H_ϵ are (I, φ) and (rather than the Cartesian variables (q, p)) the symplectic action-angle variables (w, ζ), where $w_j = (u_j^2 + v_j^2)/2$ for $j = 1, \ldots, p$ and ζ_j is the angle of the circle† $|w_j| = \text{constant}$. Indeed, Theorem 4 has, in terms of such variables, a natural reformulation, which gives a deeper insight into the structure of the invariant tori. (In reformulating Theorem 4 in terms of the variables (ζ, w) we often use the same symbols used above. The proof of Theorem 5 is not explicitly given since it follows easily from the proof of Theorem 4.)

THEOREM 5. *Let $H_\epsilon(\varphi, I, \zeta, w) = h(I) + \epsilon f(\varphi, I, \zeta, w)$ be real-analytic for $(\varphi, I, \zeta, w) \in \mathbb{T}^d \times \mathcal{B} \times \mathbb{T}^p \times \{w \in \mathbb{R}^p : 0 < |w_j| < r\} =: \mathcal{M}$ for some open set $\mathcal{B} \subset \mathbb{R}^d$ and $0 < r$; $\mathcal{M}$ is endowed with the symplectic form $d\varphi \wedge dI + d\zeta \wedge dw$. The perturbation f is of the form $f = f_0(I, \zeta, w) + f_1$ with f_1 having vanishing φ-mean value over $\mathbb{T}^d$; furthermore f_0 has the form $f_0 = f_{00}(I) + \langle \Omega(I), w\rangle + o(|w|)$. Then, if the frequency map $I \in \mathcal{B} \to (\omega, \Omega) := (\partial_I h(I), \Omega(I)) \in \mathbb{R}^d \times \mathbb{R}^p$ is R-non-degenerate, and ϵ is small enough, there exists a positive measure set of phase space points belonging to real-analytical, Lagrangian, H_ϵ-invariant tori, which have the following parametrization:*

$$\begin{cases} \varphi = \theta + \tilde{\varphi}(\theta, \psi), \\ I = \bar{I} + \tilde{I}(\theta, \psi), \\ \zeta = \psi + \tilde{\zeta}(\theta, \psi), \\ w = \bar{w} + \tilde{w}(\theta, \psi), \end{cases}$$

where $\bar{w}$ is a constant vector of norm 2ϵ and $\tilde{\varphi}$, $\tilde{I}$, $\tilde{\zeta}$ and $\tilde{w}$ real-analytic functions for $(\theta, \psi) \in \mathbb{T}^d \times \mathbb{T}^p$ (with range, respectively, in $\mathbb{T}^d$, $\mathbb{R}^d$, $\mathbb{T}^p$ and $\mathbb{R}^p$) with

$$\begin{cases} \tilde{I} = O(\epsilon (\log \epsilon^{-1})^{-(\tau_0+1)}), \\ \tilde{w} = O(\epsilon^{(\nu+1)/2}), \\ \tilde{\varphi}, \tilde{\zeta} = O(\epsilon), \end{cases}$$

for suitable $\nu \geq 4$ and (see equations (16) and (21) below) $\tau_0 \geq d + p$. Moreover, if (ω, Ω) is the above frequency map, the H_ϵ-flow on such invariant tori is conjugated to

$$(\theta, \psi) \longrightarrow (\theta + \tilde{\omega} t, \psi + \epsilon \tilde{\Omega} t)$$

for a suitable Diophantine vector $(\tilde{\omega}, \tilde{\Omega})$ satisfying

$$|\tilde{\omega} - \omega|, |\tilde{\Omega} - \Omega| = O(\epsilon).$$

2.3. *Proof of Theorem 4.* First of all, let us introduce some notation and quantitate the assumptions of Theorem 4.

† Compare this with (27) below, where w is related to ρ by $w = \rho^0 + \rho$.

- For $\delta > 0$, $d \in \mathbb{N}$, $A \subset \mathbb{R}^d$ or $\mathbb{C}^d$, we denote

$$\begin{aligned} B^d(x_0, \delta) &:= \{x \in \mathbb{R}^d : |x - x_0| < \delta\}, \quad (x_0 \in \mathbb{R}^d), \\ D^d(x_0, \delta) &:= \{x \in \mathbb{C}^d : |x - x_0| < \delta\}, \quad (x_0 \in \mathbb{C}^d), \\ \mathbb{T}^d_\delta &:= \{x \in \mathbb{C}^d : |\mathrm{Im}\, x_j| < \delta, \mathrm{Re}\, x_j \in \mathbb{T}, \ \forall j = 1 \ldots d\}, \\ A + \delta &:= \bigcup_{x \in A} D^d(x, \delta). \end{aligned} \tag{11}$$

- We may assume that H_ϵ in (7) and (8) can be holomorphically extended for

$$(\varphi, I, (u, v)) \in \mathbb{T}^d_\sigma \times (\mathcal{B} + r_0) \times (\mathcal{U} + r_1) =: \mathcal{M}_\star. \tag{12}$$

In particular, H_ϵ is real-analytic on $\mathbb{T}^d \times B^d(I_0, s) \times B^{2p}(0, r_1)$ for any I_0 in $\mathcal{B}$ and $s < r_0$. Moreover, we denote

$$M_0 := \sum_{k \in \mathbb{Z}^d} \left(\sup_{(\mathcal{B}+r_0)\times(\mathcal{U}+r_1)} |f_k(I, u, v)| \right) e^{|k|_1 \sigma} \tag{13}$$

as the 'sup-Fourier' norm of f and let

$$M_1 := \sup_{I \in \mathcal{B}+r_0} |(\omega(I), \Omega(I))|. \tag{14}$$

The proof of Theorem 4 is based on *two preliminary steps*:

(1) computation of a suitable normal form for H_ϵ;

(2) quantitative estimates on the amount of the non-degeneracy of the normal form.

2.3.1. *Step 1: Normal forms for properly degenerate systems.*

PROPOSITION 1. *Fix an integer $\nu \geq 4$. Then, there exists $m > d$ (depending on ω), and, for ϵ small enough, a point $I_0 \in \mathcal{B}$ and a real-analytic canonical transformation† Φ_ϵ such that the following holds. Let*

$$s := O((\log \epsilon^{-1})^{-m}) \tag{15}$$

then $B^d(I_0, s) \subset \mathcal{B}$ and $\Phi_\epsilon : (\vartheta, r, \zeta, \rho) \longrightarrow (\varphi, I, u, v)$ satisfies

$$\Phi_\epsilon : \mathbb{T}^d \times B^d(0, s/5) \times \mathbb{T}^p \times B^p(0, \epsilon) \longrightarrow \mathbb{T}^d \times B^d(I_0, s) \times \mathcal{U}$$

and $\hat{H}_\epsilon := H_\epsilon \circ \Phi_\epsilon$ takes the form

$$\hat{H}_\epsilon(\vartheta, r, \zeta, \rho) = N_\epsilon(r, \rho; \rho^0) + \epsilon^\nu P_\epsilon(\vartheta, r, \zeta, \rho; \rho^0) \tag{16}$$

with

$$N_\epsilon := \frac{1}{\epsilon} h(I_0 + \epsilon r) + \hat{g}(I_0 + \epsilon r) + \frac{1}{2} \hat{\Omega}(I_0 + \epsilon r) \cdot (\rho^0 + \epsilon\rho) + Q_{\epsilon, I_0 + \epsilon r}(\rho^0 + \epsilon\rho) \tag{17}$$

and ρ^0 in $(\mathbb{R}_+)^p$ is some point having Euclidean norm 2ϵ; $Q_{\epsilon, I_0+\epsilon r}$ is a polynomial of degree $\nu - 1$ starting with cubic terms; $\hat{g}$, $\hat{\Omega}$ and P_ϵ are real-analytic functions. Furthermore, one has

$$\sup_{r \in B^d(0, s/5)} |\hat{\Omega}(I_0 + \epsilon r) - \Omega(I_0 + \epsilon r)| = O(\epsilon (\log \epsilon^{-1})^{2m-1}).$$

† Symplectic up to rescalings.

https://doi.org/10.1017/S0143385708000503 Published online by Cambridge University Press

Proof of Proposition 1. We start by recalling a measure-theoretical result due to Pyartli (see [**16**] or [**18**, Theorem 17.1]).

LEMMA 2. (Pyartli) *Let $\mathcal{K} \subset \mathbb{R}^d$ be a compact set; let $\theta \in (0, 1)$ define $\mathcal{K}_\theta := \bigcup_{y\in\mathcal{K}} B^d(y, \theta)$. Let $g : \mathcal{K}_\theta \to \mathbb{R}$ be a real-analytic function satisfying*

$$\min_{y\in\mathcal{K}} \max_{0\le\nu\le\mu_0} |\partial^\nu g(y)| \ge \beta$$

for some $\beta > 0$. Then there exists $C = C(\mu_0, \beta, d, \mathcal{K}, \theta)$ such that

$$\mathrm{meas}_d \{y \in \mathcal{K} : |g(y)| \le t\} \le C|g|^{\mu_0+1}_{\mathcal{K}_\theta} t^{1/\mu_0}$$

for any $0 \le t \le \beta/(2\mu_0 + 2)$.

Pyartli's lemma implies the following.

LEMMA 3. *Let $\mathcal{K}$ be a compact set with positive d-dimensional Lebesgue measure and let $0 < \varepsilon^\star < \mathrm{meas}_d\, \mathcal{K}$; let $\omega : \mathcal{K}_\theta \to \mathbb{R}^d$ be R-non-degenerate and let μ_0 and β be its index and amount of non-degeneracy with respect to $\mathcal{K}$. Let us denote by $\mathcal{D}^d_{\gamma_0,\tau_0}$ the set of Diophantine vectors in $\mathbb{R}^d$ with Diophantine constants γ_0, τ_0, that is, the set*

$$\mathcal{D}^d_{\gamma_0,\tau_0} := \left\{\omega \in \mathbb{R}^d : |\langle\omega, k\rangle| \ge \frac{\gamma_0}{|k|^{\tau_0}}, \forall k \in \mathbb{Z} \setminus \{0\}\right\}.$$

Then, if γ_0 is sufficiently small and $\tau_0 \ge d\mu_0$ one has

$$\mathrm{meas}_d(\mathcal{K} \cap \mathcal{D}^d_{\gamma_0,\tau_0}) \ge \mathrm{meas}_d\, \mathcal{K} - \varepsilon^\star. \tag{18}$$

Proof of Lemma 3. First of all observe that for any $m \in \mathbb{Z}_+$, $a \in \mathbb{R}^d$, $k \in \mathbb{Z}^d \setminus \{0\}$ and $b \in \mathcal{K}_\theta$

$$|\partial^m \langle\omega(b), k|k|^{-1}\rangle(a^m)| = |\langle\partial^m \omega(b)(a^m), k|k|^{-1}\rangle| \le |\partial^m \omega(b)(a^m)|;$$

taking the sup over $|a| = 1$, $b \in \mathcal{K}_\theta$ and $0 \le m \le \mu$ we obtain

$$|\langle\omega, k|k|^{-1}\rangle|^\mu_{\mathcal{K}_\theta} \le |\omega|^\mu_{\mathcal{K}_\theta} < \infty$$

for any $\mu \in \mathbb{Z}_+$. Now we use this last inequality and Theorem 2, assuming $\gamma_0 \le \beta/(2\mu_0 + 2)$, to estimate

$$\begin{aligned}\mathrm{meas}_d\, (\mathcal{K} \setminus \mathcal{D}^d_{\gamma_0,\tau_0}) &= \mathrm{meas}_d \bigcup_{k\in\mathbb{Z}^d\setminus\{0\}} \left\{b \in \mathcal{K} : |\langle\omega(b), k\rangle| < \frac{\gamma_0}{|k|^{\tau_0}}\right\}\\ &\le \sum_{k\in\mathbb{Z}^d\setminus\{0\}} \mathrm{meas}_d \left\{b \in \mathcal{K} : \left|\left\langle\omega(b), \frac{k}{|k|}\right\rangle\right| < \frac{\gamma_0}{|k|^{\tau_0+1}}\right\}\\ &\le C(\mu_0, \beta, d, \mathcal{K}, \theta)|\omega|^{\mu_0+1}_{\mathcal{K}_\theta} \gamma_0^{1/\mu_0} \sum_{k\in\mathbb{Z}^d\setminus\{0\}} \frac{1}{|k|^{(\tau_0+1)/\mu_0}}.\end{aligned}$$

Since $\tau_0 \ge d\mu_0$ this last sum converges and one has

$$\mathrm{meas}_d(\mathcal{K} \setminus \mathcal{D}^d_{\gamma_0,\tau_0}) \le \bar{C}\gamma_0^{1/\mu_0}$$

for a suitable $\bar{C} = \bar{C}(\mu_0, \beta, d, \mathcal{K}, \theta, \omega, \tau_0)$. Choosing $\gamma_0 \le (\bar{C}^{-1}\epsilon^\star)^{\mu_0}$ we obtain the estimate (18). □

https://doi.org/10.1017/S0143385708000503 Published online by Cambridge University Press

Now consider the real-analytic Hamiltonian H_ϵ in (7) and (8). Let $\nu_1, \nu_2 \geq 4$ be two integers to be later determined and set

$$K_1 := \frac{6}{\sigma}(\nu_1 - 1)\log\frac{1}{\epsilon M_0} \tag{19}$$

where M_0 is defined by (13). Lemma 3 and the R-non-degeneracy of ω assure the existence of $I_0 \in \mathcal{B}$ such that $\omega(I_0)$ belongs to $\mathcal{D}^d_{\gamma_0,\tau_0}$ (for suitable γ_0 and τ_0). Then, from Taylor's formula it follows that†

$$|\omega(I)\cdot k| \geq \alpha_1 > 0 \quad \text{for all } k \in \mathbb{Z}^d, \quad 0 < |k|_1 \leq K_1 \quad \text{for all } I \in D^d(I_0, s) \tag{20}$$

with‡

$$s := O((\log\epsilon^{-1})^{-(\tau_0+1)}) \quad \text{and} \quad \alpha_1 := O((\log\epsilon^{-1})^{-\tau_0}). \tag{21}$$

Furthermore, we can assume that there exists α_2 (independent of ϵ) such that

$$|\Omega(I)\cdot k| \geq \alpha_2 > 0 \quad \text{for all } k \in \mathbb{Z}^p, \quad 0 < |k|_1 \leq \nu_2 \quad \text{for all } I \in D^d(I_0, s). \tag{22}$$

Next, we want to average H_ϵ over the 'fast angles' φ up to order ν_1. To do this we apply the following classical 'averaging lemma', whose proof can be found in [**3**, Appendix A, p. 110].

LEMMA 4. (Averaging lemma) *Let H_ϵ, M_0, σ, α_1 and s be as above. Assume that (20) holds with K_1 as in (19). Then, if ϵ is small enough, there exists a real-analytic symplectic transformation $\Phi^1_\epsilon : (\tilde\varphi, \tilde I, \tilde u, \tilde v) \to (\varphi, I, u, v)$ mapping*

$$\mathcal{M}_1 := \mathbb{T}^d_{\sigma/6} \times D^d\left(I_0, \frac{s}{2}\right) \times D^{2p}\left(0, \frac{r_1}{2}\right) \xrightarrow{\Phi^1_\epsilon} \mathcal{M}_0 := \mathbb{T}^d_\sigma \times D^d(I_0, s) \times D^{2p}(0, r_1)$$

that casts H_ϵ into the Hamiltonian

$$H^1_\epsilon := H_\epsilon \circ \Phi^1_\epsilon = h + \epsilon f_0 + \tilde g + \tilde f,$$

where $\tilde g = \tilde g(\tilde I, \tilde u, \tilde v)$ and $\tilde f$ satisfy

$$\sup_{\mathcal{M}_1} |\tilde g| \leq C\frac{(\epsilon M_0)^2}{s\alpha_1}, \quad \sup_{\mathcal{M}_1} |\tilde f| \leq (\epsilon M_0)^{\nu_1} \tag{23}$$

for a suitable $C = C(\sigma, \nu_1)$.

Lemma 4 can be immediately derived from [**3**, Proposition A.1] with the following correspondences: $\alpha_1 = \alpha$ for α_1 as in (20), $K_1 = K$ for K_1 as in (19), $\epsilon M_0 = \varepsilon$ for M_0 as in (13), $s = r, d$ for s as in (15) and (20), $\{0\} = \Lambda$ and $\epsilon f(I, \varphi, u, v)$ in (7) is just $f(u, \varphi)$ in [**3**]; as a result, one has that $\epsilon f_0 + \tilde g$ and $\tilde f$ are respectively given by g and $f_\star$ in [**3**] with estimates (23) holding in view of the previous correspondences.

Thus, if we set $\tilde g =: \epsilon^2 \bar g$ and $\tilde f =: \epsilon^{\nu_1}\bar f$, using (8) and (23) we have

$$\begin{cases} H^1_\epsilon(\tilde\varphi, \tilde I, \tilde u, \tilde v) = h(\tilde I) + \epsilon[f_0(\tilde I, \tilde u, \tilde v) + \epsilon\bar g(\tilde I, \tilde u, \tilde v)] + \epsilon^{\nu_1}\bar f(\tilde\varphi, \tilde I, \tilde u, \tilde v), \\ f_0(\tilde I, \tilde u, \tilde v) = f_{00}(\tilde I) + \displaystyle\sum_{j=1}^p \Omega_j(\tilde I)\frac{\tilde u_j^2 + \tilde v_j^2}{2} + O(|\tilde u, \tilde v|^3; \tilde I). \end{cases} \tag{24}$$

† For $k = (k_1, k_2, \ldots, k_d) \in \mathbb{Z}^d$ we denote $|k|_1 := \sum_{i=1}^d |k_i|$; recall also the definition of complex balls D^d in (11).
‡ This means that we can take $m = \tau_0 + 1$ in (15).

From equation (24) we see that the application of averaging theory may cause, in general, a shift of order ϵ of the elliptic equilibrium, which, before, was in the origin of $\mathbb{R}^{2p}$. Therefore, we focus our attention on the Hamiltonian function $f_0 + \epsilon\bar{g}$ with the aim to find a real-analytic symplectic transformation restoring the equilibrium in the origin. An application of the standard implicit function theorem yields the following result.

LEMMA 5. *Let* $\mathcal{M}_2 := \mathbb{T}^d_{(\sigma/7)} \times D^d(I_0, s/4) \times D^{2p}(0, r_1/4)$; *then, provided* ϵ *is sufficiently small, there exists a (close to the identity) real-analytic symplectic transformation*

$$\Phi^2_\epsilon : (x, y, p, q) \in \mathcal{M}_2 \longrightarrow (\tilde{\varphi}, \tilde{I}, \tilde{u}, \tilde{v}) \in \mathcal{M}_1$$

such that $H^2_\epsilon := H^1_\epsilon \circ \Phi^2_\epsilon$ *is of the form*

$$H^2_\epsilon(x, y, p, q) = h(y) + \epsilon\hat{g}(y, p, q) + \epsilon^{\nu_1}\hat{f}(x, y, p, q)$$

with $\partial_p\hat{g}(y, 0, 0) = 0 = \partial_q\hat{g}(y, 0, 0)$, $\hat{g}$ *and* $\hat{f}$ *real-analytic on* $\mathcal{M}_2$.

Since $\Omega_j \neq 0$ for every j in view of (22), we can apply the implicit function theorem to obtain, for small enough ϵ, the existence of two functions $u_0 = u_0(\tilde{I}, \epsilon)$ and $v_0 = v_0(\tilde{I}, \epsilon)$ which are real-analytic for $\tilde{I} \in D^d(I_0, s/4)$ and such that $\nabla_{\tilde{u},\tilde{v}}(f_0 + \epsilon\bar{g})(\tilde{I}, u_0, v_0) = 0$. Furthermore, using (23) together with $\tilde{g} = \epsilon^2\bar{g}$ and (21), one has $u_0, v_0 = O(\epsilon(\log\epsilon^{-1})^{2\tau_0+1})$. The symplectic transformation in Lemma 5 is then generated by $x \cdot \tilde{\varphi} + (p + u_0(x, \epsilon)) \cdot (\tilde{v} - v_0(x, \epsilon))$.

Now, we need to control the frequencies associated to the modified Hamiltonian $\hat{g}(y, 0, 0)$.

LEMMA 6. *If* ϵ *is small enough then the eigenvalues of the Hamiltonian* $\hat{g}$, *that is, the eigenvalues of*† $J_{2p}\partial^2_{(p,q)}\hat{g}(y, 0, 0)$, *are given by* $2p$ *purely imaginary functions* $\pm i\hat{\Omega}_1, \ldots, \pm i\hat{\Omega}_p$ *verifying*

$$\sup_{y\in D^d(I_0,s/4)} |\hat{\Omega}(y) - \Omega(y)| = O(\epsilon(\log\epsilon^{-1})^{2\tau_0+1}) \tag{25}$$

for τ_0 *as in (21).*

Proof of Lemma 6. Consider the quadratic part of $\hat{g}$, that is the real-analytic $2p \times 2p$ symmetric matrix $\hat{A}(y) := \partial^2_{(p,q)}\hat{g}(y, 0, 0)$. Using the construction of Φ^2_ϵ in Lemma 5, $\hat{g} = (f_0 + \epsilon\bar{g}) \circ \Phi^2_\epsilon$, estimate (23) together with $\tilde{g} = \epsilon^2\bar{g}$ and the definition of s and α_1 in (21), equation (24) for f_0 and Cauchy's estimate for derivatives of analytic functions, one has

$$\hat{A}(y) = \mathrm{diag}(\Omega_1(y), \ldots, \Omega_p(y), \Omega_1(y), \ldots, \Omega_p(y)) + O(\epsilon(\log\epsilon^{-1})^{2\tau_0+1}).$$

Since $\Omega_j \neq \Omega_k$, for $j \neq k$, an application of the implicit function theorem (cf. (22)) tells us that the eigenvalues of $\hat{g}$ (that *a priori* might have non-zero real part) are $O(\epsilon(\log\epsilon^{-1})^{2\tau_0+1})$ close to $\pm i\Omega_j$. Now, as it is well known, eigenvalues of Hamiltonians always appear in quadruplets $\pm\lambda, \pm\bar{\lambda}$; thus, from the simplicity of the eigenvalues of $\hat{g}$ (holding for ϵ small enough), one has that its eigenvalues are purely imaginary as claimed. □

† Here J_{2p} denotes the standard $2p \times 2p$ symplectic matrix.

By normal form theory (see [**2**, Corollary 8.7]) we can find a real-analytic symplectic transformation $O(\epsilon)$-close to the identity

$$\Phi^3_\epsilon : (\tilde x, \tilde y, \tilde p, \tilde q) \in \mathcal{M}_3 := \mathbb{T}^d_{\sigma/8} \times D^d\left(I_0, \frac{s}{5}\right) \times D^{2p}\left(0, \frac{r_1}{5}\right) \longrightarrow (x, y, p, q) \in \mathcal{M}_2$$

with $\tilde y = y$ and such that the transformed Hamiltonian function $H^3_\epsilon := H^2_\epsilon \circ \Phi^3_\epsilon$, which is real-analytic on $\mathcal{M}_3$, has the form

$$H^3_\epsilon(\tilde x, \tilde y, \tilde p, \tilde q) = h(\tilde y) + \epsilon \hat g_0(\tilde y) + \frac{\epsilon}{2}\sum_{j=1}^p \hat\Omega_j(\tilde y)(\tilde p_j^2 + \tilde q_j^2) + \epsilon \tilde g_3(\tilde y, \tilde p, \tilde q) + \epsilon^{\nu_1} \tilde f_3(\tilde x, \tilde y, \tilde p, \tilde q)$$

where $\hat g_0 := \hat g(\tilde y, 0, 0)$, $\tilde f_3 := \bar f \circ \Phi^3_\epsilon$ and $\tilde g_3 := \hat g_3 \circ \Phi^3_\epsilon$ verifies

$$\sup_{\tilde y \in D^d(I_0, s/5)} |\tilde g_3(\tilde y, \tilde p, \tilde q)| \le C|(\tilde p, \tilde q)|^3 \quad \text{for all } (\tilde p, \tilde q) \in D^{2p}(0, r_1/5).$$

Now let

$$\tilde g_2(\tilde y, \tilde p, \tilde q) := \frac{1}{2}\sum_{i=1}^p \hat\Omega_i(\tilde p_i^2 + \tilde q_i^2),$$

we want to put $\tilde g_2 + \epsilon \tilde g_3$ into Birkhoff's normal form up to order ν_2. In view of inequalities (22) and (25), provided that ϵ is small enough, we have

$$|\hat\Omega(\tilde y)\cdot k| \ge \frac{\alpha_2}{2} \quad \text{for all } k \in \mathbb{Z}^p, \quad 0 < |k|_1 \le \nu_2 \quad \text{for all } \tilde y \in D^d(I_0, s/5). \tag{26}$$

By Birkhoff's normal form theory†, one easily obtains the following.

LEMMA 7. *If inequality (26) is satisfied, then there exist $0 < r_\star < r_1' \le r_1/5$ and a real-analytic symplectic diffeomorphism $\Phi^4_\epsilon : (\theta, r, u, v) \to (\tilde x, \tilde y, \tilde p, \tilde q)$ mapping*

$$\mathcal{M}_4 := \mathbb{T}^d_{\sigma/8} \times D^d\left(I_0, \frac{s}{5}\right) \times D^{2p}(0, r_\star) \xrightarrow{\Phi^4_\epsilon} \mathcal{M}_3' := \mathbb{T}^d_{\sigma/8} \times D^d\left(I_0, \frac{s}{5}\right) \times D^{2p}(0, r_1')$$

leaving the origin and the quadratic part of H^3_ϵ invariant, such that $(\theta, r) = (\tilde x, \tilde y)$ and $H^4_\epsilon := H^3_\epsilon \circ \Phi^4_\epsilon$ is of the form

$$H^4_\epsilon(\theta, r, u, v) = h(r) + \epsilon \hat g_0(r) + \frac{\epsilon}{2}\sum_{j=1}^p \hat\Omega_j(r)(u_j^2 + v_j^2) + \epsilon\, Q_\star(r, u, v) + \epsilon R_\star(r, u, v) + \epsilon^{\nu_1} \tilde f_4(\theta, r, u, v)$$

where:

- *$Q_\star$ is a polynomial of degree $[\nu_2/2]$ in the variables $I = (I_1, \ldots, I_p)$ having the form*

  $$\langle \hat\Omega(r), I\rangle + \tfrac{1}{2}\langle T(r) I, I\rangle + \cdots \quad \textit{with } I_j := \tfrac{1}{2}(u_j^2 + v_j^2)$$

  with $T(r)$ a $2p \times 2p$ real-analytic matrix;

† See, e g., [**12**, Theorem 11, p. 43] or [**15**, §3.4] for a quantitative version.

- *$R_\star$ is a real-analytic function verifying $|R_\star(r, u, v)| \le C|(u, v)|^{\nu_2+1}$ for every $(u, v) \in D^{2p}(0, r_\star)$ and $r \in D^d(I_0, s/5)$;*
- *$\tilde{f}_4 := \tilde{f}_3 \circ \Phi^4_\epsilon$ is real-analytic on $\mathcal{M}_4$.*

We may conclude the proof of Proposition 1. Following [**9**, pp. 1561–1562], we pass to symplectic polar coordinates in order to move $R_\star$ to the perturbation of H^4_ϵ with the help of a rescaling by a factor ϵ. Let $\rho^0 = (\rho^0_1, \ldots, \rho^0_p)$ in $(\mathbb{R}_+)^p$ be sufficiently close to the origin; consider, for a suitable $\sigma_\star > 0$, the real-analytic symplectic transformation $\Phi^5_\epsilon : (\theta, r, \zeta, \rho) \to (\theta, I_0 + r, z)$ mapping

$$\mathcal{M}_5 := \mathbb{T}^d_{\sigma/8} \times D^d\left(0, \frac{s}{5}\right) \times \mathbb{T}^p_{\sigma_\star} \times D^p\left(0, \frac{|\rho^0|}{2}\right) \xrightarrow{\Phi^5_\epsilon} \mathcal{M}_4,$$

where

$$z_j = u_j + i v_j := \sqrt{2(\rho^0_j + \rho_j)}\, e^{-i\zeta_j}. \tag{27}$$

The transformed Hamiltonian function $H^5_\epsilon := H^4_\epsilon \circ \Phi^5_\epsilon$, real-analytic on $\mathcal{M}_5$, assumes the form

$$\begin{aligned} H^5_\epsilon(\theta, r, \zeta, \rho) = h(I_0 + r) + \epsilon \hat{g}_0(I_0 + r) + \frac{\epsilon}{2} \sum_{j=1}^{p} \hat{\Omega}_j(I_0 + r)(\rho^0_j + \rho^0_j) \\ + \epsilon Q_{I_0+r}(\rho^0 + \rho) + \epsilon R(I_0 + r, \zeta, \rho^0 + \rho) + \epsilon^{\nu_1} \tilde{f}_5(\theta, r, \zeta, \rho; \rho^0) \end{aligned}$$

where:

- $Q_{I_0+r} := Q_\star \circ \Phi^5_\epsilon$ is a polynomial of degree $[\nu_2/2]$ with respect to $\rho^0 + \rho$, depending also on $I_0 + r$;
- $R := R_\star \circ \Phi^5_\epsilon$ verifies

$$|R(I_0 + r, \zeta, \rho^0 + \rho)| \le C|\rho^0|^{(\nu_2+1)/2}$$

  for every $\rho \in D^{2p}(0, |\rho^0|/2)$, $r \in D^d(0, s/5)$ and $\zeta \in \mathbb{T}^p_{\sigma_\star}$;
- $\tilde{f}_5 := \tilde{f}_4 \circ \Phi^5_\epsilon$ is real-analytic on $\mathcal{M}_5$.

Now, let A_ϵ be the homothety given by

$$A_\epsilon : (\theta, r, \zeta, \rho) \longrightarrow (\theta, \epsilon r, \zeta, \epsilon\rho).$$

Even though A_ϵ is not a symplectic map it preserves the structure of Hamilton's equations if we consider the Hamiltonian function $H^6_\epsilon := (1/\epsilon) H^5_\epsilon \circ A_\epsilon$. Explicitly, we have

$$\begin{aligned} H^6_\epsilon(\theta, r, \zeta, \rho) = \frac{1}{\epsilon} h(I_0 + \epsilon r) + \hat{g}_0(I_0 + \epsilon r) + \frac{1}{2}\hat{\Omega}(I_0 + \epsilon r) \cdot (\rho^0 + \epsilon\rho) \\ + Q_{\epsilon, I_0+\epsilon r}(\rho^0 + \epsilon\rho) + R(I_0 + \epsilon r, \epsilon\rho, \zeta; \rho^0) \\ + \epsilon^{\nu_1 - 1} \tilde{f}_6(\theta, r, \zeta, \rho; \rho^0) \end{aligned} \tag{28}$$

where $\tilde{f}_6 := \tilde{f}_5 \circ A_\epsilon$. Now we fix $\rho^0 \in (\mathbb{R}_+)^p$ with $|\rho^0| = 2\epsilon$ so that $|R| \le C\epsilon^{(\nu_2+1)/2}$. Thus, if we choose ν_1 and ν_2 so that

$$\nu_1 - 1 = \left[\frac{\nu_2 + 1}{2}\right] := \nu \tag{29}$$

we may write

$$R(I_0+\epsilon r, \epsilon\rho, \zeta; \rho^0)+\epsilon^{\nu_1-1}\tilde{f}_6(\theta, r, \zeta, \rho) =: \epsilon^{\nu}P_\epsilon(\theta, r, \zeta, \rho) \tag{30}$$

for a suitable function P_ϵ real-analytic on $\mathbb{T}^d_{\sigma/8}\times D^d(0, s/5)\times\mathbb{T}^p_{\sigma_\star}\times D^p(0, \epsilon)$.

We have proved Proposition 1 with $\hat{H}_\epsilon = H^6_\epsilon$ as in (28), (30). □

2.3.2. *Step 2: Amounts of non-degeneracy of the normal form.*

PROPOSITION 2. *Let N_ϵ be as in (17). If ϵ is small enough, the frequency map*

$$\hat{\Psi}_\epsilon : (r, \rho)\in B^d(0, s/5)\times B^p(0, \epsilon)\longrightarrow\left(\frac{\partial}{\partial r}N_\epsilon, \frac{\partial}{\partial\rho}N_\epsilon\right)$$

is R-non-degenerate.

Moreover, let $\bar{\mu}$ and $\bar{\beta}$ denote respectively the index and the amount of non-degeneracy of the unperturbed frequency map (10) with respect to a closed ball $\bar{B}^d(I_0, t)\subset\mathcal{B}$, for some $t>0$ independent of ϵ. Then, if we define $\mathcal{K}_\epsilon := \bar{B}^d(0, s/10)\times\bar{B}^p(0, \epsilon/2)$ and let $\hat{\mu}_\epsilon$ denote the index of non-degeneracy of $\hat{\Psi}_\epsilon$ with respect to $\mathcal{K}_\epsilon$ and

$$\hat{\beta}_\epsilon := \min_{c\in\mathcal{S}^{d+p-1}}\min_{(r,\rho)\in\mathcal{K}_\epsilon}\max_{0\le\mu\le\bar{\mu}}|\partial^\mu_{(r,\rho)}|\langle c, \hat{\Psi}_\epsilon\rangle|^2|,$$

one has

$$\hat{\mu}_\epsilon\le\bar{\mu}\quad\textit{and}\quad\hat{\beta}_\epsilon\ge\frac{\epsilon^{\bar{\mu}+2}\bar{\beta}}{8}. \tag{31}$$

Proof of Proposition 2. From (17), it follows that the frequency map of N_ϵ is given by

$$\hat{\Psi}_\epsilon(r, \rho)=\left(\omega(I_0+\epsilon r)+O(\epsilon), \frac{\epsilon}{2}\hat{\Omega}(I_0+\epsilon r)+O(\epsilon^2)\right)$$

and it is real-analytic on $D^d(0, s/5)\times D^p(0, \epsilon)$. Using (25) one has

$$\hat{\Psi}_\epsilon(r, \rho)=\left(\omega(I_0+\epsilon r)+O(\epsilon), \frac{\epsilon}{2}(\Omega(I_0+\epsilon r)+O(\epsilon))\right). \tag{32}$$

Now, let $\bar{\mu}\in\mathbb{N}_+$ and $\bar{\beta}>0$ denote respectively the index and the amount of non-degeneracy of $\Psi := (\omega, \Omega)$ with respect to $\bar{B}^d(I_0, t)$, for some positive t independent of ϵ. Set

$$\Psi_0(r) := (\omega(I_0+\epsilon r), \Omega(I_0+\epsilon r)), \tag{33}$$

$\mathcal{K}_0 := \bar{B}^d(0, s/10)$ and use definition 2 to obtain

$$\min_{r\in\mathcal{K}_0}\max_{0\le\mu\le\bar{\mu}}|\partial^\mu_r|\langle c, \Psi_0(r)\rangle|^2|\ge\epsilon^{\bar{\mu}}\bar{\beta}>0$$

for every $c\in\mathcal{S}^{d+p-1}$.

Next, denote by Ψ_ϵ the real-analytic function over $D^d(0, s/5)\times D^p(0, \epsilon)$ obtained by multiplying the last p component of $\hat{\Psi}_\epsilon$ by a factor $2/\epsilon$. Then, observe that (32) and (33)

https://doi.org/10.1017/S0143385708000503 Published online by Cambridge University Press

imply $\Psi_\epsilon(r,\rho) = \Psi_0(r) + O(\epsilon)$. Therefore, denoting $\mathcal{K}_1 := \bar{B}^p(0,\epsilon/2)$ and assuming ϵ small enough, one has

$$\beta_\epsilon := \min_{(r,\rho)\in\mathcal{K}_0\times\mathcal{K}_1} \max_{0\le\mu\le\bar{\mu}} |\partial^\mu_{(r,\rho)}|\langle c, \Psi_\epsilon(r,\rho)\rangle|^2| \ge \frac{\epsilon^{\bar{\mu}}\bar{\beta}}{2} > 0$$

for every $c \in \mathcal{S}^{d+p-1}$.

Now write $\Psi_\epsilon = (\Psi^{(1)}_\epsilon, \Psi^{(2)}) \in \mathbb{R}^d \times \mathbb{R}^p$, so that

$$\hat{\Psi}_\epsilon = \left(\Psi^{(1)}_\epsilon, \frac{\epsilon}{2}\Psi^{(2)}_\epsilon\right).$$

Define for $c = (c_1, c_2) \in \mathbb{R}^d \times \mathbb{R}^p$ with $|c| = 1$ the function

$$\begin{aligned} f(r,\rho,c_1,c_2) &:= \max_{0\le\mu\le\bar{\mu}} |\partial^\mu_{(r,\rho)}|\langle c, \Psi_\epsilon\rangle|^2| \\ &= \max_{0\le\mu\le\bar{\mu}} |\partial^\mu_{(r,\rho)}|\langle c_1, \Psi^{(1)}_\epsilon\rangle + \langle c_2, \Psi^{(2)}_\epsilon\rangle|^2|; \end{aligned}$$

furthermore set

$$t_\epsilon := \sqrt{|c_1|^2 + \frac{\epsilon^2}{4}|c_2|^2}$$

and $\bar{c}_1 = c_1 t_\epsilon^{-1}$, $\bar{c}_2 = \epsilon c_2 (2t_\epsilon)^{-1}$ so that $|(\bar{c}_1, \bar{c}_2)| = 1$. Then one has

$$\begin{aligned} \max_{0\le\mu\le\bar{\mu}} |\partial^\mu_{(r,\rho)}|\langle c, \hat{\Psi}_\epsilon\rangle|^2| &= f\left(r,\rho,c_1,\frac{\epsilon}{2}c_2\right) = t_\epsilon^2 f\left(r,\rho,\frac{c_1}{t_\epsilon},\frac{\epsilon}{2}\frac{c_2}{t_\epsilon}\right) \\ &\ge \frac{\epsilon^2}{4} f(r,\rho,\bar{c}_1,\bar{c}_1) \ge \frac{\epsilon^{\bar{\mu}+2}\bar{\beta}}{8} > 0 \end{aligned}$$

and it follows immediately that

$$\min_{(r,\rho)\in\mathcal{K}_0\times\mathcal{K}_1} \max_{0\le\mu\le\bar{\mu}} |\partial^\mu_{(r,\rho)}|\langle c, \hat{\Psi}_\epsilon\rangle|^2| \ge \frac{\epsilon^{\bar{\mu}+2}\bar{\beta}}{8} > 0$$

for every $c \in \mathcal{S}^{d+p-1}$. Since $\mathcal{K}_0 \times \mathcal{K}_1 = \bar{B}^d(0, s/10) \times \bar{B}^p(0,\epsilon/2) = \mathcal{K}_\epsilon$ we have verified (31). In view of Remark 2 we also conclude that $\hat{\Psi}_\epsilon$ is R-non-degenerate on $B^d(0,s/5) \times B^p(0,\epsilon)$, provided that ϵ is small enough.

Proposition 2 is thus proved. □

2.3.3. *Conclusion of the proof of Theorem 4.* We want to apply Rüßmann's Theorem 3 to the properly degenerate case of $\hat{H}_\epsilon$ in (16). With Propositions 1 and 2 we are in a position to meet the hypothesis of R-non-degeneracy of the frequency application required in Theorem 3. However, the 'degenerate' case of $\hat{H}_\epsilon$ requires that the size of its perturbation is of a sufficiently small order in ϵ. From (16) we see that the size of the perturbation of $\hat{H}_\epsilon$ is of the order of ϵ^ν where ν can be chosen to be arbitrarily large† but independent of ϵ.

Next we provide an explicit expression for the admissible size of the perturbation in Rüßmann's theorem, that is, ϵ_0 in (4).

† Recall (29) and the fact that both ν_1 and ν_2 can be arbitrarily fixed at the beginning of the process described in §2.3.1.

LEMMA 8. (Rüßmann [18]) *Let H, $\mathcal{Y}$, $\epsilon^\star$, $\mathcal{A}$ and τ be as in Theorem 3 and let $\omega := \nabla h$ be R-non-degenerate (as in the hypotheses of Theorem 3). Consider the following quantities.*

(1) *Let $\mathcal{K} \subset \mathcal{Y}$ be any chosen compact set; let μ be any integer greater than the index of non-degeneracy of ω with respect to $\mathcal{K}$ and let β be the 'amount of non-degeneracy' corresponding to μ.*

(2) *Let $\vartheta \in (0, 1)$ be chosen such that† $\mathbb{T}^d_\vartheta \times (\mathcal{K} + 4\vartheta) \subset \mathcal{A}$ and define $C_1 := |\omega|_{\mathcal{K}+3\vartheta}$. Let d_0 be the diameter of $\mathcal{K}$, i.e. $d_0 := \sup_{x,y\in\mathcal{K}} |x - y|$.*

(3) *Let $T_0 \geq e^{(n+1)/\tau}$ such that the following inequality holds*

$$\int_{T_0}^{\infty} \frac{\log T}{T^2}\, dT \leq \vartheta. \tag{34}$$

(4) *Define*

$$C^\star := 2^{\mu+1} \frac{(\mu + 1)^{\mu+2}}{\vartheta^{\mu+1}} (C_1 + 1) \tag{35}$$

and set

$$\gamma := (d_0^n C^\star)^{-\mu/2} \beta^{(\mu+1)/2} \epsilon^{\star\mu/2}, \tag{36}$$

$$t_0 := \frac{\gamma T_0^{-(\tau+n+1)} \vartheta}{C_1 + 1}. \tag{37}$$

(5) *Finally set*

$$\begin{aligned} E_1 &:= \gamma T_0^{-(\tau+n+1)} \vartheta, \\ E_2 &:= \frac{\beta t_0^\mu}{T_0 C^\star}. \end{aligned} \tag{38}$$

Then ϵ_0 in (4) can be taken to be

$$\epsilon_0 := c_0 \frac{\vartheta}{C_1} (\min\{E_1, E_2\})^2 \tag{39}$$

for a suitable $c_0 = c_0(n, \mu)$.

Remark 5. The above result follows from [18] by considering the case of maximal tori only‡. More precisely, we have the following.

(i) The maximal case corresponds to the easier case $p = q = 0$ in [18]. Note, however, that it is not sufficient to substitute the values $p = q = 0$ into Rüßmann's estimates (as, when $p = 0$ for instance, many terms in [18, p. 171] become meaningless) but, rather, one has to go through most of [18, Theorem 18.5] to obtain the value of $C^\star$ and γ in (35) and (36) and through the first part of Lemma 13.4 in [18, pp. 158–161] to obtain the value of t_0 in (37);

(ii) (Control of small divisors) In [18, §1.4] Rüßmann introduces a so-called 'approximation function' Φ in order to control the small divisors. Our choice is to take $\Phi(T) = T^{-\tau}$ with $\tau > n\mu$. Comparing [18, §1.4], one sees that such Φ does

† Recall definition (11)

‡ Compare this, in particular, with the estimates listed in [18, p. 171]; see, also, [15, Ch. 2].

not verify property 3, that is, $T^\lambda \Phi(T) \overset{T\to\infty}{\longrightarrow} 0$ for any $\lambda \geq 0$. However, when we consider $H = \hat{H}_\epsilon$, we will see below that one has $T_0 = T_{0,\epsilon} = O(\epsilon^{-2})$; then, in [**18**, (14.10.10), (14.10.11) and (13.1.4)], this would cause $O(\epsilon^\nu)$ to be an inadmissible size for a perturbation. Nevertheless we claim that the only decay property which is actually needed in Rüßmann's Theorem 3 is

$$\lim_{T\to\infty} T^\lambda \Phi(T) = 0 \quad \text{for all } 0 \leq \lambda < n\mu$$

so that our choice is perfectly suitable.

Now we are going to analyze what happens to the estimate in Lemma 8 when we consider $\hat{H}_\epsilon$ as a Hamiltonian function. In particular, we are going to show that each of the quantities appearing in Lemma 8 can be controlled by constants involving initial parameters related only to H_ϵ in (7) and (8) times powers of ϵ.

We point out that in the application of Theorem 3 with $H = \hat{H}_\epsilon$ in (16) we have the following correspondences†:

$$\begin{aligned} &\mathbb{T}^n = \mathbb{T}^d \times \mathbb{T}^p, \quad x = (\theta, \zeta), \\ &\mathcal{Y} = \mathcal{Y}_\epsilon := B^d(0, s/5) \times B^p(0, \epsilon), \quad y = (r, \rho), \\ &\mathcal{A} = \mathcal{A}_\epsilon := \mathbb{T}^d_{\sigma/8} \times \mathbb{T}^p_{\sigma_\star} \times D^d(0, s/5) \times D^p(0, \epsilon), \\ &P = \epsilon^\nu P_\epsilon, \quad N = N_\epsilon. \end{aligned} \tag{40}$$

According to Lemma 8(1) we consider the frequency application of the integrable part of $\hat{H}_\epsilon$, that is, $\hat{\Psi}_\epsilon(r, \rho)$ as in Proposition 2. We have already proved that $\hat{\Psi}_\epsilon$ is R-non-degenerate for $(r, \rho) \in B^d(0, s/5) \times B^{2p}(0, \epsilon)$. Now, in view of Lemma 8(1) and the correspondences in (40) we need to fix a compact set $\mathcal{K} = \mathcal{K}_\epsilon \subset \mathcal{A}_\epsilon$. For convenience we take $\mathcal{K}_\epsilon := \bar{B}^d(0, s/10) \times \bar{B}^p(0, \epsilon/2)$ so that the first inequality in (31) allows us to consider‡ $\mu = \bar{\mu}$ as an integer greater than the actual index of non-degeneracy of $\hat{\Psi}_\epsilon$ with respect to $\mathcal{K}_\epsilon$. Also, in view of the second inequality in (31), we can take

$$\beta = \beta_\epsilon := \frac{\epsilon^{\bar{\mu}+2}\bar{\beta}}{8} \tag{41}$$

in (36) and (38).

Next, we choose $\vartheta = \vartheta_\epsilon := \epsilon/16$ so that, for ϵ sufficiently small and in view of (15), one has $\mathcal{K}_\epsilon + 4\vartheta_\epsilon \subset \mathcal{A}_\epsilon$ as required in Lemma 8(2). Accordingly to Theorem 3 we also need to fix a positive number $\epsilon^\star < \operatorname{meas}_{d+p} \mathcal{K}_\epsilon$. In view of our definition of $\mathcal{K}_\epsilon$ and (15) a suitable choice is given by

$$\epsilon^\star = \epsilon^{p+1} \tag{42}$$

for ϵ small enough.

Now, observe that the quantities C_1 and d_0 in Lemma 8(2) do not cause any change in the order in ϵ of the size of the admissible perturbation. In fact, using (32) and taking ϵ

† See Theorem 3, Proposition 1, (11) and (21) for the notation.

‡ Recall Proposition 2 for the definition of $\bar{\mu}$.

sufficiently small, we have

$$\begin{aligned} C_1 = C_{1,\epsilon} &:= |\hat{\Psi}_\epsilon|_{\mathcal{K}_\epsilon+3\vartheta_\epsilon} \\ &\leq \sup_{r\in D^d(0,s/5)} |\omega(I_0+\epsilon r)| + \epsilon \sup_{r\in D^d(0,s/5)} |\Omega(I_0+\epsilon r)| + O(\epsilon) \\ &\leq \sup_{r\in D^d(I_0,s/5)} |\omega(r)| + \epsilon \sup_{r\in D^d(I_0,s/5)} |\Omega(r)| + O(\epsilon) \leq M_1 \end{aligned}$$

where M_1 is defined in (14). Since the estimate for ϵ_0 is decreasing with respect to C_1, we can substitute C_1 in (35) and (37) with M_1. The estimate for ϵ_0 is also decreasing with respect to d_0 so that when we consider $\mathcal{K} = \mathcal{K}_\epsilon$, we may simply replace d_0 by 1.

Let us now analyze the quantities in Lemma 8(3) and (4). First of all observe that in view of (31) and $n = d + p$ we can *a priori* fix an exponent $\tau \geq (d+p)\bar{\mu}$ satisfying the requirement in Lemma 8(3). Furthermore, given the previous choice of ϑ_ϵ, inequality (34) becomes

$$\int_{T_0}^{\infty} \frac{\log T}{T^2}\, dT \leq \frac{\epsilon}{16},$$

which can be easily fulfilled, together with $T_0 \geq e^{(d+p+1)/\tau}$, by choosing

$$T_0 = T_{0,\epsilon} := \frac{1}{\epsilon^2} \tag{43}$$

for ϵ sufficiently small. For what concerns the quantities defined in Lemma 8(4) we see that since $\vartheta = \vartheta_\epsilon := \epsilon/16$ and the estimate for ϵ_0 is decreasing in $C^\star$, we can choose

$$C^\star = C^\star_\epsilon := 2^{5(\bar{\mu}+1)}(\bar{\mu}+1)^{\bar{\mu}+2}(M_1+1)\epsilon^{-(\bar{\mu}+1)} \tag{44}$$

having also used $C_1 = C_{1,\epsilon} \leq M_1$. From the fact that we can replace d_0 by 1 together with (41), (42) and (44), we have

$$\gamma = \gamma_\epsilon := c_1(M_1+1)^{-\bar{\mu}/2}\epsilon^{(\bar{\mu}+1)^2+((p+1)\bar{\mu}/2)}\bar{\beta}^{(\bar{\mu}+1)/2}\epsilon^{\star\bar{\mu}/2} \tag{45}$$

for a suitable constant $c_1 < 1$ depending only on $\bar{\mu}$. Moreover, given once again the previous choice of $\vartheta = \vartheta_\epsilon$ together with (43) and the above definition of γ_ϵ, we can replace t_0 in (37) by

$$t_0 = t_{0,\epsilon} = c_1(M_1+1)^{-\bar{\mu}/2}\epsilon^{(\bar{\mu}+1)^2+((p+1)\bar{\mu}/2)+2(\tau+d+p)+3} \tag{46}$$

for $c_1 < 1$ as above.

From (38) we see that E_1 and E_2 have simple polynomial dependence on the quantities γ, T_0^{-1}, ϑ, β and t_0. Our previous analysis shows that when we consider $\hat{H}_\epsilon$ as a Hamiltonian function, these quantities can be replaced, respectively, by†

$$\begin{aligned} &\gamma_\epsilon = O(\epsilon^{(\bar{\mu}+1)^2+((p+1)\bar{\mu}/2)}), \quad T_{0,\epsilon}^{-1} = O(\epsilon^2), \quad \vartheta_\epsilon = O(\epsilon) \\ &\beta_\epsilon = O(\epsilon^{\bar{\mu}+2}), \quad t_{0,\epsilon} = O(\epsilon^{(\bar{\mu}+1)^2+((p+1)\bar{\mu}/2)+2(\tau+d+p)+3}). \end{aligned}$$

† See (45), (43), (41), (46) and recall $\vartheta = \vartheta_\epsilon := \epsilon/16$.

https://doi.org/10.1017/S0143385708000503 Published online by Cambridge University Press

Therefore, in view of (39) the size of the perturbation allowed by Rüßmann's theorem when we consider $H = \hat{H}_\epsilon$, is of the order of ϵ^{ν_0} with†

$$\nu_0 := 2\bar{\mu}^3 + (p+5)\bar{\mu}^2 + [14 + 4(\tau + d + p)]\bar{\mu} + 13. \tag{47}$$

In particular, we have a condition of the form $\epsilon_0 \le \bar{c}\epsilon^{\nu_0}$ where $\bar{c}$ is some positive constant independent of ϵ and depending only on quantities related to the initial Hamiltonian H_ϵ, namely $\bar{\mu}$, $\bar{\beta}$ and $\bar{\mathcal{K}}$ as in Proposition 2, the Diophantine constant $\tau \ge (d+p)\bar{\mu}$ and M_1 as in (14) with r_0 as in (12). By Proposition 1 we know that we can assume that the size of the perturbation of $\hat{H}_\epsilon$ is of the order of ϵ^ν for any fixed integer $\nu \ge 4$ independent of ϵ. Thus, by simply taking‡ $\nu > \nu_0$, we can apply Rüßmann's theorem to $\hat{H}_\epsilon$ and obtain Theorem 4 as a consequence. □

3. *Proof of Theorem 2*

As follows from the analysis described in [**9**, §6, pp. 1563–1569], the motions of $(n+1)$ bodies (point masses) interacting only through gravitational attraction, restricted to the invariant symplectic submanifold of vanishing total linear momentum, are governed by the real-analytic Hamiltonian

$$F = H^0(\Lambda) + \epsilon(H^1(\Lambda, \xi, \eta, q, p) + H^2(\lambda, \Lambda, \xi, \eta, q, p)) \tag{48}$$

where:

(i) $(\lambda, \Lambda, \xi, \eta, q, p) \in \mathbb{T}^n \times (0,\infty)^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ are standard symplectic coordinates;

(ii) $\Lambda_j = \mu_j\sqrt{M_j a_j}$, where $a_j > 0$ are the semi major-axis of the 'instantaneous' Keplerian ellipse formed by the 'Sun' (major body) and the jth 'planet', while

$$\frac{1}{\epsilon\mu_j} = \frac{1}{m_0} + \frac{1}{\epsilon m_j}, \qquad M_j := m_0 + \epsilon m_j,$$

with m_0 and ϵm_j the mass of the Sun and the mass of the jth planet, respectively;

(iii) the phase space $\mathcal{M}$ is the open subset of $\mathbb{T}^n \times (0,\infty)^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ subject to the collisionless constraint

$$0 < a_n < a_{n-1} < \cdots < a_1$$

and endowed with the standard symplectic form

$$\sum_{j=1}^n d\lambda_j \wedge d\Lambda_j + d\xi_j \wedge d\eta_j + dq_j \wedge dp_j;$$

(iv) $H^0 := F_{\mathrm{Kep}}$ is the Keplerian integrable limit given by

$$H^0 := F_{\mathrm{Kep}} := \sum_{j=1}^n -\frac{\mu_j^3 M_j^2}{2\Lambda_j^2},$$

describing n decoupled two-body systems formed by the Sun and the jth planet;

† Using (47) we are able to define the values of ν_1 in (19) and (20) and ν_2 in (22) through (29).
‡ Note that this can be done since ν_0 only depends on d, p, τ and $\bar{\mu}$.

https://doi.org/10.1017/S0143385708000503 Published online by Cambridge University Press

(v) the 'secular' Hamiltonian H^1 has the form†

$$H^1 = C_0 + \sum_{j=1}^{n} \sigma_j \frac{\xi_j^2 + \eta_j^2}{2} + \sum_{j=1}^{n} \varsigma_j \frac{q_j^2 + p_j^2}{2} + O(4), \tag{49}$$

where C_0, σ_j and ς_j depend on Λ; '$O(4)$' denotes terms of order greater than or equal to four in (ξ, η, q, p) (and depending on Λ);

(vi) H^2 has vanishing average over $\lambda \in \mathbb{T}^n$; H^i also depend (in a regular and non-influential way) on ϵ.

Remark 6. The variables $(\lambda, \Lambda, \xi, \eta, q, p)$ are obtained from standard Poincaré variables after a rotation in (ξ, η, q, p) needed to diagonalize the quadratic part of the secular Hamiltonian; the 'eigenvalues' σ_j and ς_j are the first Birkhoff invariants of the secular Hamiltonian; cf. [**9**, pp. 1568–1569].

The frequency map of the planetary Hamiltonian F is given by

$$\{\nu_1, \ldots, \nu_n, \sigma_1, \ldots, \sigma_n, \varsigma_1, \ldots, \varsigma_n\}$$

where the ν_j are the Keplerian frequencies

$$\nu_j := \frac{\partial F_{\mathrm{Kep}}}{\partial \Lambda_j} = \frac{\sqrt{M_j}}{a_j^{3/2}} = \frac{\mu_j^3 M_j^2}{\Lambda_j^3}. \tag{50}$$

It is customary to consider the frequency map as a function of the semi-major axes a (rather than of the actions Λ); we therefore call 'planetary frequency map' the application‡

$$\alpha : a \in \mathcal{A} \longmapsto \{\nu_1, \ldots, \nu_n, \sigma_1, \ldots, \sigma_n, \varsigma_1, \ldots, \varsigma_n\} \in \mathbb{R}^{3n} \tag{51}$$

where

$$\mathcal{A} := \{(a_1, a_2, \ldots, a_n) \in \mathbb{R}^n \mid 0 < a_n < a_{n-1} < \cdots < a_1\}.$$

Clearly, the idea is to apply Theorem 4 to the real-analytic Hamiltonian F in (48) with $d = n$ and $p = 2n$: (φ, I) corresponding to (λ, Λ) here and u corresponding to (ξ, q) and v to (η, p). However, it turns out that *the main hypothesis of Theorem 4 does not hold*, namely, the planetary frequency map α is R-degenerate: in fact (up to rearranging the (q, p)-variables) one has

$$\begin{cases} \varsigma_n = 0, \\ \displaystyle\sum_{j=1}^{n} (\sigma_j + \varsigma_j) = 0. \end{cases} \tag{52}$$

The first relation is related to the rotation invariance of the system; the second relation seems to have been noticed (at least in this generality) for the first time by Herman and is therefore normally referred to as the 'Herman resonance'.

The two resonances in (52) are, however, the only linear relations satisfied identically; in fact, the following result was proved in [**9**, Proposition 78, p. 1575].

† There is a difference of a factor $\frac{1}{2}$ with the notation used in [**9**]. The computations are performed in [**13**].
‡ Obviously, the property of being R-non-degenerate can be equivalently discussed in terms of the Λ or in terms of the a.

PROPOSITION 3. *For all $n \geq 2$ there exists an open and dense set with full Lebesgue measure $U \subset \mathcal{A}$, where $\alpha_j \neq \alpha_i$ whenever $j \neq i$ and the following property holds: for any open and simply connected set $V \subset U$, the α_j define $3n$ holomorphic functions and if*

$$\alpha \cdot (c^1, c^2, c^3) = \nu \cdot c^1 + \sigma \cdot c^2 + \varsigma \cdot c^3 \equiv 0$$

for some $c^i \in \mathbb{R}^n$, then

$$\begin{cases} \text{either} & c^1 = 0, \quad c^2 = 0, \quad c^3 = (0, \ldots, 0, 1), \\ \text{or} & c^1 = 0, \quad c^2 = (1, \ldots, 1) = c^3. \end{cases} \tag{53}$$

In order to remove the secular resonances (52), we consider the following 'extended Hamiltonian' on

$$\tilde{\mathcal{M}} := \mathcal{M} \times \mathbb{T} \times \mathbb{R}$$

adding a pair of conjugate symplectic variables† $(\theta_\rho, \rho) \in \mathbb{T} \times \mathbb{R}$:

$$\tilde{F} := F + \frac{\rho^2}{2} + \epsilon \rho^2 C_z \quad \text{with} \quad C_z := \sum_{j=1}^{n} \left(\Lambda_j - \frac{1}{2}(\xi_j^2 + \eta_j^2 + q_j^2 + p_j^2) \right). \tag{54}$$

Let us make a few comments.

(vii) Here C_z is the vertical component of the total angular momentum in Poincaré variables (cf. [**14**] and also [**9**, formula (44)]); the form of C_z is unchanged in the above variables (ξ, η, q, p), which are obtained from the Poincaré variables by an orthogonal transformation.

(viii) Since C_z is an integral for F (i.e. Poisson commutes with F), F and $\tilde{F}$ Poisson commute:

$$\{F, \tilde{F}\}^{\sim} = \{F, \tilde{F}\} = 0,$$

where $\{\cdot, \cdot\}^{\sim}$ and $\{\cdot, \cdot\}$ denote, respectively, the Poisson bracket on $\tilde{\mathcal{M}}$ and on $\mathcal{M}$; clearly, since $\tilde{F}$ does not depend explicitly on the angle θ_ρ, ρ is an integral for $\tilde{F}$ (and for F).

This fact will be important later since from Lagrangian intersection theory it follows that two commuting Hamiltonians have, in general, the same Lagrangian tori (see item (x) below for the precise statement).

(ix) The extended Hamiltonian $\tilde{F}$ may be rewritten as

$$\tilde{F} = \tilde{H}^0 + \epsilon(\tilde{H}^1 + H^2)$$

with

$$\begin{aligned} \tilde{H}^0 &:= F_{\mathrm{Kep}}(\Lambda) + \frac{\rho^2}{2}, \\ \tilde{H}^1 &:= C_0(\Lambda) + \rho^2 \sum_{j=1}^{n} \Lambda_j \\ &\quad + \sum_{j=1}^{n} (\sigma_j - \rho^2) \frac{\xi_j^2 + \eta_j^2}{2} + \sum_{j=1}^{n} (\varsigma_j - \rho^2) \frac{p_j^2 + q_j^2}{2} + O(4). \end{aligned}$$

† That is, $\tilde{\mathcal{M}}$ is endowed with the symplectic form $\sum_{j=1}^{n} (d\lambda_j \wedge d\Lambda_j + d\xi_j \wedge d\eta_j + dq_j \wedge dp_j) + d\theta_\rho \wedge d\rho$.

https://doi.org/10.1017/S0143385708000503 Published online by Cambridge University Press

Thus, the 'slow' action variables‡ are $I = (\rho, \Lambda_1, \ldots, \Lambda_n)$ and the (extended) planetary frequency map is given by

$$\tilde{\alpha} : (\rho, a) \in \mathcal{A} \times \mathbb{R} \longmapsto \tilde{\alpha}(\rho, a) := ((\rho, \nu), \tilde{\sigma}, \tilde{\varsigma}) \in \mathbb{R}^{n+1} \times \mathbb{R}^n \times \mathbb{R}^n$$

with

$$\tilde{\sigma}_j := \sigma_j - \rho^2, \quad \tilde{\varsigma}_j := \varsigma_j - \rho^2.$$

Proposition 3, implies immediately that $\tilde{\alpha}$ is R-non-degenerate: suppose, in fact, that

$$\tilde{\alpha} \cdot ((c, c^1), c^2, c^3) = \rho\, c + \nu \cdot c^1 + \tilde{\sigma} \cdot c^2 + \tilde{\varsigma} \cdot c^3 \equiv 0$$

for some $c \in \mathbb{R}$ and $c^i \in \mathbb{R}^n$; such an expression is a second-order polynomial in ρ and in order to vanish identically have to vanish its coefficients, that is,

$$\nu \cdot c^1 + \sigma \cdot c^2 + \varsigma \cdot c^3 = 0, \tag{55}$$

$$c = 0,$$

$$-\sum_{j=1}^{n} c_j^2 + c_j^3 = 0. \tag{56}$$

However, then, by Proposition 3 (and because of (55)), one must have one of the alternatives listed in (53), which are incompatible with (56).

Thus, $\tilde{\alpha}$ is *R-non-degenerate* as claimed and Theorem 4 can be applied to the extended Hamiltonian† $\tilde{F}$, yielding, for ϵ small enough, a positive measure set of real-analytic $(3n + 1)$-dimensional Lagrangian tori in $\tilde{\mathcal{M}}$ invariant for $\tilde{F}$ and carrying quasi-periodic motion with Diophantine frequencies.

The fact that $\tilde{F}$ is independent of θ_ρ and that $\omega_* := \partial_\rho \tilde{F} = (\rho + 4\epsilon\rho C_z)$ is constant along $\tilde{F}$-trajectories (cf. point (viii) above) implies immediately that the tori $\mathcal{T} \subset \tilde{\mathcal{M}}$ obtained through Theorem 4 have the following parametrization

$$\mathcal{T} := \{(Z(\psi, \theta_\rho), \theta_\rho, \rho) \mid (\psi, \theta_\rho) \in \mathbb{T}^{3n} \times \mathbb{T}\}, \tag{57}$$

where $Z \in \mathcal{M}$ and with $\tilde{F}$-flow given by

$$\phi^t_{\tilde{F}}(Z(\psi, \theta_\rho), \theta_\rho, \rho) = (Z(\psi + \omega t, \theta_\rho + \omega_* t), \theta_\rho + \omega_* t, \rho),$$

for a suitable vector $\omega \in \mathbb{R}^{3n}$, so that (ω, ω_*) forms a Diophantine vector in $\mathbb{R}^{3n+1}$.

(x) In [**9**, Lemma 82, p. 1578] the following statement is proved:

if F and G are two commuting Hamiltonians and if $\mathcal{T}$ is a Lagrangian torus invariant for F and with a dense F-orbit, then it is also G-invariant.

Thus, since $\tilde{F}$ and F (viewed as a functions on $\tilde{\mathcal{M}}$) commute, the tori obtained in (ix) (on which any $\tilde{F}$-orbit is dense) are also invariant for the flow on $\tilde{\mathcal{M}}$ generated by F. Furthermore, the F-flow in $\tilde{\mathcal{M}}$ leaves *both* θ_ρ and ρ fixed so that, for any fixed $\theta_\rho \in \mathbb{T}$, the $3n$-dimensional torus

$$\mathcal{T}_{\theta_\rho} := \{(Z(\psi, \theta_\rho), \theta_\rho, \rho) \mid \psi \in \mathbb{T}^{3n}\}$$

is invariant for F. However, this means that such tori are invariant also for the F-flow in $\mathcal{M}$, finishing the proof of Theorem 2. □

‡ Corresponding in Theorem 4 to $I = (I_1, \ldots, I_d)$, $d = n + 1$; cf. also the following footnote.

† The correspondence with the notation of Theorem 4 being $d = n + 1$, $p = 2n$, $H_\epsilon = \tilde{F}$, $f = \tilde{H}^1 + H^2$, $I = (\Lambda, \rho)$, $(u, v) = ((\xi, q), (\eta, p))$, $h(I) = H^0$, $f_{00} = C_0 + \rho^2 \sum_{j=1}^n \Lambda_j$, $\omega = (\nu, \rho)$, $\Omega = (\tilde{\sigma}, \tilde{\varsigma})$.

Remark 7. The strategy followed here is similar to that followed in [**9**] with a few differences: first, in [**9**], ρ is treated as a dumb parameter and no extended phase space is introduced (but an extra argument is then needed to discuss the non-degeneracy of the frequency map with respect to parameters and to discuss the measure of the tori obtained); secondly, in [**9**], there is a restriction to a fixed vertical angular momentum submanifold, which is not needed here.

Acknowledgements. We are indebted to Jacques Féjoz for many enlightening discussions. This work was partially supported by the Italian MIUR project 'Metodi variazionali e equazioni differenziali nonlineari'.

REFERENCES

[1] V. I. Arnol'd. Small denominators and problems of stability of motion in classical and celestial mechanics. *Uspekhi Mat. Nauk.* **18**(6 (114)) (1963), 91–192.
[2] V. I. Arnol'd, V. V. Kozlov and A. I. Neishtadt (eds). *Mathematical Aspects of Classical and Celestial Mechanics*, 3rd edition *(Encyclopaedia of Mathematical Sciences, 3)*. Springer, Berlin, 2006.
[3] L. Biasco, L. Chierchia and E. Valdinoci. Elliptic two-dimensional invariant tori for the planetary three-body problem. *Arch. Ration. Mech. Anal.* **170** (2003), 91–135.
[4] F. Boigey. Élimination des nòe uds dans le problème newtonien des quatre corps. *Celestial Mech.* **27**(4) (1982), 399–414.
[5] H. Broer and F. Takens. Unicity of Kam tori. *Ergod. Th. & Dynam. Sys.* **27** (2007), 713–724.
[6] A. Celletti and L. Chierchia. Kam stability and celestial mechanics. *Mem. Amer. Math. Soc.* **187**(878) (2007), 134.
[7] A. Deprit. Elimination of the nodes in problems of n bodies. *Celestial Mech.* **30**(2) (1983), 181–195.
[8] A. Fathi, A. Giuliani and A. Sorrentino. Uniqueness of invariant Lagrangian graphs in a homology or a cohomology class, 2008. *Preprint*. Available at http://arxiv.org/PS_cache/arxiv/pdf/0801/0801.3568v1.pdf.
[9] J. Féjoz. Démonstration du 'théoréme d'Arnold' sur la stabilité du système planétaire (d'après M. Herman). *Ergod. Th. & Dynam. Sys.* **24**(5) (2004), 1521–1582.
[10] J. Féjoz. Version révisée de l'article paru dans le Michael Herman Memorial Issue. *Ergod. Th. & Dynam. Sys.* **24**(5) (2004), 1521–1582.
[11] M. R. Herman. Démonstration d'un théoréme de V.I. Arnold. *Séminaire de Systémes Dynamiques et manuscripts*, Université D. Diderot, Paris 7, 1998.
[12] H. Hofer and E. Zehnder. *Symplectic Invariants and Hamiltonian Dynamics*. Birkhäuser, Basel, 1994.
[13] J. Laskar. Analytical framework in Poincaré variables for the motion of the solar system. *Predictability, Stability and Chaos in n-Body Dynamical Systems (Nato Adv. Sci. Inst. Ser. B Phys., 272)*. Plenum, New York, 1991, pp. 93–114.
[14] H. Poincaré. *Leçons de mécanique céleste*. Gauthier-Villars, Paris, 1905 (Vol. 1); 1907 (Vol. 2); 1910 (Vol. 3); 2005 (Reprint).
[15] F. Pusateri. Analytic KAM tori for the planetary $(n+1)$-body problem. Master's Thesis, Università degli Studi Roma Tre, 2006. http://www.mat.uniroma3.it/users/chierchia/TESI.
[16] A. S. Pyartli. Diophantine approximations on Euclidean submanifolds. *Functional. Anal. Appl.* **3** (1969), 303–306.
[17] P. Robutel. Stability of the planetary three-body problem II. KAM theory and existence of quasi-periodic motions. *Celestial Mech. Dynam. Astronom.* **62**(3) (1995), 219–261.
[18] H. Rüßmann. Invariant tori in non-degenerate nearly integrable Hamiltonian systems. *R. & C. Dynamics* **2**(6) (2001), 119–203.
[19] D. Salamon. The Kolmogorov–Arnold–Moser theorem. *Math. Phys. Electron. J.* **3**(37) (2004), (electronic).
[20] M. B. Sevryuk. The classical KAM theory and the dawn of the twenty-first century. *Mosc. Math. J.* **3**(3) (2003), 1113–1144.