

The Classical KAM Theorem for Hamiltonian Systems via Rational Approximations

Abed Bounemoura^{1*} and Stéphane Fischler^{2**}

¹*CNRS — CEREMADE, Université Paris Dauphine
Place du Maréchal de Lattre de Tassigny, 75775 Paris Cedex 16, France
IMCCE, Observatoire de Paris
77 avenue Denfert-Rochereau, 75014 Paris, France*

²*Laboratoire de mathématiques d'Orsay,
Univ Paris Sud, 91405 Orsay Cedex, France*

Received January 21, 2014; accepted March 11, 2014

Abstract—In this paper, we give a new proof of the classical KAM theorem on the persistence of an invariant quasi-periodic torus, whose frequency vector satisfies the Bruno–Rüssmann condition, in real-analytic non-degenerate Hamiltonian systems close to integrable. The proof, which uses rational approximations instead of small divisors estimates, is an adaptation to the Hamiltonian setting of the method we introduced in [4] for perturbations of constant vector fields on the torus.

MSC2010 numbers: 37J25, 37J40, 70H08, 70H09

DOI: 10.1134/S1560354714020087

Keywords: perturbation of integrable Hamiltonian systems, KAM theory, Diophantine duality, periodic approximations

1. INTRODUCTION

In this paper, we consider small perturbations of integrable Hamiltonian systems, which are defined by a Hamiltonian function of the form

$$H(p, q) = h(p) + \epsilon f(p, q), \quad (p, q) \in \mathbb{R}^n \times \mathbb{T}^n, \quad 0 \leq \epsilon < 1,$$

where $n \geq 2$ is an integer and $\mathbb{T}^n = \mathbb{R}^n / \mathbb{Z}^n$: the Hamiltonian system associated to this Hamiltonian function is then given by

$$\begin{cases} \dot{p} = -\partial_q H(p, q) = -\epsilon \partial_q f(p, q), \\ \dot{q} = \partial_p H(p, q) = \nabla h(p) + \epsilon \partial_p f(p, q). \end{cases}$$

When $\epsilon = 0$, the system associated to $H = h$ is trivially integrable: all solutions are given by

$$(p(t), q(t)) = (p(0), q(0) + t \nabla h(p(0)) \ [\mathbb{Z}^n]),$$

and therefore the sets $\mathcal{T}_{p_0} = \{p_0\} \times \mathbb{T}^n$, $p_0 \in \mathbb{R}^n$, are invariant tori on which the dynamics is quasi-periodic with frequency $\omega_0 = \nabla h(p_0) \in \mathbb{R}^n$.

Now for $\epsilon > 0$, the system is in general no longer integrable, and one is interested to know whether such quasi-periodic solutions persist under an arbitrary small perturbation. It is not hard to see that if the frequency $\omega_0 \in \mathbb{R}^n$ is resonant, that is if there exists a vector $k \in \mathbb{Z}^n \setminus \{0\}$ such that $k \cdot \omega_0 = 0$, then the torus \mathcal{T}_{p_0} is destroyed by a general perturbation. This goes back to Poincaré, who initiated the study of small perturbations of integrable Hamiltonian systems in his seminal work on Celestial Mechanics.

*E-mail: abedbou@gmail.com

**E-mail: stephane.fischler@math.u-psud.fr

1.1. K for Kolmogorov

The fate of quasi-periodic solutions with non-resonant frequencies remained an open question for more than half a century, until it was solved by Kolmogorov. In [14], he proved that if H is real-analytic and if h is non-degenerate at some point $p_0 \in \mathbb{R}^n$ in the sense that ∇h is a local diffeomorphism at p_0 , then, provided $\omega_0 = \nabla h(p_0)$ satisfies a classical Diophantine condition, the torus \mathcal{T}_{p_0} survives an arbitrary small perturbation. The condition on ω_0 is a strengthening of the non-resonance condition, namely one requires the existence of constants $\gamma > 0$ and $\tau \geq n - 1$ such that for all $k \in \mathbb{Z}^n \setminus \{0\}$,

$$|k \cdot \omega_0| \geq \gamma |k|_1^{-\tau},$$

where $|k|_1 = |k_1| + \dots + |k_n|$ (this is a generalization of a condition used by Siegel in [25] for the problem of linearization of a one-dimensional holomorphic map at an elliptic fixed point). Whereas classical Hamiltonian perturbation theory, as pioneered by Poincaré, only yields formal quasi-periodic solutions through an iterative procedure which may or may not converge, Kolmogorov's main idea was to focus on such a strongly non-resonant torus in order to use a modified and rapidly converging inductive scheme, similar to a Newton method, leading to the persistence of this torus.

1.2. A and M for Arnold and Moser

Kolmogorov's fundamental theorem was later revisited and improved by Arnold and Moser, leading to what is known as KAM theory.

In [2], Arnold gave a more detailed and technically different proof, under a different non-degeneracy assumption on the integrable Hamiltonian, and in [3], he further improved the non-degeneracy assumption in order to apply the theorem to Celestial Mechanics.

In the meantime, in [16], Moser was able to replace the analyticity condition on the Hamiltonian by a mere finite differentiability condition, in the related context of invariant curves of area-preserving maps of the annulus. Moreover, in [17], he introduced a powerful formalism for the perturbation theory (not necessarily Hamiltonian) of quasi-periodic solutions, which eventually led to many applications (see [5] for instance).

1.3. R for Rüssmann

Further important contributions to KAM theory are due to Rüssmann. Indeed, following works of Arnold and Pyartli, Rüssmann was able to find the most general non-degeneracy condition for the integrable Hamiltonian h : in the analytic case, it is sufficient to require that locally, the image of the gradient map ∇h is not contained in any hyperplane of \mathbb{R}^n (it is also necessary as shown by Sevryuk in [24]). This was announced in [20], and details are given in [22]. Also, he was able to greatly relax the condition imposed on the frequency, going beyond the classical Diophantine condition (see [21, 22]). This condition, which generalizes a condition introduced by Bruno in the context of holomorphic linearization ([8, 9]), is now usually called Bruno–Rüssmann condition (see Section 2.2 for a definition), and is known to be optimal in one-dimensional problems following works of Yoccoz (see [27, 28]). This extension to more general frequency vectors also led to a different method of proof, in which no rapid convergence is involved ([21, 22], see also [23] for the latest improvement of this method). In another important work ([19]), Rüssmann was able to obtain optimal estimates for the solution of the so-called cohomological equation (which is Eq. (1.1) in Section 1.5 below), of fundamental importance in KAM theory.

1.4. Other approaches to the classical KAM theorem

Apart from Rüssmann's modified iterative scheme, a number of other proofs have appeared in the literature.

First, the classical iterative scheme of Kolmogorov has been replaced by the use of an adapted implicit function theorem in a scale of Banach spaces (or in a Fréchet space), following works of Zehnder ([29, 30]), Herman ([6]) and more recently Féjóz ([11]).

Another proof, based on the Lagrangian formalism and that avoids the composition of infinitely many coordinate transformations, was presented in [15] in the case of invariant curves for area-preserving maps of the annulus, and in [26] for Hamiltonian systems in any number of degrees of freedom.

But perhaps the most striking proof is due to Eliasson. The classical theorem of Kolmogorov shows, a posteriori and in an indirect way, that some formal solutions of classical perturbation theory do converge. In [10], Eliasson managed to prove directly the convergence of these formal solutions, by adding suitable terms in the formal series in order to exhibit subtle cancellations yielding the absolute convergence.

At last, we should also point out that using a multi-dimensional continued fraction algorithm due to Lagarias, Khanin, Lopes Dias and Marklof gave a proof of the KAM theorem with techniques closer to renormalization theory (see [13] for the case of constant vector fields on the torus, and [12] for the case of Hamiltonian systems).

1.5. Approach via Rational Approximations

The purpose of this article, which can be considered as a continuation of our previous work [4], is to present yet another proof of the classical KAM theorem for Hamiltonian systems, which differs qualitatively from all other existing proofs as it does not involve any small divisors estimates and Fourier series expansions.

First we should recall that in the classical approach to the KAM theorem, as well as the other methods of proof we just mentioned, a central role is played by the following equation:

$$\mathcal{L}_{\omega_0}g = f - [f], \quad [f] = \int_{\mathbb{T}^n} f(\theta)d\theta, \tag{1.1}$$

where g (respectively f) is the unknown (respectively known) smooth function on \mathbb{T}^n , ω_0 is non-resonant and \mathcal{L}_{ω_0} is the derivative in the direction of ω_0 . It is precisely in trying to solve Eq. (1.1) that small divisors arise: geometrically, one needs to integrate along the integral curves $t \mapsto \theta + t\omega_0 \in [\mathbb{Z}^n]$, and these curves are not closed (they densely fill the torus). Analytically, one needs to invert the operator \mathcal{L}_{ω_0} acting on the space of smooth functions, and this operator is unbounded. Indeed, this operator can be diagonalized in a Fourier basis: letting $e_k(\theta) = e^{2\pi i k \cdot \theta}$ for $k \in \mathbb{Z}^n$ and expanding in Fourier series $g = \sum_{k \in \mathbb{Z}^n} g_k e_k$ and $f = \sum_{k \in \mathbb{Z}^n} f_k e_k$, the solution is given by

$$g_0 = 0, \quad g_k = (2\pi i k \cdot \omega_0)^{-1} f_k, \quad k \in \mathbb{Z}^n \setminus \{0\}.$$

The quantities $k \cdot \omega_0$, which enter in the denominators, can be arbitrarily small if the norm $|k|_1$ is arbitrarily large: these are called the small divisors, and are the main source of complications.

In this article, we will show how one can, using rational approximations, avoid solving Eq. (1.1) and therefore avoid facing small divisors. Assume, without loss of generality, that the first component of ω_0 has been normalized to one. Using a Diophantine result deduced in [4] from classical properties of geometry of numbers, we will approximate ω_0 by n rational vectors $v_1, \dots, v_n \in \mathbb{Q}^n$, with a control on their denominators q_1, \dots, q_n in terms of the quality of the approximation, and such that the integer vectors $q_1 v_1, \dots, q_n v_n$ form a \mathbb{Z} -basis of \mathbb{Z}^n . This result will enable us to replace the study of Eq. (1.1) with the study of the equations

$$\mathcal{L}_{v_j}g_j = f_j - [f_j]_{v_j}, \quad [f_j]_{v_j}(\theta) = \int_0^1 f(\theta + tq_j v_j)dt, \quad 1 \leq j \leq n, \tag{1.2}$$

where the f_j are defined inductively by $f_1 = f$ and $f_{j+1} = [f_j]_{v_j}$ for $1 \leq j \leq n - 1$, and the g_j are the unknown. Equations (1.2) are much simpler than Eq. (1.1), they can be solved without Fourier expansions by the following simple integral formula

$$g_j(\theta) = q_j \int_0^1 (f_j - [f_j]_{v_j})(\theta + tq_j v_j)tdt$$

and there are no small divisors: geometrically, the integral curves $t \mapsto \theta + tv_j [\mathbb{Z}^n]$ are q_j -periodic hence closed, and analytically, the inverse operator of \mathcal{L}_{v_j} is bounded (by q_j , with respect to any translation-invariant norm on the space of functions on the torus).

In [4], this approach was already used in the model problem of perturbations of constant vector fields on the torus; the aim of this article is therefore to explain how to adapt the arguments of [4] to the context of real-analytic non-degenerate Hamiltonian systems close to integrable.

2. STATEMENTS

As explained in the Introduction, the result that we will prove here is not new, only the method of proof is. For convenience, we will follow the very nice survey [18] for the exposition of the statements. Compared to [18], we decided for simplicity to focus on the persistence of a single invariant torus instead of a Cantor family of invariant tori; on the other hand, our frequency will be assumed to satisfy the Bruno–Rüssmann condition, which is more general than the classical Diophantine condition.

2.1. Setting

Recall that $n \geq 2$ is an integer, $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$ and let $D \subseteq \mathbb{R}^n$ be an open domain containing the origin. For a small parameter $\epsilon \geq 0$, we consider a Hamiltonian function $H : D \times \mathbb{T}^n \rightarrow \mathbb{R}$ of the form

$$\begin{cases} H(p, q) = h(p) + \epsilon f(p, q), \\ \nabla h(0) = \omega_0 = (1, \bar{\omega}_0) \in \mathbb{R}^n, \quad \bar{\omega}_0 \in [-1, 1]^{n-1}. \end{cases} \quad (*)$$

Assuming that the vector $\nabla h(0) = \omega_0$ is non-zero, it can always be written as above, re-ordering its components and re-scaling the Hamiltonian if necessary. The integrable Hamiltonian h is said to be *non-degenerate* at the origin if the map $\nabla h : D \rightarrow \mathbb{R}^n$ is a local diffeomorphism at the origin. Upon restricting D if necessary, we may assume that ∇h is actually a global diffeomorphism. The Hamiltonian H is said to be *real-analytic* on $\bar{D} \times \mathbb{T}^n$, where \bar{D} denotes the closure of D in \mathbb{R}^n , if it is analytic on a fixed (that is, independent of ϵ) neighborhood of $\bar{D} \times \mathbb{T}^n$ in $\mathbb{R}^n \times \mathbb{T}^n$. This implies that H can be extended as a holomorphic function on a fixed complex neighborhood of $\bar{D} \times \mathbb{T}^n$ in $\mathbb{C}^n \times \mathbb{T}_{\mathbb{C}}^n$, where $\mathbb{T}_{\mathbb{C}}^n = \mathbb{C}^n/\mathbb{Z}^n$, which is real-valued for real arguments.

2.2. Bruno–Rüssmann Condition

For $Q \geq 1$, let us define the function $\Psi = \Psi_{\omega_0}$ by

$$\Psi(Q) = \sup\{|k \cdot \omega_0|^{-1} \mid k \in \mathbb{Z}^n, 0 < |k|_1 \leq Q\} \in [1, +\infty]. \quad (2.1)$$

Recall that the vector ω_0 is said to be non-resonant if $k \cdot \omega_0 = 0$ implies $k = 0 \in \mathbb{Z}^n$, which is equivalent to $\Psi(Q)$ being finite for all $Q \geq 1$. We say that ω_0 satisfies the Bruno–Rüssmann condition if it is non-resonant and

$$\int_1^{+\infty} Q^{-2} \ln(\Psi(Q)) dQ < \infty. \quad (\text{BR})$$

Now let us define two other functions $\Delta = \Delta_{\omega_0}$ and $\Delta^* = \Delta_{\omega_0}^*$ by $\Delta(Q) = Q\Psi(Q)$ for $Q \geq 1$, and $\Delta^*(x) = \sup\{Q \geq 1 \mid \Delta(Q) \leq x\}$ for $x \geq \Delta(1)$. We obviously have $\Delta^*(\Delta(Q)) = Q$ and $\Delta(\Delta^*(x)) \leq x$. Moreover, it is not hard to check (see [4], Appendix A) that Ψ satisfies (BR) if and only if Δ^* satisfies

$$\int_{\Delta(1)}^{+\infty} (x\Delta^*(x))^{-1} dx < \infty. \quad (2.2)$$

2.3. Classical KAM Theorem

Consider the map $\Theta_0 : \mathbb{T}^n \rightarrow D \times \mathbb{T}^n$ given by $\Theta_0(q) = (0, q)$: this is a real-analytic torus embedding such that $\Theta_0(\mathbb{T}^n)$ is invariant by the Hamiltonian flow of H_0 and quasi-periodic with frequency ω_0 . The classical KAM theorem states that this invariant quasi-periodic torus is preserved, being only slightly deformed, by an arbitrary small perturbation, provided h is non-degenerate, H analytic and ω_0 satisfies the Bruno–Rüssmann condition.

Theorem 1. *Let H be as in (*), with h non-degenerate and H real-analytic, and assume that $\Psi = \Psi_{\omega_0}$ satisfies (BR). For ϵ small enough, there exists a real-analytic torus embedding $\Theta_{\omega_0} : \mathbb{T}^n \rightarrow D \times \mathbb{T}^n$ such that $\Theta_{\omega_0}(\mathbb{T}^n)$ is invariant by the Hamiltonian flow of H and quasi-periodic with frequency ω_0 .*

Moreover, Θ_{ω_0} converges uniformly to Θ_0 as ϵ goes to zero.

As in [18], Theorem 1 will be deduced from a KAM theorem for a Hamiltonian “with parameters”, for which a quantitative statement is given below.

2.4. KAM Theorem with Parameters

Let us now consider a different setting. Given r, s, h real numbers such that $0 \leq r \leq 1, 0 \leq s \leq 1, 0 \leq h \leq 1$, we let

$$D_{r,s} = \{I \in \mathbb{C}^n \mid |I| < r\} \times \{\theta \in \mathbb{T}_{\mathbb{C}}^n = \mathbb{C}^n / \mathbb{Z}^n \mid |\text{Im}(\theta)| < s\}$$

and

$$O_h = \{\omega \in \mathbb{C}^n \mid |\omega - \omega_0| < h\}$$

be complex neighborhoods of respectively $\{0\} \times \mathbb{T}^n$ and ω_0 , where $|\cdot|$ stands for the supremum norm of vectors.

For a small parameter $\epsilon \geq 0$, consider a function H , which is bounded and real-analytic on $D_{r,s} \times O_h$, and of the form

$$\begin{cases} H(I, \theta, \omega) = N(I, \omega) + P(I, \theta, \omega), \\ N(I, \omega) = e(\omega) + \omega \cdot I, \quad |P|_{r,s,h} \leq \epsilon, \end{cases} \tag{**}$$

where

$$|P|_{r,s,h} = \sup_{(I, \theta, \omega) \in D_{r,s} \times O_h} |P(I, \theta, \omega)|.$$

The function H should be considered as a real-analytic Hamiltonian on $D_{r,s}$, depending analytically on a parameter $\omega \in O_h$; for a fixed parameter $\omega \in O_h$, when convenient, we will write

$$H_{\omega}(I, \theta) = H(I, \theta, \omega), \quad N_{\omega}(I) = N(I, \omega), \quad P_{\omega}(I, \theta) = P(I, \theta, \omega).$$

Let $B = \{I \in \mathbb{R}^n \mid |I| < r\}$ so that $B \times \mathbb{T}^n$ is the real part of the domain $D_{r,s}$, and $\Phi_0 : \mathbb{T}^n \rightarrow B \times \mathbb{T}^n$ be the map given by $\Phi_0(\theta) = (0, \theta)$. Then $\Phi_0(\mathbb{T}^n)$ is an embedded real-analytic torus in $B \times \mathbb{T}^n$, invariant by the Hamiltonian flow of N_{ω_0} and quasi-periodic with frequency ω_0 . The next theorem states that this quasi-periodic torus will persist, being only slightly deformed, as an invariant torus not for the Hamiltonian flow of H_{ω_0} but for the Hamiltonian flow of $H_{\tilde{\omega}}$, where $\tilde{\omega}$ is a parameter close to ω_0 , provided ϵ is sufficiently small and ω_0 satisfies the Bruno–Rüssmann condition. Here is a more precise statement.

Theorem 2. *Let H be as in (**), with $\Delta^* = \Delta_{\omega_0}^*$ satisfying (2.2). Then there exist positive constants $c_1 \leq 1, c_2 \leq 1, c_3 \leq 1, c_4 \geq 1$ and $c_5 \geq 1$ depending only on n such that if*

$$\epsilon r^{-1} \leq c_1 h \leq c_2 (Q_0 \Psi(Q_0))^{-1}, \tag{2.3}$$

where $Q_0 \geq 1$ is sufficiently large so that

$$Q_0^{-1} + (\ln 2)^{-1} \int_{\Delta(Q_0)}^{+\infty} \frac{dx}{x \Delta^*(x)} \leq c_3 s, \tag{2.4}$$

there exist a real-analytic torus embedding $\Phi_{\omega_0} : \mathbb{T}^n \rightarrow B \times \mathbb{T}^n$ and a vector $\tilde{\omega} \in \mathbb{R}^n$ such that $\Phi_{\omega_0}(\mathbb{T}^n)$ is invariant by the Hamiltonian flow of $H_{\tilde{\omega}}$ and quasi-periodic with frequency ω_0 . Moreover, Φ_{ω_0} is real-analytic on $\mathbb{T}_{s/2}^n = \{\theta \in \mathbb{T}_{\mathbb{C}}^n \mid |\text{Im}(\theta)| < s/2\}$ and we have the estimates

$$|W(\Phi_{\omega_0} - \Phi_0)|_{s/2} \leq c_4 \varepsilon (rh)^{-1}, \quad |\tilde{\omega} - \omega_0| \leq c_5 \varepsilon r^{-1}, \tag{2.5}$$

where $W = \text{diag}(r^{-1}\text{Id}, Q_0^{-1}\text{Id})$.

Theorem 1 follows directly from Theorem 2, introducing the frequencies $\omega = \nabla h(p)$ as independent parameters and compensating the shift of frequency $\tilde{\omega} - \omega_0$ using the non-degeneracy assumption on h . This deduction is classical but for completeness we will repeat the details below, following [18].

3. PROOF OF THEOREM 1 ASSUMING THEOREM 2

In this section, we assume Theorem 2 and we show how it implies Theorem 1.

Proof of Theorem 1. For $p_0 \in D$, we expand h in a small neighborhood of p_0 : writing $p = p_0 + I$ for I close to zero, we get

$$h(p) = h(p_0) + \nabla h(p_0) \cdot I + \int_0^1 (1-t) \nabla^2 h(p_0 + tI) I \cdot I dt.$$

Since $\nabla h : D \rightarrow \Omega$ is a diffeomorphism, instead of p_0 we can use $\omega = \nabla h(p_0)$ as a new variable, and we write

$$h(p) = e(\omega) + \omega \cdot I + P_h(I, \omega)$$

with $e(\omega) = h(\nabla^{-1}(\omega))$ and $P_h(I, \omega) = \int_0^1 (1-t) \nabla^2 h(\nabla^{-1}(\omega) + tI) I \cdot I dt$. Letting $\theta = q$ and

$$P_\epsilon(I, \theta, \omega) = \epsilon f(p, q) = \epsilon f(p_0 + I, \theta) = \epsilon f(\nabla^{-1}(\omega) + I, \theta),$$

we can eventually define

$$H(I, \theta, \omega) = H(p, q) = e(\omega) + \omega \cdot I + P_h(I, \omega) + P_\epsilon(I, \theta, \omega) = e(\omega) + \omega \cdot I + P(I, \theta, \omega).$$

This Hamiltonian is obviously real-analytic in (I, θ, ω) , hence we can fix some small $0 < s < 1$ and $0 < h < 1$ so that P is real-analytic on the complex domain $D_{r,s} \times O_h$ for all sufficiently small $0 < r < 1$. Moreover, as $\Psi = \Psi_{\omega_0}$ satisfies (BR), $\Delta^* = \Delta_{\omega_0}^*$ satisfies (2.2), and choosing $Q_0 = Q_0(s)$ sufficiently large so that (2.4) is satisfied, we may assume, restricting h if necessary, that the second inequality of (2.3) holds true, namely $c_1 h \leq c_2 (Q_0 \Psi(Q_0))^{-1}$. In the same way, we may also assume that the real part of O_h is contained in Ω .

Now since $P(I, \theta, \omega) = P_h(I, \omega) + P_\epsilon(I, \theta, \omega)$, we have

$$|P|_{r,s,h} \leq \varepsilon = Mr^2 + F\epsilon,$$

where

$$M = \sup_{p \in D} |\nabla^2 h(p)|, \quad F = \sup_{(p,q) \in D \times \mathbb{T}^n} |f(p, q)|.$$

Therefore we choose $r = (M^{-1}F\epsilon)^{1/2}$ to have $\varepsilon = 2F\epsilon$, and assuming $\epsilon \leq (4MF)^{-1}c_1^2 h^2$, we have

$$\varepsilon r^{-1} = 2F\epsilon(MF^{-1}\epsilon^{-1})^{1/2} = 2(MF\epsilon)^{1/2} \leq c_1 h,$$

hence the first inequality of (2.3) is satisfied. So Theorem 2 can be applied: there exist a real-analytic torus embedding $\Phi_{\omega_0} : \mathbb{T}^n \rightarrow B \times \mathbb{T}^n$ and a vector $\tilde{\omega} \in \mathbb{R}^n$ such that $\Phi_{\omega_0}(\mathbb{T}^n)$ is invariant by the Hamiltonian flow of $H_{\tilde{\omega}}$ and quasi-periodic with frequency ω_0 . Moreover, Φ_{ω_0} is real-analytic on $\mathbb{T}_{s/2}^n = \{\theta \in \mathbb{T}_{\mathbb{C}}^n \mid |\text{Im}(\theta)| < s/2\}$ and we have the estimates

$$|W(\Phi_{\omega_0} - \Phi_0)|_{s/2} \leq c_4 \varepsilon (rh)^{-1}, \quad |\tilde{\omega} - \omega_0| \leq c_5 \varepsilon r^{-1}, \tag{3.1}$$

where $W = \text{diag}(r^{-1}\text{Id}, Q_0^{-1}\text{Id})$.

Since $\tilde{\omega}$ is real and the real part of O_h is contained in Ω , there exists \tilde{I} close to zero such that $\nabla h(\tilde{I}) = \tilde{\omega}$. Now observe that an orbit $(I(t), \theta(t))$ for the Hamiltonian $H_{\tilde{\omega}}$ corresponds to an orbit $(p(t), q(t)) = (\tilde{I} + I(t), \theta(t))$ for our original Hamiltonian. Hence, if we define $T : B \times \mathbb{T}^n \rightarrow D \times \mathbb{T}^n$ by $T(I, \theta) = (\tilde{I} + I, \theta)$ and

$$\Theta_{\omega_0} = T \circ \Phi_{\omega_0} : \mathbb{T}^n \rightarrow D \times \mathbb{T}^n,$$

then Θ_{ω_0} is a real-analytic torus embedding such that $\Theta_{\omega_0}(\mathbb{T}^n)$ is invariant by the Hamiltonian flow of H and quasi-periodic with frequency ω_0 .

Moreover, as ϵ goes to zero, it follows from (3.1) that Φ_{ω_0} converges uniformly to Φ_0 and $\tilde{\omega}$ converges to ω_0 , hence T converges uniformly to the identity which eventually implies that Θ_{ω_0} converges uniformly to Θ_0 .

4. PROOF OF THEOREM 2

This section is devoted to the proof of Theorem 2, in which we will construct, by an iterative procedure, a vector $\tilde{\omega}$ close to ω_0 and a real-analytic torus embedding $\Phi_{\tilde{\omega}}$ whose image is invariant by the Hamiltonian flow of $H_{\tilde{\omega}}$. We start, in Section 4.1, by recalling the Diophantine result of [4] which will be crucial in our approach. Then, in Section 4.2, we describe an elementary step of our iterative procedure, and finally, in Section 4.3, we will show one can perform infinitely many steps to obtain a convergent scheme.

In this section, for simplicity, we shall write $u < v$ (respectively $u \cdot < v$), if, for some constant $C \geq 1$ which depends only on n and could be made explicit, we have $u \leq Cv$ (respectively $Cu \leq v$).

4.1. Approximation by Rational Vectors

Recall that we have written $\omega_0 = (1, \bar{\omega}_0) \in \mathbb{R}^n$ with $\bar{\omega}_0 \in [-1, 1]^{n-1}$. For a given $Q \geq 1$, it is always possible to find a rational vector $v = (1, p/q) \in \mathbb{Q}^n$, with $p \in \mathbb{Z}^{n-1}$ and $q \in \mathbb{N}$, which is a Q -approximation in the sense that $|q\omega_0 - qv| \leq Q^{-1}$, and for which the denominator q satisfies the upper bound $q \leq Q^{n-1}$: this is essentially the content of Dirichlet’s theorem on simultaneous rational approximations, and it holds true without any assumption on ω_0 . In our situation, since we have assumed that ω_0 is non-resonant, it is not hard to see that there exist not only one, but n linearly independent rational vectors in \mathbb{Q}^n which are Q -approximations. Moreover, one can obtain not only linearly independent vectors, but rational vectors v_1, \dots, v_n of denominators q_1, \dots, q_n such that the associated integer vectors q_1v_1, \dots, q_nv_n form a \mathbb{Z} -basis of \mathbb{Z}^n . However, the upper bound on the corresponding denominators q_1, \dots, q_n is necessarily larger than Q^{n-1} , and is given by a function of Q that we can call here $\Psi' = \Psi'_{\omega_0}$ (see [4] for more precise and general information, but note that in this reference, Ψ' was denoted by Ψ and Ψ , which we defined in (2.1), was denoted by Ψ'). A consequence of the main Diophantine result of [4] is that this function Ψ' is in fact essentially equivalent to the function Ψ .

Proposition 1. *Let $\omega_0 = (1, \bar{\omega}_0) \in \mathbb{R}^n$ be a non-resonant vector, with $\bar{\omega}_0 \in [-1, 1]^{n-1}$. For any $1 \cdot < Q$, there exist n rational vectors v_1, \dots, v_n , of denominators q_1, \dots, q_n , such that q_1v_1, \dots, q_nv_n form a \mathbb{Z} -basis of \mathbb{Z}^n and for $j \in \{1, \dots, n\}$,*

$$|\omega_0 - v_j| \cdot < (q_jQ)^{-1}, \quad 1 \leq q_j \cdot < \Psi(Q).$$

For a proof, we refer to [4], Theorem 2.1 and Proposition 2.3.

Now given a q -rational vector v and a function P defined on $D_{r,s} \times O_h$, we define

$$[P]_v(I, \theta, \omega) = \int_0^1 P(I, \theta + tqv, \omega) dt.$$

Given n rational vectors v_1, \dots, v_n , we let $[P]_{v_1, \dots, v_n} = [\dots [P]_{v_1} \dots]_{v_n}$. Finally we define

$$[P](I, \omega) = \int_{\mathbb{T}^n} P(I, \theta, \omega) d\theta.$$

A consequence of the fact that the vectors q_1v_1, \dots, q_nv_n form a \mathbb{Z} -basis of \mathbb{Z}^n is contained in the following proposition.

Proposition 2. *Let v_1, \dots, v_n be rational vectors, of denominators q_1, \dots, q_n , such that $q_1 v_1, \dots, q_n v_n$ form a \mathbb{Z} -basis of \mathbb{Z}^n , and P a function defined on $D_{r,s} \times O_h$. Then*

$$[P]_{v_1, \dots, v_n} = [P].$$

A proof of this proposition can be found in [7], Corollary 6.

4.2. KAM Step

Now we describe an elementary step of our iterative procedure. To our Hamiltonian H , we will apply transformations of the form

$$\mathcal{F} = (\Phi, \varphi) : (I, \theta, \omega) \mapsto (\Phi(I, \theta, \omega), \varphi(\omega))$$

which consist of a parameter-dependent change of coordinates Φ and a change of parameters φ . Moreover, our change of coordinates will be of the form $\Phi = (U, V)$, where U is affine in I and V is independent of I , and will be symplectic for each fixed parameter ω . It is easy to check that such transformations $\mathcal{F} = (\Phi, \varphi)$ form a group under composition.

From now on, we fix a positive constant η sufficiently small (one could take, for instance, $\eta = 1/66$).

Proposition 3. *Let H be as in (**), with $\omega_0 = (1, \bar{\omega}_0) \in \mathbb{R}^n$ non-resonant, consider $0 < \sigma < s$ and $1 \ll Q$, and assume that*

$$\varepsilon r^{-1} \ll h \ll (Q\Psi(Q))^{-1}, \quad 1 \ll Q\sigma. \tag{4.1}$$

Then there exists a real-analytic transformation

$$\mathcal{F} = (\Phi, \varphi) : D_{\eta r, s-\sigma} \times O_{h/4} \rightarrow D_{r,s} \times O_h,$$

such that $H \circ \mathcal{F} = N^+ + P^+$ with $N^+(I, \omega) = e^+(\omega) + \omega \cdot I$ and $|P^+|_{\eta r, s-\sigma, h/4} \leq \eta\varepsilon/8$. Moreover, we have the estimates

$$|W(\Phi - \text{Id})|_{\eta r, s-\sigma, h} \ll (r\sigma)^{-1}\Psi(Q)\varepsilon, \quad |W(D\Phi - \text{Id})W^{-1}|_{\eta r, s-\sigma, h} \ll (r\sigma)^{-1}\Psi(Q)\varepsilon,$$

$$|\varphi - \text{Id}|_{h/4} \ll \varepsilon r^{-1}, \quad h|D\varphi - \text{Id}|_{h/4} \ll \varepsilon r^{-1},$$

where $W = \text{diag}(r^{-1}\text{Id}, \sigma^{-1}\text{Id})$.

Proof. We divide the proof of the KAM step in six small steps. Except for the last one, everything will be uniform in $\omega \in O_h$, so for simplicity, in the first five steps we will drop the dependence on the parameter $\omega \in O_h$. Let us first notice that (4.1) implies the following five inequalities:

$$h \ll (Q\Psi(Q))^{-1}, \quad \varepsilon \ll \sigma r \Psi(Q)^{-1}, \quad \varepsilon \ll r(Q\Psi(Q))^{-1}, \quad 1 \ll Q\sigma, \quad \varepsilon r^{-1} \ll h. \tag{4.2}$$

1. Affine approximation of P .

Let \bar{P} be the linearization of P in I at $I = 0$; that is,

$$\bar{P}(I, \theta) = P(0, \theta) + \partial_I P(0, \theta) \cdot I.$$

Using Cauchy's estimate, it is easy to see that $|\bar{P}|_{r,s} \ll \varepsilon$. Moreover, using Lemma 1, we have

$$|P - \bar{P}|_{2\eta r, s} \leq (2\eta)^2(1 - 2\eta)^{-1}\varepsilon \leq \eta\varepsilon/16 \tag{4.3}$$

where we used in the second inequality that η is small enough.

2. Rational approximations of ω_0 .

Since ω_0 is non-resonant, given $1 \ll Q$, we can apply Proposition (1): there exist n rational vectors v_1, \dots, v_n , of denominators q_1, \dots, q_n , such that $q_1 v_1, \dots, q_n v_n$ form a \mathbb{Z} -basis of \mathbb{Z}^n and for $j \in \{1, \dots, n\}$,

$$|\omega_0 - v_j| \ll (q_j Q)^{-1}, \quad 1 \leq q_j \leq \Psi(Q).$$

3. Rational approximations of $\omega \in O_h$.

For any $\omega \in O_h$, using the first inequality in (4.2) and $q_j \ll \Psi(Q)$, we have

$$|\omega - v_j| \leq |\omega - \omega_0| + |\omega_0 - v_j| \ll h + (q_j Q)^{-1} \ll (Q\Psi(Q))^{-1} + (q_j Q)^{-1} \ll (q_j Q)^{-1}. \tag{4.4}$$

4. Successive rational averagings.

Let $P_1 = \bar{P}$, and define inductively $P_{j+1} = [P_j]_{v_j}$ for $1 \leq j \leq n$. Let us also define F_j , for $1 \leq j \leq n$, by

$$F_j(I, \theta) = q_j \int_0^1 (P_j - P_{j+1})(I, \theta + tq_j v_j) t dt$$

and N_j by $N_j(I) = e(\omega) + v_j \cdot I$. It is then easy to check, by a simple integration by parts, that the equations

$$\{F_j, N_j\} = P_j - P_{j+1}, \quad 1 \leq j \leq n \tag{4.5}$$

are satisfied, where $\{., .\}$ denotes the usual Poisson bracket. Moreover, we obviously have the estimates

$$|P_j|_{r,s} \leq |\bar{P}|_{r,s} \ll \varepsilon \tag{4.6}$$

and

$$|F_j|_{r,s} \leq q_j |P_j|_{r,s} \ll \Psi(Q)\varepsilon. \tag{4.7}$$

Next, for any $0 \leq j \leq n$, define $r_j = r - (2n)^{-1}jr$ and $s_j = s - n^{-1}j\sigma$. Obviously $r_j > 0$ whereas $s_j > 0$ follows from $\sigma < s$. Using (4.7) and Cauchy's estimate, we have

$$|\partial_\theta F_j|_{r_j, s_j} \leq n(j\sigma)^{-1} |F_j|_{r,s} \ll \sigma^{-1}\Psi(Q)\varepsilon, \quad |\partial_I F_j|_{r_j, s_j} \leq 2n(jr)^{-1} |F_j|_{r,s} \ll r^{-1}\Psi(Q)\varepsilon,$$

and together with the second inequality of (4.1) with a suitable implicit constant, we can ensure that

$$|\partial_\theta F_j|_{r_j, s_j} \leq (2n)^{-1}r, \quad |\partial_I F_j|_{r_j, s_j} \leq n^{-1}\sigma.$$

This implies that for $1 \leq j \leq n$, the time-one map $X_{F_j}^1$ of the Hamiltonian flow of F_j defines a symplectic real-analytic embedding

$$X_{F_j}^1 = (U_j, V_j) : D_{r_j, s_j} \rightarrow D_{r_{j-1}, s_{j-1}}.$$

Moreover, as \bar{P} is affine in I then so are the F_j , for $1 \leq j \leq n$, and therefore U_j is affine in I and V_j is independent of I , and we also have the estimates

$$|U_j - \text{Id}|_{r_j, s_j} \leq |\partial_\theta F_j|_{r_j, s_j} \ll \sigma^{-1}\Psi(Q)\varepsilon, \quad |V_j - \text{Id}|_{r_j, s_j} \leq |\partial_I F_j|_{r_j, s_j} \ll r^{-1}\Psi(Q)\varepsilon. \tag{4.8}$$

The Jacobian of $X_{F_j}^1$ is given by the matrix

$$DX_{F_j}^1 = \begin{pmatrix} \partial_I U_j & \partial_\theta U_j \\ 0 & \partial_\theta V_j \end{pmatrix}.$$

If we define $r_j^+ = r - (4n)^{-1}jr$ and $s_j^+ = s - (2n)^{-1}j\sigma$, then (4.7) and Cauchy's estimate also implies that

$$|\partial_\theta F_j|_{r_j^+, s_j^+} \ll \sigma^{-1}\Psi(Q)\varepsilon, \quad |\partial_I F_j|_{r_j^+, s_j^+} \ll r^{-1}\Psi(Q)\varepsilon,$$

and, as $r_j^+ - r_j = (4n)^{-1}jr$ and $s_j^+ - s_j = (2n)^{-1}j\sigma$, a further Cauchy's estimate proves that

$$|\partial_I U_j - \text{Id}|_{r_j, s_j} \ll (r\sigma)^{-1}\Psi(Q)\varepsilon, \quad |\partial_\theta V_j - \text{Id}|_{r_j, s_j} \ll (r\sigma)^{-1}\Psi(Q)\varepsilon, \quad |\partial_\theta U_j|_{r_j, s_j} \ll \sigma^{-2}\Psi(Q)\varepsilon.$$

The estimates (4.8) and the above estimates can be conveniently written as

$$|W(X_{F_j}^1 - \text{Id})|_{r_j, s_j} \leq (r\sigma)^{-1}\Psi(Q)\varepsilon, \quad |W(DX_{F_j}^1 - \text{Id})W^{-1}|_{r_j, s_j} \leq (r\sigma)^{-1}\Psi(Q)\varepsilon, \quad (4.9)$$

where $W = \text{diag}(r^{-1}\text{Id}, \sigma^{-1}\text{Id})$.

Let $S_j = \omega \cdot I - v_j \cdot I$ so that $N = N_j + S_j$, and let $H_j = N + P_j$. Writing $H_j = N + P_j = N_j + S_j + P_j$ and using the equality (4.5), a standard computation based on Taylor's formula with integral remainder gives

$$H_j \circ X_{F_j}^1 = N + [P_j]_{v_j} + \tilde{P}_j = N + P_{j+1} + \tilde{P}_j$$

with

$$\tilde{P}_j = \int_0^1 \{(1-t)P_{j+1} + tP_j + S_j, F_j\} \circ X_{F_j}^t dt$$

and where $X_{F_j}^t$ is the time- t map of the Hamiltonian flow of F_j . Using (4.4), (4.6), (4.7), the definition of the Poisson bracket and Cauchy's estimate, one easily obtains

$$|\tilde{P}_j|_{r_j, s_j} \leq (\sigma r)^{-1}\Psi(Q)\varepsilon^2 + (Q\sigma)^{-1}\varepsilon \leq (Q\sigma)^{-1}\varepsilon \quad (4.10)$$

where the last inequality follows from the third inequality of (4.2).

5. Change of coordinates

For $1 \leq j \leq n$, let $\Phi_j = X_{F_1}^1 \circ \dots \circ X_{F_j}^1 : D_{r_j, s_j} \rightarrow D_{r, s}$. Since $H_1 = N + \bar{P}$, by a straightforward induction we have

$$H_1 \circ \Phi_j = (N + \bar{P}) \circ \Phi_j = N + P_{j+1} + P_j^+,$$

where P_j^+ is defined inductively by $P_1^+ = \tilde{P}_1$ and $P_{j+1}^+ = \tilde{P}_{j+1} + P_j^+ \circ X_{F_{j+1}}^1$ for $1 \leq j \leq n-1$. Let $\Phi = \Phi_n$, and first note that $\Phi = (U, V)$ where U is affine in I and V is independent of I , since each $X_{F_j}^1$ is of this form. Moreover, $\Phi : D_{r/4, s-\sigma} \rightarrow D_{r, s}$, so in particular $\Phi : D_{\eta r, s-\sigma} \rightarrow D_{r, s}$ as η is small enough, and using (4.9) and a classical telescopic argument, we have the estimates

$$|W(\Phi - \text{Id})|_{\eta r, s-\sigma} \leq (r\sigma)^{-1}\Psi(Q)\varepsilon, \quad |W(D\Phi - \text{Id})W^{-1}|_{\eta r, s-\sigma} \leq (r\sigma)^{-1}\Psi(Q)\varepsilon.$$

Concerning P_n^+ , using (4.10) and the fourth equality of (4.2) with a suitable implicit constant, we can ensure that

$$|P_n^+|_{\eta r, s-\sigma} \leq (Q\sigma)^{-1}\varepsilon \leq \eta\varepsilon/16.$$

Now as $H = N + P = N + \bar{P} + P - \bar{P} = H_1 + P - \bar{P}$, this gives

$$H \circ \Phi = H_1 \circ \Phi + (P - \bar{P}) \circ \Phi = N + P_{n+1} + P_n^+ + (P - \bar{P}) \circ \Phi.$$

We finally set $P^+ = P_n^+ + (P - \bar{P}) \circ \Phi$, and as $P_{n+1} = [\dots [\bar{P}]_{v_1} \dots]_{v_n} = [\bar{P}]_{v_1, \dots, v_n}$, by Proposition 2, $P_{n+1} = [\bar{P}]$, we arrive at

$$H \circ \Phi = N + [\bar{P}] + P^+.$$

Using the second inequality of (4.2) with an appropriate implicit constant, we may assume that the image of Φ actually sends $D_{\eta r, s-\sigma}$ into $D_{2\eta r, s}$, and together with (4.3), we obtain the estimate

$$|P^+|_{\eta r, s-\sigma} \leq |P_n^+|_{\eta r, s-\sigma} + |(P - \bar{P}) \circ \Phi|_{\eta r, s-\sigma} \leq |P_n^+|_{\eta r, s-\sigma} + |P - \bar{P}|_{2\eta r, s-\sigma} \leq \eta\varepsilon/16 + \eta\varepsilon/16 = \eta\varepsilon/8.$$

6. Change of frequencies

As \bar{P} is affine in I , $[\bar{P}]$ is independent of θ and of the form $[\bar{P}](I, \omega) = c(\omega) + \nu(\omega) \cdot I$, with $c(\omega) \in \mathbb{C}$ and $\nu(\omega) \in \mathbb{C}^n$, and therefore

$$(N + [\bar{P}])(I, \omega) = e(\omega) + c(\omega) + (\omega + \nu(\omega)) \cdot I = e^+(\omega) + (\omega + \nu(\omega)) \cdot I.$$

Since $\nu = \partial_T[\bar{P}]$, Cauchy's estimate together with the fifth inequality of (4.2) implies that the map ν satisfy the estimate

$$|\nu|_h \leq \varepsilon r^{-1} \leq h/4.$$

Setting $f(\omega) = \omega + \nu(\omega)$, we can apply Lemma 2 to find a real-analytic inverse $\varphi : O_{h/4} \rightarrow O_h$ to f satisfying the estimates

$$|\varphi - \text{Id}|_{h/4} \leq \varepsilon r^{-1}, \quad h|D\varphi - \text{Id}|_{h/4} \leq \varepsilon r^{-1}.$$

Eventually, we set $N^+ = (N + [\bar{P}]) \circ \varphi$ and $\mathcal{F} = (\Phi, \varphi) : D_{\eta r, s-\sigma} \times O_{h/4} \rightarrow D_{r,s} \times O_h$, and we obtain

$$H \circ \mathcal{F} = N^+ + P^+$$

as wanted, with the desired estimates on \mathcal{F} and on P^+ .

4.3. Iterations and Convergence

Recall that η has been fixed before Proposition 3, and now we define, for $i \in \mathbb{N}$, the following decreasing geometric sequences:

$$\varepsilon_i = (\eta/8)^i \varepsilon, \quad r_i = \eta^i r, \quad h_i = (1/4)^i h.$$

Next, for a constant Q_0 to be chosen below, we define Δ_i and Q_i , $i \in \mathbb{N}$, by

$$\Delta_i = 2^i \Delta(Q_0), \quad Q_i = \Delta^*(\Delta_i) = \sup\{Q \geq 1 \mid \Delta(Q) \leq \Delta_i\},$$

and then we define σ_i , $i \in \mathbb{N}$, by

$$\sigma_i = CQ_i^{-1},$$

where $C \geq 1$ is a constant, depending only n , sufficiently large so that the last part of (4.1) is satisfied for $\sigma = \sigma_i$ and $Q = Q_i$. Finally, we define s_i , $i \in \mathbb{N}$, by $s_0 = s$ and $s_{i+1} = s_i - \sigma_i$ for $i \in \mathbb{N}$.

We claim that, assuming Δ^* satisfies (2.2), we can choose Q_0 sufficiently large so that

$$\lim_{i \rightarrow +\infty} s_i \geq s/2 \iff \sum_{i \in \mathbb{N}} \sigma_i \leq s/2.$$

Indeed, since $Q_i = \Delta^*(\Delta_i) = \Delta^*(2^i \Delta(Q_0))$, we have

$$\sum_{i \geq 1} Q_i^{-1} = \sum_{i \geq 1} \frac{1}{\Delta^*(2^i \Delta(Q_0))} \leq \int_0^{+\infty} \frac{dy}{\Delta^*(2^y \Delta(Q_0))} = (\ln 2)^{-1} \int_{\Delta(Q_0)}^{+\infty} \frac{dx}{x \Delta^*(x)} < +\infty,$$

where the last integral converges since Δ^* satisfies (2.2). Now as $\sigma_i = CQ_i^{-1}$, we have

$$\sum_{i \in \mathbb{N}} \sigma_i = C \sum_{i \in \mathbb{N}} Q_i^{-1} = CQ_0^{-1} + C \sum_{i \geq 1} Q_i^{-1} \leq CQ_0^{-1} + C(\ln 2)^{-1} \int_{\Delta(Q_0)}^{+\infty} \frac{dx}{x \Delta^*(x)} \leq s/2$$

provided we choose Q_0 sufficiently large in order to have

$$Q_0^{-1} + (\ln 2)^{-1} \int_{\Delta(Q_0)}^{+\infty} \frac{dx}{x \Delta^*(x)} \leq (2C)^{-1} s. \tag{4.11}$$

Proposition 4. *Let H be as in (**), with $\Delta^* = \Delta_{\omega_0}^*$ satisfying (2.2), and fix Q_0 sufficiently large so that (4.11) is satisfied. Assume that*

$$\varepsilon r^{-1} \leq h \leq \Delta(Q_0)^{-1}. \tag{4.12}$$

Then, for each $i \in \mathbb{N}$, there exists a real-analytic transformation

$$\mathcal{F}^i : D_{r_i, s_i} \times O_{h_i} \rightarrow D_{r,s} \times O_h,$$

such that $H \circ \mathcal{F}^i = N^i + P^i$ with $N^i(I, \omega) = e^i(\omega) + \omega \cdot I$ and $|P^i|_{r_i, s_i, h_i} \leq \varepsilon_i$. Moreover, we have the estimates

$$|\bar{W}_0(\mathcal{F}^{i+1} - \mathcal{F}^i)|_{r_{i+1}, s_{i+1}, h_{i+1}} < \varepsilon_i (r_i h_i)^{-1}$$

where $\bar{W}_0 = \text{diag}(r_0^{-1} \text{Id}, \sigma_0^{-1} \text{Id}, h_0^{-1} \text{Id})$.

Proof. For $i = 0$, we let \mathcal{F}^0 be the identity and there is nothing to prove. The general case follows by an easy induction. Indeed, assume that the statement holds true for some $i \in \mathbb{N}$, and let $H_i = H \circ \mathcal{F}^i = N^i + P^i$, defined on $D_{r_i, s_i} \times O_{h_i}$. We want to apply Proposition (3) to this Hamiltonian, with $\varepsilon = \varepsilon_i$, $r = r_i$, $s = s_i$, $h = h_i$, $\sigma = \sigma_i$ and $Q = Q_i$. First, we need to check that $0 < \sigma_i < s_i$ and $1 < Q_i$. The first condition is equivalent to $\sum_{l=0}^i \sigma_l < s$, whereas the second condition is implied by $1 < Q_0$, and it is easy to see that both conditions are satisfied by the choice of Q_0 in (4.11), as $s < 1$. Then we need to check that the conditions

$$\varepsilon_i r_i^{-1} < h_i < \Delta(Q_i)^{-1}, \quad 1 < Q_i \sigma_i$$

are satisfied. Since

$$\Delta(Q_i) = \Delta(\Delta^*(\Delta_i)) \leq \Delta_i, \tag{4.13}$$

it is sufficient to check the conditions

$$\varepsilon_i r_i^{-1} < h_i < \Delta_i^{-1}, \quad 1 < Q_i \sigma_i. \tag{4.14}$$

The second condition of (4.14) is satisfied, for all $i \in \mathbb{N}$, simply by the choice of the constant in the definition of σ_i . As for the first condition of (4.14), it is satisfied for $i = 0$ by (4.12), and since the sequences $\varepsilon_i r_i^{-1}$, h_i and Δ_i^{-1} decrease at a geometric rate with respective ratio $1/8$, $1/4$ and $1/2$, the first condition of (4.14) is therefore satisfied for any $i \in \mathbb{N}$. Hence Proposition (3) can be applied: there exists a real-analytic transformation

$$\mathcal{F}_i = (\Phi_i, \varphi_i) : D_{r_{i+1}, s_{i+1}} \times O_{h_{i+1}} \rightarrow D_{r_i, s_i} \times O_{h_i},$$

such that $H_i \circ \mathcal{F}_i = N^{i,+} + P^{i,+}$ with $N^{i,+}(I, \omega) = e^{i,+}(\omega) + \omega \cdot I$ and $|P^{i,+}|_{r_{i+1}, s_{i+1}, h_{i+1}} \leq \varepsilon_{i+1}$. Moreover, we have the estimates

$$\begin{aligned} |W_i(\Phi_i - \text{Id})|_{r_{i+1}, s_{i+1}, h_{i+1}} &< (r_i \sigma_i)^{-1} \Psi(Q_i) \varepsilon_i, & |W_i(D\Phi - \text{Id})W_i^{-1}|_{r_{i+1}, s_{i+1}, h_{i+1}} &< (r_i \sigma_i)^{-1} \Psi(Q_i) \varepsilon_i, \\ |\varphi_i - \text{Id}|_{h_{i+1}} &< \varepsilon_i r_i^{-1}, & h_i |D\varphi_i - \text{Id}|_{h_{i+1}} &< \varepsilon_i r_i^{-1}, \end{aligned}$$

where $W_i = \text{diag}(r_i^{-1} \text{Id}, \sigma_i^{-1} \text{Id})$.

We just need to set $\mathcal{F}^{i+1} = \mathcal{F}^i \circ \mathcal{F}_i$, $N^{i+1} = N^{i,+}$ and $P^{i+1} = P^{i,+}$ to have

$$H \circ \mathcal{F}^{i+1} = H_i \circ \mathcal{F}_i = N^{i+1} + P^{i+1}$$

with $N^{i+1}(I, \omega) = e^{i+1}(\omega) + \omega \cdot I$ and $|P^{i+1}|_{r_{i+1}, s_{i+1}, h_{i+1}} \leq \varepsilon_{i+1}$.

It remains to estimate $\mathcal{F}^{i+1} - \mathcal{F}^i$. Setting $\bar{W}_i = \text{diag}(r_i^{-1} \text{Id}, \sigma_i^{-1} \text{Id}, h_i^{-1} \text{Id})$, the estimates above implies

$$|\bar{W}_i(\mathcal{F}_i - \text{Id})|_{r_{i+1}, s_{i+1}, h_{i+1}} < \max\{(r_i \sigma_i)^{-1} \Psi(Q_i) \varepsilon_i, \varepsilon_i (r_i h_i)^{-1}\} < \varepsilon_i (r_i h_i)^{-1} \tag{4.15}$$

since $\sigma_i^{-1} \Psi(Q_i) = Q_i \Psi(Q_i) < h_i^{-1}$, and similarly

$$|\bar{W}_i(D\mathcal{F}_i - \text{Id})\bar{W}_i^{-1}|_{r_{i+1}, s_{i+1}, h_{i+1}} < \varepsilon_i (r_i h_i)^{-1}. \tag{4.16}$$

Using a classical telescoping argument, the fact that $|\bar{W}_{i-1} \bar{W}_i| \leq 1$ and the estimate (4.16), we get

$$|\bar{W}_0 D\mathcal{F}^i \bar{W}_i^{-1}|_{r_{i+1}, s_{i+1}, h_{i+1}} < \prod_{l=0}^i (1 + \varepsilon_l (r_l h_l)^{-1}) < 1 \tag{4.17}$$

as $\varepsilon_l (r_l h_l)^{-1}$ decreases geometrically. Then, from the mean value theorem, we have

$$|\bar{W}_0(\mathcal{F}^{i+1} - \mathcal{F}^i)|_{r_{i+1}, s_{i+1}, h_{i+1}} = |\bar{W}_0(\mathcal{F}^i \circ \mathcal{F}_i - \mathcal{F}^i)|_{r_{i+1}, s_{i+1}, h_{i+1}}$$

$$\leq |\bar{W}_0 D\mathcal{F}^i \bar{W}_i^{-1}|_{r_{i+1}, s_{i+1}, h_{i+1}} |\bar{W}_i(\mathcal{F}_i - \text{Id})|_{r_{i+1}, s_{i+1}, h_{i+1}},$$

and this estimate, together with (4.15) and (4.17), implies

$$|\bar{W}_0(\mathcal{F}^{i+1} - \mathcal{F}^i)|_{r_{i+1}, s_{i+1}, h_{i+1}} < \varepsilon_i (r_i h_i)^{-1}$$

which is the required estimate.

We can finally conclude the proof of Theorem 2.

Proof of Theorem 2. Recall that we are given ε, r, s, h and that we fixed η small enough to define the sequences ε_i, r_i, h_i , and then we chose $Q_0 \geq 1$ satisfying (4.11) to define the sequences Δ_i, Q_i, σ_i and s_i . Moreover, we have

$$\lim_{i \rightarrow +\infty} \varepsilon_i = \lim_{i \rightarrow +\infty} r_i = \lim_{i \rightarrow +\infty} h_i = 0, \quad \lim_{i \rightarrow +\infty} s_i \geq s/2. \tag{4.18}$$

Now the condition (2.3) implies the condition (4.12) and Proposition 4 can be applied: for each $i \in \mathbb{N}$, there exists a real-analytic transformation

$$\mathcal{F}^i : D_{r_i, s_i} \times O_{h_i} \rightarrow D_{r, s} \times O_h,$$

such that $H \circ \mathcal{F}^i = N^i + P^i$ with $N^i(I, \omega) = e^i(\omega) + \omega \cdot I$ and $|P_i|_{r_i, s_i, h_i} \leq \varepsilon_i$. Moreover, we have the estimates

$$|\bar{W}_0(\mathcal{F}^{i+1} - \mathcal{F}^i)|_{r_{i+1}, s_{i+1}, h_{i+1}} < \varepsilon_i (r_i h_i)^{-1}. \tag{4.19}$$

As $\varepsilon_i (r_i h_i)^{-1}$ decreases geometrically, these estimates and (4.18) show that the transformations $\mathcal{F}^i = (\Phi^i, \varphi^i)$ converge uniformly, as i goes to infinity, to a map

$$\mathcal{F} = (\Phi_{\omega_0}, \varphi) : \{0\} \times \mathbb{T}_{s/2}^n \times \{\omega_0\} = D_{0, s/2} \times O_0 \rightarrow D_{r, s} \times O_h$$

which consists of a real map $\varphi : \{\omega_0\} = O_0 \rightarrow O_h$ and a real-analytic embedding

$$\Phi_{\omega_0} : \mathbb{T}_{s/2}^n \rightarrow D_{r, s},$$

where, for simplicity, we identified $D_{0, s/2} = \{0\} \times \mathbb{T}_{s/2}^n$ with $\mathbb{T}_{s/2}^n$. Note that by reality, $\varphi(\omega_0) = \tilde{\omega} \in \mathbb{R}^n$ and $\Phi_{\omega_0}(\mathbb{T}^n) \subseteq B \times \mathbb{T}^n$. Moreover, from the estimate (4.19) and a usual telescopic argument,

$$|\bar{W}_0(\mathcal{F} - \text{Id})|_{s/2} < \varepsilon (rh)^{-1}$$

from which one deduces that

$$|W(\Phi_{\omega_0} - \Phi_0)|_{s/2} < \varepsilon (rh)^{-1}, \quad |\varphi(\omega_0) - \omega_0| < \varepsilon r^{-1},$$

where $W = \text{diag}(r^{-1}\text{Id}, Q_0^{-1}\text{Id})$, since $r_0 = r$ and $\sigma_0 = CQ_0^{-1}$.

To conclude, fix $i \in \mathbb{N}$ and $\omega \in O_{h_i}$. Then, denoting $X_{H_\omega}, X_{N_\omega^i}$ and $X_{P_\omega^i}$ the Hamiltonian vector fields associated to H_ω, N_ω^i and P_ω^i , we have

$$\begin{aligned} X_{H_{\varphi_i(\omega)}} \circ \Phi_\omega^i - D\Phi_\omega^i X_{N_\omega^i} &= D\Phi_\omega^i \left((\Phi_\omega^i)^* X_{H_{\varphi_i(\omega)}} - X_{N_\omega^i} \right) \\ &= D\Phi_\omega^i \left(X_{H_{\varphi_i(\omega)} \circ \Phi_\omega^i} - X_{N_\omega^i} \right) \\ &= D\Phi_\omega^i X_{P_\omega^i}, \end{aligned} \tag{4.20}$$

where we used the fact that Φ_ω^i is symplectic in the second equality, and the relation

$$H_{\varphi_i(\omega)} \circ \Phi_\omega^i - N_\omega^i = P_\omega^i$$

in the last equality. Using the inequality $|P_i|_{r_i, s_i, h_i} \leq \varepsilon_i$ together with Cauchy's estimate, one obtains that, as i goes to infinity, $X_{P_\omega^i}$ converges to zero uniformly on $D_{0, s/2}$, whereas $D\Phi_\omega^i$ is uniformly bounded by the estimate (4.17). Hence the right-hand side of (4.20) converges uniformly to zero, and so does the left-hand side: at the limit we obtain the equality

$$X_{H_{\tilde{\omega}}} \circ \Phi_{\omega_0} = D\Phi_{\omega_0} \cdot X_{\omega_0}$$

since $X_{N_\omega^i}$ converges to the constant vector field $X_{\omega_0} = \omega_0$ on \mathbb{T}^n . From this equality it follows that the embedded torus $\Phi_{\omega_0}(\mathbb{T}^n)$ is invariant by the Hamiltonian flow of $H_{\tilde{\omega}}$ and quasi-periodic with frequency ω_0 , and this ends the proof.

APPENDIX. TECHNICAL LEMMAS

In this appendix, we state two technical lemmas that were used in the proof of Proposition 3. The first one deals with the estimate of the remainder of the Taylor's expansion at order one of an analytic function.

Lemma 1. *Let P be an analytic function defined on $D_{r,s}$ and*

$$\bar{P}(I, \theta) = P(0, \theta) + \partial_I P(0, \theta) \cdot I.$$

Then, for any $0 < c < 1$, we have the estimate

$$|P - \bar{P}|_{c r, s} \leq c^2 (1 - c)^{-1} |P|_{r, s}.$$

For a proof (of a more general statement), we refer to [1], Lemma A.5.

Then we need a quantitative version of the implicit function theorem for a real-analytic map.

Lemma 2. *Let $f : O_h \rightarrow \mathbb{C}^n$ be a real-analytic map satisfying $|f - \text{Id}|_h \leq \delta \leq h/4$. Then f has a real-analytic inverse $\varphi : O_{h/4} \rightarrow O_h$ which satisfies*

$$|\varphi - \text{Id}|_{h/4} \leq \delta, \quad h/4 |D\varphi - \text{Id}|_{h/4} \leq \delta.$$

For a proof, we refer to [18], Lemma A.3.

ACKNOWLEDGMENTS

The first author would like to thank the organizers of the conference “Geometry, Dynamics, Integrable Systems — GDIS 2013” in Izhevsk, Russia, for the invitation to give a talk on this topic, and IMPA, Rio de Janeiro, Brazil, for its hospitality during the preparation of this work. The second author is partially supported by Agence Nationale de la Recherche (project HAMOT, ref. ANR 2010 BLAN-0115), and thanks Michel Waldschmidt for his advice.

REFERENCES

1. Albrecht, J., An Analytic KAM-Theorem, arXiv:0705.0275 (2 May 2007).
2. Arnol'd, V.I., Proof of a Theorem of A.N.Kolmogorov on the Invariance of Quasi-Periodic Motions under Small Perturbations of the Hamiltonian, *Russian Math. Surveys*, 1963, vol. 18, no. 5, pp. 9–36; see also: *Uspekhi Mat. Nauk*, 1963, vol. 18, no. 5, pp. 13–40 (Russian).
3. Arnol'd, V.I., Small Denominators and Problems of Stability of Motion in Classical and Celestial Mechanics, *Russian Math. Surveys*, 1963, vol. 18, no. 6, pp. 85–191; see also: *Uspekhi Mat. Nauk*, 1963, vol. 18, no. 6(114), pp. 91–192 (Russian).
4. Bounemoura, A. and Fischler, S., A Diophantine Duality Applied to the KAM and Nekhoroshev Theorems, *Math. Z.*, 2013, vol. 275, no. 3, pp. 1135–1167.
5. Broer, H.W., Huitema, G.B., and Sevryuk, M.B., *Quasi-Periodic Motions in Families of Dynamical Systems: Order Amidst Chaos*, Lecture Notes in Math., vol. 1645, Berlin: Springer, 1996.
6. Bost, J.-B., Tores invariants des systèmes dynamiques hamiltoniens, *Séminaire N. Bourbaki*, 1984/1985, vol. 27, exp. no. 639, pp. 113–157.
7. Bounemoura, A., Normal Forms, Stability and Splitting of Invariant Manifolds: 1. Gevrey Hamiltonians, *Regul. Chaotic Dyn.*, 2013, vol. 18, no. 3, pp. 237–260.
8. Brjuno, A.D., Analytic Form of Differential Equations: 1, *Trans. Moscow Math. Soc.*, 1971, vol. 25, pp. 131–288; see also: *Tr. Mosk. Mat. Obs.*, 1971, vol. 25, pp. 119–262 (Russian).
9. Brjuno, A.D., Analytic Form of Differential Equations: 2, *Trans. Moscow Math. Soc.*, 1972, vol. 26, pp. 199–239; see also: *Tr. Mosk. Mat. Obs.*, 1972, vol. 26, pp. 199–239 (Russian).
10. Eliasson, L.H., Absolutely Convergent Series Expansions for Quasi Periodic Motions, *Math. Phys. Electron. J.*, 1996, vol. 2, Paper 4.
11. Féjóz, J., A Proof of the Invariant Torus Theorem of Kolmogorov, *Regul. Chaotic Dyn.*, 2012, vol. 17, no. 1, pp. 1–5.
12. Khanin, K., Lopes Dias, J., and Marklof, J., Renormalization of Multidimensional Hamiltonian Flows, *Nonlinearity*, 2006, vol. 19, no. 12, pp. 2727–2753.
13. Khanin, K., Lopes Dias, J., and Marklof, J., Multidimensional Continued Fractions, Dynamical Renormalization and KAM Theory, *Comm. Math. Phys.*, 2007, vol. 270, no. 1, pp. 197–231.

14. Kolmogorov, A. N., Preservation of Conditionally Periodic Movements with Small Change in the Hamilton Function, in *Stochastic Behaviour in Classical and Quantum Hamiltonian Systems (Volta Memorial Conference, Como, 1977)*, G. Casati, J. Ford (Eds.), Lect. Notes Phys. Monogr., vol. 93, Berlin: Springer, 1979, pp. 51–56; see also: *Dokl. Akad. Nauk SSSR (N. S.)*, 1954, vol. 98, pp. 527–530 (Russian).
15. Levi, M. and Moser, J., A Lagrangian Proof of the Invariant Curve Theorem for Twist Mappings, in *Smooth Ergodic Theory and Its Applications (Univ. of Washington, Seattle, July 26–August 13, 1999)*, A. B. Katok et al. (Eds.), Proc. Sympos. Pure Math., vol. 69, Providence, R.I.: AMS, 2001, pp. 733–746.
16. Moser, J., On Invariant Curves of Area-Preserving Mappings of an Annulus, *Nach. Akad. Wiss. Göttingen, Math. Phys. Kl. II*, 1962, vol. 1962, no. 1, pp. 1–20.
17. Moser, J., Convergent Series Expansions for Quasi-Periodic Motions, *Math. Ann.*, 1967, vol. 169, pp. 136–176.
18. Pöschel, J., A Lecture on the Classical KAM Theorem, in *Smooth Ergodic Theory and Its Applications (Univ. of Washington, Seattle, July 26–August 13, 1999)*, A. B. Katok et al. (Eds.), Proc. Sympos. Pure Math., vol. 69, Providence, R.I.: AMS, 2001, pp. 707–732.
19. Rüssmann, H., On Optimal Estimates for the Solutions of Linear Partial Differential Equations of First Order with Constant Coefficients on the Torus, in *Dynamical Systems, Theory and Applications (Rencontres, Battelle Res. Inst., Seattle, Wash., 1974)*, J. Moser (Ed.), Lecture Notes in Phys., vol. 38, Berlin: Springer, 1975, pp. 598–624.
20. Rüssmann, H., Nondegeneracy in the Perturbation Theory of Integrable Dynamical Systems, in *Number Theory and Dynamical Systems: Papers from the Meeting Held at the Univ. of York (York, March 30–April 15, 1987)*, M. M. Dodson, J. A. G. Vickers (Eds.), London Math. Soc. Lecture Note Ser., col. 134, Cambridge: Cambridge Univ. Press, 1989, pp. 5–18.
21. Rüssmann, H., On the Frequencies of Quasi-Periodic Solutions of Analytic Nearly Integrable Hamiltonian Systems, in *Seminar on Dynamical Systems (St. Petersburg, 1991)*, S. Kuksin et al. (Eds.), Progr. Nonlinear Differential Equations Appl., vol. 12, Basel: Birkhäuser, 1994, pp. 160–183.
22. Rüssmann, H., Invariant Tori in Non-Degenerate Nearly Integrable Hamiltonian Systems, *Regul. Chaotic Dyn.*, 2001, vol. 6, no. 2, pp. 119–204.
23. Rüssmann, H., KAM Iteration with Nearly Infinitely Small Steps in Dynamical Systems of Polynomial Character, *Discrete Contin. Dyn. Syst. Ser. S*, 2010, vol. 3, no. 4, pp. 683–718.
24. Sevryuk, M. B., KAM-Stable Hamiltonians, *J. Dynam. Control Systems*, 1995, vol. 1, no. 3, pp. 351–366.
25. Siegel, C. L., Iteration of Analytic Functions, *Ann. of Math. (2)*, 1942, vol. 43, no. 4, pp. 607–612.
26. Salamon, D. and Zehnder, E., KAM Theory in Configuration Space, *Comment. Math. Helv.*, 1989, vol. 64, no. 1, pp. 84–132.
27. Yoccoz, J.-Ch., *Petits diviseurs en dimension 1*, Astérisque, vol. 231, Paris: Soc. Math. de France, 1995.
28. Yoccoz, J.-Ch., Analytic Linearization of Circle Diffeomorphisms, in *Dynamical Systems and Small Divisors (Cetraro, Italy, June 13–20, 1998)*, Lecture Notes in Math., vol. 1784, Berlin: Springer, 2002, pp. 125–173.
29. Zehnder, E., Generalized Implicit Function Theorems with Applications to Some Small Divisor Problems: 1, *Comm. Pure Appl. Math.*, 1975, vol. 28, no. 1, pp. 91–140.
30. Zehnder, E., Generalized Implicit Function Theorems with Applications to Some Small Divisor Problems: 2, *Comm. Pure Appl. Math.*, 1976, vol. 29, no. 1, pp. 49–111.