HAL open science

Nekhoroshev estimates for steep real-analytic elliptic equilibrium points

Abed Bounemoura, Bassam Fayad, Laurent Niederman

▶ **To cite this version:**

Abed Bounemoura, Bassam Fayad, Laurent Niederman. Nekhoroshev estimates for steep real-analytic elliptic equilibrium points. Nonlinearity, 2020, 10.1088/1361-6544/ab4c89 . hal-02074689

HAL Id: hal-02074689

https://hal.science/hal-02074689v1

Submitted on 20 Mar 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Nekhoroshev estimates for steep real-analytic elliptic equilibrium points

Abed Bounemoura* and Bassam Fayad† and Laurent Niederman‡

March 20, 2019

Abstract

We prove that steep real-analytic elliptic equilibrium points are exponentially stable, generalizing results which were known only under a convexity assumption. This proves the general case of a conjecture of Nekhoroshev. This result is also an important step in our proof that generically, both in a topological and measure-theoretical sense, equilibrium points are super-exponentially stable.

1 Introduction

Let $n \geq 1$ be an integer, $D \subset \mathbb{R}^n$ an open bounded convex domain and $\mathbb{T}^n := \mathbb{R}^n/\mathbb{Z}^n$. Consider a smooth Hamiltonian function H defined on the domain $\mathbb{T}^n \times D$ of the form

$$H(\theta, I) = h(I) + \varepsilon f(\theta, I), \quad \varepsilon \geq 0, \quad (\theta, I) = (\theta_1, \ldots, \theta_n, I_1, \ldots, I_n) \in \mathbb{T}^n \times D,$$

and its associated Hamiltonian system

$$\begin{cases} \dot{\theta}_i(t) = \partial_{I_i} H(\theta(t), I(t)) = \partial_{I_i} h(I(t)) + \varepsilon \partial_{I_i} f(\theta(t), I(t)), \\ \dot{I}_i(t) = -\partial_{\theta_i} H(\theta(t), I(t)) = -\varepsilon \partial_{\theta_i} f(\theta(t), I(t)) \end{cases} \qquad 1 \leq i \leq n.$$

For $\varepsilon = 0$, the system is stable in the sense that the action variables $I(t)$ of all solutions are constant, and all solutions are quasi-periodic. Now for $\varepsilon \neq 0$ but sufficiently small, a fundamental result of Nekhoroshev states that if the system is real-analytic and the integrable part h satisfies a steepness condition (this condition will be defined later), then the action variables $I(t)$ of all solutions are almost constant for an interval of time which is exponentially long with respect to the inverse of the perturbation. More precisely, the following estimates hold true along all solutions:

$$\sum_{i=1}^{n} |I_i(t) - I_i(0)| \leq R_0 \varepsilon^b, \quad |t| \leq T_0 \exp\left(c \left(\frac{1}{\varepsilon}\right)^a\right) \tag{1.1}$$

for some positive constants R_0, T_0, c, a and b. We refer to [Nek77] and [Nek79]. In the special case where h is quasi-convex, improved values for the exponents a and b in (1.1) were obtained

*CNRS-PSL (Université Paris-Dauphine and Observatoire de Paris), abedbou@gmail.com
†CNRS - IMJ-PRG and Centro Ennio De Giorgi, bassam.fayad@imj-prg.fr
‡Laboratoire Mathématiques d'Orsay & IMCCE/ASD, Laurent.Niederman@math.u-psud.fr

independently by Lochak on the one hand (see [Loc92], [LN92],[LNN94]) who introduced a novel approach using periodic orbits, and by Pöschel ([Pös93]) on the other hand, who improved Nekhoroshev original construction. The method of periodic approximation have been extended to the steep case in [BN12] following a strategy first proposed in [Nie07] but without good values for the exponents; improved values for the exponents in the general steep case were later obtained by Guzzo-Chierchia-Benettin in [GCB16].

The original work of Nekhoroshev left open the question of exponential stability in the neighborhood of an elliptic equilibrium point. This corresponds to a Hamiltonian H defined in a neighborhood of the origin in $\mathbb{R}^{2n}$ whose expansion is, in Cartesian coordinates $(x, y) \in \mathbb{R}^{2n}$, given by

$$H(x, y) = \sum_{j=1}^{n} \alpha_j I_j(x, y) + O_3(x, y), \quad I_j(x, y) = (x_j^2 + y_j^2)/2.$$

It is well-known ([HZ94]) that if α is non-resonant up to order K, that is

$$k \in \mathbb{Z}^n, \quad 0 < |k_1| + \cdots + |k_n| \leq K \Longrightarrow k \cdot \alpha \neq 0$$

then there exists a symplectic transformation Φ^K, well-defined in a smaller neighborhood of the origin, such that

$$H \circ \Phi^K(x, y) = \alpha \cdot I(x, y) + h^m(I(x, y)) + O_{K+1}(x, y)$$

where h^m is a polynomial of degree $m = [K/2]$ (the integer part of $K/2$) in n variables, with vanishing constant and linear terms. This transformation, usually called a Birkhoff normal form up to order K, allows us to transform our perturbation problem for a linear system to a perturbation problem for a non-linear system, and make the results of Nekhoroshev potentially applicable. Yet the perturbation problem here is singular since the distance to the equilibrium point represent the size of the perturbation, and one cannot introduce action-angle coordinates; this was the reason why Nekhoroshev could only conjecture the result. Working directly with Cartesian coordinates, the conjecture was solved in the convex case independently by Niederman ([Nie98]) on the one hand, using Lochak method of periodic approximation, and by Fasso-Guzzo-Benettin ([FGB98],[GFB98]) on the other hand, using the Nekhoroshev-Pöschel construction. Still using the method of periodic orbits, Pöschel simplified the proof in [Pös99] and made it clear that the period orbit method is intrinsic and do not rely on a choice of coordinates (as opposed to [GFB98] for instance, which is an improvement of [FGB98] based on a careful choice of Cartesian coordinates "close" to singularities and action-angle coordinates "far" from singularities). The purpose of this paper is to prove the general case of the conjecture of Nekhoroshev, using the extension of the method of periodic approximation introduced in [BN12] and [Bou11]; namely for a real-analytic Hamiltonian of the form

$$H(x, y) = h(I(x, y)) + f(x, y)$$

provided that h is steep and f is small enough, the estimates (1.1) hold true, if we let $I(t) = I(x(t), y(y))$. We refer to Theorem A below for a proper statement. For the sake of completeness, our proof is completely constructive with explicit expressions of the constants involved in our reasonings.

This work was initially developed because it is a key ingredient to prove the result of [BFN19] where we show that generically, both in a topological and measure-theoretical sense,

any solution starting sufficiently close to the equilibrium point remains close to it for an interval of time which is doubly exponentially large with respect to the inverse of the distance to an equilibrium point. From a more practical point of view, the theorem developed in the present paper admits useful applications in celestial mechanics where elliptic equilibria appear in old and important problems. In [BFG98], the stability of Lagrange equilibrium points in the restricted three body problem is studied at the light of Nekhoroshev's theory and there exists a set of masses which leads to a Birkhoff normal form which is not convex at order 4 but steep at order 6 with the so-called 3-jet non degeneracy condition introduced by Nekhoroshev ([Nek77]). This property yields exponential stability of Lagrange points for all but a few values of the masses provided a theorem about Nekhoroshev-stability of steep equilibrium point is proved. This latter result is required and announced in [BFG98] for a futur work but there is no published paper on this point up to our knowledge. Another occurrence of the same phenomenon with a non-convex but steep Birkhoff normal form appears in [Pin13] where the author study the secular planar planetary three body problem and the stability theorem proved in the present paper yields a complete proof of Nekhoroshev stability for nearly circular and coplanar trajectories in the planar planetary three body problem.

2 Main result

To state precisely our result, we introduce some notation.

- For vectors in $\mathbb{C}^{2n}$, $\|\,.\,\|$ denotes the norm defined as

$$\|z\| := \max_{1\leq j\leq n} \sqrt{|z_j|^2 + |z_{n+j}|^2}, \quad z = (z_1, \ldots, z_n, z_{n+1}, \ldots, z_{2n}) \tag{2.1}$$

  and for vectors in $\mathbb{C}^n$, $\|\,.\,\|$ denotes the usual Euclidean norm

$$\|I\| := \sqrt{|I_1|^2 + \cdots + |I_n|^2}, \quad I = (I_1, \ldots, I_n,). \tag{2.2}$$

  It will be more convenient to use these different norms for vectors in $\mathbb{C}^{2n}$ or in $\mathbb{C}^n$, and we hope that this abuse of notations will not confuse the reader. For vectors in $\mathbb{C}^n$, it will be convenient to also use the sup norm $|\,.\,|$ defined as

$$|I| := \max\{|I_1|, \ldots, |I_n|\}, \quad I = (I_1, \ldots, I_n). \tag{2.3}$$

  This norm allows an easier comparison between $z \in \mathbb{C}^{2n}$ and

$$I(z) = (I_1(z), \ldots, I_n(z)) \in \mathbb{C}^n, \quad I_j(z) = (z_j^2 + z_{n+j}^2)/2;$$

  indeed, we have $|I(z)| \leq \|z\|^2/2$ and the equality holds true if $z \in \mathbb{R}^{2n}$. Given $R > 0$, we denote by

$$\mathcal{B}_R := \{z \in \mathbb{C}^{2n} \mid \|z\| < R\} \tag{2.4}$$

  and the associated real ball will be denoted by $B_R := \mathcal{B}_R \cap \mathbb{R}^{2n}$.

- Given $r > 0$, we define the domain $\mathcal{D}_r$ to be the open ball centered at the origin in $\mathbb{C}^n$ of radius $r^2/2$ with respect to the norm $|\,.\,|$:

$$\mathcal{D}_r := \{I \in \mathbb{C}^n \mid |I| < r^2/2\}$$

  and we let $D_r := \mathcal{D}_r \cap \mathbb{R}^n$. This choice is motivated by the fact that if $I : z \in \mathbb{C}^{2n} \mapsto I(z) \in \mathbb{C}^n$, then $I(\mathcal{B}_r) \subseteq \mathcal{D}_r$ and $I(B_r) = D_r \cap \mathbb{R}^n_+$, where $\mathcal{B}_r$ and B_r have been defined in (2.4).

- We define $\| \, . \, \|_r$ to be the sup norm for functions defined on $\mathcal{B}_r$ or on $\mathcal{D}_r$. Extending the norm $\| \, . \, \|$ initially defined for vectors in $\mathbb{C}^n$ and $\mathbb{C}^{2n}$ (respectively in (2.1) and in (2.2)) to tensors in $\mathbb{C}^n$ and $\mathbb{C}^{2n}$, we extend the sup norm $\| \, . \, \|_r$ for tensor-valued functions defined on $\mathcal{B}_r$ or on $\mathcal{D}_r$. The same notation $\| \, . \, \|_r$ will be used also for the real domains B_r and D_r: this will not cause confusion as it will be clear from the context if it is the complex or the real domains that are considered.

- We consider a Hamiltonian H of the form

$$H(z) = h(I(z)) + f(z), \quad h : \mathcal{D}_r \to \mathbb{C}, \quad f : \mathcal{B}_r \to \mathbb{C} \tag{$*$}$$

  which is real analytic and such that

$$\|\nabla h\|_r \leq E, \quad \|\nabla^2 h\|_r \leq F, \quad \|X_f\|_r \leq \varepsilon \tag{2.5}$$

  where X_f is the Hamiltonian vector field associated to f

- The integrable Hamiltonian h is supposed to be *steep* on the domain D_r, as defined below.

  Definition 1. *A differentiable function $h : D_r \to \mathbb{R}$ is steep if there exist positive constants C, δ, p_l, for any integer $l \in [1, n-1]$, and κ such that for all $I \in D_r$, we have $||\nabla h(I)|| \geq \kappa$ and, for all integer $l \in [1, n-1]$, for all vector space $\Lambda \in \mathbb{R}^n$ of dimension l, letting $\lambda = I + \Lambda$ the associated affine subspace passing through I and h_λ the restriction of h to λ, the inequality*

$$\max_{0 \leq \eta \leq \xi} \; \min_{||I' - I|| = \eta, \, I' \in \lambda \cap D_r} ||\nabla h_\lambda(I') - \nabla h_\lambda(I)|| > C \xi^{p_l}$$

  holds true for all $0 < \xi \leq \delta$. We say that h is $(r, \kappa, C, \delta, (p_l)_{l=1,\ldots,n-1})$-steep and, if all the $p_i = p$, we say that h is $(r, \kappa, C, \delta, p)$-steep.

Let us point out that the definition of steepness that we use is not exactly the one given by Nekhoroshev but it is obviously equivalent to it (see [Nek73] or [Nek77]). Indeed, Nekhoroshev only requires steepness for subspaces Λ which are orthogonal to $\nabla h(I)$, in which case $\nabla h_\lambda(I) = 0$; for subspaces Λ such that $\nabla h_\lambda(I) \neq 0$, the inequality in Definition 1 is clearly satisfied (and one may even set $p_l = 0$ in this case).

Theorem A. *Let $H(z) = h(I(z)) + f(z)$ be as in $(*)$ satisfying (2.5), such that h is $(r, \kappa, C, \delta, (p_l)_{l=1,\ldots,n-1})$-steep. Then there exist $\tilde{r}^*, \tilde{c}, \tilde{c}' > 0$, which depend only on n, E, F, κ, C and p_l for $1 \leq l \leq n-1$ such that if*

$$r \leq \tilde{r}^*, \quad r\varepsilon \leq \tilde{c} \min\left\{\delta^{2na}, r^{4na}\right\} \tag{2.6}$$

where

$$a := 1 + p_1 + p_1 p_2 + \cdots + p_1 p_2 \ldots p_{n-1},$$

then for any solution $z(t)$ of the Hamiltonian flow $(*)$ with $z(0) = z_0 \in B_{r/2}$ we have

$$|I(z(t)) - I(z_0)| \leq \tilde{c}' (r\varepsilon)^{\frac{1}{2na}}, \quad |t| \leq \exp\left((r\varepsilon)^{-\frac{1}{2na}}\right).$$

We will prove in fact a slightly more general and more precise statement (but whose formulation is also more cumbersome): there exist positive constants $\tilde{c}_1$, $\tilde{c}_2$, $\tilde{c}_3$, $\tilde{c}_4$, $\tilde{c}_5$, $\tilde{c}_6$ and $\tilde{c}_7$, which depend only on n, E, F and on the steepness constants κ, C, p_l, for $1 \leq l \leq n-1$, such that for any solution $z(t)$ with $z(0) = z_0 \in B_{r/2}$, if

$$r\varepsilon \leq \min\left\{\tilde{c}_1, \tilde{c}_2\delta^{2na}, \tilde{c}_3 r^{4na}, \tilde{c}_4 r^{\frac{2a}{a-a'}}\right\} \tag{2.7}$$

where a is as above and

$$a' := 1 + p_2 + p_2p_3 + \cdots + p_2p_3\ldots p_{n-1},$$

then

$$|I(z(t)) - I(z_0)| \leq \tilde{c}_5(r\varepsilon)^{\frac{1}{2na}}, \quad |t| \leq \tilde{c}_6 r^{-1}(r\varepsilon)^{-\frac{1}{2na}} \exp\left(\tilde{c}_7 r^{-1}(r\varepsilon)^{-\frac{1}{2na}}\right).$$

This statement obviously implies the statement of Theorem A.

3 Proof of the main result

The goal of this section is to prove Theorem A, following the method introduced in [BN12] and [Bou11]. This method, which uses only periodic approximations and compositions of periodic averagings, has the advantage of being directly applicable in a neighborhood of an elliptic equilibrium point where action-angle coordinates cannot be used.

Since the proof contains some technical statements, we first give in the next Section 3.1 a long and complete heuristic description of the method that would hopefully make the reading of the proof easier. We emphasize that Section 3.1 is included only for the convenience of the reader and does not interfere with the proof strictly speaking.

3.1 Heuristic description and plan of the proof

Given an arbitrary initial condition z_0 and the associated solution $z(t)$ (that is $z(0) = z_0$) of the Hamiltonian $H = h + f$, our goal is to prove that the variation of the action $I(z(t)) - I(z_0)$ remains small (as a small power of ϵ) for an interval of time which is exponentially large with respect to the inverse of (some power of) ε, ε being the size of the perturbation f. The proof is based on an algorithm that, for $0 \leq j \leq n-1$, reduces to a space of dimension $n-j-1$ the directions in which a fast drift (before an exponentially long interval of time) may be possible at each step j, and that stops therefore after at most $j = n-1$ steps. We now describe the heuristics of this algorithm, which depends on a positive parameter $Q \geq 1$ and an integer parameter $m \geq 1$.

For the step $j = 0$, we write $H = H_0$ and given the parameter $Q \geq 1$, we use Dirichlet's box principle to approximate the unperturbed frequency $v_0 = \nabla h(I(z_0))$ by a periodic vector ω_0, that is, a vector which is a real multiple of an integer vector (this corresponds to a frequency vector which is maximally resonant, as the set of integer vectors k orthogonal to ω_0 forms a sub-module of maximal rank $n-1$). Letting T_0 be the period of ω_0, which is the smallest positive number t such that $t\omega_0 \in \mathbb{Z}^n$, the parameter Q controls the approximation as follows:

$$||v_0 - \omega_0|| = s_0 \lesssim (T_0 Q)^{-1}, \quad ||v_0||^{-1} \lesssim T_0 \lesssim ||v_0||^{-1} Q^{n-1}.$$

Then, on some small neighborhood V_0 of z_0, given the integer parameter $m \geq 1$ and assuming certain compatibility conditions between m, s_0, T_0 and ε, it is possible to construct a resonant

normal form (with respect to ω_0) up to a remainder which is exponentially small in m: more precisely, by a symplectic transformation Φ_0 which is close to the identity, the transformed Hamiltonian $H_0 \circ \Phi_0$ can be written as a perturbation of h, but this time the perturbation splits into two parts: a resonant part, which is still of order ε but has the additional property that its Hamiltonian flow commutes with the linear flow of frequency ω_0, and a non-resonant part which is of order $2^{-m}\varepsilon$. As a consequence, we have the following partial stability result for the solution of $H_0 \circ \Phi_0$ starting at $\Phi_0^{-1}(z_0)$: the variation of the action variables in the (one-dimensional) direction given by ω_0 is small for an exponentially long interval of time of order 2^m, unless the solution escapes from the domain of the normal form V_0 before. In other words, we excluded at this step $j = 0$ the direction ω_0 from the directions along which a drift in the actions may appear before an exponentially long interval of time. Since Φ_0 is symplectic and close to the identity, the same holds true for the solution $z(t) = z_0(t)$ of $H = H_0$, and we arrive at the following dichotomy: either the action variables have also small variation in the direction transverse to ω_0 for an exponentially long interval of time, or not.

In the first case, for an exponentially long interval of time with respect to m, the variation of the action variables is small: the stability condition is satisfied and the algorithm stops. Once the algorithm stops, one can determine the parameters Q and m in order to fulfill the compatibility conditions which essentially read as follows:

$$s_0 \lesssim 1, \quad mT_0\varepsilon \lesssim s_0, \quad mT_0 s_0 \lesssim 1.$$

Since $s_0 \sim (T_0 Q)^{-1}$ and $T_0 \lesssim Q^{n-1}$ (as $||\nabla h(I(z_0))|| = ||v_0||$ is of order one), these conditions are satisfied if we choose $m \sim Q \sim \varepsilon^{-\frac{1}{2n}}$ and $\varepsilon \lesssim 1$, and thus we obtain a result of exponential stability with respect to $m \sim \varepsilon^{-\frac{1}{2n}}$. If h is convex or quasi-convex it is simple to see, due to energy conservation, that this first case is automatic and it is not possible that the action variables drifts transversely to ω_0, hence exponential stability is proved with one step of the algorithm. But in general, the second case is possible and more work is needed to further reduce the drifting possibilities of the actions going from the step $j = 0$ to the step $j = 1$ of the algorithm.

In the second case, setting $H_0^+ = H_0 \circ \Phi_0$ and denoting by $z_0^+(t)$ the associated solution, we can find a positive time $\tilde{t}_0^+$, which is shorter than 2^m, such that the $I(z_0^+(\tilde{t}_0^+)) - I(z_0^+(0))$ has a small drift of order s_0 in the direction orthogonal to ω_0. Letting Π_0 be the projection onto the orthogonal of ω_0, we can define a curve

$$\gamma_0(t) = I(z_0^+(0)) + \Pi_0(I(z_0^+(t)) - I(z_0^+(0)))$$

which takes values in an affine subspace of dimension $n-1$. One can then exploit the steepness property to find a time $\tilde{t}_0 \leq t_0^+$ for which the vector $\nabla h(\gamma_0(\tilde{t}_0))$ is linearly independent from ω_0 in a quantitative way:

$$\Pi_0(\nabla h(\gamma_0(\tilde{t}_0))) \gtrsim s_0^{p_{n-1}}$$

where p_{n-1} is the steepness index in dimension $n-1$. Using Dirichlet's box principle again with the same parameter Q, we can approximate the vector $\Pi_0(\nabla h(\gamma_0(\tilde{t}_0))) = v_1$ by another periodic vector ω_1:

$$||v_1 - \omega_1|| = s_1 \lesssim (T_1 Q)^{-1}, \quad ||v_1||^{-1} \lesssim T_1 \lesssim ||v_1||^{-1} Q^{n-1} \lesssim s_0^{-p_{n-1}} Q^{n-1}.$$

First observe that since v_1 is orthogonal to ω_0, ω_1 is almost orthogonal to ω_0 and in particular it is linearly independent from ω_0. Then since $||\gamma_0(\tilde{t}_0) - I(z_0^+(\tilde{t}_0))|| = ||\Pi_0^\perp(I(z_0^+(\tilde{t}_0)) - I(z_0^+(0)))||$

is small (as $I(z_0^+(t) - I(z_0^+(0))$ has only small variation in the direction given by ω_0), we also have

$$||\Pi_0(\nabla h(I(z_0^+(\tilde{t}_0)))) - \omega_1|| \sim s_1 \lesssim (T_1 Q)^{-1}, \quad T_1 \lesssim s_0^{-p_{n-1}} Q^{n-1}.$$

Set $z_1 = z_0^+(\tilde{t}_0)$ and $H_1 = H_0^+$. On a small neighborhood V_1 of z_1 (small enough so that V_1 is still included in V_0) we can then construct, as in the first step, a resonant normal form with respect to ω_1 up to an exponentially small remainder with respect to m. Unlike the step $j = 0$ in which the perturbation was arbitrary, here the perturbation is given by the non-resonant part with respect to ω_0, and this explains why it is sufficient to have an approximation of $\Pi_0(\nabla h(I(z_1)))$ and not of the full vector $\nabla h(I(z_1))$. Moreover, a careful construction of the new normalizing transformation Φ_1 shows that the resonant part of $H_1 \circ \Phi_1$, whose flow commutes with the linear flow of frequency ω_1, also commutes with the linear flow of frequency ω_0. Exactly like in the first step, we arrive at a dichotomy which determines whether the algorithm stops (the variation of the action of the solution of $H_1 \circ \Phi_1$, and then of $H_1 = H_0 \circ \Phi_0$ and $H_0 = H$, is small and the theorem is proved) or moves to the next step with the gain that now both the directions ω_0 and ω_1, that are linearly independent, are excluded from the directions along which a drift in the actions may appear before an exponentially long interval of time. Note that if the algorithm stops, the parameters Q and m have to be chosen according to

$$s_i \lesssim 1, \quad m T_i \varepsilon \lesssim s_i, \quad m T_i s_i \lesssim 1, \quad 0 \le i \le 1.$$

Observe that

$$T_1 \lesssim s_0^{-p_{n-1}} Q^{n-1} \sim (T_0 Q)^{p_{n-1}} Q^{n-1} \lesssim Q^{n(1+p_{n-1})-1}$$

and hence the compatibility conditions are satisfied if we choose $m \sim Q \sim \varepsilon^{-\frac{1}{2na_1}}$ and $\varepsilon \lesssim 1$, with $a_1 = 1 + p_{n-1}$, and thus we obtain a result of exponential stability with respect to $m \sim \varepsilon^{-\frac{1}{2na_1}}$. In particular, as $2na_1 > 2n$, this stability result also holds true if the algorithm stopped at step $j = 0$.

We have briefly explained how to pass from the step $j = 0$ to the step $j = 1$, and how the parameters Q and m are chosen if the algorithm stops at step $j = 0$ or $j = 1$. But of course, for any $0 \le j \le n-2$, one proceeds exactly the same way to go from step j to step $j+1$. The fact that the algorithm stops after n steps (if, of course, it didn't stop before) is clear since then n linearly independent directions are excluded from the directions along which a drift in the actions may appear before an exponentially long interval of time. More formally, in the case $j = n-1$, the resonant part in the normal form $H_{n-1}^+ = H_{n-1} \circ \Phi_{n-1}$ consists of a Hamiltonian whose flow commutes with n linearly independent linear flows with frequency $\omega_0, \dots, \omega_{n-1}$; it is plain to see that such a Hamiltonian is integrable, so H_{n-1}^+ consist of an exponentially small perturbation of some integrable Hamiltonian: the first case of the dichotomy thus holds, the algorithm stops and the theorem is proved. At each step j, the compatibility conditions are given by

$$s_i \lesssim 1, \quad m T_i \varepsilon \lesssim s_i, \quad m T_i s_i \lesssim 1, \quad 0 \le i \le j,$$

and using the fact that

$$T_0 \lesssim Q^{n-1}, \quad T_j \lesssim (T_{j-1} Q)^{p_{n-j}} Q^{n-1}, \quad 1 \le j \le n-1,$$

we can choose, if the algorithm stops at step j, $m \sim Q \sim \varepsilon^{-\frac{1}{2na_j}}$ and $\varepsilon \lesssim 1$, with

$$a_0 = 1, \quad a_1 = 1 + p_{n-1}, \quad a_j = 1 + p_{n-j} + \dots + p_{n-j} \dots p_{n-1}, \quad j \ge 2,$$

leading to a result of exponential stability with exponent $2na_j$. We have

$$a_{n-1} = 1 + p_1 + p_1p_2 + \cdots + p_1p_2 \dots p_{n-1} = a,$$

and since $2na_{n-1} > 2na_{n-2} > \cdots > 2na_0$, independently of the step j at which the algorithm stops (and hence independently of the choice of Q and m), we obtain a result of exponential stability with exponent $2na$.

Let us now describe the plan of the proof. In §3.2, our aim will be to obtain a suitable normal form (Proposition 3.1) for an abstract Hamiltonian H_j, where $0 \leq j \leq n-1$, which, as we explained in the heuristic description above, will be later related to our original Hamiltonian H described in $(*)$ in the following way: $H_0 = H$ and for $j \geq 1$, $H_j = H_{j-1} \circ \Phi_{j-1} = H_0 \circ \Phi_0 \circ \cdots \circ \Phi_{j-1}$ where Φ_i, for $0 \leq i \leq j-1$, is the normalizing transformation with respect to the periodic frequency ω_i. Here the periodic frequencies $\omega_0, \dots, \omega_j$ are assumed to be known, H_j is already normalized with respect to $\omega_0, \dots, \omega_{j-1}$ and our aim is to explain the construction of the transformation Φ_j which will further normalize H_j with respect to ω_j. The proof being technical, details will be given in Appendix A. In §3.3, a partial stability result (in the direction given by the linear span of $\omega_0, \dots, \omega_j$, up to times of order 2^m) will be easily deduced from the normal form Hamiltonian $H_j \circ \Phi_j$. Using this, we will introduce a first version of the algorithm dichotomy in Proposition 3.2: either stability holds for an exponentially long interval of time and the algorithm stops, or a small drift in the action does appear in the orthocomplement of $\omega_0, \dots, \omega_j$ and the algorithm should move to the next step. We will also prove in Proposition 3.2 that if $j = n-1$, then only the first alternative can be true. In §3.4 and §3.5, we will examine the situation where the second alternative holds true (so necessarily $j \leq n-2$), and using the steepness property of h (in §3.4) and then Dirichlet's box principle (in §3.5) we will prove how to pass from the step j to the step $j+1$. In §3.6, we summarize the work done in §3.2, §3.3, §3.4 and §3.5 in Proposition 3.7 that gives one step of the algorithm, and which clearly shows that at some given step, either the algorithm stops or it yields the hypotheses that allow to apply it again. Finally, in §3.7 we conclude the proof of Theorem A, which will follow easily from Proposition 3.7 by determining the parameters $Q \geq 1$ and $m \geq 1$ in terms of our small parameter ε.

3.2 Normal form statement

In this section we fix $0 \leq j \leq n-1$, and we assume the existence of periodic vectors $\omega_0, \dots, \omega_j$, with periods $T_0, \dots T_j$, which are linearly independent. For convenience we set $\omega_{-1} = 0 \in \mathbb{R}^n$. We define the complex and real vector space

$$\tilde{\Lambda}_j := \{v \in \mathbb{C}^n \mid v \cdot \omega_{-1} = v \cdot \omega_0 = \cdots = v \cdot \omega_{j-1} = 0\}, \quad \Lambda_j := \tilde{\Lambda}_j \cap \mathbb{R}^n,$$

which are of complex (respectively real) dimension $n-j$. Then we consider three positive real numbers r_j, s_j and ξ_j, a point $z_j \in B_{r_j}$ and we define the complex domain

$$\mathcal{V}_{3s_j,3\xi_j}(z_j) := \{z \in \mathbb{C}^{2n} \mid I(z) - I(z_j) \in \tilde{\Lambda}_j,\ |I(z) - I(z_j)| < 3s_j,\ ||z|| < r_j + 3\xi_j\} \tag{3.1}$$

where, for simplicity, the dependence on r_j is omitted. We will simply write $|| \cdot ||_{3s_j,3\xi_j}$ for the supremum norm for vector fields defined on $\mathcal{V}_{3s_j,3\xi_j}(z_j)$ and, for $-1 \leq i \leq n-1$, we will denote by $l_{\omega_i}(z) := \omega_i \cdot I(z)$ and X_{ω_i} its associated Hamiltonian vector field. With our convention, the function $l_{\omega_{-1}}$ and its associated Hamiltonian vector field are identically zero.

Given a real number $0 < \varepsilon < 1$ and an integer $m \geq 1$, we can define a set of Hamiltonians as follows.

Definition 2. *The set $NF_j(\omega_{-1},\dots,\omega_{j-1},z_j,s_j,r_j,\xi_j,F,\varepsilon,m)$, or for short NF_j, consists of real-analytic Hamiltonians H_j defined on $\mathcal{V}_{3s_j,3\xi_j}(z_j)$, and of the form*

$$\begin{cases} H_j(z) := h(I(z)) + g_j(z) + f_j(z), \quad z \in \mathcal{V}_{3s_j,3\xi_j}(z_j) \\ h : \mathcal{D}_r \to \mathbb{C}, \quad ||\nabla^2 h(I)||_r \le F, \quad I(\mathcal{V}_{3s_j,3\xi_j}(z_j)) \subseteq \mathcal{D}_r, \\ ||X_{g_j}||_{3s_j,3\xi_j} \le 2^j\varepsilon, \quad ||X_{f_j}||_{3s_j,3\xi_j} \le j2^{j-1}2^{-m}\varepsilon, \\ \{l_{\omega_{-1}}, g_j\} = \{l_{\omega_0}, g_j\} = \cdots = \{l_{\omega_{j-1}}, g_j\} = 0. \end{cases}.$$

Now let us introduce another definition, taking into account the periodic frequency ω_j.

Definition 3. *The set $\widetilde{NF}_j(\omega_0,\dots,\omega_j,z_j,s_j,r_j,\xi_j,F,\varepsilon,m)$, or for short $\widetilde{NF}_j$, consists of real-analytic Hamiltonians $H_j \in NF_j(\omega_{-1},\dots,\omega_{j-1},z_j,s_j,r_j,\xi_j,F,\varepsilon,m)$ which satisfy the following additional conditions: if we denote $\tilde{\Pi}_j$ (respectively Π_j) the orthogonal projection onto $\tilde{\Lambda}_j$ (respectively Λ_j), then*

$$||\tilde{\Pi}_j \nabla h(I(z_j)) - \omega_j|| = ||\Pi_j \nabla h(I(z_j)) - \omega_j|| \le s_j \tag{3.2}$$

and

$$s_j \le (r_j + 2\xi_j)\xi_j, \quad 2^j 216(r_j + 3\xi_j)mT_j\varepsilon \le s_j, \quad 72(3F\sqrt{n}+1)\xi_j^{-1}(r_j+3\xi_j)mT_js_j \le 1. \tag{3.3}$$

The interest of the subset $\widetilde{NF}_j \subset NF_j$ is that if $H_j \in \widetilde{NF}_j$, then up to a change of coordinates Φ_j (which is real-analytic, symplectic and close to identity), we get that $H_j \circ \Phi_j \in NF_{j+1}$ which will constitute a main ingredient in our algorithm. Here's the precise statement.

Proposition 3.1. *Let $H_j \in \widetilde{NF}_j(\omega_0,\dots,\omega_j,z_j,s_j,r_j,\xi_j,F,\varepsilon,m)$. Then there exist a real-analytic symplectic embedding*

$$\Phi_j : \mathcal{V}_{2s_j,2\xi_j}(z_j) \to \mathcal{V}_{3s_j,3\xi_j}(z_j), \quad \Phi_j\left(\mathcal{V}_{2s_j,2\xi_j}(z_j)\right) \supseteq \mathcal{V}_{s_j,\xi_j}(z_j),$$

such that $H_j^+ := H_j \circ \Phi_j = h + g_j^+ + f_j^+$ with

$$\{l_{\omega_{-1}}, g_j^+\} = \{l_{\omega_0}, g_j^+\} = \cdots = \{l_{\omega_j}, g_j^+\} = 0, \tag{3.4}$$

and with the estimates

$$||X_{g_j^+}||_{2s_j,2\xi_j} \le 2^{j+1}\varepsilon, \quad ||X_{f_j^+}||_{2s_j,2\xi_j} \le (j+1)2^j 2^{-m}\varepsilon, \tag{3.5}$$

$$||\Phi_j - \mathrm{Id}||_{2s_j,2\xi_j} \le 2^{j+1}T_j\varepsilon. \tag{3.6}$$

In particular, $H_j^+ \in NF_{j+1}(\omega_{-1},\dots,\omega_j,z_{j+1},s_{j+1},r_{j+1},\xi_{j+1},F,\varepsilon,m)$ given any choice of s_{j+1}, r_{j+1}, ξ_{j+1} and $z_{j+1} \in B_{r_{j+1}}$ for which the inclusion $\mathcal{V}_{3s_{j+1},3\xi_{j+1}}(z_{j+1}) \subseteq \mathcal{V}_{2s_j,2\xi_j}(z_j)$ holds true.

The proof of Proposition 3.1, which is technical, is deferred to Appendix A. The second part of Proposition 3.1 follows easily from the first: if we define $H_{j+1} := H_j^+$, $g_{j+1} := g_j^+$ and $f_{j+1} := f_j^+$, then (3.4) read

$$\{l_{\omega_{-1}}, g_{j+1}\} = \{l_{\omega_0}, g_{j+1}\} = \cdots = \{l_{\omega_j}, g_{j+1}\} = 0,$$

whereas the inclusion $\mathcal{V}_{3s_{j+1},3\xi_{j+1}}(z_{j+1}) \subseteq \mathcal{V}_{2s_j,2\xi_j}(z_j)$ yields

$$I(\mathcal{V}_{3s_{j+1},3\xi_{j+1}}(z_{j+1})) \subseteq I(\mathcal{V}_{2s_j,2\xi_j}(z_j)) \subseteq I(\mathcal{V}_{3s_j,3\xi_j}(z_j)) \subseteq \mathcal{D}_r$$

and, together with (3.5), the estimates

$$||X_{g_{j+1}}||_{3s_{j+1},3\xi_{j+1}} \leq ||X_{g_{j+1}}||_{2s_j,2\xi_j} \leq 2^{j+1}\varepsilon,$$

$$||X_{f_{j+1}}||_{3s_{j+1},3\xi_{j+1}} \leq ||X_{f_{j+1}}||_{2s_j,2\xi_j} \leq (j+1)2^j 2^{-m}\varepsilon.$$

This exactly means that $H_{j+1} = H_j^+ \in NF_{j+1}(\omega_{-1},\ldots,\omega_j, z_{j+1}, s_{j+1}, r_{j+1}, \xi_{j+1}, F, \varepsilon, m)$.

3.3 Use of the normal form

From now on, we will mainly work on the real domains

$$V_{2s_j,2\xi_j}(z_j) := \mathcal{V}_{2s_j,2\xi_j}(z_j) \cap \mathbb{R}^{2n}, \quad 0 \leq j \leq n-1.$$

The normal form in Proposition 3.1 is used to show that given a solution $z_j^+(t)$ of the Hamiltonian system associated to $H_j^+ = H_j \circ \Phi_j$, the curve $I(z_j^+(t))$ has a small variation in the direction spanned by $\omega_0,\ldots,\omega_j$, which is nothing but $\Lambda_{j+1}^\perp$ (the orthocomplement of Λ_{j+1}), for times $|t|$ as large as the inverse of $||X_{f_j^+}||_{2s_j,2\xi_j}$. It may well happen that in the direction given by Λ_{j+1}, the curve $I(z_j^+(t))$ has also a small variation and hence $I(z_j(t))$ where $z_j(t) = \Phi_j(z_j^+(t))$, has small variation, which yields our confinement result. But if not, that is if there is a faster deviation of $I(z_j^+(t))$ from $I(z_j^+(0))$, this has to occur in the direction given by Λ_{j+1}. Here is a precise statement.

Proposition 3.2. *For $0 \leq j \leq n-1$, let $H_j \in \widetilde{NF}_j(\omega_0,\ldots,\omega_j, z_j, s_j, r_j, \xi_j, F, \epsilon, m)$ and $\Phi_j : \mathcal{V}_{2s_j,2\xi_j}(z_j) \to \mathcal{V}_{3s_j,3\xi_j}(z_j)$ given by Proposition 3.1, and let $z_j^+(t)$ be the forward solution of the Hamiltonian $H_j^+ = H_j \circ \Phi_j$ starting at $z_j^+ := \Phi_j^{-1}(z_j)$. If $0 \leq j \leq n-2$, and if we define*

$$\bar{t}_j := (r_j + \xi_j)^{-1}(j+1)^{-1}2^{-j}s_j^{-1}2^m, \tag{3.7}$$

then we have the following dichotomy:

(1) *either $z_j^+(t) \in V_{s_j,\xi_j}(z_j)$ for $0 \leq t \leq \bar{t}_j$,*

(2) *or there exists a positive time $t_j^+ < \bar{t}_j$ such that*

$$|I(z_j^+(t_j^+)) - I(z_j)| = s_j/4$$

and, for $0 \leq t \leq t_j^+$,

$$z_j^+(t) \in V_{s_j,\xi_j}(z_j), \quad |I(z_j^+(t)) - I(z_j)| \leq s_j/4, \quad |\Pi_{j+1}^\perp(I(z_j^+(t)) - I(z_j))| \leq s_j^{-1}\varepsilon. \tag{3.8}$$

If $j = n-1$, and if we define

$$\bar{t}_{n-1} := (r_{n-1} + \xi_{n-1})^{-1}n^{-1}2^{-(n-1)}s_{n-1}\varepsilon^{-1}2^m, \tag{3.9}$$

then $z_{n-1}^+(t) \in V_{s_{n-1},\xi_{n-1}}(z_{n-1})$, for $0 \leq t \leq \bar{t}_{n-1}$.

Proof. First observe that since the image of Φ_j contains $V_{s_j,\xi_j}(z_j)$, it contains z_j and so $z_j^+ = \Phi_j^{-1}(z_j)$ is well-defined, and we have, using (3.6) and the first two inequalities of (3.3),

$$||z_j^+ - z_j|| = ||z_j^+ - \Phi_j(z_j^+)|| \leq 2^{j+1}T_j\varepsilon \leq s_j(108(r_j + 3\xi_j))^{-1} \leq \xi_j(r_j + 2\xi_j)(108(r_j + 3\xi_j))^{-1}$$

which easily implies

$$||z_j^+ - z_j|| < \xi_j/108, \quad |I(z_j^+) - I(z_j)| < s_j/108. \tag{3.10}$$

Observe also that since z_j is real and H_j and Φ_j are reals, the forward solution $z_j^+(t)$ is real.

We first consider the case $0 \leq j \leq n-2$. Using (3.10), we can now define $t_j^e \in (0, +\infty]$ to be the time of first exit of $z_j^+(t)$ from $V_{s_j,\xi_j}(z_j)$. We claim that the dichotomy of the statement is implied by the following trivial dichotomy: either $\bar{t}_j < t_j^e$ or $t_j^e \leq \bar{t}_j$.

Indeed, in the first case, one obviously have

$$z_j^+(t) \in V_{r_j,\xi_j}(z_j), \quad 0 \leq t \leq \bar{t}_j.$$

In the second case, either $|I(z_j^+(t_j^e)) - I(z_j)| = s_j$ or $||z_j^+(t_j^e)|| = r_j + \xi_j$. But since the solution is real, the second possibility implies that

$$|I(z_j^+(t_j^e))| = 1/2||z_j^+(t_j^e)||^2 = 1/2(r_j^2 + \xi_j^2) + r_j\xi_j,$$

while

$$|I(z_j)| = 1/2||z_j||^2 < 1/2r_j^2$$

and therefore, using the first inequality of (3.3), we obtain

$$|I(z_j^+(t_j^e)) - I(z_j)| \geq |I(z_j^+(t_j^e))| - |I(z_j)| > \xi_j(1/2\xi_j + r_j) > s_j/4.$$

So, whether $|I(z_j^+(t_j^e)) - I(z_j)| = s_j$ or $||z_j^+(t_j^e)|| = r_j + \xi_j$, there exists a positive time $t_j^+ < t_j^e \leq \bar{t}_j$ such that

$$|I(z_j^+(t_j^+)) - I(z_j)| = s_j/4$$

and

$$|I(z_j^+(t)) - I(z_j)| \leq s_j/4, \quad 0 \leq t \leq t_j^+.$$

Since $t_j^+ < t_j^e$, $z_j^+(t) \in V_{s_j,\xi_j}(z_j)$ for $0 \leq t \leq t_j^+$. It remains to show that

$$|\Pi_{j+1}^{\perp}(I(z_j^+(t)) - I(z_j))| \leq s_j^{-1}\varepsilon, \quad 0 \leq t \leq t_j^+.$$

Since h is integrable, for $0 \leq s \leq t \leq t_j^+$,

$$\begin{aligned} \frac{d}{dt}I(z_j^+(s)) &= \{I, H_j^+\}(z_j^+(s)) := (\{I_1, H_j^+\}(z_j^+(s)), \ldots, \{I_n, H_j^+\}(z_j^+(s))) \\ &= \{I, h + g_j^+ + f_j^+\}(z_j^+(s)) = \{I, g_j^+ + f_j^+\}(z_j^+(s)). \end{aligned}$$

Then, for any $z \in V_{s_j,\xi_j}(z_j)$ and any $-1 \leq i \leq j$, using (3.4) we obtain

$$\{l_{\omega_i}, g_j^+\}(z) = \omega_i \cdot \{I, g_j^+\}(z) = 0$$

which implies that $\{I, g_j^+\}(z) \in \Lambda_{j+1}$. Therefore

$$\Pi_{j+1}^{\perp}\left(\frac{d}{dt}I(z_j^+(s))\right) = \{I, f_j^+\}(z_j^+(s))$$

and hence

$$\Pi_{j+1}^{\perp}(I(z_j^+(t)) - I(z_j)) = \int_0^t \Pi_{j+1}^{\perp}\left(\frac{d}{dt}I(z_j^+(s))\right) ds = \int_0^t \{I, f_j^+\}(z_j^+(s))ds$$

and, using the second inequality of (3.5) and the fact that

$$|\{I, f_j^+\}(z_j^+(s))| \leq ||z_j^+(s)||||X_f(z_j^+(s))||$$

we obtain

$$|\Pi_{j+1}^{\perp}(I(z_j^+(t)) - I(z_j))| \leq t(r_j + \xi_j)||X_{f_j^+}||_{2s_j,2\xi_j} \leq t(r_j + \xi_j)(j+1)2^j 2^{-m}\varepsilon.$$

Since $t_j^+ \leq t_j^e \leq \bar{t}_j = (r_j + \xi_j)^{-1}(j+1)^{-1}2^{-j}s_j^{-1}2^m$, we thus obtain

$$|\Pi_{j+1}^{\perp}(I(z_j^+(t)) - I(z_j))| \leq s_j^{-1}\varepsilon, \quad 0 \leq t \leq t_j^+,$$

which concludes the proof for the case $0 \leq j \leq n-2$.

Now for the case $j = n-1$, we have $\Lambda_n = \{0\}$ and so $\Lambda_n^{\perp} = \mathbb{R}^n$, hence for $t \leq \bar{t}_{n-1} = (r_{n-1} + \xi_{n-1})^{-1}n^{-1}2^{-(n-1)}s_{n-1}\varepsilon^{-1}2^m$, repeating the last argument we get

$$|I(z_{n-1}^+(t)) - I(z_{n-1})| \leq s_{n-1}, \quad 0 \leq t \leq \bar{t}_{n-1}.$$

As $z_{n-1}^+(t)$ is real, this implies, using also the first inequality of (3.3), that for $0 \leq t \leq \bar{t}_{n-1}$,

$$\begin{aligned} ||z_{n-1}^+(t)||^2 &= 2|I(z_{n-1}^+(t))| \\ &\leq 2|I(z_{n-1}^+(t)) - I(z_{n-1})| + 2|I(z_{n-1})| \\ &\leq 2s_{n-1} + ||z_{n-1}(t)||^2 \\ &\leq 2(r_{n-1} + 2\xi_{n-1})\xi_{n-1} + r_{n-1}^2 \\ &\leq (r_{n-1} + 2\xi_{n-1})^2 \end{aligned}$$

so $z_{n-1}^+(t) \in V_{s_{n-1},2\xi_{n-1}}(z_{n-1})$ for $0 \leq t \leq \bar{t}_{n-1}$, and this concludes the proof of the proposition. □

3.4 Use of the steepness property

Let us start by giving a geometric interpretation of the steepness property, as its definition is quite abstract. Assume that h is steep on some domain D, and consider a curve $\gamma : [0,1] \to \mathbb{R}^n$ which takes values in $\lambda \cap D$, where λ is a proper affine subspace of $\mathbb{R}^n$. It may happen that $\nabla h_\lambda(\gamma(0)) = 0$ (this is the case if $\gamma(0)$ is a resonant point for h, that is, if $k \cdot \nabla h(\gamma(0))$ for some non-zero integer vector $k \in \mathbb{Z}^n$: then $\nabla h_\lambda(\gamma(0)) = 0$ where λ is the real space generated by such integer vectors k). If this happens, the steepness property ensures that, for some time $0 < \tilde{t} \leq 1$, $\nabla h_\lambda(\gamma(\tilde{t})) \neq 0$ (informally, in terms of resonances, this means that we do not have "accumulation of resonances"). Moreover, the longer is the length of the curve γ, the farther away from zero is the vector $\nabla h_\lambda(\gamma(\tilde{t}))$. Here's a quantitative statement, which is due to Nekhoroshev.

Lemma 3.3 (Nekhoroshev). *Let h be a function which is $(r, \kappa, C, \delta, (p_l)_{l=1,\dots,n-1})$-steep, and such that*

$$||\nabla^2 h(I)||_r \leq F.$$

Let $\gamma : [0, t^+] \to \mathbb{R}^n$ be a continuous curve, λ an affine subspace of $\mathbb{R}^n$ of dimension l, where $1 \leq l \leq n-1$, and d a positive real number. Assume that

(i) for all $t \in [0, t^+]$, $\gamma(t) \in \lambda$;

(ii) for all $t \in [0, t^+]$, $||\gamma(0) - \gamma(t)|| \leq d$ and $||\gamma(0) - \gamma(t^+)|| = d$;

(iii) the ball $\{I \in \mathbb{R}^n \mid ||I - \gamma(0)|| \leq d\}$ is contained in D_r;

(iv) $d < \min\{\delta, (3F)^{-1}\kappa, 2(5\kappa(4C)^{-1})^{1/p_l}\}$,

then there exists a time $\tilde{t} \in [0, t^+]$ such that

$$||\Pi_\Lambda \nabla h(\gamma(\tilde{t}))|| > C/5(d/2)^{p_l},$$

where Λ is the vector space associated to λ, and Π_Λ the orthogonal projection onto Λ.

This is a special case of the lemma on "almost plane curves" of Nekhoroshev, stated in [Nek77] and proved in [Nek79] (our case corresponds to "plane curves").

Now assume that Alternative (2) of Proposition 3.2 holds true, and let $\gamma_j(t) := I(z_j^+) + \Pi_{j+1}(I(z_j^+(t)) - I(z_j^+))$ for $t \in [0, t_j^+]$. Since this curve takes values in a proper affine subspace, the following proposition is a simple consequence of Proposition 3.2 and Lemma 3.3.

Proposition 3.4. *For $0 \leq j \leq n-1$, let $H_j \in \widetilde{NF}_j(\omega_0, \dots, \omega_j, z_j, s_j, r_j, \xi_j, F)$ and $\Phi_j : \mathcal{V}_{2s_j, 2\xi_j}(z_j) \to \mathcal{V}_{3s_j, 3\xi_j}(z_j)$ given by Proposition 3.1, and let $z_j^+(t)$ be the forward solution of the Hamiltonian $H_j^+ = H_j \circ \Phi_j$ starting at $z_j^+ = \Phi_j^{-1}(z_j)$. Assume that h is $(r, \kappa, C, \delta, (p_l)_{l=1,\dots,n-1})$-steep. Then we have the following dichotomy for $j \leq n-2$:*

(1) *either $z_j^+(t) \in V_{s_j, \xi_j}(z_j)$ for $0 \leq t \leq \bar{t}_j$,*

(2) *or there exists a time $\tilde{t}_j \leq t_j^+ < \bar{t}_j$ such that, setting $\gamma_j(\tilde{t}_j) := I(z_j^+) + \Pi_{j+1}(I(z_j^+(\tilde{t}_j)) - I(z_j^+))$, then*

$$||\Pi_{j+1} \nabla h(\gamma_j(\tilde{t}_j))|| > \mu_j s_j^{p_{n-j-1}}, \quad \mu_j := 5^{-1} C 16^{-p_{n-j-1}}, \tag{3.11}$$

provided that

$$\begin{cases} \varepsilon < s_j^2/8, \\ s_j < 8 \min\{\delta, (3F)^{-1}\kappa, 2(5\kappa(4C)^{-1})^{1/p_{n-j-1}}\}. \end{cases} \tag{3.12}$$

If $j = n-1$, then $z_{n-1}^+(t) \in V_{s_{n-1}, \xi_{n-1}}(z_{n-1})$, for $0 \leq t \leq \bar{t}_{n-1}$.

Proof. We only have to consider the case $j \leq n-2$, and we have to prove that Alternative (2) of Proposition 3.2 implies Alternative (2) of the above proposition. So we assume the existence of a positive time $t_j^+ < \bar{t}_j$ such that

$$|I(z_j^+(t_j^+)) - I(z_j)| = s_j/4$$

and, for $0 \le t \le t_j^+$,

$$z_j^+(t) \in V_{s_j,\xi_j}(z_j), \quad |I(z_j^+(t)) - I(z_j)| \le s_j/4, \quad |\Pi_{j+1}^\perp(I(z_j^+(t)) - I(z_j))| \le s_j^{-1}\varepsilon.$$

Hence, using the first inequality of (3.12),

$$|\Pi_{j+1}(I(z_j^+(t_j^+)) - I(z_j))| \ge |I(z_j^+(t_j^+)) - I(z_j)| - |\Pi_{j+1}^\perp(I(z_j^+(t)) - I(z_j)| \ge s_j/4 - s_j^{-1}\varepsilon \ge s_j/8$$

and in particular

$$||\Pi_{j+1}(I(z_j^+(t_j^+)) - I(z_j))|| \ge s_j/8.$$

Therefore we can certainly find a positive time $\tilde{t}_j^+ \le t_j^+$ such that

$$||\Pi_{j+1}(I(z_j^+(\tilde{t}_j^+)) - I(z_j))|| = s_j/8$$

and

$$||\Pi_{j+1}(I(z_j^+(t)) - I(z_j))|| \le s_j/8, \quad 0 \le t \le \tilde{t}_j^+. \tag{3.13}$$

Now we want to apply Lemma 3.3 to the curve $\gamma_j(t) = I(z_j^+) + \Pi_{j+1}(I(z_j^+(t)) - I(z_j^+))$, for $t \in [0, \tilde{t}_j^+]$, with $d := s_j/8$ and with the affine subspace $\lambda_{j+1} := I(z_j^+) + \Lambda_{j+1}$ which has dimension $n - j - 1$. The assumptions (i) and (ii) of Lemma (3.3) are trivially satisfied, as

$$\gamma_j(\tilde{t}_j^+) - \gamma_j(0) = \Pi_{j+1}(I(z_j^+(\tilde{t}_j^+) - I(z_j)).$$

Then (iii) holds true since, by definition of H_j, we have $I(V_{2s_j,2\xi_j}(z_j)) \subseteq I(V_{3s_j,3\xi_j}(z_j)) \subseteq D_r$. Eventually, the second inequality of (3.12) clearly implies (iv) therefore Lemma 3.3 can be applied, and there exists a time $\tilde{t}_j \in [0, \tilde{t}_j^+]$ such that

$$||\Pi_{j+1}\nabla h(\gamma_j(\tilde{t}_j))|| > 5^{-1}C(d/2)^{p_{n-j-1}} = \mu_j s_j^{p_{n-j-1}}.$$

This concludes the proof. □

3.5 Use of periodic approximations

Let us first state the following simple consequence of Dirichlet's theorem on approximation of real vectors by rational vectors.

Lemma 3.5. *Let $v \in \mathbb{R}^n \setminus \{0\}$, and $Q \ge 1$ a real number. Then there exists a T-periodic vector $\omega \in \mathbb{R}^n \setminus \{0\}$ such that*

$$||v - \omega|| \le \sqrt{n-1}(TQ)^{-1}, \quad ||v||^{-1} \le T \le \sqrt{n}||v||^{-1}Q^{n-1}.$$

Proof. Fix $Q \ge 1$. Up to a re-ordering of its component, we can write $v = |v|(\pm 1, x)$ for some $x \in \mathbb{R}^{n-1}$ and by Dirichlet's approximation theorem, there exists a rational vector $p/q \in \mathbb{Q}^{n-1}$, such that

$$|qx - p| \le Q^{-1}, \quad 1 \le q \le Q^{n-1}.$$

The vector $\omega = |v|(\pm 1, p/q) \in \mathbb{R}^n$ is then T-periodic, for $T = |v|^{-1}q$, and we have

$$||v - \omega|| \le T^{-1}||qx - p||, \quad |v|^{-1} \le T \le |v|^{-1}Q^{n-1}$$

which implies

$$||v - \omega|| \le \sqrt{n-1}(TQ)^{-1}, \quad ||v||^{-1} \le T \le \sqrt{n}||v||^{-1}Q^{n-1}$$

and this was the statement to prove. □

Now assume that the conclusion of Alternative 2 of Proposition 3.4 holds true, so the vector $\Pi_{j+1}\nabla h(\gamma_j(\tilde{t}_j))$ is non-zero, where $\gamma_j(\tilde{t}_j) = I(z_j^+) + \Pi_{j+1}(I(z_j^+(\tilde{t}_j)) - I(z_j^+))$. By Lemma 3.5 this non-zero vector can be approximated by a periodic vector ω_{j+1}, and it will be easy to ensure that this new periodic vector is linearly independent from $\omega_0, \dots, \omega_j$ (as ω_{j+1} is close to $\Pi_{j+1}\nabla h(\gamma_j(\tilde{t}_j))$, the latter being, obviously, linearly independent from $\omega_0, \dots, \omega_j$ as it is orthogonal to them). Moreover, as $\gamma_j(\tilde{t}_j)$ is close to $I(z_j^+(\tilde{t}_j))$, setting $z_{j+1} := z_j^+(\tilde{t}_j)$, the vector ω_{j+1} is also an approximation of $\Pi_{j+1}\nabla h(I(z_{j+1}))$. This leads to the following proposition.

Proposition 3.6. *For $0 \le j \le n-1$, let $H_j \in \widetilde{NF}_j(\omega_0, \dots, \omega_j, z_j, s_j, r_j, \xi_j, F, \epsilon, m)$ and $\Phi_j : \mathcal{V}_{2s_j,2\xi_j}(z_j) \to \mathcal{V}_{3s_j,3\xi_j}(z_j)$ given by Proposition 3.1, and let $z_j^+(t)$ be the forward solution of the Hamiltonian $H_j^+ = H_j \circ \Phi_j$ starting at $z_j^+ = \Phi_j^{-1}(z_j)$. Assume that h is $(r, \kappa, C, \delta, (p_l)_{l=1,\dots,n-1})$-steep. Then, for $j \le n-2$, we have the following dichotomy:*

(1) *either $z_j^+(t) \in V_{s_j,\xi_j}(z_j)$ for $0 \le t \le \bar{t}_j$,*

(2) *or, given some positive parameter $Q \ge 1$, there exists a T_{j+1}-periodic vector $\omega_{j+1} \in \mathbb{R}^n \setminus \{0\}$, linearly independent from $\omega_0, \dots, \omega_j$, with the estimate*

$$(F'\sqrt{n}s_j)^{-1} < T_{j+1} < \sqrt{n}\mu_j^{-1}s_j^{-p_{n-j-1}}Q^{n-1}, \quad F' := \max\{1, F\}, \tag{3.14}$$

and a time $\tilde{t}_j \le\le \bar{t}_j$ such that if we define

$$z_{j+1} := z_j^+(\tilde{t}_j), \quad s_{j+1} := 2\sqrt{n-1}(T_{j+1}Q)^{-1}, \quad r_{j+1} := r_j + \xi_j, \quad \xi_{j+1} := \xi_j/3$$

then it holds that

$$\mathcal{V}_{3s_{j+1},3\xi_{j+1}}(z_{j+1}) \subseteq \mathcal{V}_{2s_j,2\xi_j}(z_j) \tag{3.15}$$

$$||\Pi_{j+1}\nabla h(I(z_{j+1})) - \omega_{j+1}|| \le s_{j+1}, \tag{3.16}$$

$$z_j^+(t) \in V_{s_j,\xi_j}(z_j), \quad 0 \le t \le \tilde{t}_j, \tag{3.17}$$

provided that

$$\begin{cases} \varepsilon < s_j^2/8, \\ s_j < 8\min\{\delta, (3F)^{-1}\kappa, 2(5\kappa(4C)^{-1})^{1/p_{n-j-1}}\}, \\ Q \ge 8F'\sqrt{n(n-1)}, \\ \varepsilon \le (2\sqrt{n}F)^{-1}s_js_{j+1}. \end{cases} \tag{3.18}$$

If $j = n-1$, then $z_{n-1}^+(t) \in V_{s_{n-1},\xi_{n-1}}(z_{n-1})$, for $0 \le t \le \bar{t}_{n-1}$.

Proof. Since (3.18) implies in particular (3.12), it is enough to prove that Alternative (2) of Proposition 3.4 implies Alternative (2) of the above proposition. So we assume the existence of a time $\tilde{t}_j \le t_j^+ < \bar{t}_j$ such that

$$||\Pi_{j+1}\nabla h(\gamma_j(\tilde{t}_j))|| > \mu_j s_j^{p_{n-j-1}}, \quad \mu_j = 5^{-1}C16^{-p_{n-j-1}},$$

where $\gamma_j(\tilde{t}_j) = I(z_j^+) + \Pi_{j+1}(I(z_j^+(\tilde{t}_j)) - I(z_j^+))$. Let us define $v_{j+1} := \Pi_{j+1}\nabla h(\gamma_j(\tilde{t}_j))$. We have

$$||\gamma_j(\tilde{t}_j) - I(z_j)|| \le ||\Pi_{j+1}(I(z_j^+(\tilde{t}_j)) - I(z_j^+))|| + ||I(z_j^+) - I(z_j)|| \le s_j/8 + \sqrt{n}s_j/108,$$

where we used the estimate (3.13) (as $\tilde{t}_j \leq \tilde{t}_j^+$) and the estimate (3.10). Since $n \geq 2$ this implies

$$||\gamma_j(\tilde{t_j}) - I(z_j)|| \leq (\sqrt{n} - 1)s_j$$

which implies

$$||v_{j+1} - \Pi_{j+1}\nabla h(I(z_j))|| \leq F(\sqrt{n} - 1)s_j.$$

Now recall that by definition of z_j, we have

$$||\Pi_j \nabla h(I(z_j)) - \omega_j|| \leq s_j$$

and since $\omega_j \in \Lambda_{j+1}^{\perp}$ and $\Lambda_{j+1} \subseteq \Lambda_j$, $\Pi_{j+1}\omega_j = 0$ and $\Pi_{j+1} = \Pi_{j+1}\Pi_j$, and therefore

$$||\Pi_{j+1}\nabla h(I(z_j))|| = ||\Pi_{j+1}(\Pi_j(\nabla h(I(z_j))) - \omega_j)|| \leq s_j$$

which implies that

$$||v_{j+1}|| \leq ||v_{j+1} - \Pi_{j+1}\nabla h(I(z_j))|| + ||\Pi_{j+1}\nabla h(I(z_j))|| \leq F'\sqrt{n}s_j.$$

We just proved that

$$\mu_j s_j^{p_{n-j-1}} < ||v_{j+1}|| \leq F'\sqrt{n}s_j. \tag{3.19}$$

Now, for $Q \geq 1$, we apply Lemma 3.5 to v_{j+1}: there exists a T_{j+1}-periodic vector $\omega_{j+1} \in \mathbb{R}^n \setminus \{0\}$ such that

$$||v_{j+1} - \omega_{j+1}|| \leq \sqrt{n-1}(T_{j+1}Q)^{-1}, \quad ||v_{j+1}||^{-1} \leq T_{j+1} \leq \sqrt{n}||v_{j+1}||^{-1}Q^{n-1}.$$

Using (3.19), this implies

$$(F'\sqrt{n}s_j)^{-1} \leq ||v_{j+1}||^{-1} \leq T_{j+1} \leq \sqrt{n}||v_{j+1}||^{-1}Q^{n-1} < \sqrt{n}\mu_j^{-1}s_j^{-p_{n-j-1}}Q^{n-1} \tag{3.20}$$

and also, using the lower bound on T_{j+1},

$$||v_{j+1} - \omega_{j+1}|| \leq \sqrt{n-1}(T_{j+1}Q)^{-1} \leq \sqrt{n-1}||v_{j+1}||Q^{-1} \leq F'\sqrt{n-1}\sqrt{n}s_jQ^{-1}. \tag{3.21}$$

Let us prove that ω_{j+1} is linearly independent from $\omega_0, \ldots, \omega_j$, that is ω_{j+1} does not belong to $\Lambda_{j+1}^{\perp}$. To do this, it is enough to prove that if v is an arbitrary vector in $\Lambda_{j+1}^{\perp}$, then $|\omega_{j+1} \cdot v| < ||\omega_{j+1}||||v||$: indeed, otherwise, letting $v = \omega_{j+1}$, one would get a contradiction. On the one hand, we have

$$|\omega_{j+1} \cdot v| = |(\omega_{j+1} - v_{j+1}) \cdot v| \leq ||v_{j+1} - \omega_{j+1}||||v|| \leq \sqrt{n-1}Q^{-1}||v_{j+1}||||v||$$

where we used the fact $v_{j+1} \in \Lambda_{j+1}$ and (3.21), while, on the other hand,

$$||\omega_{j+1}||||v|| \geq (||v_{j+1}|| - ||v_{j+1} - \omega_{j+1}||)||v|| \geq (1 - \sqrt{n-1}Q^{-1})||v_{j+1}||||v|| > \sqrt{n-1}Q^{-1}||v_{j+1}||||v||$$

where we used the third inequality of (3.18) and (3.21). These last two inequalities imply that $|\omega_{j+1} \cdot v| < ||\omega_{j+1}||||v||$ for an arbitrary vector $v \in \Lambda_{j+1}^{\perp}$, and so ω_{j+1} is linearly independent from $\omega_0, \ldots, \omega_j$.

Next we define

$$z_{j+1} := z_j^+(\tilde{t}_j),$$

and observe that, since $\tilde{t}_j \le t_j^+$, by (3.8), $z_{j+1} \in V_{s_j,\xi_j}(z_j)$, but also $z_j^+(t) \in V_{s_j,\xi_j}(z_j)$ for $0 \le t \le \tilde{t}_j$, which justifies (3.17). Moreover, still from (3.8),

$$|I(z_{j+1}) - \gamma_j(\tilde{t}_j)| = |\Pi_{j+1}^{\perp}(I(z_{j+1}) - I(z_j^+))| \le s_j^{-1}\varepsilon$$

so

$$||I(z_{j+1}) - \gamma_j(\tilde{t}_j)|| \le \sqrt{n}s_j^{-1}\varepsilon$$

and hence

$$||\Pi_{j+1}\nabla h(I(z_{j+1})) - v_{j+1}|| \le F\sqrt{n}s_j^{-1}\varepsilon.$$

Therefore, using the first inequality of (3.21), the definition of s_{j+1} and the last inequality of (3.18),

$$\begin{aligned} ||\Pi_{j+1}\nabla h(I(z_{j+1})) - \omega_{j+1}|| &\le ||\Pi_{j+1}\nabla h(I(z_{j+1})) - v_{j+1}|| + ||v_{j+1} - \omega_{j+1}|| \\ &\le F\sqrt{n}s_j^{-1}\varepsilon + s_{j+1}/2 \le s_{j+1}, \end{aligned}$$

which proves (3.16). It remains to check (3.15), so let us fix $z \in \mathcal{V}_{3s_{j+1},3\xi_{j+1}}(z_{j+1})$. First, we have

$$I(z) - I(z_j) = I(z) - I(z_{j+1}) + I(z_{j+1}) - I(z_j) = I(z) - I(z_{j+1}) + I(z_j^+(\tilde{t}_j)) - I(z_j) \in \tilde{\Lambda}_j$$

since $I(z) - I(z_{j+1}) \in \tilde{\Lambda}_{j+1} \subseteq \tilde{\Lambda}_j$ and $I(z_j^+(\tilde{t}_j)) - I(z_j) \in \tilde{\Lambda}_j$. Then,

$$|I(z) - I(z_j)| \le |I(z) - I(z_{j+1})| + |I(z_{j+1}) - I(z_j)| < 3s_{j+1} + s_j/4 < 2s_j$$

provided that $s_{j+1} \le 7s_j/12$: but this inequality (in fact, the stronger inequality $s_{j+1} \le s_j/2$) follows from the definition of s_{j+1}, the third inequality of (3.18) and (3.21). Eventually,

$$||z|| < r_{j+1} + 3\xi_{j+1} = r_j + \xi_j + 3\xi_{j+1} = r_j + 2\xi_j$$

and so we showed that $\mathcal{V}_{3s_{j+1},3\xi_{j+1}}(z_{j+1}) \subseteq \mathcal{V}_{2s_j,2\xi_j}(z_j)$, which concludes the proof. □

3.6 One step of the algorithm

As a straightforward application of Proposition 3.1, Proposition 3.2 and Proposition 3.6, we now describe formally one step of the algorithm that will eventually lead to the proof of Theorem A.

Proposition 3.7. *For $0 \le j \le n-1$, let $H_j \in \widetilde{NF}_j(\omega_0, \dots, \omega_j, z_j, s_j, r_j, \xi_j, F)$ and $\Phi_j : \mathcal{V}_{2s_j,2\xi_j}(z_j) \to \mathcal{V}_{3s_j,3\xi_j}(z_j)$ given by Proposition 3.1, and let $z_j^+(t)$ be the forward solution of the Hamiltonian $H_j^+ = H_j \circ \Phi_j$ starting at $z_j^+ = \Phi_j^{-1}(z_j)$. Assume that h is $(r, \kappa, C, \delta, (p_l)_{l=1,\dots,n-1})$-steep. Then, for $0 \le j \le n-2$, we have the following dichotomy:*

(1) *either $z_j^+(t) \in V_{s_j,\xi_j}(z_j)$ for $0 \le t \le \bar{t}_j$,*

(2) *or, given a real number $Q \ge 1$, there exists a T_{j+1}-periodic vector $\omega_{j+1} \in \mathbb{R}^n \setminus \{0\}$, linearly independent from $\omega_0, \dots, \omega_j$, with the estimate*

$$(F'\sqrt{n}s_j)^{-1} < T_{j+1} < \sqrt{n}\mu_j^{-1}s_j^{-p_{n-j-1}}Q^{n-1}, \quad F' = \max\{1, F\}, \tag{3.22}$$

and there exists a time $\tilde{t}_j < \bar{t}_j$ *such that,*

$$z_{j+1} = z_j^+(\tilde{t}_j) \in B_{r_{j+1}}, \quad s_{j+1} = 2\sqrt{n-1}(T_{j+1}Q)^{-1}, \quad r_{j+1} = r_j + \xi_j, \quad \xi_{j+1} = \xi_j/3$$

such that $H_j^+ \in \widetilde{NF}_{j+1}(\omega_0, \dots, \omega_{j+1}, z_{j+1}, s_{j+1}, r_{j+1}, \xi_{j+1}, F, \varepsilon, m)$ *and*

$$z_j^+(t) \in V_{s_j,\xi_j}(z_j), \quad 0 \le t \le \tilde{t}_j, \tag{3.23}$$

provided that

$$\begin{cases} s_j < 8\min\{\delta, (3F)^{-1}\kappa, 2(5\kappa(4C)^{-1})^{1/p_{n-j-1}}\}, \\ Q \ge 8F'\sqrt{n(n-1)}, \\ \varepsilon \le (2\sqrt{n}F')^{-1}s_j s_{j+1}, \\ 2^{j+1}216(r_{j+1} + 3\xi_{j+1})mT_{j+1}\varepsilon \le s_{j+1}, \\ 72(3F\sqrt{n}+1)\xi_{j+1}^{-1}(r_{j+1} + 3\xi_{j+1})mT_{j+1}s_{j+1} \le 1. \end{cases} \tag{3.24}$$

If $j = n-1$*, then* $z_{n-1}^+(t) \in V_{s_{n-1},2\xi_{n-1}}(z_{n-1})$ *for* $0 \le t \le \bar{t}_{n-1}$*.*

Proof. The case $j = n-1$ follows directly from the case $j = n-1$ of Proposition 3.2. Then, we claim that the inequalities (3.24) imply the inequalities (3.18) and the inequalities (3.3) (with j replaced by $j+1$). Assuming this claim, and using Proposition 3.6, we have the inclusion of the complex domains of (3.15), and therefore using the second part of the statement of Proposition 3.1, we can assert that $H_j^+ \in NF_{j+1}$. Moreover, in view of (3.16),and since (3.3) is satisfied (with j replaced by $j+1$), we eventually obtain that $H_j^+ \in \widetilde{NF}_{j+1}$, while (3.23) is nothing but (3.17).

It remains to prove the claim. To do this, observe that (3.24) obviously implies (3.18) and (3.3), except for the following two inequalities:

$$\varepsilon < s_j^2/8, \quad s_{j+1} \le (r_{j+1} + 2\xi_{j+1})\xi_{j+1}. \tag{3.25}$$

But using the third inequality of (3.24) and the fact that $H_j \in \widetilde{NF}_j$, we know that

$$\varepsilon \le (2\sqrt{n}F')^{-1}s_j s_{j+1}, \quad s_j \le (r_j + 2\xi_j)\xi_j. \tag{3.26}$$

Then, using the second inequality of (3.24), one easily check that $s_{j+1} \le s_j/4$, and this, together with (3.26), imply (3.25), and the proof is over. □

3.7 Proof of Nekhoroshev exponential stability

We can finally give the proof of Theorem A. Recall that we are given a Hamiltonian H as in $(*)$, which is defined on $\mathcal{B}_r$, and of the form

$$H(z) = h(I(z)) + f(z), \quad h : \mathcal{D}_r \to \mathbb{C}, \quad f : \mathcal{B}_r \to \mathbb{C}$$

and that (2.5) holds true, that is

$$||\nabla h||_r \le E, \quad ||\nabla^2 h||_r \le F, \quad ||X_f||_r \le \varepsilon.$$

We already defined $F' = \max\{1, F\}$. Recall also that h is $(r, \kappa, C, \delta, (p_l)_{l=1,\dots,n-1})$-steep. Let us now define additional parameters: for any $0 \le j \le n-1$ and $0 \le k \le j$, we set

$$\pi_j^k := \prod_{n-j \le i \le n-j+k-1} p_i, \quad a_j^k := \sum_{0 \le i \le k} \pi_j^i,$$

with the convention that the product over the empty set is one, that is, $\pi_j^0 = 1$. Observe in particular that

$$a_0^0 = 1, \quad a_1^1 = 1 + p_{n-1},$$

and at the other extreme,

$$a_{n-2}^{n-2} = 1 + p_2 + p_2 p_3 + \cdots + p_2 p_3 \dots p_{n-1} = a',$$

$$a_{n-1}^{n-1} = 1 + p_1 + p_1 p_2 + \cdots + p_1 p_2 \dots p_{n-1} = a.$$

For $0 \le j \le n-2$, recalling that the numbers μ_j have been defined in (3.11), we define

$$\eta := \min_{0 \le j \le n-2} \{(3F)^{-1}\kappa, 2(5\kappa(4C)^{-1})^{1/p_{n-j-1}}\}, \quad \nu_j := \prod_{i=0}^{j-1} \mu_i^{\pi_j^{j-1-i}}.$$

The proof of Theorem A will be a consequence of the following proposition.

Proposition 3.8. *Let $H(z) = h(I(z)) + f(z)$ be as in $(*)$ satisfying (2.5), such that h is $(r, \kappa, C, \delta, (p_l)_{l=1,\dots,n-1})$-steep. Let z_0 be an arbitrary point in $B_{r/2}$ and $z(t)$ the forward solution of H starting at z_0. Given an integer $m \ge 1$ and a real number $Q \ge 1$, we have*

$$|I(z(t)) - I(z_0)| \le s := 3E\sqrt{n-1}Q^{-1}, \quad 0 \le t \le \bar{t} := 3(2rE\sqrt{n-1})^{-1}Q2^m,$$

provided that:

$$\begin{cases} Q \ge (5r^2)^{-1}36E\sqrt{n-1}, \\ Q > E\sqrt{n-1}(8\eta)^{-1}, \\ Q > E\sqrt{n-1}(8\delta)^{-1}, \\ Q \ge 8F'\sqrt{n(n-1)}, \\ 2\sqrt{n}F'\sqrt{n}^{a+a'}\sqrt{n-1}^{-(a+a')}\nu_{n-1}^{-1}\nu_{n-2}^{-1}\kappa^{-(\pi_{n-1}^{n-1}+\pi_{n-2}^{n-2})}Q^{n(a+a')}\varepsilon \le 1, \\ 2^{n-1}27(3+3^{-n+1})r(n-1)^{-a}n^a\sqrt{n-1}\kappa^{-2\pi_{n-1}^{n-1}}\nu_{n-1}^{-2}mQ^{2na-1}\varepsilon \le 1, \\ Q \ge m216(3^n+1)(3F\sqrt{n}+1)\sqrt{n-1}. \end{cases} \tag{C}$$

Let us first prove this proposition. The fact that this proposition implies Theorem A simply follows from a suitable choice of m and Q (in terms of our given parameters) and will be detailed later.

Proof of Proposition 3.8. The proof follows from an algorithm whose inductive step is given by Proposition 3.7. But first we need to initiate the algorithm. By assumptions we have

$$\kappa \le ||\nabla h(I(z_0))|| \le E$$

and so we can apply Lemma 3.5 to $v_0 := \nabla h(I(z_0))$: there exists a T_0-periodic vector $\omega_0 \in \mathbb{R}^n \setminus \{0\}$ such that

$$||v_0 - \omega_0|| \le \sqrt{n-1}(T_0 Q)^{-1}, \quad E^{-1} \le T_0 \le \sqrt{n}\kappa^{-1}Q^{n-1}. \tag{3.27}$$

We define

$$H_0 := H, \quad s_0 := \sqrt{n-1}(T_0 Q)^{-1}, \quad r_0 := r/2, \quad \xi_0 := r_0/3 = r/6,$$

and observe that $H_0 \in NF_0(\omega_{-1}, z_0, s_0, r_0, \xi_0, F, \varepsilon, m)$. Indeed, $\tilde{\lambda}_0 = \mathbb{C}^n$, $r_0 + 3\xi_0 = r$ so that $\mathcal{V}_{3s_0,3\xi_0}(z_0) \subseteq \mathcal{B}_r$, and we can write $H_0 = h + f = h + g_0 + f_0$, with $g_0 := f$ and $f_0 := 0$,

$$||X_{g_0}||_{3r_0,3\xi_0} \leq ||X_f||_s \leq \varepsilon,$$

as the requirement $\{l_{\omega_{-1}}, g_0\} = 0$ is void since $\omega_{-1} = 0$. In fact, using the first inequality of (3.27) and assuming that

$$\begin{cases} Q \geq (5r^2)^{-1} 36E\sqrt{n-1}, \\ 216n\sqrt{n-1}^{-1}\kappa^{-2} r m Q^{2n-1}\varepsilon \leq 1, \\ Q \geq m216.3(3F\sqrt{n}+1)\sqrt{n-1}, \end{cases} \tag{C0}$$

one easily check that, using the definitions of s_0, r_0, ξ_0 (which gives in particular $r_0 + 3\xi_0 = r$ and $(r_0+3\xi_0)\xi_0^{-1} = 6$) and the second estimate of (3.27), that $H_0 \in \widetilde{NF}_0(\omega_0, z_0, s_0, r_0, \xi_0, F, \varepsilon, m)$. So Proposition 3.7 can be applied.

If Alternative (1) of Proposition 3.7 holds true, the solution $z_0^+(t)$ of $H_0^+ = H_0 \circ \Phi_0$ satisfies $z_0^+(t) \in V_{s_0,\xi_0}(z_0)$ for $0 \leq t \leq \bar{t}_0$. As Φ_0 sends $V_{2s_0,2\xi_0}(z_0)$ into $V_{3s_0,3\xi_0}(z_0)$ and $\bar{t} \leq \bar{t}_0$, then $\Phi_0(z_0^+(t)) = z_0(t) = z(t)$ satisfies in particular

$$|I(z(t)) - I(z_0)| \leq 3s_0 \leq s, \quad 0 \leq t \leq \bar{t}, \tag{3.28}$$

the proposition is proved and the algorithm stops.

If Alternative (2) of Proposition 3.7 holds true, then there exist a T_1-periodic vector $\omega_1 \in \mathbb{R}^n \setminus \{0\}$, linearly independent from ω_0 with the estimate

$$(F'\sqrt{n}s_0)^{-1} < T_1 < \sqrt{n}\mu_0^{-1} s_0^{-p_{n-1}} Q^{n-1}, \tag{3.29}$$

and

$$z_1 = z_0^+(\tilde{t}_0) \in B_{r_1}, \quad s_1 = 2\sqrt{n-1}(T_1 Q)^{-1}, \quad r_1 = r_0 + \xi_0, \quad \xi_1 = \xi_0/3$$

such that $H_0^+ \in \widetilde{NF}_1(\omega_0, \omega_1, z_1, s_1, r_1, \xi_1, F, \varepsilon, m)$ and

$$z_0^+(t) \in V_{s_0,\xi_0}(z_0), \quad 0 \leq t \leq \tilde{t}_0, \tag{3.30}$$

provided that

$$\begin{cases} Q > E\sqrt{n-1}(8\eta)^{-1}, \\ Q > E\sqrt{n-1}(8\delta)^{-1}, \\ Q \geq 8F'\sqrt{n(n-1)}, \\ 2\sqrt{n}F'\sqrt{n}^{a_1^1+a_0^0}\sqrt{n-1}^{-(a_1^1+a_0^0)}\mu_0^{-1}\kappa^{-a_1^1}Q^{n(a_1^1+a_0^0)}\varepsilon \leq 1, \\ 180r(n-1)^{-a_1^1}n^{a_1^1}\sqrt{n-1}\kappa^{-2p_{n-1}}\mu_0^{-2} m Q^{2na_1^1-1}\varepsilon \leq 1, \\ Q \geq m216.10(3F\sqrt{n}+1)\sqrt{n-1}. \end{cases} \tag{C1}$$

Indeed, using the definitions of s_0, s_1, r_0, r_1, ξ_0, ξ_1 (in particular, we use the facts that $s_1 \leq s_0$, $s_1 \geq \sqrt{n-1}(T_1 Q)^{-1}$, $r_1 + 3\xi_1 = 5r/6$ and $(r_1+3\xi_1)\xi_1^{-1} = 15$) and the estimate (3.27)

and (3.29) on respectively T_0 and T_1, one can check that (C1) imply (3.24) for $j = 1$. Setting $H_1 := H_0^+ \in \widetilde{NF}_1(\omega_0, \omega_1, z_1, s_1, r_1, \xi_1, F, \varepsilon, m)$, we can apply Proposition 3.7 again.

If Alternative (1) holds true, then the solution $z_1^+(t)$ of $H_1^+ = H_1 \circ \Phi_1 = H_0^+ \circ \Phi_1 = H_0 \circ \Phi_0 \circ \Phi_1$ starting at $z_1^+ = \Phi_1^{-1}(z_1)$ satisfies $z_1^+(t) \in V_{s_1,\xi_1}(z_1)$ for $0 \le t \le \bar{t}_1$. As Φ_1 sends $V_{2s_1,2\xi_1}(z_1)$ into $V_{3s_1,3\xi_1}(z_1)$, then $\Phi_1(z_1^+(t)) = z_1(t)$ belongs to $V_{3s_1,3\xi_1}(z_1)$ for $0 \le t \le \bar{t}_1$. By (3.15), $V_{3s_1,3\xi_1}(z_1)$ is contained in $V_{2s_0,2\xi_0}(z_0)$, and as $\bar{t} \le \bar{t}_1$, $z_1(t)$ belongs to $V_{2s_0,2\xi_0}(z_0)$ for $0 \le t \le \bar{t}$. Now observe that since $z_1 = z_0^+(\tilde{t}_0)$, by uniqueness of the solutions associated to the system defined by $H_1 = H_0^+$, we have the equality $z_1(t) = z_0^+(t + \tilde{t}_0)$ as long as the solution is defined. Using this equality, what we have proved is that

$$z_0^+(t) \in V_{2s_0,2\xi_0}(z_0), \quad \tilde{t}_0 \le t \le \tilde{t}_0 + \bar{t}$$

But recall that from (3.30), we know that

$$z_0^+(t) \in V_{s_0,\xi_0}(z_0), \quad 0 \le t \le \tilde{t}_0,$$

and therefore, since $\bar{t} < \tilde{t}_0 + \bar{t}$, we have in particular

$$z_0^+(t) \in V_{2s_0,2\xi_0}(z_0), \quad 0 \le t \le \bar{t}.$$

As before, using this and the fact that Φ_0 sends $V_{2s_0,2\xi_0}(z_0)$ into $V_{3s_0,3\xi_0}(z_0)$ we also arrive at the estimate (3.28).

If Alternative (2) holds true, then the algorithm continues. To apply Proposition 3.7 at a step j, for $1 \le j \le n-1$, it is sufficient to check that

$$\begin{cases} Q > E\sqrt{n-1}(8\eta)^{-1}, \\ Q > E\sqrt{n-1}(8\delta)^{-1}, \\ Q \ge 8F'\sqrt{n(n-1)}, \\ 2\sqrt{n}F'\sqrt{n}^{a_j^j + a_{j-1}^{j-1}}\sqrt{n-1}^{-(a_j^j + a_{j-1}^{j-1})}\nu_j^{-1}\nu_{j-1}^{-1}\kappa^{-(\pi_j^j + \pi_{j-1}^{j-1})}Q^{n(a_j^j + a_{j-1}^{j-1})}\varepsilon \le 1, \\ 2^j 27(3 + 3^{-j})r(n-1)^{-a_j^j}n^{a_j^j}\sqrt{n-1}\kappa^{-2\pi_j^j}\nu_j^{-2}mQ^{2na_j^j - 1}\varepsilon \le 1, \\ Q \ge m216(3^{j+1} + 1)(3F\sqrt{n} + 1)\sqrt{n-1}. \end{cases} \tag{Cj}$$

Indeed, (Cj) implies (3.24), using the definitions of s_i, r_i and ξ_i for $0 \le i \le j$ (which imply in particular that the s_i are decreasing, $s_i \ge \sqrt{n-1}(T_iQ)^{-1}$, $r_i + 3\xi_i = r(3 + 3^{-i})/4$ and $(r_i + 3\xi_i)\xi_i^{-1} = 3(3^{i+1} + 1)/2$), and the estimates on the period T_i that one obtains at each step using (3.22). To conclude, just observe that the conditions (C) imply the conditions (C0) and (Cj) for any $1 \le j \le n-1$. For $j = n-1$, there is only one possibility in Proposition (3.7), the algorithm stops and the statement is proved. □

Proof of Theorem A. We just need to choose m and Q in Proposition 3.8 in terms of our given parameters. First we choose m in terms of Q as follows:

$$m := [b_1Q], \quad b_1 = (216(3F\sqrt{n} + 1)(3^n + 1)\sqrt{n-1})^{-1}$$

where $[\,\cdot\,]$ denotes the integer part. Using this choice, the conditions (C) are implied by

$$Q \ge b_2, \quad Q \ge b_3\delta^{-1}, \quad Q \ge b_4r^{-2}, \quad rb_5Q^{2na}\varepsilon \le 1, \quad b_6Q^{n(a+a')}\varepsilon \le 1, \tag{3.31}$$

where

$$b_2 := \max\{8F'\sqrt{n(n-1)}, E\sqrt{n-1}(8\eta)^{-1}, b_1^{-1}\}$$
$$b_3 := E\sqrt{n-1}8^{-1}$$
$$b_4 := 5^{-1}36E\sqrt{n-1}$$
$$b_5 := 2^{n-1}27(3+3^{-n+1})n^a(n-1)^{-a}\sqrt{n-1}\nu_{n-1}^{-2}\kappa^{-2\pi_{n-1}^{n-1}}b_1$$
$$b_6 := 2\sqrt{n}F'\sqrt{n}^{a+a'}\sqrt{n-1}^{-(a+a')}\nu_{n-1}^{-1}\nu_{n-2}^{-1}\kappa^{-(\pi_{n-1}^{n-1}+\pi_{n-2}^{n-2})}.$$

Then we choose Q as follows:

$$Q := (b_5 r\varepsilon)^{-\frac{1}{2na}}$$

and observe that (3.31) becomes

$$r\varepsilon \le b_5^{-1}b_2^{-2na}, \quad r\varepsilon \le b_5^{-1}b_3^{-2na}\delta^{2na}, \quad r\varepsilon \le b_5^{-1}b_4^{-2na}r^{4na}, \quad r\varepsilon \le b_6^{-\frac{2a}{a-a'}}b_5^{\frac{a+a'}{a-a'}}r^{\frac{2a}{a-a'}}. \tag{3.32}$$

With these choices of m and Q, since $m > b_1Q - 1$ we have

$$s = 3E\sqrt{n-1}b_5^{\frac{1}{2na}}(r\varepsilon)^{\frac{1}{2na}}$$

and

$$\bar{t} \ge 3(4E\sqrt{n-1})^{-1}b_5^{-\frac{1}{2na}}r^{-1}(r\varepsilon)^{-\frac{1}{2na}}\exp\left((\ln 2)b_1 b_5^{-\frac{1}{2na}}r^{-1}(r\varepsilon)^{-\frac{1}{2na}}\right)$$

so if we define

$$\tilde{c}_1 := b_5^{-1}b_2^{-2na}, \quad \tilde{c}_2 := b_5^{-1}b_3^{-2na}, \quad \tilde{c}_3 := b_5^{-1}b_4^{-2na}, \quad \tilde{c}_4 := b_6^{-\frac{2a}{a-a'}}b_5^{\frac{a+a'}{a-a'}}$$

and

$$\tilde{c}_5 := 2E\sqrt{n-1}b_5^{\frac{1}{2na}}, \quad \tilde{c}_6 := 3(4E\sqrt{n-1})^{-1}b_5^{-\frac{1}{2na}}, \quad \tilde{c}_7 := (\ln 2)b_1 b_5^{-\frac{1}{2na}}$$

we eventually obtain that if

$$r\varepsilon \le \min\left\{\tilde{c}_1, \tilde{c}_2\delta^{2na}, \tilde{c}_3 r^{4na}, \tilde{c}_4 r^{\frac{2a}{a-a'}}\right\} \tag{3.33}$$

then

$$|I(z(t)) - I(z_0)| \le \tilde{c}_5(r\varepsilon)^{\frac{1}{2na}}, \quad 0 \le t \le \tilde{c}_6 r^{-1}(r\varepsilon)^{-\frac{1}{2na}}\exp\left(\tilde{c}_7 r^{-1}(r\varepsilon)^{-\frac{1}{2na}}\right).$$

This proves the statement for positive times, but for negative times, the proof is of course the same, so this concludes the proof. □

A Proof of the normal form statement

A.1 Technical estimates

We first derive technical estimates for real-analytic vector fields defined on certain domains in $\mathbb{C}^{2n}$. These estimates are stated and proved for Hamiltonian vector fields, even though the Hamiltonian character plays absolutely no role here.

For $0 \leq j \leq n-1$, recall that $\omega_j \in \mathbb{R}^n \setminus \{0\}$ are T_j-periodic vectors, and that $\omega_{-1} = 0 \in \mathbb{R}^n$. We write $l_{\omega_j}(z) = \omega_j \cdot I(z)$, for $1 \leq j \leq n-1$, and we define the complex vector space

$$\tilde{\Lambda}_j = \{v \in \mathbb{C}^n \mid v \cdot \omega_{-1} = v \cdot \omega_0 = \cdots = v \cdot \omega_{j-1} = 0\}.$$

Then we consider three sequences of positive real numbers r_j, ξ_j and s_j, a sequence of points $z_j \in B_{r_j}$ and we let $\tilde{\lambda}_j = I(z_j) + \tilde{\Lambda}_j$ be the complex affine subspace associated to $\tilde{\Lambda}_j$ passing through $I(z_j)$. The complex domains we consider are given by

$$\mathcal{V}_{s_j,\xi_j}(z_j) = \{z \in \mathbb{C}^{2n} \mid I(z) \in \tilde{\lambda}_j,\ |I(z) - I(z_j)| < s_j,\ ||z|| < r_j + \xi_j\}.$$

We fix $0 < \sigma_j < s_j$ and $0 < \rho_j < \xi_j$, and a real-analytic Hamiltonian vector field X_{χ_j} defined on $\mathcal{V}_{s_j,\xi_j}(z_j)$. Throughout this section, we will make the following two assumptions:

$$\sigma_j \leq (r_j + \xi_j)\rho_j, \quad \{l_{\omega_{-1}}, \chi_j\} = \{l_{\omega_{-0}}, \chi_j\} = \cdots = \{l_{\omega_{j-1}}, \chi_j\} = 0. \tag{A.1}$$

Lemma A.1. *Assume that* (A.1) *is satisfied. Then* $X^t_{\chi_j} : \mathcal{V}_{s_j-\sigma_j,\xi_j-\rho_j}(z_j) \to \mathcal{V}_{s_j,\xi_j}(z_j)$ *is a well-defined symplectic real-analytic embedding for all* $|t| \leq \tau_j = (r_j+\xi_j)^{-1}\sigma_j||X_{\chi_j}||^{-1}_{s_j,\xi_j}$, *with the estimate* $||X^t_{\chi_j} - \mathrm{Id}||_{s_j-\sigma_j,\xi_j-\rho_j} \leq |t|||X_{\chi_j}||_{s_j,\xi_j}$.

Proof. Let $z \in \mathcal{V}_{s_j-\sigma_j,\xi_j-\rho_j}(z_j)$ and $z(t) = X^t_{\chi_j}(z)$ for small $|t|$, and let $|s| \leq |t|$. Since

$$\{l_{\omega_l}, \chi_j\}(z(s)) = \omega_l \cdot (\{I_1, \chi_j\}(z(s)), \ldots, \{I_n, \chi_j\}(z(s))) := \omega_l \cdot \{I, \chi_j\}(z(s))$$

for $-1 \leq l \leq j-1$, the second part of (A.1) implies that

$$\omega_{-1} \cdot \{I, \chi_j\}(z(s)) = \omega_0 \cdot \{I, \chi_j\}(z(s)) = \cdots = \omega_{j-1} \cdot \{I, \chi_j\}(z(s)) = 0,$$

so $\{I, \chi_j\}(z(s)) \in \tilde{\Lambda}_j$ which implies that

$$\frac{d}{ds} I(z(s)) = \{I, \chi_j\}(z(s)) \in \tilde{\Lambda}_j$$

and therefore

$$I(z(t)) = I(z) + \int_0^t \frac{d}{ds} I(z(s))ds = I(z) - I(z_j) + I(z_j) + \int_0^t \frac{d}{ds} I(z(s))ds \in \tilde{\lambda}_j.$$

Then, using the first part of (A.1), for

$$|t| \leq \min\{\rho_j, (r_j + \xi_j)^{-1}\sigma_j\}||X_{\chi_j}||^{-1}_{s_j,\xi_j} = (r_j + \xi_j)^{-1}\sigma_j||X_{\chi_j}||^{-1}_{s_j,\xi_j} = \tau_j,$$

we have

$$||z(t) - z|| \leq |t|||X_{\chi_j}||_{r_j,\xi_j} \leq \rho_j$$

and, using Cauchy-Schwarz inequality,

$$|I(z(t)) - I(z)| \leq 2^{-1}||z(t) + z||||z(t) - z|| \leq (r_j + \xi_j)||z(t) - z|| \leq (r_j + \xi_j)|t|||X_{\chi_j}||_{s_j,\xi_j} \leq \sigma_j.$$

This proves that $X^t_{\chi_j} : \mathcal{V}_{s_j-\sigma_j,\xi_j-\rho_j}(z_j) \to \mathcal{V}_{s_j,\xi_j}(z_j)$ is a well-defined symplectic real-analytic embedding for $|t| \leq \tau_j$, with the estimate $||X^t_{\chi_j} - \mathrm{Id}||_{s_j-\sigma_j,\xi_j-\rho_j} \leq |t|||X_{\chi_j}||_{s_j,\xi_j}$. □

Lemma A.2. *Assume that* (A.1) *is satisfied, and let X_f be a real-analytic Hamiltonian vector field defined on $\mathcal{V}_{s_j,\xi_j}(z_j)$. Then, for $|t| \leq \tau_j/3 = (3(r_j+\xi_j))^{-1}\sigma_j||X_{\chi_j}||^{-1}_{s_j,\xi_j}$, we have*

$$||(X^t_{\chi_j})^* X_f||_{s_j-\sigma_j,\xi_j-\rho_j} \leq \left(1+3(r_j+\xi_j)\sigma_j^{-1}|t|||X_{\chi_j}||_{s_j,\xi_j}\right)||X_f||_{s_j-2\sigma_j/3,\xi_j-2\rho_j/3}$$

and therefore

$$||(X^t_{\chi_j})^* X_f||_{s_j-\sigma_j,\xi_j-\rho_j} \leq 2||X_f||_{s_j-2\sigma_j/3,\xi_j-2\rho_j/3}.$$

Proof. Let $|t| \leq \tau_j/3 = (3(r_j+\xi_j))^{-1}\sigma_j||X_{\chi_j}||^{-1}_{s_j,\xi_j}$. We have the following expression

$$(X^t_{\chi_j})^* X_f = (DX^{-t}_{\chi_j} \circ X^t_{\chi_j}).(X_f \circ X^t_{\chi_j}) = \left(DX^{-t}_{\chi_j} \circ X^t_{\chi_j} - \mathrm{Id}\right).(X_f \circ X^t_{\chi_j}) + X_f \circ X^t_{\chi_j}.$$

Lemma A.1 implies that $X^t_{\chi_j} : \mathcal{V}_{s_j-\sigma_j,\xi_j-\rho_j}(z_j) \to \mathcal{V}_{s_j-2\sigma_j/3,\xi_j-2\rho_j/3}(z_j)$ hence

$$||DX^{-t}_{\chi_j} \circ X^t_{\chi_j} - \mathrm{Id}||_{s_j-\sigma_j,\xi_j-\rho_j} \leq ||DX^{-t}_{\chi_j} - \mathrm{Id}||_{s_j-2\sigma_j/3,\xi_j-2\rho_j/3}.$$

We claim that

$$||DX^{-t}_{\chi_j} - \mathrm{Id}||_{s_j-2\sigma_j/3,\xi_j-2\rho_j/3} \leq 3(r_j+\xi_j)\sigma_j^{-1}||X^{-t}_{\chi_j} - \mathrm{Id}||_{s_j-\sigma_j/3,\xi_j-\rho_j/3}$$

while obviously, using Lemma A.1,

$$||X^{-t}_{\chi_j} - \mathrm{Id}||_{s_j-\sigma_j/3,\xi_j-\rho_j/3} = ||X^t_{-\chi_j} - \mathrm{Id}||_{s_j-\sigma_j/3,\xi_j-\rho_j/3} \leq |t|||X_{-\chi_j}||_{s_j,\xi_j} = |t|||X_{\chi_j}||_{s_j,\xi_j}.$$

Assuming this claim, using the expression for $(X^t_{\chi_j})^* X_f$ and putting all the estimates together, we arrive at

$$||(X^t_{\chi_j})^* X_f||_{s_j-\sigma_j,\xi_j-\rho_j} \leq \left(1+3(r_j+\xi_j)\sigma_j^{-1}|t|||X_{\chi_j}||_{s_j,\xi_j}\right)||X_f \circ X^t_{\chi_j}||_{s_j-\sigma_j,\xi_j-\rho_j},$$

therefore

$$||(X^t_{\chi_j})^* X_f||_{s_j-\sigma_j,\xi_j-\rho_j} \leq \left(1+3(r_j+\xi_j)\sigma_j^{-1}|t|||X_{\chi_j}||_{s_j,\xi_j}\right)||X_f||_{s_j-2\sigma_j/3,\xi_j-2\rho_j/3}$$

and also

$$||(X^t_{\chi_j})^* X_f||_{s_j-\sigma_j,\xi_j-\rho_j} \leq 2||X_f||_{s_j-2\sigma_j/3,\xi_j-2\rho_j/3}$$

since $|t| \leq \tau_j/3 = (3(r_j+\xi_j))^{-1}\sigma_j|X_{\chi_j}|^{-1}_{s_j,\xi_j}$. It remains to prove the claim. Let $F = X^{-t}_{\chi_j} - \mathrm{Id}$, $z \in \mathcal{V}_{s_j-2\sigma_j/3,\xi_j-2\rho_j/3}(z_j)$ and $v \in \mathbb{C}^{2n}$ a unit vector. The map

$$\xi \in \mathbb{C} \mapsto F_{z,v}(\xi) = F(z+\xi v) \in \mathbb{C}^{2n}$$

is holomorphic for $|\xi| \leq (3(r_j+\xi_j))^{-1}\sigma_j \leq \rho_j/3$, with $z+\xi v \in \mathcal{V}_{s_j-\sigma_j/3,\xi_j-\rho_j/3}(z_j)$. The usual Cauchy's estimate implies that

$$||DF(z)|| = \sup_{||v||=1} ||DF(z).v|| = \sup_{||v||=1} ||F'_{z,v}(0)|| \leq 3(r_j+\xi_j)\sigma_j^{-1} \sup_{|\xi|\leq(3(r_j+\xi_j))^{-1}\sigma_j} ||F_{z,v}(\xi)||$$

hence

$$||DF(z)|| \leq 3(r_j+\xi_j)\sigma_j^{-1}||F||_{s_j-\sigma_j/3,\xi_j-\rho_j/3}$$

and the claim follows since $z \in \mathcal{V}_{s_j-2\sigma_j/3,\xi_j-2\rho_j/3}(z_j)$ was arbitrary. □

Lemma A.3. *Assume that* (A.1) *is satisfied, and let* X_f *be a real-analytic Hamiltonian vector field defined on* $\mathcal{V}_{s_j,\xi_j}(z_j)$*. Then*

$$||[X_f, X_{\chi_j}]||_{s_j-2\sigma_j/3,\xi_j-2\rho_j/3} \leq 9(r_j+\xi_j)\sigma_j^{-1}||X_{\chi_j}||_{s_j,\xi_j}||X_f||_{s_j,\xi_j}.$$

Proof. We have the expression

$$[X_f, X_{\chi_j}] = \frac{d}{dt}(X_{\chi_j}^t)^* X_f\Big|_{t=0}$$

so for $z \in \mathcal{V}_{s_j-\sigma_j,\xi_j-\rho_j}(z_j)$, let us define the holomorphic map

$$t \in \mathbb{C} \mapsto F_z(t) = (X_{\chi_j}^t)^* X_f(z) \in \mathbb{C}^{2n}$$

for $|t| \leq \tau_j/3 = (3(r_j+\xi_j))^{-1}\sigma_j||X_{\chi_j}||_{s_j,\xi_j}^{-1}$. By Cauchy's estimate

$$||[X_f, X_{\chi_j}](z)|| = ||F_z'(0)|| \leq 3\tau_j^{-1}||F_z(t)|| \leq 3\tau_j^{-1}||(X_{\chi_j}^t)^* X_f||_{s_j-\sigma_j,\xi_j-\rho_j}$$

and by Lemma A.2

$$3\tau_j^{-1}||(X_{\chi_j}^t)^* X_f||_{s_j-\sigma_j,\xi_j-\rho_j} \leq 6\tau_j^{-1}||X_f||_{s_j-2\sigma_j/3,\xi_j-2\rho_j/3} \leq 6(r_j+\xi_j)\sigma_j^{-1}||X_{\chi_j}||_{s_j,\xi_j}||X_f||_{s_j,\xi_j}.$$

Since $z \in \mathcal{V}_{s_j-\sigma_j,\xi_j-\rho_j}(z_j)$ was arbitrary, this proves that

$$||[X_f, X_{\chi_j}]||_{s_j-\sigma_j,\xi_j-\rho_j} \leq 6(r_j+\xi_j)\sigma_j^{-1}||X_{\chi_j}||_{s_j,\xi_j}||X_f||_{s_j,\xi_j}$$

and the lemma follows by simply replacing σ_j and ρ_j by respectively $2\sigma_j/3$ and $2\rho_j/3$. □

Lemma A.4. *Assume that* (A.1) *is satisfied, and let* X_k *be a real-analytic Hamiltonian vector field defined on* $\mathcal{V}_{r_j,\xi_j}(z_j)$*, which is integrable, that is* k *is a function of* $I(z)$ *alone. Then*

$$||[X_k, X_{\chi_j}]||_{s_j-2\sigma_j/3,\xi_j-2\rho_j/3} \leq 9\rho_j^{-1}||X_{\chi_j}||_{s_j,\xi_j}||X_k||_{s_j,\xi_j}.$$

Proof. Here we write

$$[X_k, X_{\chi_j}] = -[X_{\chi_j}, X_k] = -\frac{d}{dt}(X_k^t)^* X_{\chi_j}\Big|_{t=0}$$

and we observe that since X_k is integrable, if we let $z(t) = X_k^t(z)$, then $I(z(t)) = I(z)$. This implies that $X_k^t : \mathcal{D}_{s_j,\xi_j-\rho_j}(z_j) \to \mathcal{D}_{s_j,\xi_j}(z_j)$ is a well-defined symplectic real-analytic embedding for all $|t| \leq \tau_j' = \rho_j||X_k||_{s_j,\xi_j}^{-1}$. The conclusion follows easily by repeating all the previous arguments with τ_j' instead of τ_j. □

A.2 Proof of the Proposition 3.1

Proposition 3.1 will be proved by iterating m times an averaging procedure, which is classical in the case $j = 0$, but more involved in the general case.

Proof of Proposition 3.1. Let us fix $0 \leq j \leq n-1$, and set $\varepsilon_j := 2^j\varepsilon$. The integer $m \geq 1$ being given, for $0 \leq i \leq m$ we define

$$\varepsilon_j^i := 2^{-i}\varepsilon_j, \quad \gamma_j^i := (1-2^{-i})2\varepsilon_j, \quad s_j^i := 3s_j - is_j/m, \quad \xi_j^i := 3\xi_j - i\xi_j/m.$$

Then we claim that for each $0 \leq i \leq m$, there exists a real-analytic symplectic embedding $\Phi_j^i : \mathcal{V}_{s_j^i,\xi_j^i}(z_j) \to \mathcal{V}_{3s_j,3\xi_j}(z_j)$ such that

$$(H_j - f_j) \circ \Phi_j^i = (h + g_j) \circ \Phi_j^i = h + g_j^i + f_j^i$$

with

$$\{l_{\omega_{-1}}, g_j^i\} = \{l_{\omega_0}, g_j^i\} = \cdots = \{l_{\omega_j}, g_j^i\} = 0, \quad \{l_{\omega_{-1}}, f_j^i\} = \{l_{\omega_0}, f_j^i\} = \cdots = \{l_{\omega_{j-1}}, f_j^i\} = 0$$

and with the estimates

$$||X_{g_j^i}||_{s_j^i,\xi_j^i} \leq \gamma_j^i, \quad ||X_{f_j^i}||_{s_j^i,\xi_j^i} \leq \varepsilon_j^i, \quad ||\Phi_j^i - \mathrm{Id}||_{s_j^i,\xi_j^i} \leq T_j\gamma_j^i.$$

Let us prove the claim by induction on $0 \leq i \leq m$.

For $i = 0$, letting Φ_j^0 be the identity, $g_j^0 := 0$ and $f_j^0 := g_j$, there is nothing to prove. Then assume that the statement holds true for some $0 \leq i \leq m-1$, and let $H_j^i = (H_j - f_j) \circ \Phi_j^i = h + g_j^i + f_j^i$. We define the functions

$$[f_j^i]_j := T_j^{-1}\int_0^{T_j} f_j^i \circ X_{\omega_j}^t dt, \quad \chi_j^i := T_j^{-1}\int_0^{T_j} t(f_j^i - [f_j^i]_j) \circ X_{\omega_j}^t dt$$

whose associated Hamiltonian vector fields are given by

$$X_{[f_j^i]_j} = T_j^{-1}\int_0^{T_j} (X_{\omega_j}^t)^* f_j^i dt, \quad X_{\chi_j^i} = T_j^{-1}\int_0^{T_j} t(X_{\omega_j}^t)^*(f_j^i - [f_j^i]_j)dt$$

with, using our inductive assumption, the following obvious estimates

$$||X_{[f_j^i]_j}||_{s_j^i,\xi_j^i} \leq ||X_{f_j^i}||_{s_j^i,\xi_j^i} \leq \varepsilon_j^i, \quad ||X_{\chi_j^i}||_{s_j^i,\xi_j^i} \leq T_j||X_{f_j^i}||_{s_j^i,\xi_j^i} \leq T_j\varepsilon_j^i. \tag{A.2}$$

It is clear that

$$\{[f_j^i]_j, l_{\omega_j}\} = 0. \tag{A.3}$$

For $-1 \leq l \leq j-1$, observe that $\{l_{\omega_l}, l_{\omega_j}\} = 0$, so $l_{\omega_l} \circ X_{\omega_j}^t = l_{\omega_l}$ and hence

$$\begin{aligned}
\{l_{\omega_l}, [f_j^i]_j\} &= T_j^{-1}\int_0^{T_j} \{l_{\omega_l}, f_j^i \circ X_{\omega_j}^t\}dt \\
&= T_j^{-1}\int_0^{T_j} \{l_{\omega_l} \circ X_{\omega_j}^t, f_j^i \circ X_{\omega_j}^t\}dt \\
&= T_j^{-1}\int_0^{T_j} \{l_{\omega_l}, f_j^i\} \circ X_{\omega_j}^t dt
\end{aligned}$$

where the last equality follows from the symplectic character of $X_{\omega_j}^t$. Using our inductive assumption, this implies

$$\{l_{\omega_{-1}}, [f_j^i]_j\} = \{l_{\omega_0}, [f_j^i]_j\} = \cdots = \{l_{\omega_{j-1}}, [f_j^i]_j\} = 0, \tag{A.4}$$

and by a completely similar argument, we also get

$$\{l_{\omega_{-1}}, \chi_j^i\} = \{l_{\omega_0}, \chi_j^i\} = \cdots = \{l_{\omega_{j-1}}, \chi_j^i\} = 0. \tag{A.5}$$

Now set

$$\sigma_j := s_j^i - s_j^{i+1} = s_j/m, \quad \rho_j := \xi_j^i - \xi_j^{i+1} = \xi_j/m.$$

Since $\xi_j^i \geq 2\xi_j$, using the first inequality of (3.3) we have

$$\sigma_j = s_j/m \leq (r_j + 2\xi_j)\xi_j/m \leq (r_j + \xi_j^i)\xi_j/m = (r_j + \xi_j^i)\rho_j$$

and therefore, using also (A.5), we can apply Lemma A.1:

$$X^t_{\chi_j^i} : \mathcal{V}_{s_j^{i+1},\xi_j^{i+1}}(z_j) = \mathcal{V}_{s_j^i-\sigma_j,\xi_j^i-\rho_j}(z_j) \to \mathcal{V}_{s_j^i,\xi_j^i}(z_j)$$

is a well-defined symplectic real-analytic embedding for all

$$|t| \leq \tau_j = (r_j + \xi_j^i)^{-1}\sigma_j ||X_{\chi_j^i}||^{-1}_{s_j^i,\xi_j^i} = (m(r_j + \xi_j^i))^{-1} s_j ||X_{\chi_j^i}||^{-1}_{s_j^i,\xi_j^i},$$

with the estimate $||X^t_{\chi_j^i} - \mathrm{Id}||_{s_j^{i+1},\xi_j^{i+1}} \leq |t|||X_{\chi_j^i}||_{s_j^i,\xi_j^i}$. Moreover, as $\xi_j^i \leq 3\xi_j$, using the second estimate of (A.2) and the second inequality of (3.3), we have

$$\tau_j \geq (m(r_j+3\xi_j)T_j\varepsilon_j^i)^{-1}s_j = 2^i(m(r_j+3\xi_j)T_j 2^j\varepsilon)^{-1}s_j \geq (m(r_j+3\xi_j)T_j 2^j\varepsilon)^{-1}s_j \geq 216 \tag{A.6}$$

so $\tau_j > 1$ and hence $X^1_{\chi_j^i} : \mathcal{V}_{s_j^{i+1},\xi_j^{i+1}}(z_j) \to \mathcal{V}_{s_j^i,\xi_j^i}(z_j)$ is well-defined, with

$$||X^1_{\chi_j^i} - \mathrm{Id}||_{s_j^{i+1},\xi_j^{i+1}} \leq ||X_{\chi_j^i}||_{s_j^i,\xi_j^i} \leq T_j\varepsilon_j^i. \tag{A.7}$$

It is easy to check, using an integration by parts, that $\{\chi_j^i, l_{\omega_j}\} = f_j^i - [f_j^i]_j$, and this equality, together with Taylor's formula with integral remainder gives

$$(h + g_j^i + f_j^i) \circ X^1_{\chi_j^i} = h + g_j^i + [f_j^i]_j + \tilde{f}_j^i$$

with

$$\tilde{f}_j^i = \int_0^1 \{(h - l_{\omega_j}) + g_j^i + f_{j,t}^i, \chi_j^i\} \circ X^t_{\chi_j^i} dt, \quad f_{j,t}^i = tf_j^i + (1-t)[f_j^i]_j.$$

We set $\Phi_j^{i+1} = \Phi_j^i \circ X^1_{\chi_j^i}$, $g_j^{i+1} = g_j^i + [f_j^i]_j$ and $f_j^{i+1} = \tilde{f}_j^i$ so that

$$(H_j - f_j) \circ \Phi_j^{i+1} = H_j^i \circ X^1_{\chi_j^i} = (h + g_j^i + f_j^i) \circ X^1_{\chi_j^i} = h + g_j^{i+1} + f_j^{i+1}.$$

First observe that $\Phi_j^{i+1} : \mathcal{V}_{s_j^{i+1},\xi_j^{i+1}}(z_j) \to \mathcal{D}_{3s_j,3\xi_j}(z_j)$ is a real-analytic symplectic embedding, and using (A.7) together with our inductive assumption, we have the estimate

$$||\Phi_j^{i+1} - \mathrm{Id}||_{s_j^{i+1},\xi_j^{i+1}} \leq ||\Phi_j^i - \mathrm{Id}||_{s_j^i,\xi_j^i} + ||X^1_{\chi_j^i} - \mathrm{Id}||_{s_j^{i+1},\xi_j^{i+1}} \leq T_j(\gamma_j^i + \varepsilon_j^i) = T_j\gamma_j^{i+1}.$$

Then

$$\{l_{\omega_{-1}}, g_j^{i+1}\} = \{l_{\omega_0}, g_j^{i+1}\} = \cdots = \{l_{\omega_j}, g_j^{i+1}\} = 0$$

follows from the definition of g_j^{i+1}, the inductive assumption, (A.3) and (A.4). Moreover, using the first estimate of (A.2) and our inductive assumption,

$$||X_{g_j^{i+1}}||_{s_j^{i+1},\xi_j^{i+1}} \leq ||X_{g_j^i}||_{s_j^i,\xi_j^i} + ||X_{[f_j^i]_j}||_{s_j^i,\xi_j^i} \leq \gamma_j^i + \varepsilon_j^i = \gamma_j^{i+1}.$$

For $-1 \leq l \leq j-1$, we already know that

$$\{l_{\omega_l}, g_j^i\} = \{l_{\omega_l}, f_j^i\} = \{l_{\omega_l}, [f_j^i]_j\} = \{l_{\omega_l}, \chi_j^i\} = 0$$

which implies $\{l_{\omega_l}, f_{j,t}\} = 0$ whereas $\{l_{\omega_l}, h - l_{\omega_j}\} = 0$ is obvious. These equalities, together with Jacobi identity, imply that

$$\{l_{\omega_l}, \{(h - l_{\omega_j}) + g_j^i + f_{j,t}^i, \chi_j^i\}\} = 0$$

and therefore

$$\{l_{\omega_l}, f_j^{i+1}\} = \int_0^1 \{l_{\omega_l}, \{(h - l_{\omega_j}) + g_j^i + f_{j,t}^i, \chi_j^i\}\} \circ X_{\chi_j^i}^t dt.$$

It follows that

$$\{l_{\omega_{-1}}, f_j^{i+1}\} = \{l_{\omega_0}, f_j^{i+1}\} = \cdots = \{l_{\omega_{j-1}}, f_j^{i+1}\} = 0.$$

To complete the proof of the claim, it remains to estimate

$$X_{f_j^{i+1}} = \int_0^1 (X_{\chi_j^i}^t)^*[X_{h-l_{\omega_j}} + X_{g_j^i} + X_{f_{j,t}^i}, X_{\chi_j^i}]dt.$$

First,

$$||X_{f_j^{i+1}}||_{s_j^{i+1},\xi_j^{i+1}} \leq \sup_{0\leq t\leq 1} ||(X_{\chi_j^i}^t)^*[X_{h-l_{\omega_j}} + X_{g_j^i} + X_{f_{j,t}^i}, X_{\chi_j^i}]||_{s_j^{i+1},\xi_j^{i+1}}$$

and since $\tau > 3$ by (A.6), we can apply Lemma A.2 to get

$$||X_{f_j^{i+1}}||_{s_j^{i+1},\xi_j^{i+1}} \leq 2||[X_{h-l_{\omega_j}} + X_{g_j^i} + X_{f_{j,t}^i}, X_{\chi_j^i}]||_{s_j^i-2\sigma_j/3,\xi_j^i-2\rho_j/3}. \tag{A.8}$$

Using Lemma A.3, we have

$$||[X_{f_{j,t}^i}, X_{\chi_j^i}]||_{s_j^i-2\sigma_j/3,\xi_j^i-2\rho_j/3} \leq 9(r_j + \xi_j^i)\sigma_j^{-1}||X_{f_{j,t}^i}||_{s_j^i,\xi_j^i}||X_{\chi_j^i}||_{s_j^i,\xi_j^i}$$

and since

$$||X_{f_{j,t}^i}||_{s_j^i,\xi_j^i} \leq \varepsilon_j^i, \quad ||X_{\chi_j^i}||_{s_j^i,\xi_j^i} \leq T_j\varepsilon_j^i,$$

we get

$$||[X_{f_{j,t}^i}, X_{\chi_j^i}]||_{s_j^i-2\sigma_j/3,\xi_j^i-2\rho_j/3} \leq (9(r_j + 3\xi_j)mT_j\varepsilon_j s_j^{-1})\varepsilon_j^i \tag{A.9}$$

since $\xi_j^i \leq 3\xi_j$ and $\varepsilon_j^i \leq \varepsilon_j$. Similarly, since $\gamma_j^i \leq 2\varepsilon_j$,

$$||[X_{g_j^i}, X_{\chi_j^i}]||_{s_j^i-2\sigma_j/3,\xi_j^i-2\rho_i/3} \leq (9(r_j + 3\xi_j)mT_j\gamma_j^i s_j^{-1})\varepsilon_j^i \leq (18(r_j + 3\xi_j)mT_j\varepsilon_j s_j^{-1})\varepsilon_j^i. \tag{A.10}$$

Concerning the last bracket, let us first prove that

$$[X_{h-l_{\omega_j}}, X_{\chi_j^i}] = [X_{h_{\tilde{\lambda}_j}-l_{\omega_j}}, X_{\chi_j^i}]$$

where we recall that $\tilde{\lambda}_j = I(z_j) + \tilde{\Lambda}_j$ and $h_{\tilde{\lambda}_j}$ is the restriction of h to $\tilde{\lambda}_j$. To do this, it is sufficient to prove $\{h - l_{\omega_j}, \chi_j^i\} = \{h_{\tilde{\lambda}_j} - l_{\omega_j}, \chi_j^i\}$, which is equivalent to $\{h, \chi_j^i\} = \{h_{\tilde{\lambda}_j}, \chi_j^i\}$. For any $z \in \mathcal{V}_{s_j^i, \xi_j^i}(z_j)$, we have

$$\{h, \chi_j^i\}(z) = \nabla h(I(z)) \cdot \{I, \chi_j^i\}(z).$$

But for any $-1 \leq l \leq j-1$, we know that

$$\{l_{\omega_l}, \chi_j^i\}(z) = \omega_l \cdot \{I, \chi_j^i\}(z) = 0$$

which means that $\{I, \chi_j^i\}(z) \in \tilde{\Lambda}_j$. Therefore, recalling that $\tilde{\Pi}_j$ denotes the orthogonal projection onto $\tilde{\Lambda}_j$, it comes that

$$\{h, \chi_j^i\}(z) = \tilde{\Pi}_j(\nabla h(I(z))) \cdot \{I, \chi_j^i\}(z) = \nabla h_{\tilde{\lambda}_j}(I(z)) \cdot \{I, \chi_j^i\}(z) = \{h_{\tilde{\lambda}_j}, \chi_j^i\}(z)$$

and therefore $\{h, \chi_j^i\} = \{h_{\tilde{\lambda}_j}, \chi_j^i\}$. Then, for any $z \in \mathcal{V}_{s_j^i, \xi_j^i}(z_j)$, we can estimate

$$\begin{aligned}
||\nabla h_{\tilde{\lambda}_j}(I(z)) - \omega_j|| &= ||\tilde{\Pi}_j(\nabla h(I(z))) - \omega_j|| \\
&\leq ||\tilde{\Pi}_j(\nabla h(I(z))) - \tilde{\Pi}_j(\nabla h(I(z_j)))|| + ||\Pi_j(\nabla h(I(z_j))) - \omega_j|| \\
&\leq ||\nabla h(I(z)) - \nabla h(I(z_j))|| + ||\Pi_j(\nabla h(I(z_j))) - \omega_j|| \\
&\leq F||I(z) - I(z_j)|| + s_j \leq F\sqrt{n}|I(z) - I(z_j)| + s_j \\
&\leq F\sqrt{n}s_j^i + s_j \leq (3F\sqrt{n} + 1)s_j
\end{aligned}$$

since $s_j^i \leq 3s_j$, and where we used the fact that

$$\sup_{z \in \mathcal{V}_{3s_j, 3\xi_j}(z_j)} ||\nabla^2 h(I(z))|| \leq F.$$

From this, we deduce that

$$\begin{aligned}
||X_{h_{\tilde{\lambda}_j} - l_{\omega_j}}||_{s_j^i, \xi_j^i} &\leq \sup_{z \in \mathcal{V}_{s_j^i, \xi_j^i}(z_j)} |\nabla(h_{\tilde{\lambda}_j} - l_{\omega_j})(I(z))|||z||| \leq (3F\sqrt{n} + 1)s_j(r_j + \xi_j^i) \\
&\leq (3F\sqrt{n} + 1)s_j(r_j + 3\xi_j)
\end{aligned}$$

and using Lemma A.4, we get

$$\begin{aligned}
||[X_{h - l_{\omega_j}}, X_{\chi_j^i}]||_{s_j^i - 2\sigma_j/3, \xi_j^i - 2\rho_j/3} &= ||[X_{h_{\tilde{\lambda}_j} - l_{\omega_j}}, X_{\chi_j^i}]||_{s_j^i - 2\sigma_j/3, \xi_j^i - 2\rho_j/3} \\
&\leq 9\rho_j^{-1}||X_{h_{\tilde{\lambda}_j} - l_{\omega_j}}||_{s_j^i, \xi_j^i}||X_{\chi_j^i}||_{s_j^i, \xi_j^i}
\end{aligned}$$

hence

$$\begin{aligned}
||[X_{h - l_{\omega_j}}, X_{\chi_j^i}]||_{s_j^i - 2\sigma_j/3, \xi_j^i - 2\rho_j/3} &\leq 9\rho_j^{-1}(3F\sqrt{n} + 1)s_j(r_j + 3\xi_j)T_j\varepsilon_j^i \\
&= (9(3F\sqrt{n} + 1)\xi_j^{-1}(r_j + 3\xi_j)mT_js_j)\varepsilon_j^i. \qquad \text{(A.11)}
\end{aligned}$$

Putting the estimates (A.8), (A.9), (A.10) and (A.11) together, and recalling that $\varepsilon_j = 2^j\varepsilon$, we arrive at

$$||X_{f_j^{i+1}}||_{s_j^{i+1}, \xi_j^{i+1}} \leq (2^j 54(r_j + 3\xi_j)mT_j\varepsilon s_j^{-1} + 18(3F\sqrt{n} + 1)\xi_j^{-1}(r_j + 3\xi_j)mT_js_j)\varepsilon_j^i.$$

Using the second and third inequality of (3.3), we obtain

$$||X_{f_j^{i+1}}||_{s_j^{i+1},s_j^{i+1}} \leq \varepsilon_j^i/2 = \varepsilon_j^{i+1}.$$

This finishes the proof of the claim.

Now let us define $\Phi_j = \Phi_j^m$, $g_j^+ = g_j^m$ and $f_j^+ = f_j^m + f_j \circ \Phi_j$. Since $s_j^m = 2s_j$ and $\xi_j^m = 2\xi_j$, Φ_j is a real-analytic symplectic embedding

$$\Phi_j : \mathcal{V}_{2s_j,2\xi_j}(z_j) \to \mathcal{V}_{3s_j,3\xi_j}(z_j)$$

such that $H_j \circ \Phi_j = h + g_j^+ + f_j^+$. We already know that $\{l_{\omega_{-1}}, g_j^+\} = \{l_{\omega_0}, g_j^+\} = \cdots = \{l_{\omega_j}, g_j^+\} = 0$, and the estimates

$$||X_{g_j^+}||_{2s_j,2\xi_j} \leq \gamma_j^m \leq 2\varepsilon_j = 2^{j+1}\varepsilon, \quad ||\Phi_0 - \mathrm{Id}||_{2s_j,2\xi_j} \leq T_j\gamma_j^m \leq 2^{j+1}T_j\varepsilon.$$

To conclude the proof of the proposition, it remains to estimate $X_{f_j^+}$. First recall that

$$||X_{f_j^m}||_{2s_j,2\xi_j} \leq \varepsilon_j^m = 2^{-m}\varepsilon_j = 2^j 2^{-m}\varepsilon. \tag{A.12}$$

Then, $f_j \circ \Phi_j = f_j \circ \Phi_j^m = f_j \circ X^1_{\chi_j^1} \circ \cdots X^1_{\chi_j^m}$ and so $X_{f_j\circ\Phi_j} = (X^1_{\chi_j^m})^* \cdots (X^1_{\chi_j^1})^* X_{f_j}$, where

$$||X^1_{\chi_j^i} - \mathrm{Id}||_{s_j^{i+1},\xi_j^{i+1}} \leq ||X_{\chi_j^i}||_{s_j^i,\xi_j^i} \leq T_j\varepsilon_j^i = 2^{-i}T_j\varepsilon_j,$$

for each $1 \leq i \leq m$. Applying Lemma A.2 inductively yields

$$\begin{aligned}
||X_{f_j\circ\Phi_j}||_{2s_j,2\xi_j} &= ||(\Phi_j)^* X_{f_j}||_{s_j^m,\xi_j^m} \leq \prod_{i=0}^{m-1}\left(1 + 3(r_j + \xi_j^i)\sigma_j^{-1}||X_{\chi_j^i}||_{s_j^i,\xi_j^i}\right)||X_{f_j}||_{2s_j,2\xi_j} \\
&= \prod_{i=0}^{m-1}\left(1 + 3(r_j + 3\xi_j)s_j^{-1}m||X_{\chi_j^i}||_{s_j^i,\xi_j^i}\right)||X_{f_j}||_{2s_j,2\xi_j} \\
&\leq \prod_{i=0}^{m-1}(1 + 2^{-i}3(r_j + 3\xi_j)s_j^{-1}mT_j\varepsilon_j)||X_{f_j}||_{2s_j,2\xi_j} \\
&\leq \exp\left(\sum_{i=0}^{m-1} 2^{-i}3(r_j + 3\xi_j)s_j^{-1}mT_j\varepsilon_j\right)||X_{f_j}||_{2s_j,2\xi_j} \\
&\leq \exp(6(r_j + 3\xi_j)s_j^{-1}mT_j\varepsilon_j)||X_{f_j}||_{2s_j,2\xi_j}.
\end{aligned}$$

The second condition of (3.3) implies in particular that $\exp(6(r_j + 3\xi_j)s_j^{-1}mT_j\varepsilon_j) \leq 2$ and therefore

$$||X_{f_j\circ\Phi_j}||_{2s_j,2\xi_j} \leq 2||X_{f_j}||_{2s_j,2\xi_j} \leq 2(j2^{j-1})2^{-m}\varepsilon = j2^j 2^{-m}\varepsilon. \tag{A.13}$$

From (A.12) and (A.13) we get

$$||X_{f_j^+}||_{2s_j,2\xi_j} \leq 2^j 2^{-m}\varepsilon + j2^j 2^{-m}\varepsilon = (j+1)2^j 2^{-m}\varepsilon,$$

and this ends the proof. □

Comment. The preprint "Double exponential stability for generic real-analytic elliptic equilibrium points" was first submitted to the Arxiv in August 2015; in order to make it more accessible, we decided to withdraw this preprint and split it into two parts. This corresponds to the first part, the second part being [BFN19].

Acknowledgements. The authors have benefited from partial funding from the ANR project Beyond KAM.

References

[BFG98] G. Benettin, F. Fassò, and M. Guzzo, *Nekhoroshev-stability of l4 and l5 in the spatial restricted three-body problem*, Regular and Chaotic Dynamics **3** (1998), no. 3, 56–72.

[BFN19] A. Bounemoura, B. Fayad, and L. Niederman, *Super-exponential stability for generic real-analytic elliptic equilibrium points*, Preprint, 2019.

[BN12] A. Bounemoura and L. Niederman, *Generic Nekhoroshev theory without small divisors*, Ann. Inst. Fourier **62** (2012), no. 1, 277–324.

[Bou11] A. Bounemoura, *Effective stability for Gevrey and finitely differentiable prevalent Hamiltonians*, Communication in Mathematical Physics **307** (2011), no. 1, 157–183.

[FGB98] F. Fassò, M. Guzzo, and G. Benettin, *Nekhoroshev-stability of elliptic equilibria of Hamiltonian systems*, Comm. Math. Phys. **197** (1998), no. 2, 347–360.

[GCB16] M. Guzzo, L Chierchia, and G. Benettin, *The steep Nekhoroshev's theorem*, Comm. Math. Phys. **342** (2016), 569–601.

[GFB98] M. Guzzo, F. Fassò, and G. Benettin, *On the stability of elliptic equilibria*, Math. Phys. Electron. J. **4** (1998), 16 pp., paper 1.

[HZ94] H. Hofer and E. Zehnder, *Symplectic invariants and hamiltonian dynamics*, Birkhauser Verlag, Basel, 1994.

[LN92] P. Lochak and A.I. Neishtadt, *Estimates of stability time for nearly integrable systems with a quasiconvex Hamiltonian*, Chaos **2** (1992), no. 4, 495–499.

[LNN94] P. Lochak, A.I. Neistadt, and L. Niederman, *Stability of nearly integrable convex Hamiltonian systems over exponentially long times.*, Kuksin, S. (ed.) et al., Seminar on dynamical systems. Basel: Birkhäuser. Prog. Nonlinear Differ. Equ. Appl. 12, 15-34 (1994)., 1994.

[Loc92] P. Lochak, *Canonical perturbation theory via simultaneous approximation.*, Russ. Math. Surv. **47** (1992), no. 6, 57–133.

[Nek73] N.N. Nekhoroshev, *Stable lower estimates for smooth mappings and for the gradients of smooth functions*, Mat. Sb. (N.S.) **90 (132)** (1973), 432–478, 480.

[Nek77] ———, *An exponential estimate of the time of stability of nearly integrable Hamiltonian systems*, Russian Math. Surveys **32** (1977), no. 6, 1–65.

[Nek79] ———, *An exponential estimate of the time of stability of nearly integrable Hamiltonian systems II*, Trudy Sem. Petrovs **5** (1979), 5–50.

[Nie98] L. Niederman, *Nonlinear stability around an elliptic equilibrium point in a Hamiltonian system*, Nonlinearity **11** (1998), no. 6, 1465–1479.

[Nie07] ———, *Prevalence of exponential stability among nearly integrable Hamiltonian systems*, Erg. Th. Dyn. Sys. **27** (2007), no. 3, 905–928.

[Pin13] G. Pinzari, *Aspects of the planetary Birkhoff normal form*, Regular and Chaotic Dynamics **18** (2013), no. 6, 860–906.

[Pös93] J. Pöschel, *Nekhoroshev estimates for quasi-convex Hamiltonian systems*, Math. Z. **213** (1993), 187–216.

[Pös99] ———, *On Nekhoroshev's estimate at an elliptic equilibrium*, Internat. Math. Res. Notices **4** (1999), 203–215.