Quantitative KAM theorem
Thibaut Castan

To cite this version:
Thibaut Castan. Quantitative KAM theorem. 2017. <hal-01514613v1>

HAL Id: hal-01514613
https://hal.archives-ouvertes.fr/hal-01514613v1
Submitted on 26 Apr 2017 (v1), last revised 21 Aug 2017 (v2)

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.
Quantitative KAM theorem
Thibaut Castan

Abstract. We revisit Pöschel’s 2001 version of the KAM theorem so as to find an explicit quantitative bound for the size of the allowed perturbation. Our theorem is applied to the plane planetary problem in a pair of other papers.

In 1963, Arnold [1] proved that Kolmogorov’s theorem [8] could be applied to the plane planetary three-body problem, thus showing the existence of quasi-periodic solution over an infinite time interval. This theorem relied on a smallness condition on the ratio of masses between the planets and the star. Hénon, in a letter to Arnold [7], gave a necessary condition for Arnold’s arguments to apply, namely that the mass ratio must be less than $10^{-320}$. This paper is part of a work to determine a sufficient condition to apply the KAM theorem to the plane planetary three-body problem [2, 4].

Quantitative conditions for the KAM stability has been established using computer-assisted proofs, for similar systems (obtained by truncating the plane circular restricted three-body problem), in the works of Celletti-Chierchia, Giorgilli-Locatelli and Robutel [5, 6, 9, 14]. In another line of thought, quantitative results on the stability of the three-body problems over exponentially long time were also computed by Niederman [12] and later improved by Castan [3].

In the present paper, we revisit the KAM theorem developed by Pöschel in [13] to derive explicit hypotheses depending on the parameters of the system, such as the analyticity widths in the actions and in the angles. We chose this theorem for simplicity of its statement and of the proof. It is a version of the KAM theorem with parameters, using the framework of Moser [10].

1 Hamiltonian with parameter

1.1 Parameterizing the Hamiltonian

Consider, for $n \geq 2$ the following analytic Hamiltonian:

$$H(p, q) = h(p) + \epsilon f(p, q, \epsilon), \quad (p, q) \in D \times T^n, \quad \epsilon \ll 1,$$

where $D \subset \mathbb{R}^n$. The Hamilton equations of $h$ give a quasi-periodic motion of frequency $\omega(p) = h'(p) \in \mathbb{R}^n$, and the perturbation $\epsilon f(p, q, \epsilon)$ have small norm compared to the Hamiltonian. Assume moreover that the unperturbed Hamiltonian is non-degenerate on the set $D$, that is $\det(h''(p)) = \det \left( \frac{\partial \omega}{\partial p} (p) \right) \neq 0$. The frequency map $h' : D \to \Omega$ is a local diffeomorphism between $D$ and the frequency set $\Omega \subset \mathbb{R}^n$. The approach of Mõser [10] consists in expanding the Hamiltonian $h$ around one particular frequency $\omega$, and work with a linear Hamiltonian parameterized with that frequency.

Let $p = p_0 + I$, with $I \in B = D - p_0$, the Hamiltonian $h$ can be written:

$$h(p) = h(p_0) + \langle h'(p_0), I \rangle + \int_0^1 (1 - t) \langle h''(p_0 + tI), I \rangle dt$$

Since the frequency map is a local diffeomorphism, it is equivalent to work not only with the action $p$, but with the two variables $(\omega, I)$. Fixing the action $p_0$ (and therefore $\omega$), one can write the equation (2):

$$h(p) = e(\omega) + \langle \omega, I \rangle + P_h(I; \omega),$$

where $P_h(I; \omega) = \int_0^1 (1 - t) \langle h''(g'(\omega) + tI), I \rangle dt$. From the action-angle coordinates, we defined new coordinates $(\omega, I, \theta) \in \Omega \times B \times T^n$, where we wrote $\theta$ instead of $q$ not to be mistaken. The term $P_h$ will be considered a part of the perturbation, since a bound on its norm can be made as small as wanted by looking...
at a sufficiently small ball in the action $I$ around the origin.

One can write $H = N + P$, where $N = \epsilon(\omega) + \langle \omega, I \rangle$, $N$ being called the normal form, and

$$P = P_h(I; \omega) + P_c(I; \omega),$$

with $P_c(I; \theta; \omega) = \epsilon f(g'(\omega) + I, \theta, \epsilon)$.

The family of Hamiltonian under normal form $N$ have equations of motions that are easy to compute. Indeed, the vector field associated to it is

$$X_N = \sum_{j=1}^{n} \omega_j \partial_{\theta_j}, \quad \omega \in \Omega.$$

The motion is quasi-periodic, and takes place on a specific torus $\{0\} \times \mathbb{T}^n$ for every $\omega \in \Omega$. These tori can be seen as a trivial embedding of $\mathbb{T}^n$ over the set $\Omega$ on the phase space given by the function

$$\Phi_0 : \mathbb{T}^n \times \Omega \to B \times \mathbb{T}^n, \quad (\theta, \omega) \mapsto (0, \theta)$$

For a generic Hamiltonian, the perturbation $P$ will limit the existence of these tori. However, under hypotheses made clear further in the paper, one can show that almost all of these tori (in the sense of the Lebesgue measure) survives a perturbation, this is the main result of the KAM theorem.

### 1.2 Sets of analyticity and other definitions

Let $\Omega$ be the set of initial frequencies we are considering. Let $\tau > n - 1, \gamma > 0$ and consider the set of Diophantine vectors:

$$D(\gamma, \tau) = \left\{ \omega \in \mathbb{R}^n : \forall k \in \mathbb{Z}^n, \ |k \cdot \omega| \geq \frac{\gamma}{\|k\|_1} \right\},$$

where $\| \cdot \|_1$ is the $l_1$-norm. Let $\Omega_{\gamma, \tau} = \Omega \cap D(\gamma, \tau)$, it is as well a Cantor set. Finally, for $\beta > 0$, let

$$\Omega_{\gamma, \tau}^\beta = \Omega_{\gamma, \tau} \setminus \{ \omega \in \Omega_{\gamma, \tau} : \exists \omega' \in \mathbb{R}^n \setminus \Omega, |\omega - \omega'| < \beta \}.$$

The last set is the set of vectors in $\Omega_{\gamma, \tau}$ that are at least at a distance $\beta$ to the boundary of $\Omega$. We will fix later the needed value for the constant $\beta$. Now let us define the various domains we use in the theorem. These sets will be polydiscs around some set. For the frequencies, define

$$O_h = \{ \omega \in \mathbb{C}^n, |\omega - \Omega_{\gamma, \tau}^\beta| < h \}.$$

Call $\mathbb{T}^n_{\mathbb{C}} = (\mathbb{T} \times \mathbb{R})^n$ the complex extension of the $n$–torus. The action-angle variable will take values in

$$D_{r, s} = \{(I, \theta) \in \mathbb{C}^n \times \mathbb{T}^n_{\mathbb{C}}, |I| < r, |\Im(\theta)| < s \}.$$

Define the norms with indices as follows, for $f : \mathbb{C}^n \times \mathbb{T}^n_{\mathbb{C}} \to \mathbb{C}$:

$$\|f\|_{r, s} = \sup_{D_{r, s}} |f|, \quad \|f\|_h = \sup_{O_h} |f|, \quad \|f\|_{r, s, h} = \sup_{D_{r, s} \times O_h} |f|.$$

For vector valued functions, we have:

$$\|f\|_{r, s} = \sup_{D_{r, s}} \|f\|,$$

where $\| \cdot \|$ is the sup norm. When considering the Diophantine condition, we will always consider the norm $|k|_1 = |k_1| + \ldots + |k_n|$ for the vectors $k \in \mathbb{Z}^n$. Finally, to state the theorem we will need the following Lipschitz norm on the frequencies:

$$|f|_{L} = \sup_{\omega \neq \omega'} \frac{|f(\omega) - f(\omega')|}{|\omega - \omega'|},$$

where $| \cdot |$ represent the supremum norm.
\section{Quantitative KAM theorem}

\subsection{Statement of the theorem}

In this section, we give the explicit statement of the KAM theorem. We chose \( \tau = n \) in the following. We introduce first some definitions to be able to state the theorem.

\[ A = \{ K \in \mathbb{N} : K \sigma \geq (2n + \nu) \log 2 \} , \]
\[ B = \left\{ K \in \mathbb{N} : 2K^{n+\nu} e^{-K \sigma} \leq \frac{1}{\delta} \right\} , \]
\[ B_+ = \{ K \in B : \forall m \in \mathbb{N}, K + m \in B \} . \]

Call \( C_1 = 4\nu(200nC_0 + 32 + 8\nu n!)^2 \) with \( C_0 = \frac{3\pi}{6n^2} \sqrt{\frac{n(2n)!}{2^n}} \). Define the important value:

\[ \epsilon = \min \left( \frac{\gamma r \sigma \nu}{4\nu C_1}, \frac{h_r}{\delta} \cdot \frac{\gamma r}{2K^n \delta} \right) . \]

Define the following (exponentially convergent) series for \( \nu \geq 1 \):

\[ S_\nu = \sum_{i=0}^{\infty} 2^{\nu(3i+2-\left(\frac{3}{2}\right)^i)} , \quad T_\nu = \sum_{i=0}^{\infty} 2^{(2\nu+1)i-\nu\left(\frac{3}{2}\right)^i} , \]

and finally

\[ \mu = \exp \left( \frac{5}{\delta} \right) , \quad \xi = \exp \left( 10C_0 \frac{C_0}{\gamma r_0 \sigma_0} \right) . \]

The quantitative statement of the KAM theorem of Pöschel is the following:

\textbf{Theorem 1.} Let \( H = N + P \) be a Hamiltonian, such that \( P \) is real analytic on the set \( D_{r,s} \times O_h \) and \( \| P \|_{r,s,h} = \epsilon_0 \leq \epsilon \). Then there exists a Lipschitz continuous map \( \varphi : \Omega_{\gamma,\tau} \rightarrow \Omega_{\gamma,0} \), with \( h_0 = \frac{\delta \epsilon_0}{r} \), and a Lipschitz continuous family of real analytic torus embeddings \( \Phi : \mathbb{T}^n \times \Omega_{\gamma,\tau} \rightarrow B \times \mathbb{T}^n \) close to \( \Phi_0 \) such that for each \( \omega \in \Omega_{\gamma,\tau} \), the embedded tori are Lagrangian and

\[ X_H |_{\varphi(\omega)} \circ \Phi = \Phi' \cdot X_N . \]

\( \Phi \) is real analytic on the set \( \{ \theta : |3\theta| < s/2 \} \) for each \( \omega \), and the following inequalities on \( \Phi \) and \( \varphi \) hold:

\[ \| W(\Phi - \Phi_0) \| \leq 6n\xi \log \xi \]
\[ \| \varphi - Id \| \leq 4hn\mu \log \mu , \]

where \( W = \text{diag}(r^{-1}Id, s^{-1}Id) \). As for their Lipschitz constant, we have:

\[ \| W(\Phi - \Phi_0) \|_L \leq 4^\nu 80 \frac{nC_0}{\gamma \sigma \delta} T_\nu \]
\[ \| \varphi - Id \|_L \leq 4nS_\nu \mu \log \mu . \]

Qualitatively, the statement may be described as follows: first, as said before, while removing the perturbation, we are slightly changing the frequencies; secondly, the Lipschitz estimates allow us to control the size of the the set \( \varphi(\Omega_{\gamma,\tau}) \). Indeed, with these estimates, one can prove that the complement of this set is of size \( O(\alpha) \) (see [13]); finally, every embedded torus is Lagrangian and is close to its associated unperturbed torus.

3
2.2 Sketch of the proof

Let us describe succinctly the general scheme of the proof, that will consist in an iteration of a KAM step. At each step, we consider a Hamiltonian \( H = N + P \) with \( N \) under normal form: \( N = e(\omega) + \langle \omega, I \rangle \). We want to find a transformation \( F \) such that \( H \circ F = N_+ + P_+ \), where \( N_+ \) is again under normal form and \( P_+ \) verifies \( \|P_+\| \leq C\|P\|^{\kappa} \), for some constants \( C \) and \( \kappa > 1 \). This transformation implies some loss of analyticity width related to the norm of \( P \), however, the constant \( \kappa \) will ensure that the scheme rapidly converge if the perturbation is small enough.

To build the transformation \( F \), instead of considering \( P \), we will consider only its linear part in the action (by truncating the Taylor expansion in the action at the order 2). Next, we truncate the Fourier series in the angle at some order \( K \). Let us call the new Hamiltonian after this two steps \( R \). The remainder \( P - R \) is either of order 2 in the actions, and therefore small looking close to the origin, or is part of the remainder of the Fourier series, which will be small as well if \( K \) is large enough.

Assume \( \omega \) is fixed. Let \( F \) be a Hamiltonian affine in the actions, and \( X_F = X \) its associated Hamiltonian vector field. Call \( \Phi^t = \Phi^t_X \) the flow associated to the previous vector field, and \( \Phi = \Phi^t|_{t=1} \) the time-1 map of this flow.

Call \( \bar{H} = N + R \), we have:

\[
\bar{H} \circ \Phi = N + \{N, F\} + \int_0^1 (1 - t) \{\{N, F\}, F\} \circ \Phi^t dt + R + \int_0^1 \{R, F\} \circ \Phi^t dt
\]

We want \( F \) such that

\[
N + \{N, F\} + R = N_+
\]  \hspace{1cm} (15)

The term under the integral is small and therefore added to the perturbation \( P - R \). The average of \( R \) not being null, we divide \( R \) into two parts: \( R = \bar{R} + \tilde{R} \) where

\[
\bar{R} = \frac{1}{(2\pi)^n} \int_{T^n} R d\theta.
\]

Formally, \( F \) is defined by

\[
F = \sum_{k \in \mathbb{Z}^n \setminus \{0\}} \frac{R_k}{i(k, \omega)} \exp(i k \cdot \theta),
\]

where the \( R_k \) are the Fourier coefficients of the Hamiltonian \( R \). As for the term \( \bar{R} \), we simply add it to the unperturbed Hamiltonian. Observe that since \( R \) was affine in the actions, \( N_+ = N + \bar{R} \) remains affine in the actions. Define

\[
N_+ = e_+(\omega) + \langle \omega + v(\omega), I \rangle.
\]

The new frequency vector is given by \( \omega_+ = \omega + v(\omega) \), and if \( v \) is small, then there exists a map \( \varphi \) close to the identity such that \( \varphi(\omega_+) = \omega \). The map \( N_+ \) can then be written \( N_+ = (N + \bar{R}) \circ \varphi \). The total transformation \( F \) then corresponds to the map \( (\Phi, \varphi) \). As for the perturbation, computing the remainders of the transformation gives:

\[
P_+ = \int_0^1 \{ (1 - t)\bar{R} + tR, F\} \circ \Phi^t dt + (P - R) \circ \Phi.
\]

If the perturbation is small enough, then \( P_+ \) will be even smaller, and we can iterate the scheme an infinite time to let \( P_+ \) go to zero.
3 Quantitative KAM step and its proof

Before proving the quantitative KAM theorem, we prove a quantitative KAM step, that will be iterated an infinite amount of times so as to obtain the complete theorem. Again, we follow almost completely the work of Pöschel in his proof. Recall as well that we chose \( \tau = n \) as Diophantine constant, for the sake of simplicity. Keeping \( \tau \) as an unknown should not be much harder to compute in this work.

3.0.1 Statement of the KAM step

**Proposition 1.** Assume that \( \|P\|_{r,s,h} \leq \epsilon \) where \( \epsilon \) is defined in (7). Then there exists a real analytic transformation

\[
\mathcal{F} = (\Phi, \varphi) : D_{\eta r,s-5\sigma} \times O_{h/4} \to D_{r,s} \times O_{h}
\]

with \( \eta = \sqrt{\frac{\epsilon}{\gamma r \sigma^\nu}} \) such that \( H \circ \mathcal{F} = N + P \) with

\[
\|P\|_{\eta r,s-5\sigma,h/4} \leq 200 n C_0 \epsilon^2 \gamma r \sigma^\nu + (32 \eta^2 + 4^n n! K' e^{-K' \sigma}) \epsilon.
\]

Moreover,

\[
2 \|W(\Phi - Id)\|, \quad \|W(D\Phi - Id) W^{-1}\| \leq \frac{40 C_0 \epsilon}{\gamma r \sigma^\nu},
\]

\[
\|\varphi - Id\|, \quad 4h \|D\varphi - Id\| \leq \frac{10 \epsilon}{r}
\]

uniformly on \( D_{\eta r,s-5\sigma} \times O_{h} \) and \( O_{h/4} \) respectively, with the weight matrix

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

**Observations on the statement** The different conditions on \( \epsilon \) arise from different parts of the proof. The first condition in the minimum is a limit due to the analyticity widths in the actions and on the angles. This limit is necessary to obtain an exponential decrease of the bound on the norm of the perturbation at each step. The second condition is related to the transformation on the frequency vector. To be able to invert the map giving the new frequency \( \omega + v(\omega) \), it is essential to have enough analyticity width in the frequencies compared to the size of \( v \). The third condition is a condition on \( K' \), the order of truncature of the Fourier series of the affine perturbation. \( K' \) needs to be big enough to allow the remainder of the Fourier series of \( Q \) to be small enough, and then for this remainder to decrease exponentially while iterating our scheme. However, we want as well \( K' \) to be small enough so that all the frequencies in \( O_{h} \) satisfy a non-resonance condition of order \( K' \).

In the KAM step, the factor \( \delta \) always appears in the denominator, and therefore always seems to add a stronger constraint when increasing it. However, when iterating the scheme, we will see that it allows to make the norm of the transformation smaller while increasing its value.

3.0.2 Proof of the proposition

**Implications of the hypotheses:** Let \( \epsilon_0 \leq \epsilon \). Define \( h_0 = \frac{\delta \epsilon_0}{r} \) and \( K_0 = \left\lfloor \frac{\gamma}{2h_0} \right\rfloor \). These two constants satisfy \( h_0 < h \) and \( K_0 > K' \). Indeed, the first inequality is clear, and for the second one, using the fact that both \( K_0 \) and \( K' \) are integers, and the definition of \( \epsilon \):

\[
K' \sigma^\nu \leq \frac{\gamma r}{2 \delta \epsilon_0} \leq \frac{\gamma r}{2 \delta \epsilon} \leq \frac{\gamma}{2h_0}.
\]

The definition of \( h_0 \) means that if we consider a smaller perturbation, we will not use all the available analyticity width \( h \) corresponding to the frequencies. These definitions allow us to compute a crucial inequality, that will be useful later in the proof:

\[
\frac{1}{2^n} K_0^n \sigma^\nu \exp(-K_0 \sigma) \leq \frac{\epsilon_0}{\gamma r} \leq \frac{h_0}{\gamma \delta} \leq \frac{1}{2^n K_0^n}.
\] (16)
The last two inequalities are straightforward given the definition of $h_0$ and $K_0$. Regarding the first one, using the definition of $B_+$:

$$
\frac{\epsilon_0}{\gamma r \sigma^\nu} = \frac{h_0}{\gamma r \sigma^\nu} \geq \frac{2K_0^{n+\nu} e^{-K_0 \sigma}}{\gamma} h_0 \geq K_0^n e^{-K_0 \sigma} \left( \sqrt{\frac{2h_0}{\gamma}} \times \left\lfloor \sqrt{\frac{\gamma}{2h_0}} \right\rfloor \right)^\nu 
\geq K_0^n e^{-K_0 \sigma} \left( \left\lfloor \sqrt{\frac{\gamma}{2h_0}} \right\rfloor^{-1} \times \left\lfloor \sqrt{\frac{\gamma}{2h_0}} \right\rfloor \right)^\nu \geq \frac{1}{2^\nu} K_0^n e^{-K_0 \sigma}.
$$

We now use the definition of $K_0$ to show the non-resonance condition that the frequency vectors of $O_{h_0}$ must satisfy. Indeed, let $k$ such that $0 < |k| \leq K_0$, and let $\omega \in O_{h_0}$, there exists $\omega^* \in \Omega_\gamma^0$ such that $|\omega - \omega^*| < h_0$, and therefore, the following inequalities hold:

$$
|\langle k, \omega - \omega^* \rangle| \leq |k| |\omega - \omega^*| \leq K_0 h_0 \leq \frac{\gamma}{2^k} \leq \frac{\gamma}{2|k|^\tau}.
$$

Since $\omega^*$ satisfy a Diophantine condition for the constant $\gamma$ and $\tau = n$, we get:

$$
|\langle k, \omega \rangle| \geq \frac{\gamma}{2|k|^n}, \quad \forall 0 < |k| \leq K_0
$$

The goal at each step of the KAM theorem is to make a change of variables that will decrease the norm of the perturbation. Yet, instead of trying to make the value of $\epsilon_0$ decrease directly, we will consider the ratio $E_0 = \frac{\omega}{\gamma r \sigma^\nu}$. If we let $r$, and $\sigma$ decrease in a polynomial way, but that we managed to let $E$ decrease exponentially, then $\epsilon_0$ will as well decrease exponentially.

**First estimates:** As introduced in the outline of the proof, we will switch from the perturbation $P$ to another perturbation $R$ in two steps. First, consider the linearization $Q$ of $P$ in the actions around the origin, secondly, truncate the Fourier series of $Q$ at order $K_0$. $R$ is a trigonometric polynomial on the angles and affine in the action. The size of the perturbation, by assumption, is smaller than $\epsilon_0$ on the set $D_{r,s} \times O_{h_0}$. Let us start by bounding the linearized function $Q$:

$$
\|Q\|_{\mathcal{F}, s, h_0} \leq \|P\|_{\mathcal{F}, s, h_0} + \frac{3r}{4} \|P\|_{\mathcal{F}, s, h_0} \leq \|P\|_{r, s, h_0} + \frac{3r}{4} \|P\|_{r, s, h_0} \leq 4\epsilon_0.
$$

Now let $\eta = \sqrt{\frac{\epsilon_0}{\gamma r \sigma^\nu}}$. Using the definition of $\epsilon$ and the fact that $\epsilon_0 \leq \epsilon$, we obtain $8\eta \leq 1$. From the Taylor expansion formula and Cauchy’s inequality, we get:

$$
\|P - Q\|_{2\eta r, s, h_0} \leq (2\eta r)^2 \frac{4\|P\|_{r, s, h_0}}{(1 - 2\eta)^2 r^2} \leq 32\eta^2 \epsilon_0.
$$

With lemma 1 of appendix A, we obtain the following estimates on the difference between $Q$ and its truncation at the order $K_0$:

$$
\|R - Q\|_{\mathcal{F}, s, -\sigma, h_0} \leq 4^n n! K_0^n \exp(-K_0 \sigma) \|Q\|_{\mathcal{F}, s, h_0} \leq 4^n n! K_0^n \exp(-K_0 \sigma) \epsilon_0.
$$

Hence:

$$
\|R\|_{\mathcal{F}, s, -\sigma, h_0} \leq \|R - Q\|_{\mathcal{F}, s, -\sigma, h_0} + \|Q\|_{\mathcal{F}, s, h_0} \leq (4 + 4^n n! K_0^n e^{-K_0 \sigma}) \epsilon_0.
$$
Since $K_0 \in A \cap B_+$, we have the following inequality:

$$4^\nu n!K_0^n e^{-K_0\sigma} = 4^\nu n! \frac{2K_0^n + \nu \sigma e^{-K_0\sigma}}{2K_0^\nu} \leq \frac{1}{\delta} \frac{4^\nu n!}{2(2n+\nu) \log 2} \leq 1.$$ 

Indeed, the term on the right depends only on $n$, and is decreasing with this variable. Since it takes a value less than 1 for $n = 1$, the result follows directly. Hence, we have

$$\|R\|_{\mathcal{F},s-\sigma,h_0} \leq 5\epsilon_0.$$

**Solving the cohomological equation:** We would now like to solve the equation (15) in $F$. Letting $\tilde{N} = N_+ - N$, we can write:

$$\{F, N\} + \tilde{N} = R.$$

Recall that in the outline, we wanted $\tilde{N} = R = \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} Rd\theta$. With the work done previously, we get the following bound on $\tilde{N}$:

$$\|\tilde{N}\|_{\mathcal{F},h_0} \leq \|R\|_{\mathcal{F},s-\sigma,h_0} \leq 5\epsilon_0.$$

Since the Fourier series of $R$ only contains terms of indices $k$ such that $|k| \leq K_0$, we can apply the theorem of Rüssmann 2 of appendix B, and solve the remainder of the cohomological equation (15). The norm of the Hamiltonian $F$ solving this equation therefore verifies:

$$\|F\|_{\mathcal{F},s-2\sigma,h_0} \leq \frac{C_0 \|R\|_{\mathcal{F},s-\sigma,h_0}}{\gamma \sigma^n} \leq \frac{5C_0 \epsilon_0}{\gamma \sigma^n}.$$ 

We multiplied $C_0$ by a factor 2, in order to get rid of this same factor in the Diophantine condition satisfied on $O_{h_0}$. With Cauchy’s inequality, we get:

$$\|F_\theta\|_{\mathcal{F},s-3\sigma,h_0} \leq \frac{5C_0 \epsilon_0}{\gamma \sigma^n}, \quad \|F_I\|_{\mathcal{F},s-3\sigma,h_0} \leq \frac{20C_0 \epsilon_0}{\gamma \sigma^n}.$$

**Estimates on the transformation $\Phi$:** After obtaining the estimates on the derivatives of $F$, we can deduce some estimates on the vector field associated to $F$, and then on the time-1 map $\Phi$. On the domain $D_{\mathcal{F},s-3\sigma,h_0}$, we get:

$$\|F_\theta\|_{\mathcal{F},s-3\sigma,h_0} \leq \sqrt{\frac{\epsilon_0}{\gamma r \sigma^n}} \leq \sqrt{\frac{5C_0}{4^\nu (200nC_0 + 32 + 4^\nu n!)}} \eta r \leq \frac{\eta r}{8},$$

$$\|F_I\|_{\mathcal{F},s-3\sigma,h_0} \leq \frac{20C_0}{4^{2^\nu (200nC_0 + 32 + 4^\nu n!)} \sigma} \leq \sigma.$$ 

With these two inequalities, the time-1 map $\Phi$ is well-defined on the domains:

$$\Phi = \Phi'|_{t=1} : D_{\mathcal{F},s-4\sigma,h_0} \rightarrow D_{\mathcal{F},s-3\sigma,h_0},$$

$$\Phi = \Phi'|_{t=1} : D_{\mathcal{F},s-5\sigma,h_0} \rightarrow D_{\mathcal{F},s-4\sigma,h_0}.$$ 

Only considering the first domain is not enough to prove the KAM step. Indeed, the estimates of the difference $P - Q$ requires to lose a lot of analyticity on the actions to keep this term small. Writing $\Phi = (U, V)$, since $F$ is linear in the actions, $V$ is independent of $I$. The Jacobian of $F$ is

$$\Phi' = \begin{pmatrix} U_I & U_\theta \\ 0 & V_\theta \end{pmatrix}.$$
On the set $D_{z,s-5\sigma,h_0}$, and hence on the set $D_{\eta r,s-5\sigma,h_0}$, the following inequalities are satisfied

$$
\|U - Id\| \leq \|F_\theta\| \leq \frac{5C_0\epsilon_0}{\gamma r\sigma^\nu}, \quad \|V - Id\| \leq \|F_\theta\| \leq \frac{20C_0\epsilon_0}{\gamma r\sigma^\nu},
$$

$$
\|U_t - Id\| \leq \frac{4C_0\epsilon_0}{\gamma r\sigma^\nu}, \quad \|U_\theta\| \leq \frac{5C_0\epsilon_0}{\gamma \sigma^\nu+1}, \quad \|V_\theta - Id\| \leq \frac{20C_0\epsilon_0}{\gamma r\sigma^\nu},
$$

whence the estimates on $\Phi$ in the proposition.

**Estimates on the new perturbation:** After the transformation, the Hamiltonian takes the form $H = N^+ + P^+$. We now work on a bound on the norm of $P^+$. First, consider $\{R,F\}$:

$$
\|\{R,F\}\|_{z,s-3\sigma,h_0} \leq n(\|R_t\|_{z,s-3\sigma,h_0} + \|\partial_t F\|_{z,s-3\sigma,h_0} + \|R_\theta\|_{z,s-3\sigma,h_0}) \leq n\left(\frac{20C_0\epsilon_0}{r\gamma r\sigma^\nu} + \frac{5C_0\epsilon_0}{\sigma} \frac{20C_0\epsilon_0}{\gamma r\sigma^\nu}\right) \leq 200nC_0\frac{\epsilon_0^2}{\gamma r\sigma^\nu}.
$$

(17)

The same inequality stays true for $\{\hat{N},F\}$. Hence:

$$
\left\| \int_0^1 \{(1-t)\hat{N} + tR,F\} \right\|_{\eta r,s-\sigma,h_0} \leq \left\| \{(1-t)\hat{N} + tR,F\} \right\|_{z,s-4\sigma,h_0} \leq 200nC_0\frac{\epsilon_0^2}{\gamma r\sigma^\nu}.
$$

(18)

It remains to find the bound of the term induced by $P - R$:

$$
\| (P - R) \Phi \|_{\eta\sigma,h_0} \leq \|P - R\|_{\eta r,s-4\sigma,h_0} \leq \|P - Q\|_{\eta r,s-4\sigma,h_0} + \|Q - R\|_{\eta r,s-4\sigma,h_0} \leq (32\eta^2 + 4\nu n!K^n e^{-K\sigma})\epsilon_0.
$$

(19)

In the end, the estimate on the norm of $P^+$ is the following:

$$
\|P^+\|_{\eta r,s-5\sigma,h_0} \leq 200nC_0\frac{\epsilon_0^2}{\gamma r\sigma^\nu} + (32\eta^2 + 4\nu n!K^n e^{-K\sigma})\epsilon_0 = \epsilon_0^+.
$$

(20)

**Exponential decrease:** As explained before, we are interested in the exponential decrease of the ratio $E = \frac{\epsilon_0^+}{\gamma r\sigma^\nu}$. Let us already think about the iterative step, and choose the variables we will use at the next step. First, let $\delta^+ = \delta/2$, $h^+ = h_0/4\nu$ and $K^+ = 4K_0$; regarding the actions, since we have to lose much more analyticity width and we let $r^+ = \eta r$. With these definitions, let us compute $E^+$:

$$
E^+ = \frac{\epsilon_0^+}{\gamma r\sigma^\nu} = \frac{2\nu \epsilon_0^+}{\gamma r\sigma^\nu} \leq 2\nu \frac{\epsilon_0^2}{\eta} \frac{\gamma r\sigma^\nu}{(\gamma r\sigma^\nu)^2} + 2\nu (32\eta^2 + 4\nu n!K^n e^{-K\sigma}) \frac{\epsilon_0}{\gamma \eta r\sigma^\nu} \leq 2\nu \frac{200nC_0}{\eta} \frac{E^2}{\eta} + (32\eta^2 + 8\nu n!K^n e^{-K\sigma})\frac{E}{\eta}.
$$

Using the fundamental inequality (16), we have $2\nu E_0^+ \geq K^n_0 e^{-K_0\sigma}$. Observe as well that $E = \eta^2$, hence:

$$
E^+ \leq 2\nu (200nC_0 + 32 + 8\nu n!E^2) = \sqrt{C_1 E^2},
$$

(21)

i.e. $C_1 E^+ \leq (C_1 E)^{\frac{1}{2}}$. The scheme converges exponentially fast if $E < C_1^{-1}$, therefore if $\epsilon_0 \leq \frac{2\nu r\sigma^\nu}{C_1}$. The initial condition on $\epsilon$, and on $\epsilon_0 \leq \epsilon$ shows that we have even better: $C_1 E \leq 4^{-\nu}$, hence the exponential decrease of $E$. 

8
Change in the frequencies: It remains to deal with the function $\varphi$, which controls the frequency shift when adding the mean of the linearized perturbation over the angles. We use lemma 2 of appendix C, so as to make explicit the domain on which this map is well-defined. Let $v = \hat{N}_I = [R_I]$, the new frequency vector is $\omega^+ = \omega + v(\omega)$. Computing the norm of $v$ gives:

$$\|v\|_{h_0} = \|N_I\|_{h_0} \leq \frac{5}{4r} \epsilon_0 \leq \frac{10\epsilon_0}{r} \leq \frac{10h_0}{\delta} \leq \frac{h_0}{4}$$

Applying lemma 2, we obtain the inverse map $\varphi : O_{h_0}^+ \rightarrow O_{h_0}$, $\omega^+ \rightarrow \omega$, satisfying:

$$\|\varphi - Id\|_{h_0} \leq \frac{10\epsilon_0}{r} \quad \text{(22)}$$

$$\|D\varphi - Id\|_{h_0} \leq \frac{5\epsilon_0}{2h_0r} \quad \text{(23)}$$

In this configuration, we let $N^+ = (N + \hat{N}) \circ \varphi$, and we obtained the new Hamiltonian $H^+ = N^+ + P^+$.

4 End of proof of the theorem

Soundness of the iteration:

To iterate the KAM step, the hypotheses at a step $j + 1$ need to be fulfilled knowing that they are at a step $j$. We therefore use the new value of each variables obtained after one KAM step, and check if they satisfy the hypotheses of the KAM step. Recall that after a step $j$, we have: $K_{j+1} = 4K_j$, $\sigma_{j+1} = \sigma_j/2$, $\eta_j = \sqrt{\frac{\epsilon_j}{\gamma_j \sigma_j}}$, $r_{j+1} = \eta_j r_j$, $h_{j+1} = h_j/4^\nu$.

It is necessary to notice that the constant $\delta$ correlating the value of $\epsilon_{j+1}$ to $h_{j+1}$ is not conserved after one step, however it is increasing and still satisfies $\delta \geq 40$. This value fixes the size of the transformation: the bigger it is, the smaller the norms $\varphi - Id$ and $\hat{\Phi} - \Phi_0$. Obviously, after one step, since the new perturbation is much smaller, the new transformation will also be smaller and therefore we can let $\delta$ increase.

Regarding the equalities that need to be fulfilled:

- The equality $K_{j+1} \sigma_{j+1} = 2K_j \sigma_j$ shows that $K_{j+1}$ belongs again to the set $A_{j+1}$.

- Let us check that $K_{j+1}$ belongs to $B_{+,j+1}$ as well:

$$2K_{j+1}^{n+\nu} \sigma_{j+1}^{\nu} \exp(-K_{j+1} \sigma_{j+1}) = 2(4^{n+\nu} K_j^{n+\nu}) \left(\frac{\sigma_j^{\nu}}{2^{\nu}}\right) \exp(-2K_j \sigma_j) \leq \frac{4^{n+\nu} \exp(-K_j \sigma_j)}{\delta} = 2^{2n+\nu} \exp(-K_j \sigma_j).$$

The condition to have $K_{j+1} \in B_{+,j+1}$ is therefore the following:

$$2^{2n+\nu} \exp(-K_j \sigma_j) \leq 1.$$

Since $K_j \in A_j$ it is satisfied.

- In the KAM step, we defined the limit value $\epsilon = \epsilon^-$. Define now

$$\epsilon^+ = \min \left( \frac{\gamma r_{j+1}^{\nu} \sigma_{j+1}^{\nu}}{4^\nu C_1}, \frac{h_{j+1} r_{j+1}}{\delta}, \frac{\gamma}{2K_{j+1}^{\nu} \delta} \right).$$

With the hypotheses, we have in fact

$$\epsilon^+ = \min \left( \frac{\eta_j}{2^{\nu}}, \frac{\eta_j}{4^{\nu}}, \frac{1}{4^{\nu}} \right) \epsilon^- = \frac{\eta_j \epsilon^-}{4^\nu}. $$
$\epsilon^+$ is therefore the new limit of the application of the KAM step. We have to check the condition $\epsilon_{j+1} \leq \epsilon^+$. 

$$
\epsilon_{j+1} = \gamma r_j \sigma_j^+ E_j+1
\leq \frac{\eta_j}{2\nu} \gamma r_j \sigma_j^+ \sqrt{C_1} E_j^3
\leq \frac{\eta_j}{2\nu} \gamma r_j \sigma_j^+ \sqrt{C_1} \left(\frac{\epsilon_j^3}{r_j \sigma_j^+}\right)^2 \leq \frac{\eta_j \epsilon_j \sqrt{C_1}}{2\nu} \sqrt{\frac{\epsilon_j}{\gamma r_j \sigma_j^+}}.
$$

By assumption, $\epsilon_j \leq \epsilon^- \leq \frac{\gamma r_j \sigma_j^+}{\epsilon^+}$. Hence $\epsilon_{j+1} \leq \eta_j \epsilon^- / 4\nu = \epsilon^+$. The condition $\epsilon \leq \frac{\gamma r_j \sigma_j^+}{4\nu}$ that we imposed in the definition of $\epsilon$, corresponding to the control of the transformation among the actions and angles, allows us to iterate the KAM step here. The fundamental inequality (16) holds at step $j+1$, in particular $K_{j+1}^n \exp(-K_{j+1} \sigma_{j+1}) \leq 2\nu E_{j+1}$. Let us express now $E_{j+1}$ using $\epsilon^+$. 

$$
K_{j+1}^n \exp(-K_{j+1} \sigma_{j+1}) = 4^n (K_j^n \exp(-K_j \sigma_j)) \exp(-K_j \sigma_j)
\leq 2^{2\nu} \epsilon^- \frac{\epsilon_j^2}{\gamma r_j \sigma_j^+} 4^n \exp(-K_j \sigma_j)
\leq 2^{2\nu} 4^{\nu} \epsilon^+ \frac{\eta_j}{2^\nu \gamma r_j \sigma_j^+} 4^n \exp(-K_j \sigma_j)
\leq 2^{2\nu} E_{j+1} 2^{2\nu+\nu} \exp(-K_j \sigma_j).
$$

Since $K_j \in A_j$, we have: $K_{j+1}^n \exp(-K_{j+1} \sigma_{j+1}) \leq 2^{2\nu} E_{j+1}$.

- It remains to check if the quantity $\frac{\epsilon_j}{h_j r_j}$ is decreasing:

$$
\frac{\epsilon_{j+1}}{h_{j+1} r_{j+1}} = \frac{\gamma \sigma_j E_{j+1}}{h_{j+1} r_{j+1}} = \frac{2^{2\nu} \gamma \sigma_j E_{j+1}}{h_j} = \frac{2^{2\nu} E_{j+1}}{h_j} \frac{\epsilon_j}{h_j r_j} \leq \frac{2^{2\nu} C_1}{h_j} \frac{\epsilon_j}{h_j r_j} \leq \frac{\epsilon_j}{h_j r_j}.
$$

We have checked all the inequalities required to iterate the KAM step an infinite amount of time. The scheme is well-defined; we can now compute the size of the transformations.

Transformations involved and their estimates:

The initial Hamiltonian is $H = N + P$. At each KAM step, we define two transformations: $\Phi_j$ which modifies the action-angle coordinates, and $\varphi_j$ which modifies the frequencies. We let $s_{j+1} = s_j - 2\sigma_j$, with $s_0 = s$, $r_0 = r$, $\eta_0 = \eta$ and $\eta_j = \sqrt{\frac{\nu}{\gamma r_j \sigma_j^+}} = \sqrt{E_j}$. Define $F_0 = Id$, and for $j > 0$:

$$
F_{j+1} : D_{j+1} \times O_{j+1} \rightarrow D_j \times O_j
(I, \theta, \omega) \mapsto (\Phi_{j+1}(I, \theta, \omega), \varphi_{j+1}(\omega)),
$$

with

$$
D_j = \{I \in \mathbb{C}^n : |I| < r_j\} \times \{\theta \in \mathbb{T}^n : |\Im(\theta)| \leq s_j\},
O_j = \{\omega \in \mathbb{R}^n : |\omega - \Omega^\beta| < h_j\}.
$$

Call $F^j = F_0 \circ \cdots \circ F_{j-1}$. We then have:

$$
F^j : D_j \times O_j \rightarrow D_0 \times O_0
$$

Thereafter, we will give some estimates on the transformation $F^j$, and show its convergence.
**Preliminaries:** The map $\mathcal{F}$ transforms a torus associated to a frequency vector belonging to the set $\Omega_{\gamma, \tau}$ to a deformed torus where the motion has frequencies belonging to the set $\Omega_\gamma^3$. The action $p_0$ on the first torus are entirely and uniquely determined by the frequency vector. The uniqueness comes from the hypothesis of non-degeneracy of the unperturbed Hamiltonian. In order to be precise, we define the following mapping:

$$\Psi: T^n \times \Omega_{\gamma, \tau} \to D \times T^n.$$ 

Define as well the map:

$$\Xi: B \times T^n \times \Omega_{\gamma, \tau} \to D \times T^n$$

where again it is necessary to check the increase of $\Phi_j$.

Then, one can define $\Psi$ as follows:

$$\Psi: T^n \times \Omega_{\gamma, \tau} \to D \times T^n$$

$$\{I, \theta, \omega\} \mapsto (h_p^{-1}(\omega) + I, \theta).$$

Assume $\Phi: \{0\} \times T^n \to B \times T^n$ and $\varphi: \Omega_\gamma^3 \to \Omega_{\gamma, \tau}$ exist as a limit of $\Phi_j$ and $\varphi_j$. Then, one can define $\Psi$ as follows:

$$\Psi: T^n \times \Omega_{\gamma, \tau} \to D \times T^n$$

$$\{\theta, \omega\} \mapsto \Xi(\Phi(0, \theta), \varphi(\omega))$$

The KAM theorem shows that, on $T^* \times \Omega_{\gamma, \tau}$, $H \circ \Psi = N'$, where $N' = \lim_{j \to \infty} N_j$.

**Estimates on the transformations:** In order to simplify the formulas, we introduce the weight matrix $W_j = \text{diag}(r_j^{-1}I_d, \sigma_j^{-1}I_d)$. Recall the size of the transformation obtained previously on the set $D_j \times O_j$:

$$\|W_j(\Phi_j - I_d)\| \leq \frac{20C_0\epsilon_j}{\gamma_j r_j\sigma_j}, \quad \|\varphi_j - I_d\| \leq \frac{10\epsilon_j}{r_j},$$

$$\|W_j(\Phi_j' - I_d)W_j^{-1}\| \leq \frac{40C_0\epsilon_j}{\gamma_j r_j\sigma_j}, \quad \|\varphi_j' - I_d\| \leq \frac{5\epsilon_j}{2h_j r_j}.$$ 

We can estimate the norm of the difference between to consecutive transformations $F^j$.

$$\|W_0(\Phi^{j+1} - \Phi^j)\| = \|W_0(\Phi^j \circ \Phi_j - \Phi^j)\|$$

$$\leq 2n \|W_0(\Phi^j)W_j^{-1}\| \|W_j(\Phi_j - I_d)\|$$

$$\leq 2n\epsilon_j \|W_j(\Phi_j - I_d)\|$$

$$\leq \xi_j \frac{40nC_0\epsilon_j}{\gamma_j r_j\sigma_j},$$

it is well-defined when $j$ goes to infinity if the variable $\xi_j = \|W_0D\Phi^jW_j^{-1}\|$ does not increase too fast on $D_j$.

In the same way, we compute:

$$\|h_0^{-1}(\varphi^{j+1} - \varphi^j)\| = \|h_0^{-1}(\varphi^j \circ \varphi_j - \varphi^j)\|$$

$$\leq n \|h_0^{-1}(\varphi^j)h_j^{-1}\| \|h_j(\varphi_j - I_d)\|$$

$$\leq n\mu_j \|h_j(\varphi_j - I_d)\|$$

$$\leq \mu_j \frac{10n\epsilon_j}{\gamma_j r_j},$$

where again it is necessary to check the increase of $\mu_j = \|h_0^{-1}(\varphi^j)h_j^{-1}\|$ on $O_j$.

On $D_j \times O_j$, we have in fact $(\Phi^j) = (\Phi_0)^j \cdots (\Phi_{j-1})^j$, where the differentials are estimated on different points, that are not important to explicit as we have a bound on their whole set of definition. With the decrease of the variables $r$ and $\sigma$, we get $\|W_jW_{j+1}\| \leq 1/2$. Hence,

$$\xi_j = \|W_0(\Phi^j)W_j^{-1}\| = \|W_0(\Phi_0)^j \cdots (\Phi_{j-1})^jW_j^{-1}\|$$

$$\leq \|W_0(\Phi_0)^jW_0^{-1}\| \|W_0W_j^{-1}\| 2n \|W_0W_j^{-1}\| 2n \cdots \times 2n \|W_j^{-1}(\Phi_{j-1})W_j^{-1}\| 2n \|W_{j-1}W_j^{-1}\|$$

$$\leq (2n)^{2j} \left( \frac{1}{2} \right)^j \prod_{i=1}^j \left( 1 + \frac{40C_0\epsilon_j}{\gamma_j r_j\sigma_j} \right) \leq (2n^2)^j \prod_{i=1}^j \left( 1 + \frac{40C_0\epsilon_j}{\gamma_j r_j\sigma_j} \right),$$

11
Indeed, we have a product of $2j$ matrices, and the presence of $j$ matrices of the form $W_j W_{j+1}^{-1}$, hence the factor $(2n^2)^j$. Likewise, for $\mu_j$:

$$
\mu_j \leq \left( \frac{n^2}{4\nu} \right)^j \prod_{i=1}^{j} \left( 1 + \frac{5\epsilon_j}{2h_j r_j} \right) \leq \prod_{i=1}^{j} \left( 1 + \frac{5\epsilon_j}{2h_j r_j} \right).
$$

This time we can get rid of the factor depending on $n$ because of the factor $4\nu$ between $h_j$ and $h_{j+1}$. Since the variables $\epsilon_j$ decreases exponentially fast towards 0, and that the terms $h_j$ and $r_j$ do not decrease as fast, the products in the formulas will converge when $j$ goes to infinity. We can bound them using the estimates we obtained in the KAM step. First, recall that $40C_0\epsilon_j < \gamma r_j \sigma_i^\nu$, whence, using the logarithm for $j \geq 1$:

$$
\log \left( (2n^2)^{-j} \xi_j \right) \leq \sum_{i=1}^{j} \log \left( 1 + \frac{40C_0\epsilon_i}{\gamma r_i \sigma_i^\nu} \right) \leq \sum_{i=1}^{j} \frac{40C_0\epsilon_i}{\gamma r_i \sigma_i^\nu} \leq \sum_{i=1}^{j} 40C_0 E_i.
$$

Using the exponential decrease of $E_j$:

$$
E_j \leq \sqrt{C_4 E_{j-1}^2} \leq \ldots \leq C_4^{\frac{j}{2}} \sum_{i=1}^{j} E_i^i (\xi_j)^i \leq (C_4 E_0) (\xi_j)^{j-1} E_0 \leq 4^{-\nu} (\xi_j)^{j-1} E_0.
$$

Finally:

$$
(2n^2)^{-j} \xi_j \leq \exp \left( \sum_{i=1}^{\infty} 40C_0 E_0 \left( 4^{-\nu} (\xi_j)^i \right) \right) \\
\leq \exp \left( 40C_0 E_0 \sum_{i=1}^{\infty} 4^{-\nu} (\xi_j)^i \right) \\
\leq \exp \left( 10C_0 E_0 \right) = \exp \left( 10C_0 \frac{\epsilon_0}{\gamma \sigma_0^\nu} \right) \equiv \xi.
$$

In the same way, we get for $\mu_j$:

$$
\mu_j \leq \exp \left( \frac{5}{2} \sum_{i=1}^{\infty} \frac{\epsilon_i}{r_i h_i} \right) = \exp \left( \frac{5}{2} \sum_{i=1}^{\infty} \frac{\gamma E_i \sigma_i^\nu}{h_i} \right) \\
\leq \exp \left( \frac{5}{2} \sum_{i=1}^{\infty} \frac{\gamma E_i \sigma_i^\nu 2^\nu}{h_0} \right) \\
\leq \exp \left( \frac{5}{2} \frac{\gamma \sigma_0^\nu}{h_0} \sum_{i=1}^{\infty} E_0 2^{\nu-2\nu (\xi_j)^i + 2\nu} \right) \\
\leq \exp \left( \frac{5}{2} \frac{\epsilon_0}{r_0 h_0} \right) \leq \exp \left( \frac{5}{6} \right) \equiv \mu.
$$

With this computation, we can continue towards our aim of estimating $\mathcal{F}^j$ for all $j \geq 1$.

$$
\| W_0 (\Phi^j - \Phi_0) \| \leq \sum_{i=0}^{j-1} \| W_0 (\Phi^{i+1} - \Phi^i) \| \\
\leq \sum_{i=0}^{j-1} \xi_i \frac{40nC_0\epsilon_i}{\gamma r_i \sigma_i^\nu} \leq 40nC_0 \xi \sum_{i=0}^{\infty} (2n^2)^i E_i \\
\leq 40nC_0 E_0 \xi \sum_{i=0}^{\infty} (2n^2)^i 4^{-\nu (\xi_j)^i + \nu} \\
\leq 60nC_0 E_0 \xi = 6n \xi \log \xi
$$
As well, for all \( j \geq 1 \),
\[
\|h_0^{-1}(\varphi^j - Id)\| \leq \sum_{i=0}^{j-1} \|h_0^{-1}(\varphi^{i+1} - \varphi^i)\| \leq \sum_{i=0}^{j-1} \mu_i \frac{10n\epsilon_i}{h_0 r_i} \leq 4n\mu \log \mu.
\]

Therefore, with these uniform bounds, we can let \( j \) go to infinity. The transformation \( \mathcal{F} \) is well-defined on \( T^* \times \Omega_{\gamma,\tau} \). The set \( \Omega^3 \), defined while constructing \( \varphi^j \), depends on \( \Omega_{\gamma,\tau} \) and on \( h_0 \). More precisely, recall that for all \( \omega' \in \Omega^3 \), there exists \( \omega \in \Omega_{\gamma,\tau} \) such that \( |\omega - \Omega^3| < h_0 \). Therefore, we can let \( \beta = h_0 \) so that the set \( O_h \subset \Omega \).

**First conclusion on the transformation:** Before computing the Lipschitz norm of the transformation, we are going to give some conclusion on the transformation we built. First, we have the relation \( H \circ \Xi \circ \mathcal{F}^j - N^j = P^j \) on the set \( D_j \times O_j \) for all \( j \geq 1 \). With this equality, we can describe the difference between the vector field associated to \( H \) and the one associated to \( N^j \). Formally, by derivation on the action-angle coordinates, we have
\[
\|H \circ \Xi \circ \mathcal{F}^j - N^j\|_{0,s_j-\sigma_j} = \max \left( \frac{\epsilon_j}{r_j}, \frac{\epsilon_j}{\sigma_j} \right).
\]

For a fixed \( \omega \), the map \( \Xi \) is constant and linear in the actions and the angles, which simplifies the computation. Therefore, using Cauchy’s inequality on the action coordinates, we can bound the previous derivative.
\[
\|\| (\Phi^j)' \cdot \nabla H \circ \Xi \circ \mathcal{F}^j - \nabla N^j \|_{0,s_j-\sigma_j} \leq \max \left( \frac{\epsilon_j}{r_j}, \frac{\epsilon_j}{\sigma_j} \right).
\]

Using the weighted matrix \( W_j \) which was useful to give the estimates on \( (\Phi^j)' \), and the symplectic matrix \( J \),
\[
J = \begin{pmatrix} 0 & Id \\ -Id & 0 \end{pmatrix}, \quad W_j \times J = \begin{pmatrix} 0 & \sigma_j^{-1} \\ -r_j^{-1} & 0 \end{pmatrix},
\]
and multiplying on the left our relation by the latter matrix, we obtain:
\[
\|W_j \times X \times (\Phi^j)' \cdot \nabla H \circ \Xi \circ \mathcal{F}^j - W_j \times J \times \nabla N^j\|_{0,s_j-\sigma_j} \leq \frac{\epsilon_j}{r_j \sigma_j}.
\]

The map \( \Phi^j \) being symplectic, it satisfies \( J^t \Phi^j = (\Phi^j)^{-1} J \). Hence, multiplying the last inequality by \( W_0 \Phi^j W_j^{-1} \), we get:
\[
\|W_0 \cdot (J \nabla H \circ \Xi \circ \mathcal{F}^j - \Phi^j J \nabla N^j)\|_{0,s_j-\sigma_j} \leq \|W_0 \Phi^j W_j^{-1}\|_{0,s_j-\sigma_j} \leq \frac{\epsilon_j}{r_j \sigma_j} \leq (2n^2)^j \frac{\epsilon_j}{r_j \sigma_j}.
\]

Looking at the vector field, this inequality becomes:
\[
\|W_0\| \times \|X_H \circ \Xi \circ \mathcal{F}^j - (\Phi^j)' \cdot X_N\|_{0,s_j-\sigma_j} \leq (2n^2)^j \frac{\epsilon_j}{r_j \sigma_j}.
\]

Letting \( j \) go to infinity, we get the equality of these two vector fields on the set \( T^* \times \Omega_{\gamma,\tau} \), i.e., for some \( \omega \in \Omega_{\gamma,\tau} \), with \( N = (e(\omega) + (\omega,I)) \),
\[
X_H \circ \Psi = (\Phi^j) \circ X_N \quad \text{(24)}
\]

Observe that we wrote \( \Psi \) instead of \( \Phi \), it expresses the fact that we consider the "origin" of the action \( I \) at the point \( \varphi(\omega) \).

**Lipschitz norm of the transformation:** On the Cantor set \( \Omega^3_\beta \subset \Omega \), the formulas of the derivatives of \( (\varphi^j - Id) \) converge. Indeed, the existence of a uniform constant for \( \epsilon_j/(2n^2) \) and \( \mu_j \), and the exponentially fast convergence of the terms \( \frac{\epsilon_j}{r_j \sigma_j} \) for all \( m \) implies that the norm of every derivative of the mapping \( \mathcal{F}^j \) (more precisely its difference to \( \Phi_0, Id \)) converges, whatever the order of the derivative.
We compute the Lipschitz norm of the transformation $\mathcal{F}$ to finish proving the theorem. First we evaluate the derivative of the map $\varphi - Id$ with respect to $\omega$. Although the map $\varphi$ is defined on a Cantor set, it is possible to extend it in a Lipschitz way, and even in a $C^1$ map, using Whitney’s extension theorem (see article [16] or the statement of the theorem 4 in appendix D.3). We will not extend further on these notions, and will just compute an estimate on its norm.

$$
\| (\varphi^j - Id)' \|_{L^2} \leq \sum_{i=0}^{j-1} \| (\varphi^{i+1})' - (\varphi^i)' \|_{L^2} \leq \sum_{i=0}^{j-1} \| (\varphi^{i+1} - \varphi^i)' \|_{L^2} \leq \sum_{i=0}^{j-1} \frac{2\| \varphi^{i+1} - \varphi^i \|_{h_i}}{h_i} \leq h_0 \times 20n\mu \sum_{i=0}^{j-1} \frac{e_i}{r_i h_i^2}.
$$

As done before, we can sum these terms:

$$\| (\varphi^j)' - I_n \|_{L^2} \leq \frac{20n\mu h_0}{\delta h_0} \sum_{i=0}^{j-1} 2^\nu \left( 3i+2-\left( \frac{j}{2} \right)^\nu \right) \leq \frac{20n\mu}{\delta} S_\nu.$$

Hence, letting $j$ goes to infinity, we obtain the Lipschitz norm:

$$\| \varphi - Id \|_{L^2} \leq \frac{20n\mu}{\delta} S_\nu. \quad (25)$$

As for $\Phi - \Phi_0$, computing the Lipschitz estimate in the exact same way, we get:

$$\| W_0(\Phi - \Phi_0) \|_{L^2} \leq 4^{14} 80 n C_0 \frac{\log \mu + 6n \log \xi}{\gamma \sigma \delta} T_\nu.$$

5 Estimates in the initial actions

After giving the estimates on the map $\mathcal{F} = (\Phi, \varphi)$, we determine some estimates on the map $\Psi = \Xi \circ \mathcal{F}$. Indeed, using this transformation we shift the torus back to its original place around the action $p_0$.

Consider the difference $\Psi - \Psi_0$. We have:

$$(\Psi - \Psi_0)(0, \theta, \omega) = (g'((\varphi(\omega)) - g'(\omega) + (\Phi_1(0, \theta, \varphi(\omega)) - \Phi_{0,1}(0, \theta, \omega)), \Phi_2(0, \theta, \varphi(\omega)) - \Phi_{0,2}(0, \theta, \omega)),$$

where $\Phi(I, \theta) = (\Phi_1(I, \theta), \Phi_2(I, \theta))$. Hence, the norm of $\Psi - \Psi_0$ on the set $T^* \times \Omega_{\gamma, \tau}$ verifies:

$$\|W_0(\Psi - \Psi_0)\| \leq \frac{1}{r} \| g' \circ \varphi - g' \| + \| W_0(\Phi - Id) \| \leq n \sup_{\Omega_{\gamma, \tau}} \| g' \| \frac{\| \varphi - Id \|}{r} + \| W_0(\Phi - Id) \|,$$

$$\|W_0(\Psi - \Psi_0)\| \leq n \sup_{\Omega_{\gamma, \tau}} \| g' \| \times \frac{4n^2 h_0}{r} \log \mu + 6n \log \xi. \quad (26)$$

We can also compute an estimate on the Lipschitz norm of $\Psi$ with respect to $\omega$. The estimate on the Lipschitz norm of $\Phi - \Phi_0$ being known, we are interested in the map $\Upsilon(\omega) = g'(\varphi(\omega)) - g'(\omega)$. Let $\omega, \omega' \in \Omega_{\gamma, \tau}$, we have:

$$| \Upsilon(\omega) - \Upsilon(\omega') | = \left| \int_0^1 [g'(t \varphi(t) - Id)(\omega) - (\varphi - Id)(\omega)] dt - \int_0^1 [g'(t \varphi(t) - Id)(\omega') - (\varphi - Id)(\omega')] dt \right|.$$

In order to compute this norm, we need to add some intermediate terms under the integral. For the sake of simplicity, we use $v = \varphi - Id$, we have:

$$g''(\omega + tv(\omega)) \cdot v(\omega) - g''(\omega + tv(\omega)) \cdot v(\omega) =$$

$$g''(\omega + tv(\omega)) \cdot [v(\omega) - v(\omega')] + \left[ g''(\omega + tv(\omega)) - g''(\omega' + tv(\omega')) \right] \cdot v(\omega').$$
The first term of this sum is bounded by
\[ |g''(\omega + tv(\omega)) \cdot [v(\omega) - v(\omega')]| \leq n \sup_{\Omega_{\gamma,\tau}} |g''| \times |\varphi - Id|_L \times |\omega - \omega'|. \]

As for the second term, we write:
\[
|g''(\omega + tv(\omega)) - g''(\omega' + tv(\omega'))| \\
= \left| \int_0^1 g^{(3)}((1-s)(\omega + tv(\omega)) - s(\omega' + tv(\omega')))ds \right| |(\omega + tv(\omega)) - (\omega' + tv)| \\
\leq n \sup_{\Omega_{\gamma,\tau}} |g^{(3)}| \times (|\omega - \omega'| + t|v(\omega) - v(\omega'))| \\
\leq n \sup_{\Omega_{\gamma,\tau}} |g^{(3)}| \times (1 + t|\varphi - Id|_L) |\omega - \omega'|. 
\]

Injecting these bounds in the previous inequality, we obtain
\[
|\Upsilon - \Upsilon_0|_L \leq n \sup_{\Omega_{\gamma,\tau}} |g''| \times |\varphi - Id|_L + n^2 \sup_{\Omega_{\gamma,\tau}} |g^{(3)}| \times |\varphi - Id|_L \times \left( 1 + \frac{|\varphi - Id|_L}{2} \right). \tag{27}
\]

Finally,
\[
|\Psi - \Psi_0|_L \leq |\Upsilon - \Upsilon_0|_L + |\Phi - \Phi_0|_L. \tag{28}
\]

A Remainder of the truncated Fourier series

Let \( A_s \) be the set of functions defined on \( T^n \) that are bounded and analytic on the set \( T^n_s = \{ \theta \in T^n, |3\theta| < s \} \). Let \( f \in A_s \), for \( \theta \in T^n_s \), we can write
\[ f(\theta) = \sum_{k \in \mathbb{Z}^n} f_k e^{i k \cdot \theta}. \]

For all \( k \in \mathbb{Z}^n \), we have \( |f_k| \leq |f_s|e^{-|k|s} \). Indeed, this result is straightforward using the fact that \( f \) is \( 2\pi \)-periodic in each variable, and analytic and bounded on its set of definition.

Let us consider the truncation of order \( K \in \mathbb{N} \) of \( f \):
\[ T_K f = \sum_{k \leq K} f_k e^{i k \cdot \theta}. \]

**Lemma 1.** Let \( s > 0 \) and \( \sigma < s \). If \( f \in A_s \), and \( K \sigma \geq n - 1 \) then
\[ |f - T_K f|_{s - \sigma} \leq 4^n n! K^n e^{-K \sigma} |f|_s, \quad 0 \leq \sigma \leq s \]

**Proof.** We have:
\[
|f - T_K f|_{s - \sigma} \leq \sum_{k \in \mathbb{Z}^n, |k|_1 > K} |f_k| \exp(|k|_1(s - \sigma)) \\
\leq |f|_s \sum_{k \in \mathbb{Z}^n, |k|_1 > K} \exp(-|k|\sigma) \\
\leq 4^n |f|_s \sum_{l \in \mathbb{N}, l > K} l^{n-1} \exp(-l\sigma),
\]

where we used the fact that the number of \( k \in \mathbb{Z}^n \) such that \( |k|_1 = l \) is less than \( 4^n l^{n-1} \). As for the last sum,
since the general term is strictly decreasing, it can be bounded by the incomplete gamma function:

$$\sum_{l \in \mathbb{N}, l > K} l^{n-1} \exp(-l \sigma) \leq \int_{K}^{\infty} x^{n-1} \exp(-x \sigma) dx \leq \frac{1}{\sigma^n} \int_{K \sigma}^{\infty} x^{n-1} \exp(-x) dx$$

$$\leq \frac{(n-1)!}{\sigma^n} \exp(-K \sigma) \sum_{k=0}^{n-1} \frac{(K \sigma)^k}{k!} \leq \frac{(n-1)!}{\sigma^n} \exp(-K \sigma)n \times (K \sigma)^{n-1}$$

$$\leq n! \frac{K^{n-1}}{\sigma} \exp(-K \sigma) \leq n! K^n \exp(-K \sigma).$$

Injecting this result in the previous inequation, the lemma is proved. \qed

B Rüssmann optimal estimate on the cohomological equation

We recall here the result obtained by Rüssmann in [15], which gives an optimal estimate on the norm of the solution for generic analytic functions. Let $n \geq 1$ and $s > 0$, define:

$$T^n_s = \{ \theta \in T^n_{\mathbb{C}}, \forall i \in [1, n] : |\Im \theta_i| < s \}$$

$$A^s = \{ f : T^n_s \rightarrow \mathbb{C}, f \mathbb{C} - \text{analytic} \}$$

$$A^s_0 = \{ f \in A^s, \text{ s.t. } \int_{T^n_s} f = 0 \}.$$

Writing $|f|_s = \sup_{T^n_s} |f|$, we have the following theorem:

**Theorem 2** (Rüssmann). Let $\omega \in D_{\gamma, \tau}$ a Diophantine vector, and $g \in A^s_0$. Then the equation

$$\partial_\omega f = g$$

has a unique solution $f$ in $\bigcup_{0 < \sigma < s} A^{s-\sigma}$, and we have the following bound on the norm of $f$ for $0 < \sigma < s$:

$$|f|_{s-\sigma} \leq C_0 \frac{|g|_s}{\gamma \sigma^2}.$$

where $C_0 = \frac{3 \pi}{2} 6^{\frac{3}{2}} \sqrt{\frac{\Gamma(2\tau)}{2\tau}}$.

C Inversion of analytic map close to the identity

Recall that we defined the set $O_h$ as the open complex neighborhood of radius $h$ of the subset of frequencies $\Omega_{\gamma, \tau}$ for some $\gamma > 0$. Using as usual the sup-norm for maps and vectors, Pöschel proves the following lemma on the inversion of the frequency vector:

**Lemma 2.** Let $f : O_h \subset \mathbb{C}^n \rightarrow \mathbb{C}^n$ be analytic, such that $|f - Id| \leq \delta \leq h/4$ on the set $O_h$. Then $f$ has an analytic inverse $g$ on $O_{h/4}$, and it satisfies:

$$|g - Id|_{h/4} \leq \frac{h}{4} |g' - Id| \leq \delta$$

See the appendix of Pöschel’s paper [13] for the proof.

D Classical formulas for analytic multivariate functions

In this section, we recall Taylor’s theorem and Cauchy’s formula for multivariate functions (see [11] for more details on the latter). As well, we state the Whitney extension theorem (for the demonstration, see [16]).
Define the following notations for $\alpha \in \mathbb{N}^n$ and $x \in \mathbb{R}^n$ with $n > 0$:

$$|\alpha| = \alpha_1 + ... + \alpha_n,$$

$$\alpha! = \alpha_1! \alpha_2! ... \alpha_n!,$$

$$x^\alpha = x_1^{\alpha_1} ... x_n^{\alpha_n}.$$

Introduce as well for an analytic function $f$:

$$f^{(\alpha)} = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} ... \partial x_n^{\alpha_n}}.$$

### D.1 Taylor expansion of analytic function

**Theorem 3.** Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a function analytic at the point $a \in \mathbb{R}^n$. Then, for $k \geq 0$, there exists a function $R_k : \mathbb{R}^n \rightarrow \mathbb{R}$ such that:

$$f(x) - \sum_{|\alpha| \leq k} \frac{f^{(\alpha)}}{\alpha!} (x-a)^\alpha = R_k(x) = o((x-a)^k).$$

Moreover, on a closed ball $B$ around $a$, we have for $x \in B$:

$$R_k(x) = \sum_{|\beta| = k+1} R_{k,\beta}(x)(x-a)^\beta,$$

with the bound

$$\max_{x \in B} |R_{k,\beta}(x)| \leq \frac{1}{\beta!} \max_{|\alpha| = |\beta|} \left( \max_{x' \in B} f^{(\alpha)}(x') \right).$$

### D.2 Cauchy Formula

**Proposition 2** (Cauchy’s formula). Let $\Omega$ be an open set in $\mathbb{C}^n$, $f$ a function holomorphic on $\Omega$, $a \in \Omega$ and let $\rho = (\rho_1, ..., \rho_n)$ with $\rho_i > 0$ be such that $\overline{P(a,\rho)} \subset \Omega$. Then, for $z \in P(a,\rho)$, we have

$$f(z) = \frac{1}{(2\pi i)^n} \int_{|z_i - a_i| = \rho_i} \left( \int_{|z_n - a_n| = \rho_n} \frac{f(z_1, ..., z_n)}{(z_1 - a_1)(z_2 - a_2)(z_n - a_n)} dz_1 ... dz_n \right).$$

**Corollary 1** (Cauchy’s inequality). If $f$ is holomorphic on $\Omega$ and $\overline{P(a,\rho)} \subset \Omega$, we have

$$f^{(\alpha)}(a) \leq \left( \sup_{|z_i - a_i| = \rho_i} |f(z)| \right)^{\alpha!} \rho^{-\alpha}.$$

### D.3 Whitney theorem

Define, for some function $f$ defined on $\mathbb{R}^n$, some $m \in \mathbb{N}$, and some $\alpha \in \mathbb{N}^n$ such that $|\alpha| \leq n$, the following functions:

$$f_\alpha(x) = \sum_{|k| \leq m - |\alpha|} \frac{f_{k+\alpha}(x')}{k!} (x-x')^k + R_\alpha(x, x').$$

Now let $A$ be a closed subset of $\mathbb{R}^n$. We will need some condition of smoothness in this set.
**Definition 1** (*C^m* in the Whitney sense). Let *f* be a function defined in the set *A* and let *m* be a positive integer. *f* is said to be of class *C^m* in *A* in the Whitney sense if the functions *f_α* (*|α| ≤ m*) are defined in *A* and the remainders *R_α* are such that for any point *x* of *A*, any ε > 0, there exists δ > 0 such that if *x'*, *x''* are any two points of *A* ∩ *B*(*x*, δ) then

\[ |R_α(x', x'')| ≤ |x' − x''|^{m−|α|}ε. \]

With this definition, we have the following statement:

**Theorem 4** (Whitney). Let *A* be a closed subset of *R^n* and let *f* be of class *C^m* (*m* finite or infinite) in the Whitney sense. Then there is a function *F* of class *C^m* (in the ordinary sense) in *R^n* such that

1. \( F^{(α)}(x) = f_α(x) \) in *A* for \(|α| ≤ m\),
2. \( F(x) \) is analytic in *R^n*.

In particular, \( f = F_{1A} \).

**References**


