NORMAL FORM À LA MOSER FOR DIFFEOMORPHISMS AND A GENERALIZATION OF RÜSSMANN'S TRANSLATED CURVE THEOREM TO HIGHER DIMENSION

JESSICA ELISA MASSETTI

Abstract. We prove a discrete time analogue of 1967 Moser's normal form of real analytic perturbations of vector fields possessing an invariant, reducible, Diophantine torus; in the case of diffeomorphisms too, the persistence of such an invariant torus is a phenomenon of finite co-dimension. Under convenient non-degeneracy assumptions on the diffeomorphisms under study (torsion property for example), this co-dimension can be reduced. As a by-product we obtain generalizations of Rüssmann's translated curve theorem in any dimension, by a technique of elimination of parameters.

Contents

1. Introduction and results

Let $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$, $a, b \in \mathbb{R}, a < b$ and consider the twist map

$$P : \mathbb{T} \times [a, b] \to \mathbb{T} \times \mathbb{R}, \quad (\theta, r) \mapsto (\theta + \alpha(r), r),$$

where $\alpha'(r) > 0$: P preserves circles $r = r_0$, $r_0 \in [a, b]$, and rotates them by an angle which *increases* as r does (this is the twist property).
Moser in [24] proved that for any $r_0 \in (a, b)$ such that $\alpha(r_0)$ is Diophantine, if Q is an exact area preserving diffeomorphism sufficiently close to P, it has an invariant curve near $r = r_0$ on which the dynamics is conjugated to the rotation $\theta \mapsto \theta + \alpha(r_0)$.
In 1970, Rüssmann generalized this fundamental result to non-conservative twist

Date: August, 15th 2017.

diffeomorphisms of the annulus [3, 27, 32]. He showed that the persistence of a Diophantine invariant circle is a phenomenon of co-dimension 1: in general the invariant curve does not persist but it is translated in the normal direction. It is the "theorem of the translated curve" (see below for a precise statement).
As in Kolmogorov's theorem [20], the dynamics on the translated curve can be conjugated to the same initial Diophantine rotation because of the non degeneracy (twist) of the map. Herman gave a proof of the translated curve theorem for diffeomorphisms with rotation number of constant type [18], then generalized Rüssmann's result in higher dimension to diffeomorphisms of $\mathbb{T}^n \times \mathbb{R}$ ($\mathbb{T}^n = \mathbb{R}^n/2\pi\mathbb{Z}^n$) close enough to the rotation $(\theta, r) \mapsto (\theta + \alpha, r)$, α being a Diophantine vector, without assuming any twist hypothesis but introducing an external parameter in order to tune the frequency on the translated torus, yet breaking the dynamical conjugacy to the Diophantine rotation, see [32].
Up to our knowledge no further generalization in $\mathbb{T}^n \times \mathbb{R}^m$ of Rüssmann's theorem has been given so far.

The first purpose of this work is to prove a discrete-time analogue of Moser's 1967 normal form [26] of real analytic perturbations of vector fields on $\mathbb{T}^n \times \mathbb{R}^m$ possessing a quasi-periodic Diophantine, reducible, invariant torus. The normal form will then be used to deduce "translated torus theorems" under convenient non-degereracy assumptions. As a by-product, Rüssmann's classical theorem will be a particular case of small dimension. While Rüssmann and Herman consider smooth or finite differentiable diffeomorphisms, we focus here on the analytic category. Let us state the main results.

A normal form for diffeomorphisms. Let $\mathbb{T}^n = \mathbb{R}^n/2\pi\mathbb{Z}^n$ be the n-dimensional torus. Let V be the space of germs along $\mathbb{T}^n \times \{0\}$ in $\mathbb{T}^n \times \mathbb{R}^m = \{(\theta, r)\}$ of real analytic diffeomorphisms. Fix $\alpha \in \mathbb{R}^n$ and $A \in \mathrm{GL}_m(\mathbb{R})$, assuming that A is diagonalizable with (possibly complex) eigenvalues $a_1, \dots, a_m \in \mathbb{C}$.
Let $U(\alpha, A)$ be the affine subspace of V of diffeomorphisms of the form

$$P(\theta, r) = (\theta + \alpha + O(r), A \cdot r + O(r^2)), \tag{1.1}$$

where $O(r^k)$ are terms of order $\geq k$ in r which may depend on θ. For these diffeomorphisms $\mathrm{T}^n_0 = \mathbb{T}^n \times \{0\}$ is an invariant, reducible, α-quasi-periodic torus whose normal dynamics at the first order is characterized by $a_1, \dots, a_m$. We will collectively refer to $\alpha_1, \dots, \alpha_n$ and $a_1, \dots, a_m$ as the *characteristic frequencies* or *characteristic numbers* of T^n_0.
Let now $a_1, \dots, a_q \in \mathbb{C}$ be the pairwise distinct eigenvalues of A. We will impose the following Diophantine conditions for some $\gamma > 0$ and $\tau \geq 1$

$$\begin{aligned}
&\forall i = 1, \dots, q : |a_i| = 1 \\
&|k \cdot \alpha + \arg a_i - 2\pi l| \geq \frac{\gamma}{|k|^\tau} \qquad \forall (k, l) \in \mathbb{Z}^n \setminus \{0\} \times \mathbb{Z}, \\
&\forall i, j = 1, \dots, q : |a_i| = |a_j| \\
&|k \cdot \alpha + \arg a_i - \arg a_j - 2\pi l| \geq \frac{\gamma}{|k|^\tau} \qquad \forall (k, l) \in \mathbb{Z}^n \setminus \{0\} \times \mathbb{Z},
\end{aligned} \tag{1.2}$$

where $\arg a_i \in [0, 2\pi[$ denotes the argument of the i-th eigenvalue $a_i = |a_i| e^{i \arg a_i}$.

Remark 1.1. Note that A being in $\mathrm{GL}_m(\mathbb{R})$, the possible complex eigenvalues come in couples and that conditions (1.2) imply the classical Diophantine condition on α when $i = j$.

Let $\mathcal{G}$ be the space of germs of real analytic isomorphisms of $\mathbb{T}^n \times \mathbb{R}^m$ of the form

$$G(\theta, r) = (\varphi(\theta), R_0(\theta) + R_1(\theta) \cdot r), \tag{1.3}$$

where φ is a diffeomorphism of the torus fixing the origin and R_0, R_1 are functions defined on the torus $\mathbb{T}^n$ with values in $\mathbb{R}^n$ and $\mathrm{GL}_m(\mathbb{R})$ respectively and such that $\Pi_{\mathrm{Ker}(A-I)} R_0(0) = 0$ and $\Pi_{\mathrm{Ker}[A,\cdot]}(R_1(0) - I) = 0$, having denoted I the identity matrix in $\mathrm{Mat}_m(\mathbb{R})$ and Π the projection on the indicated subspace.
Let us define the "correction map"

$$T_\lambda : \mathbb{T}^n \times \mathbb{R}^m \to \mathbb{T}^n \times \mathbb{R}^m, \quad (\theta, r) \mapsto (\beta + \theta, b + (I + B) \cdot r),$$

where $\beta \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ and $B \in \mathrm{Mat}_m(\mathbb{R})$ are such that

$$(A - I) \cdot b = 0, \quad [A, B] = 0. \tag{1.4}$$

We will refer to translating parameters $\lambda = (\beta, b + B \cdot r)$ as *corrections* or *counter terms*, and denote with Λ the space of such λ's

$$\Lambda = \{\lambda = (\beta, b + B \cdot r) : (A - I) \cdot b = 0, [A, B] = 0\}.$$

Theorem A (Normal form)**.** *Let* (α, A) *satisfy the Diophantine condition* (1.2)*. If* Q *is sufficiently close to* $P^0 \in U(\alpha, A)$*, there exists a unique triplet* $(G, P, \lambda) \in \mathcal{G} \times U(\alpha, A) \times \Lambda$*, close to* $(\mathrm{id}, P^0, 0)$*, such that*

$$Q = T_\lambda \circ G \circ P \circ G^{-1}.$$

In the neighborhood of $(\mathrm{id}, P^0, 0)$, the $\mathcal{G}$-orbit of all $P's \in U(\alpha, A)$ has finite co-dimension. The proof is based on a relatively general inverse function theorem in analytic class (Theorem A.1 of the Appendix).
The idea of proving the finite co-dimension of a set of conjugacy classes of a diffeomorphism or of a vector field has been successfully exploited by many authors. Arnold at first proved a normal form for diffeomorphisms of $\mathbb{T}^n$ [1], followed by Moser's normal forms for vector fields [21, 22, 25, 26, 31]. Among other authors we recall Calleja-Celletti-deLaLlave work on conformally symplectic systems [4], Chenciner's study on the bifurcation of elliptic fixed points [5–7], Herman's twisted conjugacy for Hamiltonians [12, 13] (a generalization of Arnold's work [1]) or Eliasson-Fayad-Krikorian work around the stability of KAM tori [10].

This technique allows us to study the persistence of an invariant torus in two steps: first, prove a normal form that does not depend on any non-degeneracy hypothesis (but that contains the hard analysis); second, reduce or eliminate the (finite dimensional) corrections by the usual implicit function theorem, using convenient non degeneracy assumptions on the system under study.

A generalization of Rüssmann's theorem. From the normal form of Theorem A, we see that when $\lambda = 0$, $Q = G \circ P \circ G^{-1}$: the torus $G(\mathrm{T}_0^n)$ is invariant for Q and the first order dynamics is given by $P \in U(\alpha, A)$. Conversely, whenever $\lambda = (\beta, b)$, the torus is *translated* and the α-quasi-periodic tangential dynamics is *twisted* by the correction β:

$$Q(\varphi(\theta), R_0(\theta)) = (\beta + \varphi(\theta + \alpha), b + R_0(\theta + \alpha)).$$

We will loosely say that the torus T_0^n

- persists up to twist-translation, when $\lambda = (\beta, b)$
- persists up to translation, when $\lambda = (0, b)$

We stress the fact that Theorem A not only gives the tangential dynamics to the torus, but also the normal one, of which Rüssmann's original statement is regardless:

Theorem (Rüssmann)**.** *Let $\alpha \in \mathbb{R}$ be Diophantine and $P^0 : \mathbb{T} \times [-r_0, r_0] \to \mathbb{T} \times \mathbb{R}$ be of the form*

$$P^0(\theta, r) = (\theta + \alpha + t(r) + O(r^2), A^0 r + O(r^2)),$$

where $A^0 \in \mathbb{R} \setminus \{0\}$, $t(0) = 0$ and $t'(r) > 0$.
If Q is close enough to P^0 there exists a unique analytic curve $\gamma : \mathbb{T} \to \mathbb{R}$, close to $r = 0$, an analytic diffeomorphism φ of $\mathbb{T}$ close to the identity and $b \in \mathbb{R}$, close to 0, such that

$$Q(\theta, \gamma(\theta)) = (\varphi \circ R_\alpha \circ \varphi^{-1}(\theta), b + \gamma(\varphi \circ R_\alpha \circ \varphi^{-1}(\theta))).$$

Note that $t(r)$ may depend on the angles as well. In the original statement $A^0 = 1$; to consider this case with general A^0 does not add any difficulty to the proof.
We will generalize Rüssmann's theorem on $\mathbb{T}^n \times \mathbb{R}^n$. At the expense of loosing the control on the final normal dynamics and conjugating $T_\lambda^{-1} \circ Q$ to a diffeomorphism P whose invariant torus has a normal dynamics given by a different A, under convenient non-degeneracy conditions we can prove the existence of a twisted-translated or translated α-quasi-periodic Diophantine torus by application of the classical implicit function theorem in finite dimension. The following results will be proved in section 5, where a more functional statement will be given (Theorem 5.1 and Theorem 5.2).

On $\mathbb{T}^n \times \mathbb{R}^n$, let $P \in U(\alpha, A)$, defined in expression (1.1), be such that A is invertible and has simple, real eigenvalues $a_1, \dots, a_n$. This hypothesis clearly implies that the only frequencies that can cause small divisors are the tangential ones $\alpha_1, \dots, \alpha_n$, so that we only need to require the standard Diophantine hypothesis on α.

Theorem B. *Let α be Diophantine and let $A \in \mathrm{GL}_n(\mathbb{R})$ have simple, real eigenvalues. If Q is sufficiently close to $P^0 \in U(\alpha, A)$, there exists A' close to A such that the torus T_0^n persists up to twist-translation and its normal dynamics is given by A'.*

If, in addition, Q has a torsion property we can prove the following theorem.

Theorem C. *Let α be Diophantine and let A be invertible with simple, real eigenvalues. Let also*

$$P^0(\theta, r) = (\theta + \alpha + p_1(\theta)\cdot r + O(r^2), A\cdot r + O(r^2))$$

be such that

$$\det\left(\int_{\mathbb{T}^n} p_1(\theta)\,d\theta\right) \neq 0.$$

If Q is sufficiently close to P^0, there exists A' close to A such that the torus T^n_0 persists up to translation and the normal dynamics is given by A'.

The paper is organized as follows: in sections 2-3 we introduce the normal form operator, define conjugacy spaces and present the difference equations that will be solved to linearize the dynamics on the perturbed torus; in section 4 we will prove Theorem A while in section 5 we will prove Theorems B and C.

2. The normal form operator

We will show that the operator

$$\phi : \mathcal{G}\times U(\alpha, A)\times \Lambda \to V, \quad (G, P, \lambda)\mapsto T_\lambda\circ G\circ P\circ G^{-1}$$

is a local diffeomorphism (in the sense of scales of Banach spaces) in a neighborhood of $(\mathrm{id}, P^0, 0)$. Note that ϕ is formally defined on the whole space but $\phi(G, P, \lambda)$ is analytic in the neighborhood of T^n_0 only if G is close enough to the identity with respect to the width of analyticity of P. See section 2.3.

Although the difficulty to overcome in the proof is rather standard for conjugacy problems of this kind (proving the fast convergence of a Newton-like scheme), the procedure relies on a relatively general inverse function theorem (Theorem A.1 of section A), following a strategy alternative to Zehnder's in [33]. Both Zehnder's approach and ours rely on the fact that the fast convergence of the Newton' scheme is somewhat independent of the internal structure of the variables.

2.1. **Complex extensions.** Let us extend the tori

$$\mathbb{T}^n = \mathbb{R}^n/2\pi\mathbb{Z}^n \qquad \text{and} \qquad \mathrm{T}^n_0 = \mathbb{T}^n\times\{0\}\subset \mathbb{T}^n\times\mathbb{R}^m,$$

as

$$\mathbb{T}^n_{\mathbb{C}} = \mathbb{C}^n/2\pi\mathbb{Z}^n \qquad \text{and} \qquad \mathrm{T}^n_{\mathbb{C}} = \mathbb{T}^n_{\mathbb{C}}\times\mathbb{C}^m$$

respectively, and consider the corresponding s-neighborhoods defined using ℓ^∞-balls (in the real normal bundle of the torus):

$$\mathbb{T}^n_s = \left\{\theta\in\mathbb{T}^n_{\mathbb{C}} : \max_{1\le j\le n}|\mathrm{Im}\,\theta_j|\le s\right\} \quad \text{and} \quad \mathrm{T}^n_s = \{(\theta, r)\in \mathrm{T}^n_{\mathbb{C}} : |(\mathrm{Im}\,\theta, r)|\le s\},$$

where $|(\mathrm{Im}\,\theta, r)| := \max\left(\max_{1\le j\le n}|\mathrm{Im}\,\theta_j|, \max_{1\le j\le m}|r_j|\right)$.

Let now $f : \mathrm{T}^n_s \to \mathbb{C}$ be holomorphic on the interior of T^n_s, continuous on T^n_s, and consider its Fourier expansion $f(\theta, r) = \sum_{k\in\mathbb{Z}^n} f_k(r)\, e^{i\,k\cdot\theta}$, denoting by

$k \cdot \theta = k_1\theta_1 + \ldots + k_n\theta_n$. In this context we introduce the so called "weighted norm":

$$|f|_s := \sum_{k\in\mathbb{Z}^n} |f_k|\, e^{|k|s}, \quad |k| = |k_1| + \ldots + |k_n|,$$

where $|f_k| = \sup_{|r|<s} |f_k(r)|$. Whenever $f : \mathrm{T}^n_s \to \mathbb{C}^n$, $|f|_s = \max_{1\le j\le n}(|f_j|_s)$, f_j being the j-th component of $f(\theta, r)$.
It is a trivial fact that the classical sup-norm is bounded from above by the weighted norm:

$$\sup_{z\in\mathrm{T}^n_s} |f(z)| \le |f|_s$$

and that $|f|_s < +\infty$ whenever f is analytic on its domain, which necessarily contains some $\mathrm{T}^n_{s'}$ with $s' > s$. In addition, the following useful inequalities hold if f, g are analytic on $\mathrm{T}^n_{s'}$

$$|f|_s \le |f|_{s'} \text{ for } 0 < s < s',$$

and

$$|fg|_{s'} \le |f|_{s'}|g|_{s'}.$$

Moreover, one can show that if f is analytic on $\mathrm{T}^n_{s+\sigma}$ and G is a diffeomorphism of the form (1.3) sufficiently close to the identity, then $|f\circ G|_s \le C_G|f|_{s+\sigma}$, where C_G is a positive constant depending on $|G - \mathrm{id}|_s$ (see Appendix C). For more details about the weighted norm, see for example [9, 23].
In general for complex extensions U_s and $V_{s'}$, we will denote by $\mathcal{A}(U_s, V_{s'})$ the set of holomorphic functions from U_s to $V_{s'}$ and $\mathcal{A}(U_s)$, endowed with the s-weighted norm, the Banach space $\mathcal{A}(U_s, \mathbb{C})$.

Eventually, let E and F be two Banach spaces,

- We indicate contractions with a dot " $\cdot$ ", with the convention that if $l_1, \ldots, l_{k+p} \in E^*$ and $x_1, \ldots, x_p \in E$

  $$(l_1 \otimes \ldots \otimes l_{k+p}) \cdot (x_1 \otimes \ldots \otimes x_p) = l_1 \otimes \ldots \otimes l_k \langle l_{k+1}, x_1\rangle \ldots \langle l_{k+p}, x_p\rangle.$$

  In particular, if $l \in E^*$, we simply write $l^n = l \otimes \ldots \otimes l$.

- If f is a differentiable map between two open sets of E and F, $f'(x)$ is considered as a linear map belonging to $F \otimes E^*$, $f'(x) : \zeta \mapsto f'(x)\cdot\zeta$; the corresponding norm will be the standard operator norm

  $$|f'(x)| = \sup_{\zeta\in E, |\zeta|_E = 1} |f'(x)\cdot\zeta|_F.$$

2.2. **Spaces of conjugacies.**

- We consider the neighborhood of the identity $\mathcal{G}^\sigma_s$ in the space of germs of real holomorphic diffeomorphisms on T^n_s, defined by

  $$|\varphi - \mathrm{id}|_s \le \sigma$$

  and

  $$|R_0 + (R_1 - I)\cdot r|_s \le \sigma,$$

where $\varphi(0)=0$, R_0 and R_1 satisfy $\Pi_{\ker(A-I)}R_0(0)=0$ and $\Pi_{\ker([A,\cdot])}(R_1(0)-I)=0$.
The tangent space at the identity $T_{\mathrm{id}}\mathcal{G}^\sigma_s$, consists of maps $\dot{G}\in\mathcal{A}(\mathrm{T}^n_s,\mathbb{C}^{n+m})$

$$\dot{G}(\theta,r)=(\dot{\varphi}(\theta),\dot{R}_0(\theta)+\dot{R}_1(\theta)\cdot r),$$

where $\dot{\varphi}\in\mathcal{A}(\mathbb{T}^n_s,\mathbb{C}^n)$, $\dot{R}_0\in\mathcal{A}(\mathbb{T}^n_s,\mathbb{C}^m)$ and $\dot{R}_1\in\mathcal{A}(\mathbb{T}^n_s,\mathrm{Mat}_m(\mathbb{C}))$. We endow it with the norm

$$\left|\dot{G}\right|_s=\max_{1\le j\le n+m}\left(\left|\dot{G}_j\right|_s\right).$$

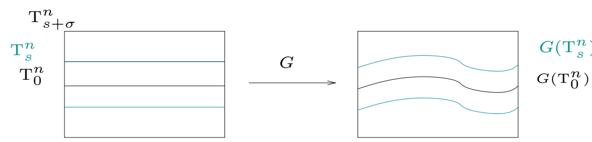

FIGURE 1. Deformed complex domain

- Let V_s be the subspace of $\mathcal{A}(\mathrm{T}^n_s,\mathbb{T}^n_{\mathbb{C}}\times\mathbb{C}^m)$ of diffeomorphisms

$$Q:(\theta,r)\mapsto(f(\theta,r),g(\theta,r)),$$

where $f\in\mathcal{A}(\mathrm{T}^n_s,\mathbb{C}^n),g\in\mathcal{A}(\mathrm{T}^n_s,\mathbb{C}^m)$, endowed with the norm

$$|Q|_s=\max\left(|f|_s,|g|_s\right).$$

- Let $U_s(\alpha,A)$ be the affine subspace of V_s of those diffeomorphisms P of the form

$$P(\theta,r)=(\theta+\alpha+O(r),A\cdot r+O(r^2)).$$

We will indicate with p_i and P_i the coefficients of the order-i term in r, in the θ and r-directions respectively.
- If $G\in\mathcal{G}^\sigma_s$ and P is a diffeomorphism over $G(\mathrm{T}^n_s)$ we define the following deformed norm

$$|P|_{G,s}:=|P\circ G|_s,$$

depending on G; this in order not to shrink artificially the domains of analyticity. The problem, in a smooth context, may be solved without changing the domain, by using plateau functions.

2.3. **The normal form operator.** By Theorem B.1 and Corollary B.1 the following operator

$$\phi:\quad \begin{array}{lll}\mathcal{G}^{\sigma/n}_{s+\sigma}\times U_{s+\sigma}(\alpha,A)\times\Lambda & \to & V_s\\ (G,P,\lambda) & \mapsto & T_\lambda\circ G\circ P\circ G^{-1}\end{array}\tag{2.1}$$

is now well defined. It would be more appropriate to write $\phi_{s,\sigma}$ but, since these operators commute with source and target spaces, we will refer to them simply as ϕ. We will always assume that $0<s<s+\sigma<1$ and $\sigma<s$.

3. Difference equations

We present here three lemmata that we will use in the following in order to linearize the tangent and the normal dynamics of the torus (see section 4).
Let $\alpha \in \mathbb{R}^n$ and let $M \in \mathrm{GL}_m(\mathbb{R})$ have pairwise distinct eigenvalues $\mu_1, \dots, \mu_m$. We assume the following Diophantine conditions on α and M:

(3.1)

$$|k \cdot \alpha - 2\pi l| \geq \frac{\gamma}{|k|^\tau}, \qquad \forall k \in \mathbb{Z}^n \setminus \{0\}, \forall l \in \mathbb{Z}$$

(3.2)

$$|k \cdot \alpha - \arg \mu_j - 2\pi l| \geq \frac{\gamma}{|k|^\tau}, \qquad \forall (k,l) \in \mathbb{Z}^n \setminus \{0\} \times \mathbb{Z}, \quad \forall j = 1, \dots, m : |\mu_j| = 1$$

(3.3)

$$|k \cdot \alpha + \arg \mu_i - \arg \mu_j - 2\pi l| \geq \frac{\gamma}{|k|^\tau}, \qquad \forall (k,l) \in \mathbb{Z}^n \setminus \{0\} \times \mathbb{Z}, \quad \forall i,j = 1, \dots, m : |\mu_i| = |\mu_j|$$

(3.4)

$$||\mu_i| - |\mu_j|| \geq \gamma, \quad \forall i,j = 1, \dots, m \quad i \neq j : |\mu_i| \neq |\mu_j|$$

$$|1 - |\mu_j|| \geq \gamma, \qquad \text{if} \quad |\mu_j| \neq 1$$

(3.5)

$$|\mu_i - \mu_j| \geq \gamma, \qquad \forall i,j = 1, \dots, m \quad i \neq j : |\mu_i| = |\mu_j|$$

$$|1 - \mu_j| \geq \gamma, \qquad \text{if } |\mu_j| = 1 \text{ and } \mu_j \neq 1$$

(3.6)

$$\min_{1 \leq j \leq m} (|\mu_j|) \geq \gamma.$$

We first prove the following fundamental lemma, which is the heart of the proof of Theorem A and, more generally, of many stability results related to Diphantine rotations on the torus.

Lemma 1. *Let $\alpha \in \mathbb{R}^n$ be Diophantine in the sense of (3.1) and let $a, b \in \mathbb{C} \setminus \{0\}$.*

(1) If $a = b$ and $|a| \geq \gamma$, for any $g \in \mathcal{A}(\mathbb{T}^n_{s+\sigma})$, there exists a unique f of zero average which is complex analytic on $\mathbb{T}^n_s$ and a unique $\lambda \in \mathbb{R}$ such that

$$\lambda + af(\theta + \alpha) - af(\theta) = g(\theta), \quad \lambda = \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} g \, d\theta,$$

satisfying

$$|f|_s \leq \frac{C}{\gamma^2 \sigma^{\tau+n}} |g|_{s+\sigma},$$

C being a constant depending only on n and τ.

(2) Let $a \neq b$.

(i) If $|a| = |b|$ *and*

$$\left\{\begin{array}{l} |a-b| \geq \gamma \\ |a| \geq \gamma \\ |k\cdot\alpha + \arg a - \arg b - 2\pi l| \geq \frac{\gamma}{|k|^\tau} \quad \forall (k,l) \in \mathbb{Z}^n \setminus \{0\} \times \mathbb{Z} \end{array}\right. \tag{3.7}$$

for any $g \in \mathcal{A}(\mathbb{T}^n_{s+\sigma})$*, there exists a unique* f *which is complex analytic on* $\mathbb{T}^n_s$ *such that*

$$a f(\theta+\alpha) - b f(\theta) = g(\theta), \tag{3.8}$$

satisfying

$$|f|_s \leq \frac{C}{\gamma^2 \sigma^{\tau+n}} |g|_{s+\sigma},$$

C *being a constant depending only on* n, τ*.*

(ii) If $|a| \neq |b|$ *and* $||a|-|b|| \geq \gamma$*, for any* $g \in \mathcal{A}(\mathbb{T}^n_{s+\sigma})$*, there exists a unique* f *which is complex analytic on* $\mathbb{T}^n_{s+\sigma}$ *such that*

$$a f(\theta+\alpha) - b f(\theta) = g(\theta),$$

satisfying

$$|f|_{s+\sigma} \leq \gamma^{-1} |g|_{s+\sigma}.$$

Proof. (1) Developing in Fourier series the equation yields

$$\lambda + a \sum_k (e^{i\,k\cdot\alpha} - 1) f_k e^{i\,k\cdot\theta} = \sum_k g_k e^{i\,k\theta};$$

letting $\lambda = g_0$ we formally have

$$f(\theta) = \frac{1}{a} \sum_{k\neq 0} \frac{g_k}{e^{i\,k\,\alpha} - 1} e^{i\,k\theta}.$$

First note that the coefficients g_k decay exponentially:

$$|g_k| = \left| \int_{\mathbb{T}^n} g(\theta) e^{-i\,k\cdot\theta}\, \frac{d\theta}{2\pi} \right| \leq |g|_{s+\sigma} e^{-|k|(s+\sigma)},$$

by deforming the path of integration to $\operatorname{Im}\theta_j = -\operatorname{sgn}(k_j)(s+\sigma)$.
Second, remark that for any $x, y \in \mathbb{R}^+$, $\varphi \in [0, 2\pi[$

$$\begin{aligned} \left|x\, e^{i\varphi} - y\right|^2 &= (x-y)^2 \cos^2\frac{\varphi}{2} + (x+y)^2 \sin^2\frac{\varphi}{2} \\ &\geq (x+y)^2 \sin^2\frac{\varphi}{2} = (x+y)^2 \sin^2\frac{\varphi - 2\pi l}{2}, \end{aligned} \tag{3.9}$$

with $l \in \mathbb{Z}$. By choosing $l \in \mathbb{Z}$ such that $-\frac{\pi}{2} \leq \frac{\varphi - 2\pi l}{2} \leq \frac{\pi}{2}$ we get

$$\left|x\, e^{i\varphi} - y\right| \geq \frac{2}{\pi}(x+y) \frac{|\varphi - 2\pi l|}{2}, \tag{3.10}$$

by the classical inequality $|\sin\delta| \geq \frac{2}{\pi}|\delta|$, whenever $-\frac{\pi}{2} \leq \delta \leq \frac{\pi}{2}$.
In our case $x = y = 1$, $\varphi = k\cdot\alpha$ and $\forall k$, by choosing $l \in \mathbb{Z}$ such that $-\frac{\pi}{2} \leq \frac{k\cdot\alpha - 2\pi l}{2} \leq \frac{\pi}{2}$ we get

$$\left|e^{i\,k\cdot\alpha} - 1\right| \geq \frac{4}{\pi} \frac{|k\cdot\alpha - 2\pi l|}{2} \geq \frac{2}{\pi} \frac{\gamma}{|k|^\tau},$$

by inequality (3.10) and the Diophantine condition (3.1).
We thus have

$$
\begin{aligned}
|f|_s &\le \frac{\pi\, |g|_{s+\sigma}}{|a|\gamma} \sum_k |k|^\tau e^{-|k|\sigma} \le \frac{\pi\, 2^n |g|_{s+\sigma}}{|a|\gamma} \sum_{\ell\ge 1} \binom{\ell+n+1}{\ell} e^{-\ell\sigma} \ell^\tau \\
&\le \frac{\pi\, 4^n |g|_{s+\sigma}}{|a|\gamma (n-1)!} \sum_{\ell \ge 1} (n+\ell-1)^{n-1+\tau} e^{-\ell\sigma} \\
&\le \frac{\pi\, 4^n |g|_{s+\sigma}}{|a|\gamma (n-1)!} \int_1^\infty (\ell+n-1)^{n+\tau-1} e^{-(\ell-1)\sigma}\, d\ell.
\end{aligned}
$$

The integral is equal to

$$
\begin{aligned}
&\sigma^{-\tau-n} e^{n\sigma} \int_{n\sigma}^\infty \ell^{\tau+n-1} e^{-\ell}\, d\ell \\
&< \sigma^{-\tau-n} e^{n\sigma} \int_0^\infty \ell^{\tau+n-1} e^{-\ell}\, d\ell = \sigma^{-\tau-n} e^{n\sigma} \Gamma(\tau+n).
\end{aligned}
$$

Hence f, of zero average, is complex analytic on $\mathbb{T}^n_s$ and, since $|a| \ge \gamma$ it satisfies the claimed estimate.

(2) Point (i). Let $a = |a| e^{i \arg a}$ and $b = |b| e^{i \arg b}$ with the convention that $\arg z = \pi$ ($\arg z = 0$) if $z \in \mathbb{R}^-$ (if $z \in \mathbb{R}^+$).
The Fourier's expansion gives

$$f_0 = \frac{g_0}{a-b}$$

and $\forall k \ne 0$

$$f_k = \frac{g_k}{e^{i\arg b}\big(|a| e^{i(\,k\,\alpha + \arg a - \arg b)} - |b|\big)} e^{i\,k\cdot\theta}.$$

In order to bound the divisors we apply the same inequalities as in (3.9)-(3.10), with $\varphi = k\cdot\alpha + \arg a - \arg b$. Since $|a| = |b|$, by conditions (3.7) we proceed as in the proof of point (1) to get the stated estimate. In the case where a (or b) is real and $\arg a$ (or $\arg b$) is equal to π, we shall englobe it with l by choosing $\hat{l} = 2l-1$ (or $\hat{l} = 2l+1$) such that $-\frac{\pi}{2} \le \frac{k\cdot\alpha + \arg a - \pi \hat{l}}{2} \le \frac{\pi}{2}$ to conclude the estimate as in (3.10).

Point (ii) follows directly from the triangular inequality.

We address the reader interested to optimal estimates (with σ^τ instead of $\sigma^{\tau+n}$) to [28]. □

Let now $\alpha \in \mathbb{R}^n$ and $M \in \mathrm{GL}_m(\mathbb{R})$ have simple eigenvalues such that[1] $\mu_i \ne 1, \forall i = 1, \dots, m$, and consider the following operator

$$L_{1,M} : \mathcal{A}(\mathbb{T}^n_{s+\sigma}, \mathbb{C}^m) \to \mathcal{A}(\mathbb{T}^n_s, \mathbb{C}^m), \quad f \mapsto f(\theta+\alpha) - M\cdot f(\theta).$$

[1]In order not to burden the following statements, we suppose that M has simple spectrum and 1 does not belong to it. Just note that in the general case, one should introduce the corrections λ meant to absorb the average of the given term in the homological equations when it is the case, as in Lemma 1 point (1) (cf. conditions (1.4)).

Lemma 2 (Relocating the torus)**.** *Let $\alpha \in \mathbb{R}^n$ and $M \in \mathrm{GL}_m(\mathbb{R})$, a diagonalizable matrix with simple eigenvalues distinct from 1, satisfy the Diophantine conditions (3.1)-(3.2)-(3.4)-(3.5)-(3.6). For every $g \in \mathcal{A}(\mathbb{T}^n_{s+\sigma}, \mathbb{C}^m)$, there exists a unique preimage $f \in \mathcal{A}(\mathbb{T}^n_s, \mathbb{C}^m)$ by $L_{1,A}$. Moreover the following estimate holds*

$$|f|_s \leq \frac{C_2}{\gamma^2} \frac{1}{\sigma^{n+\tau}} |g|_{s+\sigma},$$

C_2 being a constant depending only on the dimension n and the exponent τ.

Proof. In the scalar case $m = 1$ and $M = \mu \in \mathbb{R}$. By expanding both sides of $L_{1,M} f = g$ the formal preimage is given by

$$f_k = \frac{g_k}{e^{i\,k\,\alpha} - \mu}$$

and the proof is recovered from Lemma 1 point (2) ii). The diagonal case follows readily by working component wise and taking into account condition (3.4).

Eventually, if M is diagonalizable let $P \in \mathrm{GL}_m(\mathbb{C})$ be the diagonalizing matrix such that $PMP^{-1} = \mathrm{diag}(\mu_1, \cdots, \mu_m)$, $\mu_i \in \mathbb{C}$. By left multiplying both sides of $f(\theta + \alpha) - M \cdot f(\theta) = g$ by P, we get

$$\tilde{f}(\theta + \alpha) - PMP^{-1} \tilde{f}(\theta) = \tilde{g},$$

where we have set $\tilde{g} = Pg$ and $\tilde{f} = Pf$. By Lemma 1 point (2) and the Diophantine conditions (3.1)-(3.2)-(3.5)-(3.6), $\tilde{f}$ satisfies the wanted estimates, and $f = P^{-1}\tilde{f}$. $\square$

Eventually, consider a holomorphic function F on $\mathbb{T}^n_{s+\sigma}$ with values in $\mathrm{Mat}_m(\mathbb{C})$ and define the operator

$$\begin{array}{rlcl} L_{2,M}: & \mathcal{A}(\mathbb{T}^n_{s+\sigma}, \mathrm{Mat}_m(\mathbb{C})) & \to & \mathcal{A}(\mathbb{T}^n_s, \mathrm{Mat}_m(\mathbb{C})) \\ & F & \mapsto & F(\theta + \alpha) \cdot M - M \cdot F(\theta) \end{array}.$$

Lemma 3 (Straighten the first order dynamics)**.** *Let $\alpha \in \mathbb{R}^n$ and $M \in \mathrm{GL}_m(\mathbb{R})$, a diagonalizable matrix with simple eigenvalues distinct from 1, satisfy the Diophantine conditions (3.1)-(3.3)-(3.4)-(3.5)-(3.6). For every $G \in \mathcal{A}(\mathbb{T}^n_{s+\sigma}, \mathrm{Mat}_m(\mathbb{C}))$, such that $\int_{\mathbb{T}^n} G^i_i \frac{d\theta}{(2\pi)^n} = 0$, there exists a unique $F \in \mathcal{A}(\mathbb{T}^n_s, \mathrm{Mat}_m(\mathbb{C}))$, having zero average diagonal elements, such that the matrix equation*

$$F(\theta + \alpha) \cdot M - M \cdot F(\theta) = G(\theta)$$

is satisfied; moreover the following estimate holds

$$|F|_s \leq \frac{C_3}{\gamma^2} \frac{1}{\sigma^{n+\tau}} |G|_{s+\sigma},$$

C_3 being a constant depending only on the dimension n and the exponent τ.

Proof. Let $M = \mathrm{diag}(\mu_1, \ldots, \mu_m) \in \mathbb{R}^m$ and $F \in \mathrm{Mat}_m(\mathbb{C})$ be given, expanding $L_{2,M} F = G$ we get m equations of the form

$$\mu_j \Big(F^j_j(\theta + \alpha) - F^j_j(\theta) \Big) = G^j_j, \quad j = 1, \ldots, m$$

and $m^2 - m$ equations of the form

$$\mu_j F^i_j(\theta + \alpha) - \mu_i F^i_j(\theta) = G^i_j(\theta), \quad \forall i \neq j, \quad i, j = 1, \ldots, m.$$

where we denoted by F_j^i the element corresponding to the i-th line and j-th column of the matrix $F(\theta)$. Taking into account the Diophantine conditions (3.1)-(3.4), the thesis follows from the same computations as Lemma 1 point (1) for the m-diagonal equations and point (2)-ii) for the (m^2-m)-out diagonal ones.
Eventually, to recover the general case, we consider the transition matrix $P \in \mathrm{GL}_m(\mathbb{C})$ such that $PMP^{-1} = \mathrm{diag}(\mu_1, \cdots, \mu_m)$, $\mu_i \in \mathbb{C}$, and the equation

$$(PF(\theta+\alpha)P^{-1}PMP^{-1}) - PMP^{-1}PF(\theta)P^{-1} = PGP^{-1};$$

letting $\tilde{F} = PFP^{-1}$ and $\tilde{G} = PGP^{-1}$, the equation is of the previous kind and by the Diophantine conditions (3.1)-(3.3)-(3.4)-(3.5)-(3.6), $\tilde{F}$ satisfies the wanted estimates, and $F = P^{-1}\tilde{F}P$. □

Remark 3.1. The real analytic character of the solutions in Lemmata 2 and 3 follows from their uniqueness and the fact that the matrix M has real entries.

4. Inversion of the operator ϕ

The following theorem represents the main result of this first part, from which the normal form Theorem A follows.
Let us fix $P^0 \in U_s(\alpha, A)$ and note $V_s^\sigma = \left\{Q \in V_s : \left|Q - P^0\right|_s < \sigma\right\}$ the ball of radius σ centered at P^0.

Theorem 4.1. *The operator ϕ is a local diffeomorphism in the sense that for any $0 < \eta < s < s+\sigma < 1$ there exists $\varepsilon > 0$ and a unique C^∞-map ψ*

$$\psi : V_{s+\sigma}^\varepsilon \to \mathcal{G}_s^\eta \times U_s(\alpha, A) \times \Lambda$$

such that $\phi \circ \psi = \mathrm{id}$. Moreover ψ is Whitney-smooth with respect to (α, A).

This result will follow from the inverse function theorem A.1 and regularity propositions A.2-A.1-A.3.
In order to solve locally $\phi(x) = y$, we use the remarkable idea of Kolmogorov and find the solution by composing infinitely many times the operator

$$x = (g, u, \lambda) \mapsto x + \phi'^{-1}(x) \cdot (y - \phi(x)),$$

on extensions $\mathrm{T}_{s+\sigma}^n$ of shrinking width.
At each step of the induction, it is necessary that $\phi'^{-1}(x)$ exists at an unknown x (not only at x_0) in a whole neighborhood of x_0 and that ϕ'^{-1} and ϕ'' satisfy a suitable estimate, in order to control the convergence of the iterates.
The main step is to check the existence of a right inverse for

$$\phi'(G, P, \lambda) : T_G\mathcal{G}_{s+\sigma}^{\sigma/n} \times \overrightarrow{U}_{s+\sigma} \times \Lambda \to V_{G,s},$$

if G is close to the identity. We indicated with $\overrightarrow{U}$ the vector space directing $U(\alpha, A)$.

Proposition 4.1. *If (G, P, λ) is close enough to $(\mathrm{id}, P^0, 0)$, for all $\delta Q \in V_{G,s+\sigma} = G^*\mathcal{A}(\mathrm{T}_{s+\sigma}^n, \mathbb{C}^{n+m})$, there exists a unique triplet $(\delta G, \delta P, \delta\lambda) \in T_G\mathcal{G}_s \times \overrightarrow{U}_s \times \Lambda$ such that*

$$\phi'(G, P, \lambda) \cdot (\delta G, \delta P, \delta\lambda) = \delta Q. \tag{4.1}$$

Moreover we have the following estimates

$$\max(|\delta G|_s, |\delta P|_s, |\delta\lambda|) \le \frac{C'}{\sigma^{\tau'}} |\delta Q|_{G,s+\sigma}, \tag{4.2}$$

C' being a constant possibly depending on $|((G - \mathrm{id}), P - (\theta + \alpha, A \cdot r))|_{s+\sigma}$.

Proof. Let a vector field $\delta Q \in V_{G,s+\sigma}$ be given. Differentiating with respect to $x = (G, P, \lambda)$, we have

$$\delta(T_\lambda \circ G \circ P \circ G^{-1}) = T_{\delta\lambda} \circ (G \circ P \circ G^{-1}) + T'_\lambda \circ (G \circ P \circ G^{-1}) \cdot \delta(G \circ P \circ G^{-1})$$

hence

$$M \cdot (\delta G \circ P + G' \circ P \cdot \delta P - G' \circ P \cdot P' \cdot G'^{-1} \cdot \delta G) \circ G^{-1} = \delta Q - T_{\delta\lambda} \circ (G \circ P \circ G^{-1}),$$

where $M = \begin{pmatrix} I & 0 \\ 0 & I + B \end{pmatrix}$.

The data is δQ while the unknowns are the "tangent vectors" $\delta P \in O(r) \times O(r^2)$, δG (geometrically, a vector field along G) and $\delta\lambda \in \Lambda$.

Pre-composing by G we get the equivalent equation between germs along the standard torus T_0^n (as opposed to $G(\mathrm{T}_0^n)$):

$$M \cdot (\delta G \circ P + G' \circ P \cdot \delta P - G' \circ P \cdot P' \cdot G'^{-1} \cdot \delta G) = \delta Q \circ G - T_{\delta\lambda} \circ G \circ P;$$

multiplying both sides by $(G'^{-1} \circ P)M^{-1}$, we finally obtain

$$\dot G \circ P - P' \cdot \dot G + \delta P = G'^{-1} \circ P \cdot M^{-1} \delta Q \circ G + G'^{-1} \circ P \cdot M^{-1} T_{\delta\lambda} \circ G \circ P, \tag{4.3}$$

where $\dot G = G'^{-1} \cdot \delta G$.

Remark that the term containing $T_{\delta\lambda}$ is not constant; expanding along $r = 0$, it reads

$$T_{\dot\lambda} = G'^{-1} \circ P \cdot M^{-1} \cdot T_{\delta\lambda} \circ G \circ P = (\dot\beta + O(r), \dot b + \dot B \cdot r + O(r^2)).$$

The vector field $\dot G$ (geometrically, a germ along T_0^n of tangent vector fields) reads

$$\dot G(\theta, r) = (\dot\varphi(\theta), \dot R_0(\theta) + \dot R_1(\theta) \cdot r).$$

The problem is now: G, λ, P, Q being given, find $\dot G, \delta P$ and $\dot\lambda$, hence $\delta\lambda$ and δG.

We are interested in solving the equation up to the 0-order in r in the θ-direction, and up to the first order in r in the action direction; hence we consider the Taylor expansions along T_0^n up to the needed order.

We remark that since $\delta P = (O(r), O(r^2))$, it will not intervene in the cohomological equations given out by (4.3), but will be uniquely determined by identification of the reminders.

Let us proceed to solve the equation (4.3); taking its jet at the wanted order, it splits into the following three

$$\begin{aligned} \dot\varphi(\theta + \alpha) - \dot\varphi(\theta) + p_1 \cdot \dot R_0 &= \dot q_0 + \dot\beta \\ \dot R_0(\theta + \alpha) - A \cdot \dot R_0(\theta) &= \dot Q_0 + \dot b \\ \dot R_1(\theta + \alpha) \cdot A - A \cdot \dot R_1(\theta) &= \dot Q_1 - (2P_2 \cdot \dot R_0 + \dot R'_0(\theta + \alpha) \cdot p_1) + \dot B. \end{aligned} \tag{4.4}$$

The first equation is the one straightening the tangential dynamics, while the second and the third ones are meant to relocate the torus and straighten the normal

dynamics.
For the moment we solve the equations "modulo $\dot{\lambda}$"; eventually $\delta\lambda$ will be uniquely chosen to kill the average of the equation determining $\dot{\varphi}$ and the constant component of the given terms in the second and third equation that belong to the kernel of $A - I$ and $[A, \cdot]$ respectively, and solve the cohomological equations.
In the following we will repeatedly apply Lemmata 1-2-3 and Cauchy's inequality. Furthermore, we do not keep track of constants - just note that they may only depend on n and τ (from the Diophantine condition) and on $|G - \mathrm{id}|_{s+\sigma}$ and $|P - ((\theta + \alpha), A \cdot r))|_{s+\sigma}$, and refer to them as C.

- First, second equation has a solution
$$\dot{R}_0 = L_{1,A}^{-1}(\dot{Q}_0 + \dot{b} - \bar{b}),$$
where $\bar{b} = \prod_{\mathrm{Ker}\,(A-I)} \int_{\mathbb{T}^n} \dot{Q}_0 + \dot{b}\, \frac{d\theta}{(2\pi)^n}$, and
$$\left|\dot{R}_0\right|_s \leq \frac{C}{\gamma^2 \sigma^{\tau+n}} \left|\dot{Q}_0 + \dot{b}\right|_{s+\sigma}.$$
- Second, we have
$$\dot{\varphi}(\theta + \alpha) - \dot{\varphi}(\theta) + p_1 \cdot \dot{R}_0 = \dot{q}_0 + \dot{\beta} - \bar{\beta},$$
where $\bar{\beta} = \int_{\mathbb{T}^n} \dot{q}_0 - p_1 \cdot R_0 + \dot{\beta}\, \frac{d\theta}{(2\pi)^n}$, hence
$$\dot{\varphi} = L_\alpha^{-1}(\dot{q}_0 + \dot{\beta} - \bar{\beta}),$$
satisfying
$$|\dot{\varphi}|_{s-\sigma} \leq \frac{C}{\gamma^3 \sigma^{2(\tau+n)}} \left|\dot{q}_0 + \dot{\beta}\right|_{s+\sigma}$$
- Third, the solution of equation in $\dot{R}_1$ is
$$\dot{R}_1 = L_{2,A}^{-1}(\widetilde{Q}_1 + \dot{B} - \bar{B}),$$
hiving denoted $\widetilde{Q}_1 = \dot{Q}_1 - (2P_2 \cdot \dot{R}_0 + \dot{R}_0'(\theta + \alpha) \cdot p_1)$, and $\bar{B} = \prod_{\mathrm{Ker}[A,\cdot]} \int_{\mathbb{T}^n} \widetilde{Q}_1 + \dot{B}\, \frac{d\theta}{(2\pi)^n}$. It satisfies
$$\left|\dot{R}_1\right|_{s-2\sigma} \leq \frac{C}{\gamma^2 \sigma^{n+\tau}} \left|\widetilde{Q}_1 + \dot{B}\right|_{s+\sigma}.$$

We now handle the unique choice of the correction $\delta\lambda = (\delta\beta, \delta b + \delta B \cdot r)$ given by $T_{\delta\lambda}$. Letting $\bar{\lambda} = (\bar{\beta}, \bar{b} + \bar{B} \cdot r)$, the map $f : \Lambda \to \Lambda$, $\delta\lambda \mapsto -\bar{\lambda}$ is well defined in the neighborhood of $\delta\lambda = 0$. In particular $f' = -\,\mathrm{id}$ at $G = \mathrm{id}$, and it will remain bounded away from 0 if G stays sufficiently close to the identity. In particular, $\delta\lambda \mapsto -\bar{\lambda}$ is affine: the system to solve being linear of the form $\int_{\mathbb{T}^n} a(G, \dot{Q}) + A(G) \cdot \delta\lambda = 0$, with diagonal close to 1 when G is close to the identity, f is invertible. Thus, there exists a unique $\delta\lambda$ such that $f(\delta\lambda) = 0$, satisfying
$$|\delta\lambda| \leq \frac{C}{\gamma^2 \sigma^{\tau+n+1}} |\delta Q|_{G,s+\sigma}.$$
We finally have
$$\left|\dot{G}\right|_{s-2\sigma} \leq \frac{C}{\gamma^3}\, \frac{1}{\sigma^{2(\tau+n)+1}} |\delta Q|_{G,s+\sigma}.$$

Now, from the definition of $\dot{G} = G'^{-1} \cdot \delta G$ we get $\delta G = G' \cdot \dot{G}$. The unique solutions such that $\delta\varphi(0) = 0$, $\delta R_0(0) = 0$ and $\delta R_1(0) = 0$ are easily determined, since G is close to the identity and similar estimates hold for δG:

$$|\delta G|_{s-2\sigma} \leq \sigma^{-1}(1 + |G - \mathrm{id}|_s)\frac{C}{\sigma^{2(\tau+n)+1}}|\delta Q|_{G,s+\sigma}.$$

Eventually, equation (4.3) uniquely determines δP.
Letting $\tau' = 2(\tau + n) + 2$, up to redefining $\sigma' = \sigma/3$ and $s' = s + \sigma$, we have the stated estimates for all $s', \sigma' : s' < s' + \sigma'$. □

Proposition 4.2 (Boundness of ϕ'')**.** *The bilinear map* $\phi''(x)$

$$\phi''(x) : (T_G\mathcal{G}^{\sigma/n}_{s+\sigma} \times \overrightarrow{U}_{s+\sigma}(\alpha, A) \times \Lambda)^{\otimes 2} \to \mathcal{A}(\mathrm{T}^n_s, \mathrm{T}^n_{\mathbb{C}}),$$

satisfies the estimates

$$\left|\phi''(x) \cdot \delta x^{\otimes 2}\right|_{G,s} \leq \frac{C''}{\sigma^{\tau''}}|\delta x|^2_{s+\sigma},$$

where $\tau'' \geq 1$ *and* C'' *is a constant depending on* $|x|_{s+\sigma}$.

Proof. Differentiating twice $\phi(x)$, yields

$$\begin{aligned}
&- M\Big\{\Big[\delta G' \circ P \cdot \delta P + \delta G' \circ P \cdot \delta P + G'' \circ P \cdot \delta P^2 - (\delta G' \circ P + G'' \circ P \cdot \delta P) \cdot P' \cdot G'^{-1} \cdot \delta G \\
&- G' \circ P \cdot \big(\delta P' \cdot (-G'^{-1} \cdot \delta G' \cdot G'^{-1}) \cdot \delta G\big)\Big] \circ G^{-1} \\
&+ \Big[\delta G' \circ P \cdot \delta P + \delta G' \circ P \cdot \delta P + G'' \circ P \cdot \delta P^2 - (\delta G' \circ P + G'' \circ P \cdot \delta P) \cdot P' \cdot G'^{-1} \cdot \delta G \\
&- G' \circ P \cdot \big(\delta P' \cdot (-G'^{-1} \cdot \delta G' \cdot G'^{-1}) \cdot \delta G\big)\Big]' \circ G^{-1} \cdot (-G'^{-1} \cdot \delta G) \circ G^{-1}\Big\}.
\end{aligned}$$

Once we precompose with G, the estimate follows. □

Hypothesis of Theorem A.1 are satisfied, hence the existence of (G, P, λ) such that $Q = T_\lambda \circ G \circ P \circ G^{-1}$ is proved. Uniqueness and smoothness of the normal form follows from Propositions A.1-A.2-A.3. Theorem 4.1 follows, hence Theorem A.

5. A generalization of Rüssmann's theorem

Theorem A provides a normal form that does not rely on any non-degeneracy assumption; thus, the existence of a translated Diophantine, reducible torus will be subordinated to eliminating the "parameters in excess" (β, B) using a non-degeneracy hypothesis. We will implicitly solve $B = 0$ and $\beta = 0$ by using the normal frequencies as *free* parameters and a torsion hypothesis respectively. Rüssmann's classical result will be the immediate small dimensional case.

Elimination of B**.** Let $\Delta^s_m(\mathbb{R}) \subset \mathrm{GL}_m(\mathbb{R})$ be the open set of invertible matrices with simple, real eigenvalues. On $\mathbb{T}^n \times \mathbb{R}^m$, let us define

$$\widehat{U} = \bigcup_{A \in \Delta^s_m(\mathbb{R})} U(\alpha, A).$$

We recall that those $P's \in U(\alpha, A)$ are diffeomorphisms of the form

$$P(\theta, r) = (\theta + \alpha + O(r), A \cdot r + O(r^2)),$$

on a neighborhood of $\mathbb{T}^n \times \{0\}$.

The following theorem is an intermediate, yet fundamental result to prove the translated torus Theorem C and holds without requiring any torsion assumption on the class of diffeomorphisms.

Theorem 5.1 (Twisted Torus of co-dimension 1)**.** *For every $P^0 \in U_{s+\sigma}(\alpha, A^0)$ with α Diophantine, and $A^0 \in \Delta^s_m(\mathbb{R})$, there is a germ of C^∞-maps*

$$\psi : V_{s+\sigma} \to \mathcal{G}_s \times \widehat{U}_s \times \Lambda(\beta, b), \quad Q \mapsto (G, P, \lambda),$$

at $P^0 \mapsto (\mathrm{id}, P^0, 0)$, such that $Q = T_\lambda \circ G \circ P \circ G^{-1}$, where $\lambda = (\beta, b) \in \mathbb{R}^{n+1}$.

Corollary 5.1 (Twisted torus)**.** *If 1 does not belong to the spectrum of A^0, the translation correction $b = 0$.*

Proof. Denote ϕ_A the operator ϕ, as now we want A to vary. Let define the map

$$\hat{\psi} : \Delta^s_m(\mathbb{R}) \times V_{s+\sigma} \to \mathcal{G}_s \times \hat{U}_s \times \Lambda,\ (A, Q) \mapsto \hat{\psi}_A(Q) := \phi_A^{-1}(Q) = (G, P, \lambda)$$

in the neighborhood of (A^0, P^0), such that $Q = T_\lambda \circ G \circ P \circ G^{-1}$ where $\lambda = (\beta, b, B \cdot r)$, $\beta \in \mathbb{R}^n, b \in \mathbb{R}^m$ such that $(A - I) \cdot b = 0$ and $B \in \mathrm{Mat}_m(\mathbb{R})$ satisfies $[B, A] = 0$. Equivalently, B is simultaneously diagonalizable with A, since A has simple spectrum; we can thus restrict our analysis to a neighborhood of A^0 in the subspace of those matrices commuting with A^0. Note that we can choose such a neighborhood so that it is contained in $\Delta^s_m(\mathbb{R})$. Then we study the dependence of B on A in their diagonal form.
Without loss of generality, let A^0 be in its canonical form, let Δ_{A^0} be the subspace of diagonal matrices, namely the matrices which commute with A^0. Consider the restriction of $\hat{\psi}$ to Δ_{A^0}. Let $A \in \Delta_{A^0}$ be close to A^0, let $\delta A := A^0 - A$ and write P^0 as

$$P^0(\theta, r) = (\theta + \alpha + O(r), (A^0 - \delta A) \cdot r + \delta A \cdot r + O(r^2));$$

we remark that $P^0 = T_\lambda \circ P_A$, where

$$\lambda = \big(0, B(A) = (A^0 - A) \cdot A^{-1}\big), \quad [B(A), A] = 0$$

and $P_A = (\theta + \alpha + O(r), A \cdot r + O(r^2))$, $A = A^0 - \delta A$.[2] Remark that, since $A \in \Delta_{A^0}$ has simple spectrum, B is indeed in Δ_{A^0}.
According to Theorem A, $\phi_A(\mathrm{id}, P_A, \lambda) = P^0$, thus locally for all $A \in \Delta_{A^0}$ close to A^0 we have

$$\hat{\psi}(A, P^0) = (\mathrm{id}, P_A, B \cdot r), \quad B(A, P^0) = (A^0 - A) \cdot A^{-1} = \delta A \cdot (A^0 - \delta A)^{-1}$$

and, in particular $B(A^0, P^0) = 0$ and

$$\frac{\partial B}{\partial A}\Big|_{A = A^0} = -{A^0}^{-1},$$

[2]The terms $O(r^2)$ contain a factor $(I + \delta A \cdot A^{-1})^{-1}$.

which is invertible, since A^0 is so. Hence $A \mapsto B(A)$ is a local diffeomorphism on Δ_{A^0} and by the implicit function theorem (in finite dimension) locally for all Q close to P^0 there exists a unique $\bar{A}$ such that $B(\bar{A}, Q) = 0$. It remains to define $\psi(Q) = \hat{\psi}(\bar{A}, Q)$. $\square$

The proof of Corollary 5.1 is immediate, by conditions (1.4).

Remark 5.1. This twisted-torus theorem relies on the peculiarity of the normal dynamics of the torus T_0^n. The direct applicability of the implicit function theorem is subordinated to the fact that no arithmetic condition is required on the characteristic (normal) frequencies so that the correction $A^0 + \delta A$ is well defined; beyond that, since having simple, real eigenvalues is an open property, the needed counter term B is indeed a diagonal matrix, so that the number of free frequencies (parameters) is enough to solve, implicitly, $B(A) = 0$. The generic case of complex eigenvalues is more delicate since one should guarantee that corrections $A^0 + \delta A$ at each step satisfy the Diophantine condition (1.2). It seems reasonable to think that one would need more parameters to control this issue, using the Whitney smoothness of ϕ on A, and verify that the measure of such stay positive (see [12]).

Elimination of β. If Q satisfy a torsion hypothesis, the existence of a translated Diophantine torus can be proved.

Theorem 5.2 (Translated Diophantine torus)**.** *Let α be Diophantine. On a neighborhood of $\mathbb{T}^n \times \{0\} \subset \mathbb{T}^n \times \mathbb{R}^n$, let $P^0 \in U(\alpha, A^0)$ be a diffeomorphism of the form*

$$P^0(\theta, r) = (\theta + \alpha + p_1(\theta) \cdot r + O(r^2), A^0 \cdot r + O(r^2)),$$

where A^0 is invertible and has simple, real eigenvalues and such that

$$\det \left(\int_{\mathbb{T}^n} p_1(\theta)\, d\theta \right) \neq 0.$$

If Q is close enough to P^0 there exists a unique A', close to A^0, and a unique $(G, P, b) \in \mathcal{G} \times U(\alpha, A') \times \mathbb{R}^n$ such that $Q = T_b \circ G \circ P \circ G^{-1}$.

Phrasing the thesis, the graph of $\gamma = R_0 \circ \varphi^{-1}$ is a translated torus on which the dynamics is conjugated to R_α by φ (remember the form of $G \in \mathcal{G}$ given in (1.3)). Before proceeding with the proof of Theorem 5.2, let us consider a parameter $c \in B_1^n(0)$ (the unit ball in $\mathbb{R}^n$) and the family of maps defined by $Q_c(\theta, r) := Q(\theta, c + r)$ obtained by translating the action coordinates. Considering the corresponding normal form operators ϕ_c, the parametrized version of Theorem A follows readily. Now, if Q_c is close enough to P_c^0, Theorem 5.1 asserts the existence of $(G_c, P_c, \lambda_c) \in \mathcal{G} \times U(\alpha, A) \times \Lambda(\beta, b)$ such that

$$Q_c = T_\lambda \circ G_c \circ P_c \circ G_c^{-1}.$$

Hence we have a family of tori parametrized by $\tilde{c} = c + \int_{\mathbb{T}^n} \gamma \, \frac{d\theta}{(2\pi)^n}$,

$$Q(\theta, \tilde{c} + \tilde{\gamma}(\theta)) = (\beta(c) + \varphi \circ R_\alpha \circ \varphi^{-1}(\theta), b(c) + \tilde{c} + \tilde{\gamma}(\varphi \circ R_\alpha \circ \varphi^{-1}(\theta))),$$

where $\gamma := R_0 \circ \varphi^{-1}$ and $\tilde{\gamma} = \gamma - \int_{\mathbb{T}} \gamma \, \frac{d\theta}{2\pi}$.

Proof. Let $\hat{\varphi}$ be the function defined on $\mathbb{T}^n$ taking values in $\mathrm{Mat}_n(\mathbb{R})$ that solves the (matrix of) difference equation

$$\hat{\varphi}(\theta+\alpha)-\hat{\varphi}(\theta)+p_1(\theta)=\int_{\mathbb{T}^n} p_1(\theta)\frac{d\theta}{(2\pi)^n},$$

and let $F:(\theta,r)\mapsto(\theta+\hat{\varphi}(\theta)\cdot r, r)$. The diffeomorphism F restricts to the identity at T^n_0. At the expense of substituting P^0 and Q with $F\circ P^0\circ F^{-1}$ and $F\circ Q\circ F^{-1}$ respectively, we can assume that

$$P^0(\theta,r)=(\theta+\alpha+p_1\cdot r+O(r^2), A^0\cdot r+O(r^2)),\quad p_1=\int_{\mathbb{T}^n} p_1(\theta)\frac{d\theta}{(2\pi)^n}.$$

The germs so obtained from the initial P^0 and Q are close to one another.
The proof will follow from Theorem 5.1 and the elimination of the parameter $\beta\in\mathbb{R}^n$ obstructing the rotation conjugacy.
In line with the previous reasoning, we want to show that the map $c\mapsto\beta(c)$ is a local diffeomorphism. It suffices to show this for the trivial perturbation P^0_c. The Taylor expansion of P^0_c directly gives the normal form. In particular $b(c)=A^0\cdot c+O(c^2)$, while the map $c\mapsto\beta(c)=p_1\cdot c+O(c^2)$ is such that $\beta(0)=0$ and $\beta'(0)=p_1$which is invertible by twist hypothesis, thus a local diffeomorphism. Hence, the analogous map for Q_c, which is a small C^1-perturbation, is a local diffeomorphism too and, together with Theorem 5.1, there exists unique $c\in\mathbb{R}^n$ and $A\in\mathrm{Mat}_n(\mathbb{R})$, such that $(\beta,B)=(0,0)$. □

Remark 5.2. The theorem holds also on $\mathbb{T}^n\times\mathbb{R}^m$, with $m\geq n$, requiring that

$$\mathrm{rank}\left(\int_{\mathbb{T}^n} p_1(\theta)\,d\theta\right)=n.$$

This guarantees that $c\mapsto\beta(c)$ is submersive, but c solving $\beta(c)=0$ would no more be uniquely determined.

Remark 5.3. Theorem 5.2 generalizes the classical translated curve theorem of Rüssmann in higher dimension, in the case of normally hyperbolic systems such that A has simple, real, non 0 eigenvalues, for general perturbations.
We stress the fact that if P^0 was of the form

$$P^0(\theta,r)=(\theta+\alpha+O(r), I\cdot r+O(r^2)),$$

like in the original frame studied by Rüssmann, we would need a whole matrix $B\in\mathrm{Mat}_n(\mathbb{R})$ in order to solve the homological equations, and, having just n characteristic frequencies at our disposal, it is hopeless to completely solve $B=0$ and eliminate the whole obstruction. The torus would not be just translated.

Appendix A. Inverse function theorem & regularity of ϕ

We state here the implicit function theorem we use to prove Theorem A as well as the regularity statements needed to guarantee uniqueness and smoothness of the normal form. These results follow from Féjoz [13, 14]. Remark that we endowed functional spaces with weighted norms and bounds appearing in propositions 4.1-4.2 may depend on $|x|_s$ (as opposed to the analogue statements in [13, 14]); for the corresponding proofs taking account of these (slight) differences we send the reader to [21, 22] and the proof or Moser's theorem therein.

Let $E = (E_s)_{0<s<1}$ and $F = (F_s)_{0<s<1}$ be two decreasing families of Banach spaces with increasing norms $|\cdot|_s$ and let $B_s^E(\sigma) = \{x \in E : |x|_s < \sigma\}$ be the ball of radius σ centered at 0 in E_s.

On account of composition operators, we additionally endow F with some deformed norms which depend on $x \in B_s^E(s)$ such that

$$|y|_{0,s} = |y|_s \qquad \text{and} \qquad |y|_{\hat{x},s} \leq |y|_{x,s+|x-\hat{x}|_s}.$$

Consider then operators commuting with inclusions $\phi : B_{s+\sigma}^E(\sigma) \to F_s$, with $0 < s < s+\sigma < 1$, such that $\phi(0) = 0$.

We then suppose that if $x \in B_{s+\sigma}^E(\sigma)$ then $\phi'(x) : E_{s+\sigma} \to F_s$ has a right inverse $\phi'^{-1}(x) : F_{s+\sigma} \to E_s$ (for the particular operators ϕ of this work, ϕ' is both left and right invertible).

ϕ is supposed to be at least twice differentiable.

Let $\tau := \tau' + \tau''$ and $C := C'C''$.

Theorem A.1 (Inverse function theorem)**.** *Further assume*

$$\left|\phi'^{-1}(x) \cdot \delta y\right|_s \leq \frac{C'}{\sigma^{\tau'}} |\delta y|_{x,s+\sigma} \tag{A.1}$$

$$\left|\phi''(x) \cdot \delta x^{\otimes 2}\right|_{x,s} \leq \frac{C''}{\sigma^{\tau''}} |\delta x|_{s+\sigma}^2, \quad \forall s, \sigma : 0 < s < s+\sigma < 1 \tag{A.2}$$

C' and C'' depending on $|x|_{s+\sigma}$, $\tau', \tau'' \geq 1$.

For any s, σ, η with $\eta < s$ and $\varepsilon \leq \eta \frac{\sigma^{2\tau}}{2^{8\tau} C^2}$ ($C \geq 1, \sigma < 3C$), ϕ has a right inverse $\psi : B_{s+\sigma}^F(\varepsilon) \to B_s^E(\eta)$. In other words, ϕ is locally surjective:

$$B_{s+\sigma}^F(\varepsilon) \subset \phi(B_s^E(\eta)).$$

Proposition A.1 (Lipschitz continuity of ψ)**.** *Let $\sigma < s$. If $y, \hat{y} \in B_{s+\sigma}^F(\varepsilon)$ with $\varepsilon = 3^{-4\tau} 2^{-16\tau} \frac{\sigma^{6\tau}}{4C^3}$, the following inequality holds*

$$|\psi(y) - \psi(\hat{y})|_s \leq L|y - \hat{y}|_{x,s+\sigma},$$

with $L = 2C'/\sigma^{\tau'}$. In particular, ψ being the unique local right inverse of ϕ, it is also its unique left inverse.

Proposition A.2 (Smooth differentiation of ψ)**.** *Let $\sigma < s < s+\sigma$ and ε as in proposition A.1. There exists a constant K such that for every $y, \hat{y} \in B_{s+\sigma}^F(\varepsilon)$ we have*

$$\left|\psi(\hat{y}) - \psi(y) - \phi'^{-1}(\psi(y))(\hat{y} - y)\right|_s \leq K(\sigma)|\hat{y} - y|_{x,s+\sigma}^2,$$

and the map $\psi' : B_{s+\sigma}^F(\varepsilon) \to L(F_{s+\sigma}, E_s)$ defined locally by $\psi'(y) = \phi'^{-1}(\psi(y))$ is continuous. In particular ψ has the same degree of smoothness as ϕ.

It is sometimes convenient to extend ψ to non-Diophantine characteristic frequencies (α, A). Whitney smoothness guarantees that such an extension exists. Let suppose that $\phi(x) = \phi_\nu(x)$ depends on some parameter $\nu \in B^k$ (the unit ball of $\mathbb{R}^k$) and that it is C^1 with respect to ν and that estimates on $\phi_\nu'^{-1}$ and ϕ_ν'' are uniform with respect to ν over some closed subset D of $\mathbb{R}^k$.

Proposition A.3 (Whitney differentiability). *Let us fix ε, σ, s as in proposition A.1. The map $\psi : D \times B^F_{s+\sigma}(\varepsilon) \to B^E_s(\eta)$ is C^1-Whitney differentiable and extends to a map $\psi : \mathbb{R}^{2n} \times B^F_{s+\sigma}(\varepsilon) \to B^E_s(\eta)$ of class C^1. If ϕ is C^k, $1 \le k \le \infty$, with respect to ν, this extension is C^k.*

Appendix B. Inversion of a holomorphism of $\mathbb{T}^n_s$

We present here a classical result and a lemma that justify the well definition of the normal form operator ϕ defined in section 2.3.
Complex extensions of manifolds are defined at the help of the ℓ^∞-norm.
Let

$$\mathbb{T}^n_{\mathbb{C}} = \mathbb{C}^n / 2\pi\mathbb{Z}^n \qquad \text{and} \qquad \mathrm{T}^n_{\mathbb{C}} = \mathbb{T}^n_{\mathbb{C}} \times \mathbb{C}^m,$$

$$\mathbb{T}^n_s = \left\{ \theta \in \mathbb{T}^n_{\mathbb{C}} : |\theta| := \max_{1 \le j \le n} |\operatorname{Im} \theta_j| \le s \right\}, \quad \mathrm{T}^n_s = \left\{ (\theta, r) \in \mathrm{T}^n_{\mathbb{C}} : \ |(\operatorname{Im} \theta, r)| \le s \right\},$$

where $|(\operatorname{Im}\theta, r)| := \max_{1\le j\le n} \max(|\operatorname{Im}\theta_j|, |r_j|)$.
Let also define $\mathbb{R}^n_s := \mathbb{R}^n \times (-s, s)$ and consider the universal covering of $\mathbb{T}^n_s$, $p : \mathbb{R}^n_s \to \mathbb{T}^n_s$.

Theorem B.1. *Let $v : \mathbb{T}^n_s \to \mathbb{C}^n$ be a vector field such that $|v|_s < \sigma/n$. The map $\mathrm{id} + v : \mathbb{T}^n_{s-\sigma} \to \mathbb{R}^n_s$ induces a map $\varphi = \mathrm{id} + v : \mathbb{T}^n_{s-\sigma} \to \mathbb{T}^n_s$ which is a biholomorphism and there is a unique biholomorphism $\psi : \mathbb{T}^n_{s-2\sigma} \to \mathbb{T}^n_{s-\sigma}$ such that $\varphi \circ \psi = \mathrm{id}_{\mathbb{T}^n_{s-2\sigma}}$.*
In particular the following hold:

$$|\psi - \mathrm{id}|_{s-2\sigma} \le |v|_{s-\sigma}$$

and, if $|v|_s < \sigma/2n$

$$|\psi' - \mathrm{id}|_{s-2\sigma} \le \frac{2}{\sigma} |v|_s.$$

For the proof we send again to [21, 22].

Corollary B.1 (Well definition of the normal form operator ϕ). *For all s, σ if $G \in \mathcal{G}^{\sigma/n}_{s+\sigma}$, then $G^{-1} \in \mathcal{A}(\mathrm{T}^n_s, \mathrm{T}^n_{s+\sigma})$.*

Proof. We recall the form of $G \in \mathcal{G}^{\sigma/n}_{s+\sigma}$:

$$G(\theta, r) = (\varphi(\theta), R_0(\theta) + R_1(\theta) \cdot r).$$

G^{-1} reads

$$G^{-1}(\theta, r) = (\varphi^{-1}(\theta), R_1^{-1} \circ \varphi^{-1}(\theta) \cdot (r - R_0 \circ \varphi^{-1}(\theta))).$$

Up to rescaling norms by a factor 1/2 like $\|x\|_s := \frac{1}{2}|x|$, the statement is straightforward and follows from theorem B.1. By abuse of notations, we keep on indicating $\|x\|_s$ with $|x|_s$. □

Appendix C. Fourier norms

Let $\mathcal{A}(\mathrm{T}^n_s, \mathbb{C})$ be the space of holomorphic functions on T^n_s with values in $\mathbb{C}$, endowed with the norm

$$\|f\|_s = \sum_k \sup_{|r|<s} |f_k(r)|\, e^{|k|s}, \quad |k| = |k_1| + \cdots + |k_n|.$$

If $f \in \mathcal{A}(\mathrm{T}^n_s, \mathrm{Mat}_m(\mathbb{C}))$, the definition of the norm is adapted in the obvious way and the expression $|f_k(r)|$ denotes the standard operator norm $\sup_{|\xi|=1} |f_k(r)\xi|$. If $f : \mathrm{T}^n_s \to \mathbb{C}^n$, $\|f\|_s = \max_{1\le j\le n}(\|f^j\|_s)$.

Lemma 4. *Let $f \in \mathcal{A}(\mathrm{T}^n_{s+\sigma}, \mathbb{C})$ and let $h \in \mathcal{A}(\mathrm{T}^n_s, \mathbb{C}^n)$ be such that $\|h\|_s < \frac{\sigma}{e}$, then*

$$\|f(\theta, r + h(\theta, r))\|_s \le \frac{1}{1 - e\frac{\|h\|_s}{\sigma}} \|f\|_{s+\sigma}.$$

Proof. Let $f(\theta, r + h(\theta, r)) = \sum_n \frac{D^n f(\theta,r) h^n(\theta,r)}{n!}$ be the Taylor expansion of f.

$$\|f(\theta, r + h(\theta, r))\|_s \le \sum_k \sup_{|r|<s} \left(\sum_n \frac{1}{n!} \sum_{\ell + k^1 + \cdots + k^n = k} |D^n f_\ell(r)||h_{k^1}(r)| \cdots |h_{k^n}(r)| \right) e^{|k|s},$$

where $k^i \in \mathbb{Z}^n$, $i = 1, \dots, n$ are the Fourier indexes. Since $|k| \le |\ell| + |k^1| + \cdots |k^n|$

$$\begin{aligned}
&\le \sum_k \sup_{|r|<s} \left(\sum_n \frac{1}{n!} \sum_{\ell + k^1 + \cdots + k^n = k} |D^n f_\ell(r)| e^{|\ell|s} |h_{k^1}(r)| e^{|k^1|s} \cdots |h_{k^n}(r)| e^{|k^n|s} \right) \\
&\le \sum_k \left(\sum_n \frac{1}{n!} \sum_{\ell + k^1 + \cdots + k^n = k} \sup_{|r|<s} |D^n f_\ell(r)| e^{|\ell|s} |h_{k^1}(r)| e^{|k^1|s} \cdots |h_{k^n}(r)| e^{|k^n|s} \right) \\
&\le \sum_k \left(\sum_n \frac{1}{n!} \sum_{\ell + k^1 + \cdots + k^n = k} \sup_{|r|<s} |D^n f_\ell(r)| e^{|\ell|s} |h_{k^1}(r)| e^{|k^1|s} \cdots |h_{k^n}(r)| e^{|k^n|s} \right) \\
&\le \sum_n \frac{1}{n!} \sum_\ell \sup_{|r|<s} |D^n f_\ell(r)| e^{|\ell|s} \sum_{k^1} \sup_{|r|<s} |h_{k^1}(r)| e^{|k^1|s} \cdots \sum_{k^n} \sup_{|r|<s} |h_{k^n}(r)| e^{|k^n|s} \\
&\le \sum_\ell \left(\sum_n \frac{n^n}{n!} \sup_{|r|<s+\sigma} |f_\ell(r)| \frac{\|h\|_s^n}{\sigma^n} \right) e^{|\ell|(s+\sigma)},
\end{aligned}$$

where the last estimate follows from the fact that $(D^n f)_\ell = D^n(f_\ell)$ and the classical Cauchy's estimate by observing that for all $|r| < s$ letting $\mathbb{R}^n_{s+\sigma} \ni \xi \ne 0$, the analytic function $\varphi(t) = f(r + t\xi)$ on the complex disc $|t| < \sigma/|\xi|$ satisfies $\frac{d^n\varphi}{dt^n}\big|_{t=0} = D^n f(r)\xi^n$. The factor n^n coming from the classical bound on the norm of a symmetric multilinear mapping by the associated homogeneous polynomial, see for example [15].

It thus follows that

$$\begin{aligned}\|f(\theta, r+h(\theta,r))\|_s &\leq 2\sum_{\ell} \sup_{|r|<s+\sigma} |f_\ell(r)|_{s+\sigma} \sum_{n\geq 1} \frac{e^n}{\sqrt{2\pi n}} \left(\frac{\|h\|_s}{\sigma}\right)^n e^{|\ell|(s+\sigma)} \\ &\leq 2\|f\|_{s+\sigma} \frac{1}{2\Big(1-e^{\frac{\|h\|_s}{\sigma}}\Big)},\end{aligned}$$

hence the stated bound. □

Lemma 5. *Let* $f \in \mathcal{A}(\mathrm{T}^n_{s+\sigma}, \mathbb{C})$ *and* $h \in \mathcal{A}(\mathbb{T}^n_s, \mathbb{C})$ *be such that* $\|h\|_s < \sigma$*, then*

$$\|f(\theta + h(\theta), r)\|_s \leq \|f\|_{s+\sigma}.$$

For the proof see [9, Appendix B] for example.

Acknowledgments. *The author has been supported by the ERC grant HamPDE's. The author is grateful for the mathematical (and moral) support and advises of A. Chenciner and J. Féjoz all the way through her Ph.D, at the Astronomie et Systèmes Dynamiques group of the IMCCE, Observatoire de Paris. The contained work owes a lot to the meaningful talks of those times. Thank you also to L. Biasco, T. Castan, B. Fayad, J. Laskar, J.-P. Marco, P. Mastrolia, L. Niederman, P. Robutel, A. Pousse, and M. Procesi for the enlightening discussions we had and their constant interest in this (and future) work.*

References

[1] V. I. Arnolđ. Small denominators. I. Mapping the circle onto itself. Izv. Akad. Nauk SSSR Ser. Mat., 25:21–86, 1961.

[2] V. I. Arnold. Proof of a theorem of A. N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian. Usp. Mat. Nauk SSSR, 18(13), 1963.

[3] J.-B. Bost. Tores invariants des systèmes dynamiques Hamiltoniens [d'après Kolmogorov, Arnold, Moser, Rüssmann, Zender, Herman, Pöschel, ...]. Number 639 in Séminaires BOURBAKI, pages 113–157, 1985.

[4] Renato C. Calleja, Alessandra Celletti, and Rafael de la Llave. A KAM theory for conformally symplectic systems: efficient algorithms and their validation. J. Differential Equations, 255(5):978–1049, 2013.

[5] A. Chenciner. Bifurcations de points fixes elliptiques. I. Courbes invariantes. Inst. Hautes Études Sci. Publ. Math., 61:67–127, 1985.

[6] A. Chenciner. Bifurcations de points fixes elliptiques. II. Orbites periodiques et ensembles de Cantor invariants. Invent. Math., 80(1):81–106, 1985.

[7] A. Chenciner. Bifurcations de points fixes elliptiques. III. Orbites périodiques de "petites" périodes et élimination résonnante des couples de courbes invariantes. Inst. Hautes Études Sci. Publ. Math., 66:5–91, 1988.

[8] C. Q. Cheng and Y. S. Sun. Existence of periodically invariant curves in 3-dimensional measure-preserving mappings. Celestial Mech. Dynam. Astronom., 47(3):293–303, 1989/90.

[9] L. Chierchia. KAM lectures. In Dynamical systems. Part I, Pubbl. Cent. Ric. Mat. Ennio Giorgi, pages 1–55. Scuola Norm. Sup., Pisa, 2003.

[10] L. H. Eliasson, B. Fayad, and R. Krikorian. Around the stability of KAM tori. Duke Math. J., 164(9):1733–1775, 2015.

[11] B. Fayad and R. Krikorian. Herman's last geometric theorem. Ann. Sci. Éc. Norm. Supér. (4), 42(2):193–219, 2009.

[12] J. Féjoz. Démonstration du "théorème d'Arnold" sur la stabilité du système planétaire (d'après Michael Herman). Michael Herman Memorial Issue, Ergodic Theory Dyn. Syst, (24:5):1521–1582, 2004.

[13] J. Féjoz. Mouvements périodiques et quasi-périodiques dans le problème des n corps, *Mémoire d'habilitation à diriger des recherches*, 2010. UPMC.

[14] J. Féjoz. Introduction to KAM theory and application to the trhee-body problem, 2015. To appear.

[15] L. A. Harris. Bounds on the derivatives of holomorphic functions of vectors. pages 145–163. Actualités Aci. Indust., No. 1367, 1975.

[16] M. R. Herman. Sur un thórème d'Arnold et Kolmogorov. C. R. Acad. Sc. Paris Śrie A, (273):409–411, 1971.

[17] M. R. Herman. Sur la conjugaison diffŕentiable des diffómorphismes du cercle à des rotations. Inst. Hautes Études Sci. Publ. Math., 49(1):5–233, 1979.

[18] M.-R. Herman. Sur les courbes invariantes par les difféomorphismes de l'anneau. Vol. 1, volume 103 of Astérisque. Société Mathématique de France, Paris, 1983. With an appendix by Albert Fathi, With an English summary.

[19] M.-R. Herman. Séminaires de systèmes dynamiques. 1997.

[20] A. N. Kolmogorov. On conservation of conditionally periodic motions for a small change in Hamilton's function. Dokl. Akad. Nauk SSSR (N.S.), 98:527–530, 1954.

[21] J. E. Massetti. Quasi-périodicité et quasi-conservativité. *Ph.D thesis*, Observatoire de Paris, 2015.

[22] J. E. Massetti. Normal forms for perturbations of systems possessing a diophantine invariant torus. 2016. Submitted.

[23] K. R. Meyer. The implicit function theorem and analytic differential equations. In Dynamical systems—Warwick 1974 (Proc. Sympos. Appl. Topology and Dynamical Systems, Univ. Warwick, Coventry, 1973/1974; presented to E. C. Zeeman on his fiftieth birthday), pages 191–208. Lecture Notes in Math., Vol. 468. Springer, Berlin, 1975.

[24] J. Moser. On invariant curves of area-preserving mappings of an annulus. Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II, 1962:1–20, 1962.

[25] J. Moser. A rapidly convergent iteration method and non-linear differential equations. II. Ann. Scuola Norm. Sup. Pisa (3), 20:499–535, 1966.

[26] J. Moser. Convergent series expansions for quasi-periodic motions. Math. Annalen, (169):136–176, 1967.

[27] H. Rüssmann. Kleine Nenner. I. Über invariante Kurven differenzierbarer Abbildungen eines Kreisringes. Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II, 1970:67–105, 1970.

[28] H. Rüssmann. On optimal estimates for the solutions of linear difference equations on the circle. In Proceedings of the Fifth Conference on Mathematical Methods in Celestial Mechanics (Oberwolfach, 1975), Part I, volume 14, pages 33–37, 1976.

[29] M. B. Sevryuk. The lack-of-parameters problem in the KAM theory revisited. In Hamiltonian systems with three or more degrees of freedom (S'Agaró, 1995), volume 533 of NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., pages 568–572. Kluwer Acad. Publ., Dordrecht, 1999.

[30] M. B. Sevryuk. Partial preservation of frequencies in KAM theory. Nonlinearity, 19(5):1099–1140, 2006.

[31] F. Wagener. A parametrised version of Moser's modifying terms theorem. Discrete Contin. Dyn. Syst. Ser. S, 3(4):719–768, 2010.

[32] J-C. Yoccoz. Travaux de Herman sur les tores invariants. Astérisque, (206):Exp. No. 754, 4, 311–344, 1992. Séminaire Bourbaki, Vol. 1991/92.

[33] E. Zehnder. Generalized implicit function theorem with applications to some small divisor problems, i. XXVIII:91–140, 1975.

Università degli Studi Roma Tre - Dipartimento di Matematica e Fisica, Sezione Matematica, Largo San L. Murialdo 1, I-00146 Roma, Italy e-mail: jmassetti@mat.uniroma3.it