Ergod. Th. & Dynam. Sys. (2019), **39**, 2176–2222
doi:10.1017/etds.2017.116

Normal forms for perturbations of systems possessing a Diophantine invariant torus

JESSICA ELISA MASSETTI

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

(*Received* 12 *April* 2016 *and accepted in revised form* 18 *September* 2017)

Abstract. We give a new proof of Moser's 1967 normal-form theorem for real analytic perturbations of vector fields possessing a reducible Diophantine invariant quasi-periodic torus. The proposed approach, based on an inverse function theorem in analytic class, is flexible and can be adapted to several contexts. This allows us to prove in a unified framework the persistence, up to finitely many parameters, of Diophantine quasi-periodic normally hyperbolic reducible invariant tori for vector fields originating from dissipative generalizations of Hamiltonian mechanics. As a byproduct, generalizations of Herman's twist theorem and Rüssmann's translated curve theorem are proved.

Contents

1. *Introduction*

1.1. *Moser's normal form.* The starting point of this article is Moser's 1967 theorem [**19**] which, although it has been used by various authors, has remained relatively unnoticed for several years. We present an alternative proof of this result, relying on a more geometrical and conceptual construction, and we use it as inspiration in order to prove new normal-form theorems. We believe that such theorems will be useful in proving other results regarding the persistence of Diophantine tori. Although the difficulties contained in this proof are the same as in the original one (proving the fast convergence of a Newton-like scheme), it relies on a relatively general inverse function theorem (Theorem A.1, unlike in Moser's approach), following an alternative strategy with respect to the one proposed by Zehnder in [**30**, **31**]. Recently Wagener in [**28**] generalized the theorem to vector fields of different regularity, focusing on possible applications in the context of bifurcation theory. We focus here on the analytic category.

Let us introduce Moser's normal form and the framework in which the results will be stated. Let $\mathcal{V}$ be the space of germs of real analytic vector fields along $\mathbb{T}^n \times \{0\}$ in $\mathbb{T}^n \times \mathbb{R}^m$ (where $\mathbb{T}^n = \mathbb{R}^n/2\pi\mathbb{Z}^n$). Let us *fix* $\alpha \in \mathbb{R}^n$ and $A \in \mathrm{Mat}_m(\mathbb{R})$ a diagonalizable matrix of possibly multiple or possibly zero eigenvalues $a_1, \ldots, a_m \in \mathbb{C}^m$. The focus of our interest is on the affine subspace of $\mathcal{V}$ consisting of vector fields of the form

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

where $O(r^k)$ stands for terms of order k or higher which may depend on θ as well. We will denote this subset by $\mathcal{U}(\alpha, A)$.

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

Vector fields in $\mathcal{U}(\alpha, A)$ possess a reducible invariant quasi-periodic torus $\mathrm{T}^n_0 := \mathbb{T}^n \times \{0\}$ of Floquet exponents $a_1, \ldots, a_m$.

We will refer to $\alpha_1, \ldots, \alpha_n, a_1, \ldots, a_m$ as the *characteristic numbers* or characteristic *frequencies*.

Let Λ be the subspace of $\mathcal{V}$ of vector fields of the form

$$\lambda(\theta, r) = (\beta, b + B \cdot r), \quad \beta \in \mathbb{R}^n, b \in \mathbb{R}^m, B \in \mathrm{Mat}_m(\mathbb{R})$$
$$\text{such that } A \cdot b = 0, \quad [A, B] = 0. \tag{1.2}$$

In the following we will refer to λ as *(external) parameters* or *counter terms*.

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

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

φ being a diffeomorphism of the torus fixing the origin and R_0, R_1 being respectively an $\mathbb{R}^m$-valued and a $\mathrm{GL}_m(\mathbb{R})$-valued function defined on $\mathbb{T}^n$ such that† $\Pi_{\ker A} R_0(0) = 0$ and $\Pi_{\ker[A,\cdot]}(R_1(0) - \mathrm{I}) = 0$.

In order to avoid resonances and small divisors, we impose the following Diophantine conditions on the characteristic numbers, for some real positive γ, τ:

$$|k \cdot \alpha| \geq \frac{\gamma}{|k|^\tau} \quad \text{for all } k \in \mathbb{Z}^n \backslash \{0\},$$
$$|\imath k \cdot \alpha - a_i| \geq \frac{\gamma}{(1 + |k|)^\tau} \quad \text{for all } k \in \mathbb{Z}^n \text{ and } 1 \leq i \leq m \text{ such that } a_i \neq 0,$$
$$|\imath k \cdot \alpha + a_i - a_j| \geq \frac{\gamma}{(1 + |k|)^\tau} \quad \text{for all } k \in \mathbb{Z}^n \text{ and } 1 \leq i, j \leq m, i \neq j \text{ such that } a_i \neq a_j, \tag{1.4}$$

where $|k| = |k_1| + \cdots + |k_n|$.

It is a known fact that if τ is large enough and γ small enough, the measure of the set of 'good frequencies' tends to the full measure as γ tends to 0 (see [**20**, **21**] and references therein). Also, remark that only the pure imaginary parts of the Floquet exponents may interfere and create small divisors, due to the factor $\imath$ in front of $k \cdot \alpha$. We will denote by $\mathcal{D}_{\gamma,\tau}$ the set of characteristic numbers satisfying the Diophantine conditions (1.4).

In §2.1 we will introduce complex extensions of manifolds and define the corresponding spaces of real analytic vector fields having such an extension. We will endow such spaces with a Banach norm. All the closeness conditions appearing in the statements of this section thus have to be understood as referring to that norm (see formula (2.1)).

THEOREM 1.1. (Moser, 1967) *If $v \in \mathcal{V}$ is close enough to $u^0 \in \mathcal{U}(\alpha, A)$, there exists a unique triplet $(g, u, \lambda) \in \mathcal{G} \times \mathcal{U}(\alpha, A) \times \Lambda$, in the neighborhood of* $(\mathrm{id}, u^0, 0)$, *such that* $v = g_* u + \lambda$.

The notation $g_* u$ indicates the push-forward of u by g: $g_* u = (g' \cdot u) \circ g^{-1}$.

Although the presence of the counter term $\lambda = (\beta, b + B \cdot r)$ breaks the dynamical conjugacy down, it is a *finite-dimensional* obstruction: geometrically, in the neighborhood of $(\mathrm{id}, u^0, 0)$ the $\mathcal{G}$-orbit of $\mathcal{U}(\alpha, A)$ is a submanifold of $\mathcal{V}$ of finite co-dimension $N \leq n + m + m^2$ (see Figure 1). This co-dimension depends on $\beta \in \mathbb{R}^n$ and the dimension of the kernels of A and $[A, \cdot]$.

† We denote by Π_* the projection on the vector space indicated in subscript and by I the identity matrix.

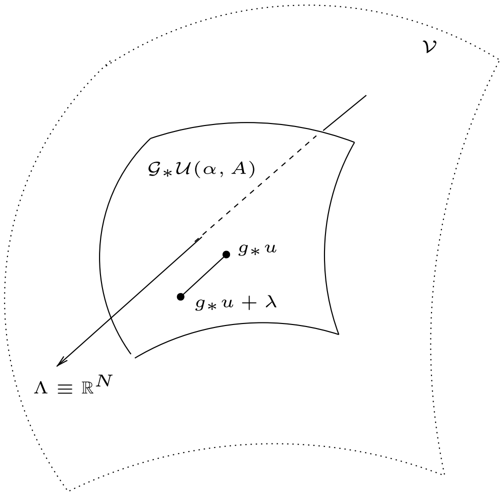

FIGURE 1. Geometrical interpretation of Moser's theorem.

Zhender's approach and ours differ for the following reason, although both rely on the fact that the convergence of the Newton scheme is somewhat independent of the internal structure of variables.

Inverting the operator

$$\phi : (g, u, \lambda) \mapsto g_* u + \lambda = v,$$

as we will in §2, is equivalent to solving implicitly the pulled-back equation ($g^* = g_*^{-1}$)

$$\Phi(g, u, \lambda; v) = g^*(v - \lambda) - u = 0,$$

with respect to u, g and λ, as Zehnder did.

The problem is that whereas ϕ is a local diffeomorphism (in the sense of scales of Banach spaces), the linearization of Φ,

$$\frac{\partial \Phi}{\partial(g, u, \lambda)}(g, u, \lambda; v) \cdot (\delta g, \delta u, \delta\lambda) = [g^*(\lambda - v), g'^{-1} \cdot \delta g] + g^* \delta\lambda + \delta u,$$

where $[\cdot, \cdot]$ is the Lie bracket, is invertible in no neighborhood of $\Phi = 0$. It is invertible in a whole neighborhood of $\Phi = 0$ only up to a second-order term (see Zehnder [**30**, §5]), which prevents us from using a Newton scheme in a straightforward manner. In §2 we give the functional setting in which we prove Moser's theorem.

1.2. *Persistence of tori: elimination of parameters.* The fact that the submanifold $\mathcal{G}_*\mathcal{U}(\alpha, A)$ has finite co-dimension leaves the possibility that in some cases the obstructions represented by the counter terms can be totally eliminated: if the system depends on a sufficient number of free parameters—either internal or external—and λ smoothly depends on them we can try to tune the parameters so that $\lambda = 0$.

When $\lambda = 0$ we have $g_* u = v$: the image $g(\mathrm{T}_0^n)$ is invariant for v and u determines the first- order dynamics along this torus.

When $g_* u + \lambda = v$, we will loosely say that:

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

- T_0^n persists *up to twist*, if $b = 0$ and $B = 0$;
- T_0^n persists *up to translation*, if $\beta = 0$ and $B = 0$;
- T_0^n persists *up to twist-translation*, if $B = 0$.

The *infinite*-dimensional conjugacy problem is reduced to a *finite*-dimensional one. In some cases the crucial point is to allow frequencies $(\alpha_1, \ldots, \alpha_n, a_1, \ldots, a_m)$ to vary, using the fact that λ is Whitney smooth with respect to them. Herman understood the power of this reduction in the 1980s (see [**25**]) and other authors, (Rüssmann, Sevryuk, Chenciner, Broer *et al*, Féjoz, …) adopted this technique of 'elimination of parameters' to prove invariant tori theorems in multiple contexts, at various level of generality, contributing to clarification of this procedure (see, for instance, [**4**, **6**, **7**, **25**, **26**]).

1.3. *Main results.* The proposed geometrization of Moser's result raises different questions about the equivariance of the correction with respect to the groupoid $\mathcal{G}$ and its canonical sub-groupoids. In §§3 and 4 we study some of these equivariance properties in some particular cases issuing from Hamiltonian dynamics and its dissipative versions issued from celestial mechanics. As a by-product, several twisted-torus and translated-torus theorems are given (see §5).

1.3.1. *Hamiltonian dissipative systems.* We start by recalling the classic Hamiltonian counter part of Moser's theorem (see §3).

On $\mathbb{T}^n \times \mathbb{R}^n$, if $\mathcal{U}^{\mathrm{Ham}}(\alpha, 0) \subset \mathcal{U}(\alpha, A)$ is the space (of germs) of Hamiltonian vector fields of the form (1.1) (hence α is Diophantine and $A = 0$), contained in the space $\mathcal{V}^{\mathrm{Ham}} \subset \mathcal{V}$ of Hamiltonian vector fields, and if $\mathcal{G}^{\mathrm{Ham}} \subset \mathcal{G}$ is the space of germs of exact-symplectic isomorphisms of the form

$$g(\theta, r) = (\varphi(\theta),\ {}^t\varphi'(\theta)^{-1} \cdot (r + S'(\theta))),$$

where φ is an isomorphism of $\mathbb{T}^n$ fixing the origin and S a function on $\mathbb{T}^n$ fixing the origin, the space of counter terms is reduced to the set of $\lambda = (\beta, 0)$, $\beta \in \mathbb{R}^n$: we have Herman's 'twisted conjugacy' theorem (see [**10**, **11**, **13**]).

THEOREM. (Herman) *Let α be Diophantine and $u^0 \in \mathcal{U}^{\mathrm{Ham}}(\alpha, 0)$. If $v \in \mathcal{V}^{\mathrm{Ham}}$ is sufficiently close to u^0, the torus T_0^n persists up to twist. In particular, the conjugacy (up to twist) is given by an exact-symplectic transformation.*

In §4 we prove a first dissipative generalization of this classic result by considering the affine spaces†

$$\mathcal{U}^{\mathrm{Ham}}(\alpha, -\eta) := \mathcal{U}^{\mathrm{Ham}}(\alpha, 0) \oplus (-\eta r\, \partial_r) \subset \mathcal{V}^{\mathrm{Ham}} \oplus (-\eta r\, \partial_r),$$

where $\eta \in \mathbb{R}$, by extending the normal direction with the constant linear term $-\eta r$ (when $\eta > 0$ we speak of 'radial dissipation'), but keeping the same space of exact-symplectic isomorphisms $\mathcal{G}^{\mathrm{Ham}}$ and Hamiltonian corrections $\lambda = (\beta, 0)$.

THEOREM A. *Fix $\eta_0 > 0$ and α Diophantine. There exists $\varepsilon > 0$ such that for any $\eta \in [-\eta_0, \eta_0]$, letting $u^0 \in \mathcal{U}^{\mathrm{Ham}}(\alpha, -\eta)$, if $v \in \mathcal{V}^{\mathrm{Ham}} \oplus (-\eta r\, \partial_r)$ is ε-close to u^0 the torus T_0^n persists up to twist and its final normal dynamics is always given by $-\eta$.*

† We denote $\partial_r = (\partial_{r_1}, \ldots, \partial_{r_n})$ and omit the tensor product sign $r \otimes \partial_r$.

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

Moreover, the conjugacy (up to twist) is given by an exact-symplectic transformation.

Obviously, Theorem A reduces to Herman's theorem when $\eta = 0$.

We stress the fact that the number of counter terms breaking the dynamical conjugacy is *the same* as in the purely Hamiltonian context (a twisting term $\beta\partial_\theta$, $\beta \in \mathbb{R}^n$, in the angle's direction). Moreover, we control both the tangent and the normal dynamics of the torus, which survive perturbations (up to twist) *uniformly* with respect to dissipation (as opposed to the classic normally hyperbolic frame). See Remark 4.1 in the proof of Proposition 4.1.

In the general non-symplectic case, if A has simple non-zero eigenvalues, the corrections space is immediately given by the set of $\lambda = (\beta, B \cdot r)$, with B a diagonal matrix.

A diagram summarizing these results is given at the end of §4.

In $\mathcal{U}^{\mathrm{Ham}}(\alpha, -\eta)$ let $\widehat{\mathcal{U}}^{\mathrm{Ham}}(\alpha, -\eta)$ be the space of those vector fields that satisfy a *torsion hypothesis* (coming from Hamiltonians with non-degenerate quadratic term). In this case, we can consider the space of perturbations

$$\mathcal{V}^{\mathrm{Ham}} \oplus (-\eta r + \zeta)\partial_r,$$

where $\zeta \in \mathbb{R}^n$, but extend the space of transformations to the space $\mathcal{G}^\omega$ of symplectic isomorphisms of the form

$$g(\theta, r) = (\varphi(\theta),\ {}^t\varphi'(\theta)^{-1} \cdot (r + S'(\theta) + \xi)), \quad \xi \in \mathbb{R}^n.$$

The space of counter terms becomes the set of translations *in action* $\lambda = (0, b)$.

THEOREM B. (Vector fields *à la* Rüssmann) *Fix $\eta_0 > 0$ and α Diophantine. There exists $\varepsilon > 0$ such that for any $\eta \in [-\eta_0, \eta_0]$, letting $u^0 \in \widehat{\mathcal{U}}^{\mathrm{Ham}}(\alpha, -\eta)$, if $v \in \mathcal{V}^{\mathrm{Ham}} \oplus (-\eta r + \zeta)\partial_r$ is ε-close to u^0 the torus T^n_0 persists up to translation and its final normal dynamics is always given by $-\eta$. Moreover, the conjugacy (up to translation) is given by a symplectic transformation.*

As in Theorem A, the bound on admissible perturbations is uniform with respect to η and the translated torus $g(\mathrm{T}^n_0)$ is dynamically characterised by the same initial frequencies $(\alpha, -\eta)$.

Theorem B can be seen as a multidimensional generalization for vector fields in this class of Rüssmann's translated curve theorem [**23**].

1.3.2. *General dissipative systems.* At the cost of changing its final (constant) normal dynamics (by conjugating $v - \lambda$ to a vector field u characterized by a different A), we can prove that an α-quasi-periodic Diophantine torus resists general perturbations. The following results will be proved in §5, where a more functional statement will be given (Theorems 5.1 and 5.2).

On $\mathbb{T}^n \times \mathbb{R}^m$, let $u \in \mathcal{U}(\alpha, A)$, defined in expression (1.1), be such that A has simple, real, non-zero eigenvalues $a_1, \ldots, a_m$. This hypothesis of course implies that the only frequencies that can cause small divisors are the tangential ones $\alpha_1, \ldots, \alpha_n$, so that we need only require the standard Diophantine condition on α.

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

THEOREM C. (Twisted torus) *Let α be Diophantine, let $A \in \mathrm{Mat}_m(\mathbb{R})$ have real, simple, non-zero eigenvalues and let $u^0 \in \mathcal{U}(\alpha, A)$. There exists $\varepsilon > 0$ such that, if $v \in \mathcal{V}$ is ε-close to u^0, there exists A' close to A such that the torus T_0^n persists up to twist and its final normal dynamics is given by A'.*

Let $m \geq n$ and consider $u \in \mathcal{U}(\alpha, A)$. Here we loosely say that u has twist if the matrix term $u_1 : \mathbb{T}^n \to \mathrm{Mat}_{n\times m}(\mathbb{R})$ in

$$u(\theta, r) = (\alpha + u_1(\theta)\cdot r + O(r^2),\ A\cdot r + O(r^2)),$$

is such that $\int_{\mathbb{T}^n} u_1(\theta)\, d\theta$ has maximal rank n.

THEOREM D. (Translated torus) *Let α be Diophantine, let $A \in \mathrm{Mat}_m(\mathbb{R})$ have real, simple eigenvalues and let $u^0 \in \mathcal{U}(\alpha, A)$ have twist. There exists $\varepsilon > 0$ such that, if $v \in \mathcal{V}$ is ε-close to u^0, there exists A' close to A such that the torus T_0^n persists up to translation and its final normal dynamics is given by A'.*

1.4. *An application to celestial mechanics.* The motivation of the previous geometric results on normal forms for dissipative systems comes from celestial mechanics. These normal forms provide ready-to-use theorems that in some cases fit very well with concrete problems issuing from celestial mechanics. Besides, if on the one hand these theorems clarify in a neat way the 'lack of parameters' problem, on the other hand the procedure of elimination of parameters highlights relations between physical parameters and the existence of invariant tori in the system.

To give a major example, we conclude the paper with an application of Theorem B to the problem of persistence of quasi-periodic attractors in the spin–orbit system; this astronomical problem amounts to studying the dynamics of the rotation about its spin axis of a non-rigid and non-elastic body whose center of mass revolves in a given elliptic Keplerian orbit around a fixed massive point (see §6.2 for the precise formulations of the model). A study of this problem using a partial-differential-equation approach was given in [**5**], while a generalization in higher dimension was presented in [**27**], but using Lie series techniques instead.

For the $2n$-dimensional model on $\mathbb{T}^n \times \mathbb{R}^n$ we consider the n-parameter family of vector fields of the form

$$\hat{v} = v - \eta(r - \Omega)\partial_r$$

where $v \in \mathcal{V}^{\mathrm{Ham}}$ is a perturbation of $u^0 \in \widehat{\mathcal{U}}^{\mathrm{Ham}}(\alpha, 0)$ with non-degenerate torsion, $\eta \in \mathbb{R}$ a friction constant and $\Omega \in \mathbb{R}^n$ a vector of external free parameters. By simple application of (the translated torus) Theorem B and the implicit function theorem in finite dimension, the persistence result is phrased as follows (see Theorem 6.1 in §6.1.1).

THEOREM. (Spin–orbit in n degrees of freedom) *Fix $\eta_0 > 0$, $\alpha \in \mathbb{R}^n$ Diophantine, and let $u^0 \in \widehat{\mathcal{U}}^{\mathrm{Ham}}(\alpha, 0)$. There exists $\varepsilon > 0$ such that, for any $\eta \in [-\eta_0, \eta_0]$, if v is ε-close to u^0, there exist a unique frequency adjustment $\Omega \in \mathbb{R}^n$ close to 0, a unique $u \in \widehat{\mathcal{U}}^{\mathrm{Ham}}(\alpha, -\eta)$ and a unique $g \in \mathcal{G}^\omega$ such that $\hat{v}$ satisfies $g_* u = \hat{v}$. Hence $\hat{v}$ possesses an invariant α-quasi-periodic torus. This torus is η-normally attractive (respectively, repulsive) if $\eta > 0$ (respectively, $\eta < 0$).*

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

This result is finally applied to the astronomical spin–orbit problem. This problem is modeled by the following one-parameter family of non-autonomous real analytic vector fields on $\mathbb{T} \times \mathbb{R}$:

$$v = (\alpha + r, -\eta r + \eta(\nu - \alpha) - \varepsilon \partial_\theta f(\theta, t)),$$

where $\nu \in \mathbb{R}$ is a free parameter. By extending the phase space in the usual way, we get the autonomous Hamiltonian dissipative system whose corresponding Hamiltonian is

$$H(\theta, r) = \alpha r_1 + r_2 + \tfrac{1}{2} r_1^2 + \varepsilon f(\theta_1, \theta_2).$$

By applying Theorem B and the elimination of the translation parameter, the result can be stated as follows.

THEOREM. (Surfaces of invariant tori) *Let ε_0 be the maximal value that the perturbation can attain. Every Diophantine α identifies a surface $(\varepsilon, \eta) \mapsto \nu(\varepsilon, \eta)$ in the space $(\varepsilon, \eta, \nu) = [0, \varepsilon_0] \times [-\eta_0, \eta_0] \times \mathbb{R}$, which is analytic in ε, smooth in η, for which the following holds: for any parameters $(\varepsilon, \eta, \nu(\varepsilon, \eta))$, $\hat{v}$ admits an invariant α-quasi-periodic torus. This torus is η-normally attractive (respectively, repulsive) if $\eta > 0$ (respectively, $\eta < 0$).*

Having made all the appropriate reductions (see Corollary 6.1), the proof is a particular case of Theorem B *à la* Rüssmann and the elimination of the translation parameter (see Theorem 6.2 and Corollary 6.2).

2. *Moser's normal form*

Theorem 1.1 will be deduced by the abstract inverse function theorem (Theorem A.1) and the regularity propositions (Propositions A.1–A.3) in Appendix A.

2.1. *Complex extensions.*

Let us extend the tori

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

as

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

and, letting $s > 0$, 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 $|(\mathrm{Im}\, \theta, r)| := \max\, (\max_{1 \le j \le n} |\mathrm{Im}\, \theta_j|, \max_{1 \le j \le m} |r_j|)$.

Let $f : \mathrm{T}^n_s \to \mathbb{C}$ now be holomorphic, and consider its Fourier expansion $f(\theta, r) = \sum_{k \in \mathbb{Z}^n} f_k(r)\, e^{i\, k \cdot \theta}$, noting that $k \cdot \theta = k_1 \theta_1 + \cdots + k_n \theta_n$. In this context we introduce the so-called 'weighted norm':

$$|f|_s := \sum_{k \in \mathbb{Z}^n} |f_k|\, e^{|k| s}, \qquad |k| = |k_1| + \cdots + |k_n|, \tag{2.1}$$

where $|f_k| = \sup_{|r| < s} |f_k(r)|$. Whenever $f : \mathrm{T}^n_s \to \mathbb{C}^n$, we have $|f|_s = \max_{1 \le j \le n}(|f_j|_s)$, f_j being the jth component of $f(\theta, r)$.

It is a trivial fact that the classical supremum 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 $\mathrm{T}^n_{s'}$:

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

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 $|g - \mathrm{id}|_s$. For more details about the weighted norm, see, for example, [**8**, **18**].

In general, for complex extensions U_s and $V_{s'}$ of $\mathbb{T}^n \times \mathbb{R}^m$, we will denote by $\mathcal{A}(U_s, V_{s'})$ the set of real holomorphic functions from U_s to $V_{s'}$ and by $\mathcal{A}(U_s)$, endowed with the s-weighted norm, the Banach space $\mathcal{A}(U_s, \mathbb{C})$. 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 $x_1, \ldots, x_p \in E$, then

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

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

If f is a differentiable map between two open sets of E and F, then $f'(x)$ 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. *Notation.* In the course of this paper we will define subsets of the three spaces of germs of transformations $\mathcal{G}$, $\mathcal{G}^{\mathrm{Ham}}$, $\mathcal{G}^{\omega}$ introduced in §§1.1–1.3, by adding superscripts $*^\sigma$ and subscripts $*_s$. A superscript $\sigma > 0$ denotes the set of diffeomorphisms σ-close to the identity, while a subscript $s > 0$ denotes the s-width of analyticity of the diffeomorphism in the set considered.

2.3. *Space of conjugacies.* Let

$$\chi_s := \{v \in \mathcal{A}(\mathbb{T}^n_s, \mathbb{C}^n) : v(0) = 0\}$$

be the space of vector fields on the torus vanishing at the origin, endowed with the norm $|\cdot|_s$.

Let $\mathcal{D}_s$ be the space of maps $\varphi \in \mathcal{A}(\mathbb{T}^n_s, \mathbb{T}^n_{\mathbb{C}})$ which are real holomorphic isomorphisms from the interior of $\mathbb{T}^n_s$ to its image and let $\mathcal{D}^\sigma_s$ be the neighborhood of the identity, identified with maps $\varphi = \mathrm{id} + v : \mathbb{T}^n_s \to \mathbb{T}^n_{\mathbb{C}}$, where $v = \varphi - \mathrm{id} \in \chi_s$ is such that $|v|_s < \sigma$.

If σ is small enough, according to Theorem B.1, such a map is a biholomorphism on $\mathbb{T}^n_{s'}$ for some $s' > 0$.

Let $\mathcal{G}_s$ be the affine space passing through the identity and directed by $\{(\varphi - \mathrm{id},\, R_0 + (R_1 - I)\cdot r)\}$, where $\varphi - \mathrm{id} \in \chi_s$, while $R_0 \in \mathcal{A}(\mathbb{T}^n_s, \mathbb{C}^m)$ and $R_1 \in \mathcal{A}(\mathbb{T}^n_s, \mathrm{GL}_m(\mathbb{C}))$ are such that $\Pi_{\ker A} R_0(0) = 0$ and $\Pi_{\ker[A,\cdot]}(R_1(0) - I) = 0$.

Now let $\mathcal{G}^\sigma_s$ be the neighborhood of the identity in $\mathcal{G}_s$, consisting of the maps

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

such that

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

and

$$|R_0(\theta) + R_1(\theta)\cdot r - r|_s < \sigma.$$

The 'Lie algebra' $\mathrm{T}_{\mathrm{id}}\mathcal{G}^\sigma_s$ of $\mathcal{G}^\sigma_s$, consists of maps

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

Here $\dot{g}$ lies in $\mathcal{A}(\mathrm{T}^n_s, \mathbb{C}^{n+m})$; more specifically, $\dot{\varphi} \in \chi_s$, $\dot{R}_0 \in \mathcal{A}(\mathbb{T}^n_s, \mathbb{C}^m)$ and $\dot{R}_1 \in \mathcal{A}(\mathbb{T}^n_s, \mathrm{Mat}_m(\mathbb{C}))$. We endow this space with the norm

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

2.4. *Spaces of vector fields.* Let $\mathcal{V}_s = \mathcal{A}(\mathrm{T}^n_s, \mathbb{C}^{n+m})$, endowed with the norm

$$|v|_s := \max_{1\le j\le n+m}(|v_j|_s),$$

and $\mathcal{V} = \bigcup_s \mathcal{V}_s$.

For $\alpha \in \mathbb{R}^n$ and $A \in \mathrm{Mat}_m(\mathbb{R})$, let $\mathcal{U}_s(\alpha, A)$ be the affine subspace of $\mathcal{V}_s$ consisting of vector fields of the form

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

2.5. *The normal-form operator ϕ.* According to Theorem B.1 and Corollary B.1, the family of operators

$$\phi : \mathcal{G}^{\sigma/n}_{s+\sigma} \times \mathcal{U}_{s+\sigma}(\alpha, A) \times \Lambda \to \mathcal{V}_s,\ (g, u, \lambda) \mapsto g_* u + \lambda, \tag{2.2}$$

is now defined. We will always assume that $0 < s < s+\sigma < 1$ and $\sigma < s$. We want to solve

$$\phi(g, u, \lambda) = v,$$

for v close to $\phi(\mathrm{id}, u^0, 0) = u^0$.

The drawback of focusing on ϕ is that we will need the germs of $g_* u + \lambda$ and v to match on the unknown torus $g(\mathrm{T}^n_0)$ and need to pay attention to composition operators in order not to shrink artificially the domains of analyticity, because of the rigidity of analytic maps. See Figure 2. For this purpose, given a diffeomorphism $g \in \mathcal{G}^\sigma_s$ and a real analytic vector field v on $g(\mathrm{T}^n_s)$, we define the deformed norm

$$|v|_{g,s} := |g^* v|_s$$

depending on g, where the notation $g^* = g_*^{-1}$ stands for the pull-back of v.

In the following we do not intend to be optimal.

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

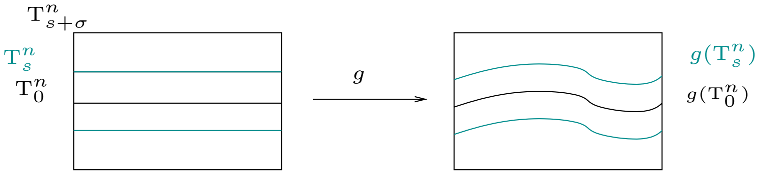

FIGURE 2. Deformed complex domain.

2.6. *Cohomological equations.* Here we present three derivative operators and estimate the solutions of the three associated cohomological equations which, in the proof of Proposition 2.1, will allow us to straighten the first-order dynamics of the torus at the infinitesimal level.

A vector $\alpha \in \mathbb{R}^n$ is identified with a constant vector field on the torus $\mathbb{T}^n$, thus with the derivation operator

$$L_\alpha : \mathcal{A}(\mathbb{T}^n_{s+\sigma}) \to \mathcal{A}(\mathbb{T}^n_s), \quad f \mapsto L_\alpha f = f' \cdot \alpha := \sum_{j=1}^{n} \alpha_j \frac{\partial f}{\partial \theta_j}.$$

Now let $\alpha \in \mathbb{R}^n$ and $M \in \mathrm{Mat}_m(\mathbb{R})$, a diagonalizable matrix of simple, non-zero eigenvalues $\mu = (\mu_1, \ldots, \mu_m) \in \mathbb{C}^m$, satisfy the Diophantine conditions

$$|k \cdot \alpha| \geq \frac{\gamma}{|k|^\tau} \quad \text{for all } k \in \mathbb{Z}^n \backslash \{0\}, \tag{2.3}$$

$$|\imath k \cdot \alpha + \mu_j| \geq \frac{\gamma}{(1+|k|)^\tau} \quad \text{for all } k \in \mathbb{Z}^n,\ j = 1, \ldots, m, \tag{2.4}$$

$$|\imath k \cdot \alpha + l \cdot \mu| \geq \frac{\gamma}{(1+|k|)^\tau} \quad \text{for all } (k, l) \in \mathbb{Z}^n \times \mathbb{Z}^m \backslash \{0\},\ |l| = 2. \tag{2.5}$$

LEMMA 1. (Straightening the dynamics on the torus) *Let $\alpha \in \mathbb{R}^n$ satisfy condition (2.3). For every $g \in \mathcal{A}(\mathbb{T}^n_{s+\sigma})$ of zero average, there exists a unique pre-image $f \in \mathcal{A}(\mathbb{T}^n_s)$ by L_α of zero average satisfying the estimate*

$$|f|_s = |L_\alpha^{-1} g|_s \leq \frac{C_1}{\gamma} \frac{1}{\sigma^{n+\tau}} |g|_{s+\sigma},$$

C_1 being a constant depending only on the dimension n and the exponent τ.

Proof. Let

$$g(\theta) = \sum_{k \in \mathbb{Z}^n \backslash \{0\}} g_k e^{i\, k \cdot \theta}$$

be the Fourier expansion of g. 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 $\operatorname{Im} \theta_j = -\operatorname{sgn}(k_j)(s+\sigma)$. Expanding the term $L_\alpha f$ too, we see that a formal solution of $L_\alpha f = g$ is given by

$$f = \sum_{k \in \mathbb{Z}^n \backslash \{0\}} \frac{g_k}{i\, k \cdot \alpha} e^{i\, k \cdot \theta}. \tag{2.6}$$

Taking into account the Diophantine condition (2.3), we have

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

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, belongs to $\mathcal{A}(\mathbb{T}^n_s)$ and satisfies the claimed estimate. □

Let us define

$$L_\alpha + M : \mathcal{A}(\mathbb{T}^n_{s+\sigma}, \mathbb{C}^m) \to \mathcal{A}(\mathbb{T}^n_s, \mathbb{C}^m), \quad f \mapsto L_\alpha f + M\cdot f = f'\cdot\alpha + M\cdot f.$$

LEMMA 2. (Relocating the torus) *Let $\alpha \in \mathbb{R}^n$ and $M \in \mathrm{GL}_m(\mathbb{R})$ satisfy the Diophantine condition (2.4). For every $g \in \mathcal{A}(\mathbb{T}^n_{s+\sigma}, \mathbb{C}^m)$, there exists a unique pre-image $f \in \mathcal{A}(\mathbb{T}^n_s, \mathbb{C}^m)$ by $L_\alpha + M$. Moreover, the following estimate holds:*

$$|f|_s = |(L_\alpha + M)^{-1}g|_s \le \frac{C_2}{\gamma}\frac{1}{\sigma^{n+\tau}}|g|_{s+\sigma},$$

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

Proof. Let us start for simplicity with the scalar case $g \in \mathcal{A}(\mathbb{T}^n_{s+\sigma})$ and $M = \mu \ne 0 \in \mathbb{R}$. Expanding both sides of $L_\alpha f + \mu f = g$, we see that the Fourier coefficients of the formal pre-image f are given by

$$f_k = \frac{g_k}{ik\cdot\alpha + \mu},$$

hence

$$f = (L_\alpha + \mu)^{-1}g = \sum_{k\in\mathbb{Z}^n}\frac{g_k}{ik\cdot\alpha+\mu}e^{ik\cdot\theta}. \tag{2.7}$$

Taking into account the Diophantine condition (2.4) and doing the same sort of calculations as in Lemma 1, we get the desired estimate.

The case where M is a diagonal matrix can be recovered from the scalar case by working componentwise.

When M is diagonalizable, let $P \in \mathrm{GL}_n(\mathbb{C})$ be such that PMP^{-1} is diagonal. Considering $f'\cdot\alpha + M\cdot f = g$, and left-multiplying both sides by P, we get

$$\tilde{f}'\cdot\alpha + PMP^{-1}\tilde{f} = \tilde{g},$$

where we have set $\tilde{g} = Pg$ and $\tilde{f} = Pf$. This equation has a unique solution with the desired estimates. We just need to put $f = P^{-1}\tilde{f}$. □

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

Let us consider the space of real analytic functions on $\mathbb{T}^n_{s+\sigma}$ with values in $\mathrm{Mat}_m(\mathbb{C})$ and define the operator

$$\begin{aligned} L_\alpha + [M,\cdot]: \quad \mathcal{A}(\mathbb{T}^n_{s+\sigma}, \mathrm{Mat}_m(\mathbb{C})) &\to \mathcal{A}(\mathbb{T}^n_s, \mathrm{Mat}_m(\mathbb{C})) \\ F &\mapsto L_\alpha F + [M, F]. \end{aligned}$$

With the notation $L_\alpha F$ (or $F'\cdot\alpha$) we mean that we are applying the Lie derivative operator to each component F^i_j of the matrix F; $[M, F]$ is the usual commutator.

LEMMA 3. (Straightening the first-order dynamics) *Let $\alpha \in \mathbb{R}^n$ and $M \in \mathrm{GL}_m(\mathbb{R})$ satisfy the Diophantine conditions (2.3) and (2.5). For every $G \in \mathcal{A}(\mathbb{T}^n_{s+\sigma}, \mathrm{Mat}_m(\mathbb{C}))$ whose diagonal elements have zero average $\int_{\mathbb{T}^n} G^i_i \,(d\theta/(2\pi)^n) = 0$, there exists a unique $F \in \mathcal{A}(\mathbb{T}^n_s, \mathrm{Mat}_m(\mathbb{C}))$ with $\int_{\mathbb{T}^n} F^i_i \,(d\theta/(2\pi)^n) = 0$, such that the matrix equation*

$$L_\alpha F + [M, F] = G$$

is satisfied. Moreover, the following estimate holds:

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

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

Proof. Let us start with the diagonal case. Let $M = \mathrm{diag}(\mu_1, \ldots, \mu_m)$, where $\mu_1, \ldots, \mu_m \in \mathbb{R}$, and $F \in \mathrm{Mat}_m(\mathbb{C})$ be given. The commutator $[M, F]$ reads

$$\begin{pmatrix} 0 & (\mu_1-\mu_2)F^1_2 & (\mu_1-\mu_3)F^1_3 & \cdots & (\mu_1-\mu_m)F^1_m \\ (\mu_2-\mu_1)F^2_1 & 0 & (\mu_2-\mu_3)F^2_3 & \cdots & (\mu_2-\mu_m)F^2_m \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ (\mu_m-\mu_1)F^m_1 & (\mu_m-\mu_2)F^m_2 & \cdots & \cdots & 0 \end{pmatrix}, \tag{2.8}$$

where we denoted by F^i_j the element corresponding to the ith line and jth column of the matrix $F(\theta)$. Using the components notation, the matrix reads

$$([A, F]^i_j) = ((\mu_i - \mu_j)F^i_j),$$

and shows all zeros along the diagonal. Adding it now up with the matrix $L_\alpha F$, which reads

$$\begin{pmatrix} L_\alpha F^1_1 & \cdots & L_\alpha F^1_m \\ \vdots & L_\alpha F^i_i & \vdots \\ L_\alpha F^n_1 & \cdots & L_\alpha F^m_m \end{pmatrix}, \tag{2.9}$$

we see that to solve the equation $L_\alpha F + [M, F] = G$, we need to solve m equations of the type of Lemma 1 and $m^2 - m$ equations of the type of Lemma 2. Expanding every element in Fourier series, we see that the formal solution is given by a matrix F whose diagonal elements are of the form

$$F^j_j = \sum_{k\in\mathbb{Z}^n\setminus\{0\}} \frac{G^j_{j,k}}{ik\cdot\alpha}\, e^{ik\cdot\theta},$$

while the non-diagonal terms are of the form

$$F^i_j = \sum_{k\in\mathbb{Z}^n} \frac{G^i_{j,k}}{ik\cdot\alpha + (\mu_i - \mu_j)}\, e^{ik\cdot\theta}.$$

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

Recall that the eigenvalues of M are simple and non-zero. By the Diophantine conditions (2.3)–(2.5), via the same kind of calculations as we did in the previous lemmas, we get the desired estimate.

To recover the general case, we consider $P \in \mathrm{GL}_m(\mathbb{C})$ such that PMP^{-1} is diagonal and the equation

$$L_\alpha(PFP^{-1}) + P[M, F]P^{-1} = PGP^{-1},$$

and observe that we can see $P[M, F]P^{-1}$ as

$$P[M, F]P^{-1} = PMP^{-1}PFP^{-1} - PFP^{-1}PMP^{-1} = [PMP^{-1}, PFP^{-1}].$$

Letting $\tilde{F} = PFP^{-1}$ and $\tilde{G} = PGP^{-1}$, $\tilde{F}$ satisfies the desired estimates, and $G = P^{-1}\tilde{G}P$. $\square$

We direct the reader looking for optimal estimates to the paper of Rüssmann [**24**].

Remark that in case of real eigenvalues, condition (2.4) is redundant. Condition (2.3) suffices, with the choice $\gamma < \min_j(|\mathrm{Re}\, \mu_j|)$.

2.7. *Inversion of the operator ϕ: estimates on ϕ'^{-1} and ϕ''.* The following theorem represents the main result of this first part from which Moser's Theorem 1.1 follows.

Let us fix $u^0 \in \mathcal{U}_s(\alpha, A)$ and denote by $\mathcal{V}_s^\sigma = \{v \in \mathcal{V}_s : |v - u^0|_s < \sigma\}$ the ball of radius σ centered at u^0.

THEOREM 2.1. *The operator ϕ is a local diffeomorphism in the sense that for any $\eta < s < s + \sigma < 1$ there exist $\varepsilon > 0$ and a unique C^∞-map ψ,*

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

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 (Theorem A.1 and the regularity propositions (Propositions A.1–A.3), provided that we appropriately estimate ϕ'^{-1} and ϕ''.

Let us sketch the proof. In order to solve $\phi(x) = y$ locally we use Kolmogorov's idea and find the solution by composing the operator

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

infinitely many times 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 suitable estimates, in order to control the convergence of the iterates.

Let $\overrightarrow{\mathcal{U}}$ be the vector space directing $\mathcal{U}(\alpha, A)$. We start to check the invertibility of

$$\phi'(g, u, \lambda) : T_g\mathcal{G}_{s+\sigma}^{\sigma/n} \times \overrightarrow{\mathcal{U}}_{s+\sigma} \times \Lambda \to \mathcal{V}_{g,s}$$

if g is close to the identity, by solving

$$\phi'(g, u, \lambda) \cdot (\delta g, \delta u, \delta\lambda) = \delta v. \tag{2.10}$$

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

PROPOSITION 2.1. *If g is close enough to the identity, for any δv in $\mathcal{V}_{g,s+\sigma}$ there exists a unique triplet $(\delta g, \delta u, \delta\lambda)$ such that equation (2.10) is satisfied. Moreover, there exist $\tau', C' > 0$ such that*

$$\max(|\delta g|_s, |\delta u|_s, |\delta\lambda|) \leq \frac{C'}{\sigma^{\tau'}} |\delta v|_{g,s+\sigma}, \tag{2.11}$$

where C' is a constant that depends only on n, τ and $|(g - \mathrm{id}, u - (\alpha, A \cdot r))|_{s+\sigma}$.

Proof. Let a vector field δv in $\mathcal{V}_{g,s+\sigma}$ be given. We want to solve the equation

$$\phi'(g, u, \lambda) \cdot (\delta g, \delta u, \delta\lambda) = [g_* u, \delta g \circ g^{-1}] + g_*\delta u + \delta\lambda = \delta v, \tag{2.12}$$

where the Lie bracket comes from the differentiation of the map $g \mapsto g_* u$ at g, since $(g_* u)' \cdot \delta g = [g_* u, \delta g \circ g^{-1}]$, by classical derivation rules on compositions and inverse mappings (see [**15**], for example).

In (2.12), δv is the data, and the unknowns are $\delta u \in O(r) \times O(r^2)$, δg (geometrically a vector field along g) and $\delta\lambda \in \Lambda$.

Both sides are supposed to belong to $\mathcal{V}_{g,s+\sigma}$; in order to solve the equation we pull it back, obtaining the equivalent equation between germs along the standard torus T^n_0 (as opposed to the g-dependent torus $g(\mathrm{T}^n_0)$). By naturality of the Lie bracket with respect to the pull-back operator, we thus obtain the equivalent system in $\mathcal{V}_{s+\sigma}$:

$$[u, g^*\delta g \circ g^{-1}] + \delta u + g^*\delta\lambda = g^*\delta v.$$

For ease of notation we denote the new terms by†

$$\dot\lambda := g^*\delta\lambda, \quad \dot v := g^*\delta v, \quad \dot g := g^*\delta g \circ g^{-1} = g'^{-1} \cdot \delta g, \tag{2.13}$$

where $\delta g = (\delta\varphi, \delta R_0 + \delta R_1 \cdot r)$, $\delta\lambda = (\delta\beta, \delta b + \delta B \cdot r)$, and

$$[u, \dot g] + \delta u + \dot\lambda = \dot v. \tag{2.14}$$

The unknowns are now $\dot g$ (geometrically a germ of vector fields along T^n_0), δu and $\dot\lambda$. The new infinitesimal vector field of counter terms $\dot\lambda$ is no longer constant in general. On the other hand, we can take advantage of u in its 'straight' form.

Let us expand the vector fields along T^n_0:

$$\begin{cases} u(\theta, r) = (\alpha + u_1(\theta) \cdot r + O(r^2), A \cdot r + U_2(\theta) \cdot r^2 + O(r^3)), \\ \dot g(\theta, r) = (\dot\varphi(\theta), \dot R_0(\theta) + \dot R_1(\theta) \cdot r), \\ \dot\lambda(\theta, r) = (\dot\lambda_0(\theta), \dot\Lambda_0(\theta) + \dot\Lambda_1(\theta) \cdot r), \\ \dot v(\theta, r) = (\dot v_0(\theta) + O(r), \dot V_0(\theta) + \dot V_1(\theta) \cdot r + O(r^2)). \end{cases} \tag{2.15}$$

We are interested in normalizing the dynamics tangentially at order zero with respect to r, and up to the first order in the normal direction. We then consider the 'mixed jet'

$$j^{0,1}\dot v = (\dot v_0(\theta), \dot V_0(\theta) + \dot V_1(\theta) \cdot r).$$

† Here and in the following we shall use the notation $\dot x$ for all the pulled-back tangent vectors, which are tangent vector fields defined (again) along a neighborhood of the standard T^n_0.

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

Since

$$[u, \dot{g}] = (\dot{\varphi}' \cdot \alpha - u_1 \cdot \dot{R}_0 + O(r^2))\frac{\partial}{\partial\theta} + (\dot{R}_0' \cdot \alpha - A \cdot \dot{R}_0 + ([A, \dot{R}_1] + \dot{R}_1' \cdot \alpha + \dot{R}_0' \cdot u_1 - 2U_2 \cdot \dot{R}_0) \cdot r + O(r^2))\frac{\partial}{\partial r},$$

by taking the image by $j^{0,1}$ of equation (2.14) and identifying terms of the same order, we obtain

$$\dot{\varphi}' \cdot \alpha - u_1 \cdot \dot{R}_0 = \dot{v}_0 - \dot{\lambda}_0, \tag{2.16}$$

$$\dot{R}_0' \cdot \alpha - A \cdot \dot{R}_0 = \dot{V}_0 - \dot{\Lambda}_0, \tag{2.17}$$

$$[A, \dot{R}_1] + \dot{R}_1' \cdot \alpha + \dot{R}_0' \cdot u_1 - 2U_2 \cdot \dot{R}_0 = \dot{V}_1 - \dot{\Lambda}_1, \tag{2.18}$$

where (2.16) concerns the tangent direction and (2.17)–(2.18) the normal direction. This is a triangular system that, starting from (2.17), we are able to solve; actually these equations are of the same type as those we already solved in Lemmas 1–3 (in the sense of their projection on the image of the operator $j^{0,1}[u, \cdot]$).

We remark that since $\delta u = (O(r), O(r^2))$, $j^{0,1}\delta u = 0$ and δu gives no contribution to the previous equations. Once we have solved them, we will determine δu by identifying the remainders.

Remark 2.1. Every equation contains two unknowns, $\dot{g}$ and $\dot{\lambda}$, while $\dot{v}$ is given. Note that the operator $j^{0,1}[u, \cdot]$ has a kernel of finite dimension $N = n + \dim \operatorname{Ker} A + \dim \operatorname{Ker}[A, \cdot]$, which consists of elements $\{(\bar{\beta}, \bar{b} + \bar{B} \cdot r)\}$ such that $A \cdot \bar{b} = 0$ and $[A, \bar{B}] = 0$; the solutions of equations (2.16)–(2.18) are thus determined up to a constant term belonging to this kernel.

We shall start to solve equations modulo $\dot{\lambda}$. Eventually $\delta\lambda$ will be uniquely chosen to kill the component of the known terms belonging to the kernel of $j^{0,1}[u, \cdot]$ and solve the cohomological equations.

Let us proceed with solving the system. We will repeatedly apply Lemmas 1–3 and Cauchy's inequality. Furthermore, we do not keep track of constants—just know that they depend only on $n, \tau > 0$ (from the Diophantine condition), $|g - \mathrm{id}|_{s+\sigma}$ and $|(u - (\alpha, A \cdot r))|_{s+\sigma}$—and hence refer to them as C.

Firstly, consider (2.17). Defining $\bar{b} = \prod_{\operatorname{Ker} A} \int_{\mathbb{T}^n} \dot{V}_0 - \dot{\Lambda}_0 \, (d\theta/(2\pi)^n)$, we have

$$\dot{R}_0 = (L_\alpha + A)^{-1}(\dot{V}_0 - \dot{\Lambda}_0 - \bar{b})$$

and

$$|\dot{R}_0|_s \le \frac{C}{\gamma} \frac{1}{\sigma^{n+\tau}} |\dot{V}_0 - \dot{\Lambda}_0|_{s+\sigma}.$$

Secondly, consider equation (2.16). Calling the average

$$\bar{\beta} = \int_{\mathbb{T}^n} \dot{v}_0 + u_1 \cdot \dot{R}_0 - \dot{\lambda}_0 \frac{d\theta}{(2\pi)^n},$$

the solution reads

$$\dot{\varphi} = L_\alpha^{-1}(\dot{v}_0 + u_1 \cdot \dot{R}_0 - \dot{\lambda}_0 - \bar{\beta}),$$

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

with

$$|\dot{\varphi}|_{s-\sigma} \leq \frac{C}{\gamma} \frac{1}{\sigma^{n+\tau}} |\dot{v}_0 + u_1 \cdot \dot{R}_0 - \dot{\lambda}_0|_s.$$

Thirdly, the $\mathrm{Mat}_m(\mathbb{C})$-valued solution of (2.18) reads

$$\dot{R}_1 = (L_\alpha + [A, \cdot])^{-1}(\dot{\tilde{V}}_1 - \dot{\Lambda}_1 - \bar{B}),$$

having defined $\dot{\tilde{V}}_1 = \dot{V}_1 - \dot{R}_0' \cdot u_1 + 2U_2 \cdot \dot{R}_0$, and $\bar{B} = \prod_{\mathrm{Ker}[A,\cdot]} \int_{\mathbb{T}^n} \dot{\tilde{V}}_1 - \dot{\Lambda}_1 (d\theta/(2\pi)^n)$.

It now remains to handle the choice of $\delta\lambda$ such that

$$\bar{\lambda}(\theta, r) := (\bar{\beta}, \bar{b} + \bar{B} \cdot r) = 0.$$

Note that the defined vector field $\bar{\lambda}(\theta, r) = (\bar{\beta}, \bar{b} + \bar{B} \cdot r)$ lies in Λ (recall definition (1.2)). We define a map

$$F_g : \Lambda \to \Lambda, \quad \delta\lambda \mapsto -\bar{\lambda}$$

in the neighborhood of $\delta\lambda = 0$. It is affine and, when g is sufficiently close to the identity, invertible†.

Hence, there exists a unique $\delta\lambda$ such that $F_g(\delta\lambda) = 0$, satisfying

$$|\delta\lambda| \leq \frac{C}{\gamma\sigma^{\tau+n+1}} |\dot{v}|_{s+\sigma}.$$

We finally have

$$|\dot{g}|_{s-2\sigma} \leq \frac{C}{\gamma^2\sigma^{2(\tau+n)+1}} |\delta v|_{g,s+\sigma}.$$

By definition of $\dot{g}$, we have $\delta g = g' \cdot \dot{g}$. In order to uniquely determine δg we shall fix its constant term so to meet the conditions $\delta\varphi(0) = 0$, $\Pi_{\ker A} \delta R_0(0) = 0$ and $\Pi_{\ker[A,\cdot]} \delta R_1(0) = 0$ (recall Remark 2.1). Since g is close to the identity, these equations have a unique solution and similar estimates hold for δg:

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

We see that δu is actually well defined in $\overrightarrow{\mathcal{U}}_{s-3\sigma}$ and have

$$|\delta u|_{s-3\sigma} \leq \frac{C}{\gamma^2\sigma^{2(\tau+n)+3}} |\delta v|_{g,s+\sigma}.$$

Letting $\tau' = 2(\tau + n) + 3$ and $C' = C/\gamma^2$, up to defining $\sigma' = \sigma/4$ and $s' = s + \sigma$, the proposition is proved for all indexes s' and σ' with $s' < s' + \sigma'$. □

LEMMA 4. (Bounding ϕ'') *The bilinear map*

$$\phi''(x) : (T_g \mathcal{G}^{\sigma/n}_{s+\sigma} \times \overrightarrow{\mathcal{U}}_{s+\sigma} \times \Lambda)^{\otimes 2} \to \mathcal{V}_s,$$

where $x = (g, u, \lambda)$*, satisfies the estimate*

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

C'' *being a constant depending on* $|x|_s$.

† More specifically, the system in Λ that solves $\bar{\lambda} = 0$ is a linear system of N equations in N unknowns $(\delta\beta, \delta b, \delta B)$, with diagonal close to 1 if g is close to the identity.

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

Proof. For simplicity write $x = (g, u, \lambda)$ and $\delta x = (\delta g, \delta u, \delta\lambda)$. Recall that $\phi'(x) \cdot \delta x = [g_* u, \delta g \circ g^{-1}] + g_* \delta u + \delta\lambda$. Differentiating again with respect to x yields

$$[[g_* u, \delta g \circ g^{-1}] + g_* \delta u, \delta g \circ g^{-1}] - [g_* u, \delta g' \circ g^{-1} \cdot \delta g^{-1}] + [g_* \delta u, \delta g \circ g^{-1}].$$

Since $\delta g^{-1} = -(g'^{-1} \cdot \delta g) \circ g^{-1}$,

$$g^* \phi''(x) \cdot \delta x^{\otimes 2} = 2[\delta u, \dot{g}] + [[u, \dot{g}], \dot{g}] + [u, g^*(\delta g' \cdot g'^{-1} \cdot \delta g) \circ g^{-1}],$$

where the last term simplifies to

$$[u, g'^{-1} \cdot (\delta g' \cdot g'^{-1} \cdot \delta g)].$$

The desired bound follows from repeatedly applying Cauchy's inequality, the triangular inequality and Lemma 15. □

2.8. *Proof of Moser's theorem.* Proposition 2.1 and Lemma 4 guarantee to apply Theorem A.1, which provides the existence of (g, u, λ) such that $g_* u + \lambda = v$. Uniqueness and smooth differentiation follow from Propositions A.1–A.3, once $|v - u^0|_{s+\sigma}$ satisfies the required bound. The only brick it remains to add is the log-convexity of the weighted norm. Let $x \in E_s$. To prove that $s \mapsto \log |x|_s$ is convex one can easily show that

$$|x|_s \leq |x|_{s_1}^{1-\mu} |x|_{s_0}^{\mu}, \quad \mu \in [0, 1] \quad \text{for all } s = (1 - \mu)s_1 + s_0 \mu$$

by Hölder's inequality with conjugates $1 - \mu$ and μ, with the counting measure on $\mathbb{Z}^n$, observing that $|x|_s$ coincides with the ℓ^1-norm of the sequence $(|x_k| e^{|k|s})$. Theorem 2.1 follows, hence Theorem 1.1.

3. *Hamiltonian systems and Herman's twisted conjugacy theorem*

The Hamiltonian analogue of Moser's theorem was presented by Michael Herman at a colloquium held in Lyon in 1990. It is also an extension of Arnold's normal-form theorem for vector fields on $\mathbb{T}^n$ (see [**2**]).

In what follows we rely on the formalism developed by Féjoz in his papers [**10**, **12**, **13**]. This framework will be also used in §4, for generalizing Herman's result. Vector fields will be defined on $\mathbb{T}^n \times \mathbb{R}^n$. As always, the standard identification $\mathbb{R}^{n*} \equiv \mathbb{R}^n$ will be used.

3.1. *Spaces of vector fields.* Let $\mathcal{H}$ be the space of germs of real analytic Hamiltonians defined on some neighborhood of $\mathrm{T}_0^n = \mathbb{T}^n \times \{0\} \subset \mathbb{T}^n \times \mathbb{R}^n$, and $\mathcal{V}^{\mathrm{Ham}}$ the corresponding set of germs along T_0^n of real analytic Hamiltonian vector fields.

In this and the following sections we will only need to consider the standard Diophantine condition (2.3), for some $\gamma, \tau > 0$.

Fixing $\alpha \in \mathcal{D}_{\gamma,\tau} \subset \mathbb{R}^n$, consider the following affine subspace of $\mathcal{H}$:

$$\mathcal{K}^\alpha = \{K \in \mathcal{H} : K(\theta, r) = c + \alpha \cdot r + O(r^2), \ c \in \mathbb{R}\}.$$

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

$\mathcal{K}^\alpha$ is the set of Hamiltonians K for which T^n_0 is invariant by the flow of u^K and α-quasi-periodic:

$$u^{\mathrm{K}} = \begin{cases} \dot\theta = \dfrac{\partial K}{\partial r}(\theta, r) = \alpha + O(r), \\ \dot r = -\dfrac{\partial K}{\partial \theta}(\theta, r) = O(r^2). \end{cases} \tag{3.1}$$

We define

$$\mathcal{U}^{\mathrm{Ham}}(\alpha, 0) = \{u^{\mathrm{K}} \in \mathcal{V}^{\mathrm{Ham}} : K \in \mathcal{K}^\alpha\}$$

and introduce the set of counter terms

$$\Lambda^{\mathrm{Ham}} = \{\lambda \in \mathcal{V}^{\mathrm{Ham}} : \lambda(\theta, r) = (\beta, 0)\} \equiv \mathbb{R}^n.$$

We define the complex extension of width s of $\mathbb{T}^n \times \mathbb{R}^n$ as in §2.1, and denote by $\mathcal{H}_s = \mathcal{A}(\mathbb{T}^n_s)$ the space of Hamiltonians defined on this extension. $\mathcal{K}^\alpha_s$ is the affine subspace consisting of those $K \in \mathcal{H}_s$ of the form $K(\theta, r) = c + \alpha \cdot r + O(r^2)$.

3.2. *Spaces of conjugacies.* As introduced in §2.3, let $\mathcal{D}^\sigma_s$ be the space of real holomorphic invertible maps $\varphi = \mathrm{id} + v : \mathbb{T}^n_s \to \mathbb{T}^n_{\mathbb{C}}$, fixing the origin such that

$$|v|_s = \max_{1 \le j \le n} (|v_j|_s) < \sigma.$$

We consider the contragredient action of $\mathcal{D}^\sigma_s$ on T^n_s, with values in $T^n_{\mathbb{C}}$:

$$\varphi(\theta, r) = (\varphi(\theta),\ {}^t\varphi'^{-1}(\theta) \cdot r).$$

This is intended to linearize the dynamics on the tori.

Let $\mathcal{B}^\sigma_s$ be the space of exact 1-forms $\rho(\theta) = dS(\theta)$ on $\mathbb{T}^n_s$ (where S is a map $\mathbb{T}^n_s \to \mathbb{C}$, vanishing at the origin) such that

$$|\rho|_s = \max_{1 \le j \le n} (|\rho_j|_s) < \sigma;$$

we hence consider the space $\mathcal{G}^{\mathrm{Ham},\sigma}_s = \mathcal{D}^\sigma_s \times \mathcal{B}^\sigma_s$ of those Hamiltonian transformations $g = (\varphi, \rho)$ acting this way,

$$g(\theta, r) = (\varphi(\theta),\ {}^t\varphi'^{-1}(\theta) \cdot (r + \rho(\theta))),$$

that is identified, locally in the neighborhood of the identity, with an open set of the affine space passing through the identity and directed by $\{(\varphi - \mathrm{id}), S\}$. The form $\rho = dS$ being exact, it does not change the cohomology class of the torus†.

Let χ_s be the space of vector fields on $\mathbb{T}^n_s$ which fix the origin. The tangent space at the identity of $\mathcal{G}^{\mathrm{Ham}}_s$, $T_{\mathrm{id}}\mathcal{G}^{\mathrm{Ham}}_s = \chi_s \times \mathcal{B}_s$, is endowed with the norm

$$|\dot g|_s = \max(|\dot\varphi|_s, |\dot\rho|_s).$$

THEOREM 3.1. (Herman) *Let $\alpha \in \mathcal{D}_{\gamma,\tau}$ and $K^0 \in \mathcal{K}^\alpha_{s+\sigma}$. If $H \in \mathcal{H}_{s+\sigma}$ is close enough to K^0, there exists a unique $(g, K, \beta) \in \mathcal{G}^{\mathrm{Ham}}_s \times \mathcal{K}^\alpha_s \times \Lambda^{\mathrm{Ham}}$ close to* $(\mathrm{id}, K^0, 0)$ *such that*

$$H = K \circ g + \beta \cdot r.$$

Moreover, the normal form is Whitney smooth with respect to α.

† In this work we indicate derivations sometimes by ‘ $'$ ’, ‘d’ or ‘D’ to avoid cumbersome notation.

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

Here, too, the presence of $\beta \cdot r$ breaks the dynamical conjugacy between H and K: the orbits of $K \in \mathcal{K}^\alpha$, under the action of diffeomorphisms in $\mathcal{G}^{\mathrm{Ham}}$, form a subspace of co-dimension n.

For a proof of this result, known also as 'twisted conjugacy theorem', see [**11**, **16**], and see [**10**] for an analogue in the context of Hamiltonians with both tangent and normal frequencies.

We now phrase the theorem in terms of vector fields.

THEOREM 3.2. (Herman) *Let $\alpha \in \mathcal{D}_{\gamma,\tau}$ and $u^{K^0} \in \mathcal{U}^{\mathrm{Ham}}_{s+\sigma}(\alpha, 0)$. If $v^{\mathrm{H}} \in \mathcal{V}^{\mathrm{Ham}}_{s+\sigma}$ is close enough to u^{K^0}, there exists a unique $(g, u^K, \beta) \in \mathcal{G}^{\mathrm{Ham}}_s \times \mathcal{U}^{\mathrm{Ham}}_s(\alpha, 0) \times \Lambda^{\mathrm{Ham}}$, close to* $(\mathrm{id}, u^{K_0}, 0)$ *such that*

$$g_* u^K + \beta\, \partial_\theta = v^{\mathrm{H}}.$$

4. *Hamiltonian dissipative systems: Generalization of Herman's theorem and translated tori à la Rüssmann*

4.1. *A generalization of Herman's theorem.*

Here we generalize Herman's theorem to a particular class of dissipative vector fields.

4.1.1. *Spaces of vector fields.*

Let $\mathcal{H}_s = \mathcal{A}(\mathrm{T}^n_s)$ and $\mathcal{V}^{\mathrm{Ham}}_s$ be the space of Hamiltonian vector fields corresponding to Hamiltonians $H \in \mathcal{H}_s$. For any $\eta \in \mathbb{R}$, let us extend $\mathcal{V}^{\mathrm{Ham}}_s$ as $(\mathcal{V}^{\mathrm{Ham}} \oplus (-\eta r \partial_r))_s$. The corresponding affine subspace becomes

$$\mathcal{U}^{\mathrm{Ham}}_s(\alpha, -\eta) = \{u \in (\mathcal{V}^{\mathrm{Ham}} \oplus (-\eta r \partial_r))_s : u(\theta, r) = (\alpha + O(r), -\eta r + O(r^2))\}.\dagger$$

When $\eta > 0$ (respectively, $\eta < 0$) the invariant quasi-periodic torus T^n_0 of u is η-normally attractive (respectively, repulsive).

The class $\mathcal{V}^{\mathrm{Ham}} \oplus (-\eta r \partial_r)$, $\eta \in \mathbb{R}$, is mathematically peculiar: it is invariant under the Hamiltonian transformations in $\mathcal{G}^{\mathrm{Ham}}$. Physically, when $\eta \neq 0$ the system described undergoes a constant linear friction (respectively, amplification) which is the same in every direction.

According to Theorem B.1 and Corollary B.1, the operators

$$\begin{aligned} \phi : \mathcal{G}^{\mathrm{Ham},\sigma^2/2n}_{s+\sigma} \times \mathcal{U}^{\mathrm{Ham}}_{s+\sigma}(\alpha, -\eta) \times \Lambda^{\mathrm{Ham}} &\to (\mathcal{V}^{\mathrm{Ham}} \oplus (-\eta r \partial_r))_s, \\ (g, u, \beta) &\mapsto g_* u + \beta \partial_\theta, \end{aligned} \tag{4.1}$$

commuting with inclusions, are well defined.

THEOREM 4.1. ('Dissipative Herman') *Fix $\eta_0 > 0$ and $\alpha \in \mathcal{D}_{\gamma,\tau}$. There exists $\varepsilon > 0$ such that for any $\eta \in [-\eta_0, \eta_0]$, letting $u^0 \in \mathcal{U}^{\mathrm{Ham}}_{s+\sigma}(\alpha, -\eta)$, if $v \in (\mathcal{V}^{\mathrm{Ham}} \oplus (-\eta r \partial_r))_{s+\sigma}$ is such that $|v - u^0|_{s+\sigma} < \varepsilon$, there exists a unique triplet $(g, u, \beta) \in \mathcal{G}^{\mathrm{Ham}}_s \times \mathcal{U}^{\mathrm{Ham}}_s(\alpha, -\eta) \times \Lambda^{\mathrm{Ham}}$, close to* $(\mathrm{id}, u^0, 0)$, *such that*

$$g_* u + \beta \partial_\theta = v.$$

The case $\eta = 0$ corresponds to Herman's theorem.

The key point for such a result to be true relies on the following two technical lemmas.

† We recall that the notation $r \partial_r$ is short for $\sum_j^n r_j \partial_{r_j}$.

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

LEMMA 5. *If $g \in \mathcal{G}^{\mathrm{Ham}}$ and $v \in \mathcal{V}^{\mathrm{Ham}} \oplus (-\eta r \partial_r)$, the vector field $g_* v$ is given by*

$$g_* v = \begin{cases} \dot{\Theta} = \dfrac{\partial \hat{H}}{\partial R}, \\ \dot{R} = -\dfrac{\partial \hat{H}}{\partial \Theta} - \eta R, \end{cases} \tag{4.2}$$

where

$$\hat{H}(\Theta, R) = H \circ g^{-1}(\Theta, R) - \eta (S \circ \varphi^{-1}(\Theta)).$$

The fact that $\eta \in \mathbb{R}$ is fundamental to maintain the Hamiltonian structure, which would be broken even if η were a diagonal matrix. Geometrically, the action of g on H is 'twisted' by the dissipation.

Proof. $g(\theta, r) = (\Theta, R)$, that is,

$$\begin{cases} \Theta = \varphi(\theta), \\ R = {}^t\varphi'^{-1}(\theta) \cdot (r + dS(\theta)). \end{cases}$$

We have, in the tangent direction,

$$\dot{\Theta} = \varphi'(\theta) \cdot \dot{\theta} = \frac{\partial (H \circ g^{-1})}{\partial R}.$$

The derivation of $\dot{R}$ requires a little more attention:

$$\begin{aligned} \dot{R} = {} & \underbrace{({}^t\varphi'^{-1}(\theta))' \cdot r \cdot \dot{\theta}}_{A} + \underbrace{{}^t\varphi'^{-1}(\theta) \cdot \dot{r}}_{B} + \underbrace{{}^t\varphi'^{-1}(\theta) \cdot D^2 S(\theta) \cdot \dot{\theta}}_{C} \\ & + \underbrace{({}^t\varphi'^{-1}(\theta))' \cdot dS(\theta) \cdot \dot{\theta}}_{D}, \end{aligned}$$

where, expanding and composing with g^{-1},

$$\begin{aligned} A &= (-{}^t\varphi'^{-1} \cdot {}^t\varphi'' \cdot {}^t\varphi'^{-1}) \circ \varphi^{-1}(\Theta) \cdot ({}^t\varphi' \circ \varphi^{-1}(\Theta) \cdot R - dS \circ \varphi^{-1}(\Theta)) \cdot \frac{\partial H}{\partial r}, \\ B &= -{}^t\varphi'^{-1} \circ \varphi^{-1}(\Theta) \cdot \frac{\partial H}{\partial \theta} - \eta R + \eta\, {}^t\varphi'^{-1} \circ \varphi^{-1}(\Theta) \cdot dS \circ \varphi^{-1}(\Theta), \\ C &= {}^t\varphi'^{-1}(\theta) \cdot D^2 S(\theta) \cdot \frac{\partial H}{\partial r} \\ &= {}^t\varphi'^{-1} \circ \varphi^{-1}(\Theta) \cdot D^2 S \circ \varphi^{-1}(\Theta) \cdot \frac{\partial H}{\partial r}, \\ D &= -({}^t\varphi'^{-1} \cdot {}^t\varphi'' \cdot {}^t\varphi'^{-1}) \circ \varphi^{-1}(\Theta) \cdot dS \circ \varphi^{-1}(\Theta) \cdot \frac{\partial H}{\partial r}. \end{aligned}$$

Remark that if

$$H \circ g^{-1}(\Theta, R) = H(\varphi^{-1}(\Theta),\ {}^t\varphi' \circ \varphi^{-1}(\Theta) \cdot R - dS \circ \varphi^{-1}(\Theta)),$$

we have

$$\begin{aligned} \frac{\partial H}{\partial \Theta} = {} & \frac{\partial H}{\partial \theta} \cdot \varphi'^{-1} \circ \varphi^{-1}(\Theta) \\ & + \frac{\partial H}{\partial r} \cdot [\,{}^t\varphi'' \circ \varphi^{-1}(\Theta) \cdot \varphi'^{-1} \circ \varphi^{-1}(\Theta) \cdot R - D^2 S \circ \varphi^{-1}(\Theta) \cdot \varphi'^{-1} \circ \varphi^{-1}(\Theta)]. \end{aligned}$$

Summing terms, we get

$$\dot{R} = -\frac{\partial H \circ g^{-1}}{\partial \Theta} - \eta R + \eta({}^t\varphi'^{-1} \circ \varphi^{-1}(\Theta) \cdot dS \circ \varphi^{-1}(\Theta)).$$

Introducing the modified Hamiltonian $\hat{H}$ as in the statement, the transformed system has the claimed form (4.2). □

The same is true for the pull-back of such a v.

LEMMA 6. *If $g \in \mathcal{G}^{\mathrm{Ham}}$ and $v \in \mathcal{V}^{\mathrm{Ham}} \oplus (-\eta r \partial_r)$, the vector field $g^* v = g_*^{-1} v$ is given by*

$$g^* v = \begin{cases} \dot{\theta} = \dfrac{\partial \hat{H}}{\partial r}, \\ \dot{r} = -\dfrac{\partial \hat{H}}{\partial \theta} - \eta r, \end{cases} \tag{4.3}$$

$\hat{H}$ being $\hat{H}(\theta, r) = H \circ g\,(\theta, r) + \eta S(\theta)$.

4.1.2. *The linearized problem.* Theorem 4.1 will follow—again—from the inverse function Theorem A.1, once we check the existence of a right (and left) inverse† for

$$\phi' : T_g \mathcal{G}^{\mathrm{Ham}}_{s+\sigma} \times \overrightarrow{\mathcal{U}}^{\mathrm{Ham}}_{s+\sigma} \times \Lambda^{\mathrm{Ham}} \to \mathcal{V}^{\mathrm{Ham}}_{g,s},$$

when g is close to the identity, and bounds on it and ϕ''.

Except for a minor difference, the system that solves the linearized problem is the same as the one in the purely Hamiltonian context.

PROPOSITION 4.1. *If g is close enough to the identity, for every δv in $\mathcal{V}^{\mathrm{Ham}}_{g,s+\sigma}$ there exists a unique triplet $(\delta g, \delta u, \delta\beta)$ in $T_g \mathcal{G}^{\mathrm{Ham}}_s \times \overrightarrow{\mathcal{U}}^{\mathrm{Ham}}_s \times \Lambda^{\mathrm{Ham}}$ such that*

$$\phi'(g, u, \beta) \cdot (\delta g, \delta u, \delta\beta) = \delta v. \tag{4.4}$$

Moreover, there exist τ', $C > 0$ such that

$$\max\,(|\delta g|_s, |\delta u|_s, |\delta\beta|) \leq \frac{C}{\sigma^{\tau'}} |\delta v|_{g,s+\sigma}, \tag{4.5}$$

where C is a constant that depends only on n, τ and $(|g - \mathrm{id}|_{s+\sigma}, |u - (\alpha, -\eta r)|_{s+\sigma})$.

Proof. The proof is recovered from that of Proposition 2.1, with the additional requirement that the transformation is Hamiltonian and that the vector fields belong to this particular class 'Hamiltonian + dissipation'.

Calculating $\phi'(x) \cdot \delta x$ and pulling back equation (4.4), we get

$$[u, \dot{g}] + \delta u = \dot{v} - \dot{\lambda},$$

where we used the same notation as in (2.13) for $\dot{g} = g'^{-1} \cdot \delta g$, $\dot{v} = g^* \delta v$ and $\dot{\lambda} = g^* \delta\lambda$. Here $\dot{g}$ has the form $\dot{g} = (\dot{\varphi}, -r \cdot \dot{\varphi}' + \dot{\rho})$, where $\dot{\varphi} \in \chi_s$ and $\dot{\rho} = d\dot{S} \in \mathcal{B}_s$ (see §3.2 where these spaces were defined). By Lemma 6, $\dot{v}$ is a Hamiltonian vector field too.

† As in Proposition 2.1, we denote the tangent space of $(\mathcal{V}^{\mathrm{Ham}} \oplus (\eta r \partial_r))_s$ (which coincides with $\mathcal{V}^{\mathrm{Ham}}_s$) by $\mathcal{V}^{\mathrm{Ham}}_{g,s}$, the subscript $*_g$ indicating the presence of the g-dependent norm.

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

Identifying terms of the equation in the first-order jet $j^{0,1}\dot{v} = (\dot{v}_0^H, \dot{V}_0^H + \dot{V}_1^H \cdot r)$, the system corresponding to (2.16)–(2.18) translates into

$$\begin{aligned} \dot{\varphi}' \cdot \alpha - u_1 \cdot d\dot{S} &= \dot{v}_0^H - \dot{\lambda}_0, \\ d\dot{S}' \cdot \alpha + \eta\, d\dot{S} &= \dot{V}_0^H - \dot{\Lambda}_0, \\ -{}^t D\dot{\varphi}' \cdot \alpha + {}^t D(u_1 \cdot d\dot{S}) &= \dot{V}_1^H - \dot{\Lambda}_1, \end{aligned}$$

where $\dot{\lambda}_0 = \varphi'^{-1} \cdot \delta\beta$, $\dot{\Lambda}_0 = -D({}^t\varphi'^{-1} \cdot dS(\theta)) \cdot \delta\beta$ and $\dot{\Lambda}_1 = -{}^t D\dot{\lambda}_0$, while u_1 is the coefficient of the linear term of u in the $\dot{\theta}$-direction (where we denote by D the derivative with respect to θ). In particular, $\dot{V}_0^H$ and $\dot{V}_1^H$ are of zero average and, according to the symmetry of a Hamiltonian system, $\dot{V}_1^H = -{}^t D\dot{v}_0^H$ so the first two equations determine the whole systems.

Coherently, the term $\dot{\Lambda}_0$ also has zero average and the $d\dot{S}$-equation can readily be solved.

Remark 4.1. Since

$$d\dot{S}(\theta) = 0 + \sum_{k\neq 0} \frac{\dot{V}_{0,k}^H}{i\, k \cdot \alpha + \eta}\, e^{i\, k\cdot\theta}$$

has zero average, we can bound the divisors uniformly with respect to η, by $|i\, k \cdot \alpha + \eta| > |i\, k \cdot \alpha|$. Thus only the standard Diophantine condition (2.3) on α is required. As a consequence, the bound $|v - u^0|_{s+\sigma} < \varepsilon$ entailed in Theorem A.1 holds uniformly with respect to η. This is fundamental for the results in the last section.

Solutions and inequalities follow readily from Lemmas 1–2 and Cauchy's inequality. □

Remark 4.2. When $\eta = 0$ the system above is the one that solves the infinitesimal problem of the 'twisted conjugacy' theorem presented in [**11**, §1.1]. Hence, up to the slight difference in the equation determining $d\dot{S}$, the proof of Theorem 4.1 follows the same steps and difficulties as in [**11**] (application of Theorem A.1 in the framework of Remark A.1 and Propositions A.1–A.3).

4.1.3. *A first portrait.* If the eigenvalues a_i of A are all distinct and different from zero, it is immediate to see that the external parameters are of the form $\lambda = (\beta, B \cdot r)$, with B a diagonal matrix as well (recall definition (1.2) in relation to Lemma 3).

COROLLARY 4.1. (Of Moser's theorem) *Let $A \in \mathrm{Mat}_m(\mathbb{R})$ be diagonalizable with simple, non-zero eigenvalues and let (α, a) satisfy the Diophantine condition (1.4). If v is sufficiently close to $u^0 \in \mathcal{U}(\alpha, A)$, there exists a unique $(g, u, \lambda) \in \mathcal{G} \times \mathcal{U}(\alpha, A) \times \Lambda(\beta, B \cdot r)$, close to* $(\mathrm{id}, u^0, 0)$*, such that*

$$g_* u + \lambda = v,$$

λ being of the form $\lambda = (\beta, \mathrm{diag}\, B \cdot r)$.

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

Here is a diagram that summarizes our results, from the most general to the purely Hamiltonian. We emphasize the required counter terms in the Λ notation.

$$\begin{array}{ll}
\text{Moser:} & \mathcal{G} \times \mathcal{U}(\alpha, A) \times \Lambda(\beta, b + B \cdot r) \xrightarrow[\simeq \text{ loc.}]{} \mathcal{V} \\[2em]
\text{General dissip. (diag } A\text{):} & \mathcal{G} \times \mathcal{U}(\alpha, A) \times \Lambda(\beta, \operatorname{diag} B \cdot r) \xrightarrow[\simeq \text{ loc.}]{} \mathcal{V} \\
& \qquad\qquad \uparrow \qquad\qquad\qquad\qquad\qquad\qquad \uparrow \\
\text{Herman dissip.:} & \mathcal{G}^{\text{Ham}} \times \mathcal{U}^{\text{Ham}}(\alpha, \eta) \times \Lambda(\beta, 0) \xrightarrow[\simeq \text{ loc.}]{} \mathcal{V}^{\text{Ham}} \oplus (\eta r\, \partial_r) \\[2em]
\text{Herman } (\eta = 0)\text{:} & \mathcal{G}^{\text{Ham}} \times \mathcal{U}^{\text{Ham}}(\alpha, 0) \times \Lambda(\beta, 0) \xrightarrow[\simeq \text{ loc.}]{} \mathcal{V}^{\text{Ham}}
\end{array}$$

4.2. *Normal form 'à la Rüssmann'.* In the context of the diffeomorphisms of the cylinder $\mathbb{T} \times \mathbb{R}$, Rüssmann proved a result that admits large applications in the study of dynamical systems: the 'theorem of the translated curve' (see [**3**, **17**, **23**, **29**], for instance).

We give here an extension of this theorem to vector fields.

If the Hamiltonians considered in §4.1 are non-degenerate (see below for a formal definition), we can define a 'hybrid normal form' that relies on both the peculiar structure of the vector fields and this torsion property; this makes unnecessary the introduction of *all* the counter terms *a priori* needed if we attacked the problem in the pure spirit of Moser.

4.2.1. *Twisted vector fields.* The starting context is that of §4.1 and notation is the same.

Let $\alpha \in \mathbb{R}^n$. We are interested in those $K \in \mathcal{K}^\alpha$ of the form

$$K(\theta, r) = c + \alpha \cdot r + \tfrac{1}{2} Q(\theta) \cdot r^2 + O(r^3), \tag{4.6}$$

where Q is a non-degenerate quadratic form on $\mathbb{T}^n_s$: $\det\, (1/(2\pi)^n) \int_{\mathbb{T}^n} Q(\theta)\, d\theta \neq 0$.

There exist s_0 and ε_0 such that, for all $s > s_0$, $K^0 \in \mathcal{H}_s$, and for all $H \in \mathcal{H}_s$ such that $|H - K^0|_{s_0} < \varepsilon_0$,

$$\left| \det \int_{\mathbb{T}^n} \frac{\partial^2 H}{\partial r^2}(\theta, 0)\, \frac{d\theta}{(2\pi)^n} \right| \geq \frac{1}{2} \left| \det \int_{\mathbb{T}^n} \frac{\partial^2 K^0}{\partial r^2}(\theta, 0)\, \frac{d\theta}{(2\pi)^n} \right| \neq 0.$$

From now on, we assume that $s \geq s_0$ and define

$$\mathcal{K}^\alpha_s = \{K \in \mathcal{K}^\alpha_s : |K - K^0|_{s_0} \leq \varepsilon_0\}.$$

Hence we consider the corresponding space of vector fields with twist and call it

$$\widehat{\mathcal{U}}^{\text{Ham}}_s(\alpha, 0) = \{u^K(\theta, r) = (\alpha + \tfrac{1}{2} Q(\theta) \cdot r + O(r^2),\, O(r^2))\}, \tag{4.7}$$

affine subspace of $\widehat{\mathcal{V}}^{\text{Ham}}_s = \{v^{\text{H}} \in \mathcal{V}^{\text{Ham}}_s : |H - K^0|_{s^0} \leq \varepsilon_0\}$.

Now, let $\eta \in \mathbb{R}$ and consider the extended spaces

$$\widehat{\mathcal{U}}^{\text{Ham}}_s(\alpha, -\eta) \quad \text{and} \quad (\widehat{\mathcal{V}}^{\text{Ham}} \oplus (-\eta r + \eta \mathbb{R}^n)\partial_r)_s. \tag{4.8}$$

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

Remark 4.3. We enlarged the target space with the translations in actions

$$\zeta \mapsto v^{\mathrm{H}} \oplus (-\eta r + \eta\zeta)\partial_r$$

in order to handle symplectic transformations and guarantee that the normal-form operator is well defined (see below). Note that the constant η multiplying ζ is not essential; it just eases the notation in calculations and makes the results below ready to use in the application presented in §6.

As in the previous section, $\mathcal{D}_s^\sigma$ is the space of holomorphic invertible maps $\varphi = \mathrm{id} + v : \mathbb{T}_s^n \to \mathbb{T}_{\mathbb{C}}^n$, fixing the origin with $|v|_s < \sigma$.

Let $\mathcal{Z}_s^\sigma$ be the space of closed 1-forms on $\mathbb{T}_s^n$ $\rho(\theta) = dS(\theta) + \xi$, where $S : \mathbb{T}_s^n \to \mathbb{C}$ and $\xi \in \mathbb{R}^n$ are uniquely determined assuming $S(0) = 0$, such that

$$|\rho|_s := \max(|\xi|, |dS|_s) < \sigma.$$

Let $\mathcal{G}_s^{\omega,\sigma} = \mathcal{D}_s^\sigma \times \mathcal{Z}_s^\sigma$, whose elements $g = (\varphi, \rho)$ define symplectic transformations

$$g(\theta, r) = (\varphi(\theta),\ {}^t\varphi'^{-1}(\theta) \cdot (r + dS(\theta) + \xi)). \tag{4.9}$$

The tangent space at the identity $T_{\mathrm{id}}\mathcal{G}_s^\omega = \chi_s \times \mathcal{Z}_s$ is endowed with the norm

$$|\dot g|_s = \max(|\dot\varphi|_s, |\dot\rho|_s).$$

Concerning the space of constant counter terms, we define the space of translations in action as

$$\Lambda_b = \{\lambda = (0, b), b \in \mathbb{R}^n\}.$$

According to the following lemmas and Corollary B.1, the normal-form operators

$$\begin{aligned} \phi : \mathcal{G}_{s+\sigma}^{\omega,\sigma^2/2n} \times \widehat{\mathcal{U}}_{s+\sigma}^{\mathrm{Ham}}(\alpha, -\eta) \times \Lambda_b &\to (\widehat{\mathcal{V}}^{\mathrm{Ham}} \oplus (-\eta r + \eta\mathbb{R}^n)\partial_r)_s, \\ (g, u, b) &\mapsto g_* u + b \end{aligned} \tag{4.10}$$

are well defined.

LEMMA 7. *If $g \in \mathcal{G}^\omega$ and $v \in \mathcal{V}^{\mathrm{Ham}} \oplus (-\eta r + \eta\zeta)\partial_r$, the push-forward $g_* v$ is given by*

$$g_* v = \begin{cases} \dot\Theta = \dfrac{\partial \hat H}{\partial R}, \\ \dot R = -\dfrac{\partial \hat H}{\partial \Theta} - \eta(R - \hat\zeta), \quad \hat\zeta = \zeta + \xi, \end{cases}$$

where $\hat H(\Theta, R) = H \circ g^{-1}(\Theta, R) - \eta(S \circ \varphi^{-1}(\Theta) + \hat\zeta \cdot (\varphi^{-1}(\Theta) - \Theta))$.

The proof is the same as for Lemma 5, taking care of the additional term $\eta\, {}^t\varphi'^{-1} \circ \varphi^{-1} \cdot (\xi + \zeta)$ coming from the non-exactness of $\rho(\theta)$ and the translation ζ.

LEMMA 8. *If $g \in \mathcal{G}^\omega$ and $v \in \mathcal{V}^{\mathrm{Ham}} \oplus (-\eta r + \eta\zeta)\partial_r$, the pull-back $g^* v$ is given by*

$$g^* v = \begin{cases} \dot\theta = \dfrac{\partial \hat H}{\partial r}, \\ \dot r = -\dfrac{\partial \hat H}{\partial \theta} - \eta(r - \hat\zeta), \quad \hat\zeta = \zeta - \xi, \end{cases} \tag{4.11}$$

where $\hat H(\theta, r) = H \circ g(\theta, r) + \eta(S(\theta) - \zeta \cdot (\varphi(\theta) - \theta))$.

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

If $v \in \mathcal{V}^{\mathrm{Ham}} \oplus (\eta\zeta)\partial_r$,

$$g^* v = \begin{cases} \dot\theta = \dfrac{\partial \hat H}{\partial r}, \\ \dot r = -\dfrac{\partial \hat H}{\partial \theta} + \eta\zeta, \end{cases}$$

where $\hat H(\theta, r) = H \circ g(\theta, r) + \eta\zeta \cdot (\varphi(\theta) - \theta)$.

The proofs of these results are immediate from the definition of g and follow that of Lemma 5.

THEOREM 4.2. (Translated torus) *Fix* $\eta_0 > 0$ *and* $\alpha \in \mathcal{D}_{\gamma,\tau}$. *There exists* $\varepsilon > 0$ *such that for any* $\eta \in [-\eta_0, \eta_0]$, *letting* $u^0 \in \widehat{\mathcal{U}}^{\mathrm{Ham}}_{s+\sigma}(\alpha, -\eta)$ *if* $v \in (\widehat{\mathcal{V}}^{\mathrm{Ham}} \oplus (-\eta r + \eta\mathbb{R})\partial_r)_{s+\sigma}$ *is such that* $|v - u^0|_{s+\sigma} < \varepsilon$, *there exists a unique* $(g, u, b) \in \mathcal{G}^{\omega}_s \times \mathcal{U}^{\mathrm{Ham}}_s(\alpha, -\eta) \times \Lambda_b$, *close to* $(\mathrm{id}, u^0, 0)$, *such that*

$$g_* u + b\, \partial_r = v.$$

From the normal form, the image $g(\mathrm{T}^n_0)$ is not invariant by the flow of v, but translated in the action direction during each infinitesimal time interval.

The proof can still be recovered from Theorem A.1 (in the framework of Remark A.1) and Propositions A.1–A.3.

Proof. The main part involves checking the invertibility of ϕ'. Let

$$\phi : \mathcal{G}^{\omega,\sigma^2/2n}_{s+\sigma} \times \widehat{\mathcal{U}}^{\mathrm{Ham}}_{s+\sigma}(\alpha, -\eta) \times \Lambda_b \to (\widehat{\mathcal{V}}^{\mathrm{Ham}} \oplus (-\eta r + \eta\mathbb{R}^n)\partial_r)_s,$$
$$(g, u, b) \mapsto g_* u + b = v.$$

We want to solve

$$\phi'(g, u, b) \cdot (\delta g, \delta u, \delta b) = [g_* u, \delta g \circ g^{-1}] + g_* \delta u + \delta b = \delta v, \tag{4.12}$$

for any $\delta v = \delta v^{\mathrm{H}} + \eta\delta\zeta \in (\widehat{\mathcal{V}}^{\mathrm{Ham}} \oplus (\eta\mathbb{R}^n)\partial_r)_{g,s+\sigma}$, when g is close to the identity. As in Proposition 2.1, we pull (4.12) back and expand the pulled-back vector fields along T^n_0.

Using the same notation as in (2.13)–(2.15), since g is of the form (4.9),

$$\dot g = g'^{-1} \cdot \delta g = (\dot\varphi, -{}^t\dot\varphi' \cdot r + d\dot S + \dot\xi),$$

where $\dot S \in \mathcal{A}(\mathbb{T}^n_s)$ and $\dot\varphi \in \chi_s$ fix the origin, and $\dot\xi \in \mathbb{R}^n$.

According to Lemma 8, the vector field $\dot v = g^* \delta v$ is given by a Hamiltonian vector field translated by $+\eta\delta\zeta\,\partial_r$ in the normal direction. Thus, the image by the jet $j^{0,1}$ of the pull-back of equation (4.12) splits into the equations

$$\dot\varphi' \cdot \alpha - Q(\theta) \cdot (d\dot S + \dot\xi) = \dot v^H_0, \tag{4.13}$$
$$d\dot S' \cdot \alpha + \eta(d\dot S + \dot\xi) = \dot V^H_0 + \eta\delta\zeta - \dot b, \tag{4.14}$$
$$-{}^t D\dot\varphi' \cdot \alpha + {}^t D(Q(\theta) \cdot (d\dot S + \dot\xi)) = \dot V^H_1, \tag{4.15}$$

where $\dot b$ is of the form ${}^t\varphi' \cdot \delta b = (\mathrm{id} + {}^t v') \cdot \delta b$ (remember that $\varphi = \mathrm{id} + v$). Furthermore, as in the Hamiltonian dissipative system in the proof of Proposition 4.1, $\dot V^H_1 = -{}^t D\dot v^H_0$ and the whole system of equations is satisfied once we determine $d\dot S$, $\dot\xi$, $\dot\varphi$ and $\dot b$ from (4.13)–(4.14).

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

We will now repeatedly apply Lemmas 1 and 2 and Cauchy's estimates. As before, we do not keep track of constants.

By averaging equation (4.14) on the torus, we determine

$$\delta b = \eta(\delta\zeta - \dot{\xi}),$$

and solve the average-free

$$d\dot{S}' \cdot \alpha + \eta\, d\dot{S} = \dot{V}_0^{\mathrm{H}} - {}^t v' \cdot \delta b.$$

Denoting $\dot{V}_0 = \dot{V}_0^{\mathrm{H}} - \eta^t v' \cdot \delta\zeta$, the solution can be written as

$$d\dot{S}(\theta) = \sum_{k\neq 0} \frac{\dot{V}_{0,k}}{i\, k \cdot \alpha + \eta} e^{i\, k\theta} + \eta M(\theta) \cdot \dot{\xi}, \tag{4.16}$$

where $M(\theta)$ is the matrix whose (i, j)th component reads $(\sum_{k\neq 0} ({}^t v'^{i}_{j,k}/(i\, k \cdot \alpha + \eta))\, e^{i\, k\cdot\theta})$. Note that the Fourier coefficients smoothly depend on $\eta \in [-\eta_0, \eta_0]$ and Remark 4.1 holds. Moreover, $|\eta M|_s \leq |\eta_0|\sigma^{-\tau'} C|v|_{s+\sigma}$, which will remain small in all the iterates not affecting the torsion term (see below).

Remark 4.4. In the proof we never use the fact that $\eta \neq 0$. Taking the limit $\eta \to 0$, the system (4.13)–(4.15) corresponds to the linearized problem of the classical Kolmogorov theorem: vector fields of both sides are entirely Hamiltonian (with twist), $\delta b = 0$, $d\dot{S} = L_\alpha^{-1} \dot{V}_0^{\mathrm{H}}$, and the torus persists (see [**11**–**13**]).

Let $S_0 = \sum_{k\neq 0} (\dot{V}_{0,k}/(i\, k \cdot \alpha + \eta))e^{i\, k\theta}$ (the first part of (4.16)). Averaging on the torus equation (4.13), by the torsion hypotheses we can determine

$$\dot{\xi} = -\left(\frac{1}{(2\pi)^n} \int_{\mathbb{T}} Q \cdot (\eta M + \mathrm{id})\, d\theta\right)^{-1} \cdot \left(\frac{1}{(2\pi)^n} \int_{\mathbb{T}} \dot{v}_0^{\mathrm{H}} + Q \cdot S_0\, d\theta\right), \tag{4.17}$$

with

$$|\dot{\xi}| \leq \frac{C}{\gamma\sigma^{\tau+n}} |\delta v|_{g,s+\sigma},$$

hence

$$|d\dot{S}|_s \leq \frac{C}{\gamma\sigma^{\tau+n}} |\delta v|_{g,s+\sigma} \quad \text{and} \quad |\delta b| \leq \frac{C}{\gamma\sigma^{\tau+n}} |\delta v|_{g,s+\sigma}.$$

It remains to solve equation (4.13); the average-free part determines $\delta\varphi$ with

$$|\dot{\varphi}|_{s-\sigma} \leq \frac{C}{\gamma^2\sigma^{2\tau+2n}} |\delta v|_{g,s+\sigma}.$$

The same sort of estimates hold for the desired δg:

$$|\delta g|_{s-2\sigma} \leq \frac{C}{\gamma^2\sigma^{2\tau+2n+1}} |\delta v|_{g,s+\sigma}.$$

Again, $[u, \dot{g}] + \delta u = \dot{v} - \dot{b}$ determines δu explicitly, and we have

$$|\delta u|_{s-2\sigma} \leq \frac{C}{\gamma^2\sigma^{2\tau+2n+2}} |\delta v|_{g,s+\sigma}.$$

Up to defining $\sigma' = \sigma/3$ and $s' = s + \sigma$, we have proved the following lemma for all s', σ' such that $s' < s' + \sigma'$.

LEMMA 9. *If g is close enough to the identity, for every δv in $(\widehat{\mathcal{V}}^{\mathrm{Ham}} \oplus (\eta\mathbb{R}^n)\partial_r)_{g,s+\sigma}$, there exists a unique triplet $(\delta g, \delta u, \delta b) \in T_g\mathcal{G}_s^{\omega} \times \overrightarrow{\widehat{\mathcal{U}}_s^{\mathrm{Ham}}} \times \Lambda_b$ such that*

$$\phi'(g, u, b) \cdot (\delta g, \delta u, \delta b) = \delta v.$$

Moreover, there exist $\tau', C' > 0$ such that

$$\max\left(|\delta g|_s, |\delta u|_s, |\delta b|\right) \leq \frac{C'}{\sigma^{\tau'}}|\delta v|_{g,s+\sigma},$$

where C' depends only on n, τ, $|g - \mathrm{id}|_{s+\sigma}$ and $|u - (\alpha, -\eta r)|_{s+\sigma}$.

Concerning the bound on ϕ'', the analogue of Lemma 4 readily follows.

It just remains to apply Theorem A.1 and Propositions A.1–A.3 and complete the proof for the chosen v in $(\widehat{\mathcal{V}}^{\mathrm{Ham}} \oplus (-\eta r + \eta\mathbb{R}^n)\partial_r)_{s+\sigma}$. □

We conclude the section with a second diagram.

$$\begin{array}{lrcl} \text{Moser:} & \mathcal{G} \times \mathcal{U}(\alpha, A) \times \Lambda(\beta, b + B \cdot r) & \xrightarrow[\simeq \text{ loc.}]{} & \mathcal{V} \\ & & & \uparrow \\ \text{‘}\textit{à la}\text{ Rüssmann’:} & \mathcal{G}^{\omega} \times \widehat{\mathcal{U}}^{\mathrm{Ham}}(\alpha, -\eta) \times \Lambda_b & \xrightarrow[\simeq \text{ loc.}]{} & \mathcal{V}^{\mathrm{Ham}} \oplus (-\eta r + \eta\mathbb{R}^n)\partial_r \end{array}$$

5. *Extension of Herman's and Rüssmann's theorems to simple normally hyperbolic tori*

The peculiarity of the normal forms proved in the previous section is that the translated (or twisted) α-quasi-periodic torus $g(\mathrm{T}_0^n)$ of the perturbed v keeps its η-normally attractive (respectively, repulsive, if $\eta < 0$) dynamics, the reason for such a result relying on the Hamiltonian nature of perturbations.

On $\mathbb{T}^n \times \mathbb{R}^m$, let $u \in \mathcal{U}(\alpha, A)$. We will say that T_0^n is *simple normally hyperbolic* if A has simple, real, non-zero eigenvalues.

Note that the space of matrices $A \in \mathrm{Mat}_m(\mathbb{R})$ with simple, non-zero real eigenvalues is open in $\mathrm{Mat}_m(\mathbb{R})$, thus it provides a consistent interesting set of frequencies to work on.

We show here that for general perturbations, at the expense of conjugating $v - \lambda$ to a vector field u with different (opportunely chosen) normally hyperbolic dynamics, we can show that a translated or twisted reducible α-quasi-periodic Diophantine torus exists (Theorems C and D stated in the introduction).

Notation is the same as in §2. Let $\Delta_m^{si}(\mathbb{R}) \subset \mathrm{Mat}_m(\mathbb{R})$ be the space of $m \times m$ matrices with simple, non-zero, real eigenvalues and let

$$\mathbf{U}_s = \bigcup_{A \in \Delta_m^{si}(\mathbb{R})} \mathcal{U}_s(\alpha, A) = \{u(\theta, r) = (\alpha + O(r), A \cdot r + O(r^2)),\ A \in \Delta_m^{si}(\mathbb{R})\}.$$

THEOREM 5.1. (Twisted torus) *For every $u^0 \in \mathcal{U}_{s+\sigma}(\alpha, A^0)$ with α Diophantine and $A^0 \in \Delta_m^{si}(\mathbb{R})$, there exist an $\varepsilon > 0$ and a germ of C^{∞}-maps*

$$\psi : \mathcal{V}_{s+\sigma} \to \mathcal{G}_s \times \mathbf{U}_s \times \Lambda_{\beta}, \quad v \mapsto (g, u, \beta),$$

at $u^0 \mapsto (\mathrm{id}, u^0, 0)$, such that, for all v satisfying $|v - u^0|_{s+\sigma} < \varepsilon$, $v = g_* u + \beta\,\partial_\theta$.

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

Proof. We denote by ϕ_A the operator ϕ, as now we want A to vary. Locally, in the neighborhood of (A^0, u^0), let us define the map

$$\hat{\psi} : \Delta_m^{si}(\mathbb{R}) \times \mathcal{V}_{s+\sigma} \to \mathcal{G}_s \times \mathbf{U}_s \times \Lambda, \quad \hat{\psi}_A(v) := \phi_A^{-1}(v) = (g, u, \lambda),$$

associating to any (A, v) the triplet given by Moser's theorem such that $g_* u + \lambda = v$, where $\lambda = (\beta, B \cdot r)$, with $\beta \in \mathbb{R}^n$ and $B \in \mathrm{Mat}_m(\mathbb{R})$ satisfying $[B, A] = 0$†. 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 $\Delta_m^{si}(\mathbb{R})$. 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, and let Δ_{A^0} be the subspace of diagonal matrices, namely the matrices which commute with A^0. Let us consider the restriction of $\hat{\psi}$ to Δ_{A^0} and write u^0 as

$$u^0 = (\alpha + O(r), A \cdot r + (A^0 - A) \cdot r + O(r^2)),$$

where $A \in \Delta_{A^0}$ is close to A^0.

Since

$$\phi_A(\mathrm{id}, u^0 + (0, (A - A^0) \cdot r), (0, (A^0 - A) \cdot r)) \equiv u^0,$$

locally for all A close to A^0 we have

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

by local uniqueness of the normal form, where $[B, A] = 0$. Remark that, since $A \in \Delta_{A^0}$ has simple spectrum, B is indeed in Δ_{A^0}. In particular,

$$\frac{\partial B}{\partial A} = -\mathrm{id},$$

hence $A \mapsto B(A)$ is a local diffeomorphism on Δ_{A^0}; thus by the implicit function theorem locally for all v there exists a unique $\bar{A}$ such that $B(\bar{A}, v) = 0$. Remark that since $\bar{A}$ is invertible, the unique counter term needed is in fact $\beta \in \mathbb{R}^n$. All that remains is to define $\psi(v) = \hat{\psi}(\bar{A}, v)$. □

Remark 5.1. The fact that A^0 has real eigenvalues makes the correction $A = A^0 - \delta A$ of A^0 (provided by the implicit function theorem) well defined. If we had considered possibly complex eigenvalues, submitted to Diophantine condition (1.4), the procedure would have been more delicate, using the Whitney dependence of ϕ in A. In this line of thought see [**10**] and the 'hypothetical conjugacy' theorem therein.

Remark 5.2. If we allow the possibility of having an eigenvalue equal to zero, the torus would be twisted-translated, due to the presence of $b \in \ker A \equiv \mathbb{R}$, providing a generalization in higher dimension, for vector fields, of Herman's translated-torus theorem for perturbations of smooth embeddings of $F : \mathbb{T}^n \times [-r_0, r_0] \to \mathbb{T}^n \times \mathbb{R}$ satisfying $F(\theta, 0) = (\theta + \alpha, 0)$ (see [**29**]).

† Since any $A \in \Delta_m^{si}(\mathbb{R})$ is invertible, the counter term $b \in \mathbb{R}^m$ is automatically 0; recall conditions (1.2).

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

On $\mathbb{T}^n \times \mathbb{R}^m$, with $m \geq n$, let $\widehat{\mathcal{U}}(\alpha, A) \subset \mathcal{U}(\alpha, A)$ be the space of vector fields with twist in the following sense: the coefficient $u_1 : \mathbb{T}^n \to \mathrm{Mat}_{n\times m}(\mathbb{R})$ in

$$u(\theta, r) = (\alpha + u_1(\theta) \cdot r + O(r^2), A \cdot r + O(r^2))$$

is such that $\int_{\mathbb{T}^n} u_1(\theta)(d\theta/(2\pi)^n)$ has maximal rank n.

THEOREM 5.2. (Translated torus) *Let α be Diophantine, let $A^0 \in \mathrm{Mat}_m(\mathbb{R})$ have real, simple, non-zero eigenvalues and let $u^0 \in \widehat{\mathcal{U}}_{s+\sigma}(\alpha, A^0)$. There exists an $\varepsilon > 0$ such that every v satisfying $|v - u^0|_{s+\sigma} < \varepsilon$ possesses a translated simple normally hyperbolic torus, of constant normal dynamics close to A^0 and α-quasi-periodic tangential dynamics.*

As in Theorem 4.2, if on the one hand we take advantage of the twist hypothesis in order to avoid the twist term β *à la* Herman, on the other hand one needs to keep track of the average of equation (2.17) (like we did for equation (4.14)) by introducing a translation parameter.

Proof. Without loss of generality let us suppose that A^0 is a diagonal matrix and let $\hat{\varphi}$ be the function defined on $\mathbb{T}^n$ taking values in $\mathrm{Mat}_{n\times m}(\mathbb{R})$ which solves the homological equation†

$$L_\alpha \hat{\varphi}(\theta) + \hat{\varphi}(\theta) \cdot A^0 + u_1(\theta) = \int_{\mathbb{T}^n} u_1(\theta) \frac{d\theta}{(2\pi)^n},$$

and let $F : (\theta, r) \mapsto (\theta + \hat{\varphi}(\theta) \cdot r, r)$. The diffeomorphism F restricts to the identity at $\mathrm{T}^n_0 = \mathbb{T}^n \times \{0\}$. At the cost of substituting u^0 and v with $F_* u^0$ and $F_* v$ respectively, we can assume that

$$u^0(\theta, r) = (\alpha + u_1 \cdot r + O(r^2), A^0 \cdot r + O(r^2)), \quad u_1 = \int_{\mathbb{T}^n} u_1(\theta) \frac{d\theta}{(2\pi)^n}.$$

The germs so obtained are close to one another.

Let us define

$$u^0_A(\theta, r) := (\alpha + u_1 \cdot r + O(r^2), A \cdot r + (A^0 - A) \cdot r + O(r^2)),$$

where A is a diagonal matrix close to A^0 in the subspace Δ_{A^0} of those matrices commuting with A^0, and consider the family of trivial perturbations obtained by translating u^0_A in actions as $u^0_{c,A}(\theta, r) := u^0_A(\theta, c + r)$ where $c \in \mathbb{R}^m$ is close to 0. So consider $u^0_{c,A}$ and $v_{c,A} := v_A(\theta, c + r)$.

By Moser's theorem there exists a triple $(g_{c,A}, u_{c,A}, \lambda_{c,A})$ such that

$$v_{c,A} = g_{c,A*} u_{c,A} + \lambda_{c,A}$$

where $\lambda_{c,A} = (\beta_{c,A}, B_{c,A} \cdot r)$, with $B_{c,A}$ commuting with A. In order to prove the theorem we have to show that there exists $(c, A) \in \mathbb{R}^m \times \Delta_{A^0}$ such that the counter terms $(\beta_{c,A}, B_{c,A})$ vanish.

We claim that the map $(c, A) \mapsto (\beta_{c,A}, B_{c,A})$ is a local submersion. Since this is an open property, and $v_{c,A}$ is close to $u^0_{c,A}$, it suffices to show it for the trivial perturbation $u^0_{c,A}$.

† Each component reads as an equation of the scalar case in Lemma 2.

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

By taking its Taylor expansion and the approximation obtained by cutting it from terms $O(c)$ that possibly depend on angles, we immediately read its normal form where $\beta^0_{c,A} = u_1 \cdot c + O(c^2)$ and $B^0_{c,A} = A^0 - A$. In particular, the torus T^n_0 is translated by $b^0_{c,A} = A^0 \cdot c + O(c^2)$. The map $(c, A) \mapsto (\beta^0_{c,A}, B^0_{c,A})$ is indeed a submersion by the twist hypothesis on $u^0_{c,A}$, since $\partial_c \beta^0_{c,A}|_{c=0} = u_1$ and $\partial_A B^0_{c,A} = -\mathrm{id}$.

The analogous map for $v_{c,A}$, being its small C^1 perturbation, is thus submersive too. Hence there exists (c, A) close to $(0, A^0)$ such that $v_{c,A} = g_{c,A*} u_{c,A} + b_{c,A}$. When $n = m$, $c \in \mathbb{R}^n$ is unique. □

6. *An application to celestial mechanics*

The normal forms constructed in §4 fit well in the dissipative spin–orbit problem. We deduce here the central results of [**27**, Theorem 3.1] and [**5**, Theorem 1] by easy application of (the translated torus) Theorem 4.2 and the elimination of the translation parameter.

6.1. *Spin–orbit in n degrees of freedom.*

6.1.1. *Normal form and elimination of b.* Let $\Omega \in \mathbb{R}^n$ and consider a vector field on $\mathbb{T}^n \times \mathbb{R}^n$ of the form

$$\hat{v} = v^{\mathrm{H}} \oplus (-\eta(r - \Omega)\partial_r) \tag{6.1}$$

where v^{H} is a Hamiltonian vector field whose Hamiltonian H is close to the Hamiltonian in Kolmogorov normal form with non-degenerate quadratic part introduced in §4.2.1:

$$K^0(\theta, r) = \alpha \cdot r + \tfrac{1}{2} Q(\theta) \cdot r^2 + O(r^3).$$

The vector field $\hat{v}$ is thus close to the corresponding unperturbed $\hat{u}$:

$$\hat{u} = u^{K^0} \oplus (-\eta(r - \Omega)\partial_r). \tag{6.2}$$

$\Omega \in \mathbb{R}^n$ is a vector of free parameters representing some ‘external frequencies’ on which the vector field depends when $\eta \neq 0$. In the spin–orbit model, when no dissipation occurs ($\eta = 0$), equations are Hamiltonian and parameter-free [**5**, **9**, **27**]. We will better see in §6.2 the physical meaning of Ω, in the low-dimensional astronomical case of the spin–orbit problem.

We will denote by v and u^0 the part of $\hat{v}$ and $\hat{u}$ with $\Omega = 0$.

THEOREM 6.1. (Dynamical conjugacy) *Fix $\eta_0 > 0$, $\alpha \in \mathbb{R}^n$ Diophantine and let $u^{K^0} \in \widehat{\mathcal{U}}^{\mathrm{Ham}}(\alpha, 0)$. There exists $\varepsilon > 0$ such that, if v^{H} is ε-close to u^{K^0}, for any $\eta \in [-\eta_0, \eta_0]$, there exists a unique† $\Omega \in \mathbb{R}^n$ close to 0, a unique $u \in \widehat{\mathcal{U}}^{\mathrm{Ham}}(\alpha, -\eta)$ and a unique $g \in \mathcal{G}^{\omega}$ such that $\hat{v} = v + \eta\Omega\partial_r$ (close to $\hat{u} = u^0 + \eta\Omega\partial_r$) is conjugated to u by g: $\hat{v} = g_* u$. Hence $\hat{v}$ possesses an invariant α-quasi-periodic torus. This torus is η-normally attractive (respectively, repulsive) if $\eta > 0$ (respectively, $\eta < 0$).*

Proof. Let us write the non-perturbed $\hat{u}$:

$$\hat{u} = \begin{cases} \dot{\theta} = \alpha + O(r), \\ \dot{r} = -\eta r + \eta\Omega + O(r^2). \end{cases} \tag{6.3}$$

† To be rigorous, the uniqueness of Ω formally loses its sense when $\eta = 0$, since η is a factor of Ω in the system; nevertheless since Ω admits a continuous extension as $\eta \to 0$, we agree to take its limit as the stated unique value.

We remark that $\eta\Omega$ is the first term in the Taylor expansion of the counter term b appearing in the normal form of Theorem 4.2, applied to $\hat{v}$ close to $\hat{u}$. In particular, $\hat{u} = \mathrm{id}_* u^0 + \eta\Omega\partial_r$ by uniqueness of the normal form and, if $\Omega = 0$, T_0^n is invariant for (6.3).

Hence consider the family of maps

$$\begin{array}{rccc} \psi : & (\widehat{\mathcal{V}}^{\mathrm{Ham}} \oplus (-\eta(r-\Omega)\partial_r, \hat{u}) & \to & (\mathcal{G}^\omega \times \widehat{\mathcal{U}}^{\mathrm{Ham}}(\alpha, -\eta) \times \Lambda_b, (\mathrm{id}, u^0, \eta\Omega)) \\ & \hat{v} & \mapsto & \psi(\hat{v}) := \phi^{-1}(\hat{v}) = (g, u, b) \end{array}$$

(which are continuous, differentiable and Whitney differentiable by Theorem A.1 and Propositions A.1–A.3) associating to $\hat{v}$ the unique triplet provided by (the translated torus) Theorem 4.2.

In order to prove that the equation $b = 0$ implicitly defines Ω, it suffices to show that $\Omega \mapsto b(\Omega)$ is a local diffeomorphism; since this is an open property with respect to the C^1-topology, and $\hat{v}$ is close to $\hat{u}$, it suffices to show it for $\hat{u}$, which is immediate.

Note, in particular, that $b = \eta\Omega + \sum_{k\geq 1} \delta b_k$ where δb_k, uniquely determined at each step of the Newton scheme, is of the form† $\delta b_k = -\eta\delta\xi_k$ (recall Lemma 8 and system (4.13)–(4.15)).

Hence $b = \eta\Omega +$ (perturbations $\ll \eta\Omega$). So there exists a unique value of Ω, close to 0, such that $b(\Omega) = 0$. □

Note that the distance $|\hat{v} - \hat{u}|_{s+\sigma} = |v^{\mathrm{H}} - u^{\mathrm{H}}|_{s+\sigma}$ is independent of Ω and that constants C' and C'' (appearing in (A.1) and (A.2) in the proof of Theorem A.1) are uniform with respect to Ω and η.

Remark 6.1. When $\eta \neq 0$, Ω is the value that compensates the 'total translation' of the torus, the sum of the successive translations by ξ at each step of the Newton algorithm.

When $\eta = 0$, the standard Kolmogorov, Arnold and Moser (KAM) theory applies since (6.1)–(6.2) are entirely Hamiltonian.

6.2. *Spin–orbit problem of celestial mechanics.* Applying Theorems 4.2 and 6.1, the elimination of the obstructing translation parameter b provides here a picture of the space of parameters proper to this physical system (see Theorem 6.2).

A satellite (or a planet) is said to be in $n : k$ spin–orbit resonance when it accomplishes n complete rotations about its spin axis while revolving exactly k times around its planet (or star). There are various examples of such motion in astronomy, among them the Moon (1 : 1) and Mercury (3 : 2).

The 'dissipative spin–orbit problem' of celestial mechanics can be modeled by the following equation of motion in $\mathbb{R}$:

$$\ddot{\theta} + \eta(\dot{\theta} - \nu) + \varepsilon\partial_\theta f(\theta, t) = 0, \tag{6.4}$$

where $(\theta, t) \in \mathbb{T}^2$, the angular variable θ determines the position of an oblate satellite (modeled as an ellipsoid) whose center of mass revolves on a given elliptic Keplerian orbit around a fixed massive major body, $\eta > 0$ is a dissipation constant depending on the internal non-rigid structure of the body that responds in a non-elastic way to the

† Because of the form of g and the fact that $\xi \in \mathbb{R}^n$, the terms $\delta\xi$ and $\dot{\xi}$ appearing in δg and $\dot{g} = g'^{-1} \cdot \delta g$ coincide.

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

gravitational forces, $\varepsilon > 0$ measures the oblateness of the satellite, while $\nu \in \mathbb{R}$ is an external free parameter proper to the physical problem. We suppose that the potential function f is real analytic in all its variables.

See [**9**] and references therein for a complete physical discussion of the model and deduction of the equation.

Now let α be a fixed Diophantine frequency. In the coordinates $(\theta, r = \dot\theta - \alpha)$ the system associated to (6.4) is

$$\begin{cases} \dot\theta = \alpha + r, \\ \dot r = -\eta r + \eta(\nu - \alpha) - \varepsilon \partial_\theta f(\theta, t). \end{cases} \tag{6.5}$$

We immediately see that when $\varepsilon = 0$ and $\eta \neq 0$, $r = 0$ is an invariant torus provided that $\nu = \alpha$. Furthermore, the general solution of $\ddot\theta + \eta(\dot\theta - \nu) = 0$ is given by

$$\theta(t) = \nu t + \theta^0 + \frac{r_0 - (\nu - \alpha)}{\eta}(1 - e^{-\eta t}),$$

showing that the rotation tends asymptotically to a ν-quasi-periodic behavior. Here the meaning of ν is revealed: ν is the frequency of rotation to which the satellite tends because of the dissipation, if no 'oblate-shape effects' are present.

On the other hand when $\varepsilon \neq 0$ and $\eta = 0$ we are in the conservative regime, and the classical KAM theory applies.

The main question then is: fixing α Diophantine, does there exist a value of the proper rotation frequency ν such that the perturbed system possesses an α-quasi-periodic invariant η-attractive torus?†

6.2.1. *Extending the phase space.* In order to apply our general scheme to the non-autonomous system (6.5), as usual we extend the phase space by introducing time (or its translates) as a variable. The phase space becomes $\mathbb{T}^2 \times \mathbb{R}^2$ with variable θ_2 corresponding to time and r_2 its conjugate.

Hence consider the family of vector fields (parametrized by $\Omega \in \mathbb{R}$)

$$v = v^{\mathrm{H}} \oplus (-\eta r + \eta \Omega)\partial_r,$$

where $\Omega = (\nu - \alpha, 0)$, $\nu \in \mathbb{R}$, and v^{H} corresponds to

$$H(\theta, r) = \alpha \cdot r_1 + r_2 + \tfrac{1}{2} r_1^2 + \varepsilon f(\theta_1, \theta_2).$$

The following objects are essentially the ones introduced in §4.2.1, taking into account the introduction of the time variable $\theta_2 = t$ and its conjugate r_2. Let $\bar\alpha = (\alpha, 1)$ satisfy

$$|k_1 \alpha + k_2| \geq \frac{\gamma}{|k|^\tau} \quad \text{for all } k \in \mathbb{Z}^2 \backslash \{0\}. \tag{6.6}$$

Let $\bar{\mathcal{H}}$ be space of real analytic Hamiltonians defined in a neighborhood of $\mathrm{T}_0 = \mathbb{T}^2 \times \{0\}$ such that, for $H \in \bar{\mathcal{H}}$, $\partial_{r_2} H \equiv 1$. For these Hamiltonians the frequency $\dot\theta_2 = 1$ (corresponding to time) is fixed. Let $\bar\alpha = (\alpha, 1)$ and $\bar{\mathcal{K}} = \bar{\mathcal{H}} \cap \mathcal{K}^{\bar\alpha}$. Also let $\bar{\mathcal{G}}^\omega$ be the subset of $\mathcal{G}^\omega$ such that $\bar\xi = (\xi, 0)$, $\varphi(\theta) = (\varphi_1(\theta), \theta_2)$. The corresponding

† In [**5**] Celletti and Chierchia look for a function $u : \mathbb{T}^2 \to \mathbb{R}$ of the form $\theta(t) = \alpha t + u(\alpha t, t)$ satisfying (6.4) for a particular value ν. This function is found as the solution of an opportune partial differential equation.

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

$\dot{g} \in T_{\mathrm{id}}\bar{\mathcal{G}}^{\omega}$ are $\dot{g} = (\dot{\varphi}, -{}^t\dot{\varphi}' \cdot r + d\dot{S} + \dot{\xi})$ with $\dot{\varphi} = (\dot{\varphi}_1, 0)$ and $\dot{\xi} = (\dot{\xi}_1, 0)$, and $\bar{\Lambda} = \{\lambda : \lambda(\theta, r) = b(\partial/\partial r_1)\} \equiv \mathbb{R}$.

By restriction, the normal-form operator

$$\bar{\phi} : \bar{\mathcal{G}}^{\omega,\sigma^2/2n}_{s+\sigma} \times \widehat{\mathcal{U}}^{\mathrm{Ham}}_{s+\sigma}(\bar{\alpha}, -\eta) \times \bar{\Lambda} \to (\widehat{\mathcal{V}}^{\mathrm{Ham}} \oplus (-\eta r + \eta\mathbb{R}\partial_{r_1}))_s,$$
$$(g, u, b) \mapsto g_* u + b\partial_{r_1},$$

and the corresponding

$$\bar{\phi}'(g, u, \lambda) : T_g \mathcal{G}^{\omega,\sigma^2/2n}_{s+\sigma} \times \overrightarrow{\widehat{\mathcal{U}}^{\mathrm{Ham}}_{s+\sigma}} \times \bar{\Lambda} \to (\widehat{\mathcal{V}}^{\mathrm{Ham}} \oplus (\eta\mathbb{R})\partial_{r_1})_{g,s},$$

are now defined. The notation $\overrightarrow{\widehat{\mathcal{U}}^{\mathrm{Ham}}_{s+\sigma}}$ always stands for the vector space directing $\widehat{\mathcal{U}}^{\mathrm{Ham}}_{s+\sigma}(\bar{\alpha}, -\eta)$.

COROLLARY 6.1. (Normal form for time-dependent perturbations) *The operator*

$$\bar{\phi} : \bar{\mathcal{G}}^{\omega,\sigma^2/2n}_{s+\sigma} \times \widehat{\mathcal{U}}^{\mathrm{Ham}}_{s+\sigma}(\bar{\alpha}, -\eta) \times \bar{\Lambda} \to (\widehat{\mathcal{V}}^{\mathrm{Ham}} \oplus (-\eta r + \eta\mathbb{R}\partial_{r_1}))_s$$

is a local diffeomorphism.

The proof is recovered from that of (the translated torus) Theorem 4.2, taking into account that the perturbation belongs to the particular class $\bar{\mathcal{H}}$.

LEMMA 10. (Inversion of $\bar{\phi}'$) *If g is close to the identity, for every $\delta v \in (\mathcal{V}^{\mathrm{Ham}} \oplus (\eta\mathbb{R})\partial_{r_1})_{g,s+\sigma}$ there exists a unique triplet $(\delta g, \delta u, \delta\lambda)$ such that $\bar{\phi}'(g, u, b) \cdot (\delta g, \delta u, \delta b) = \delta v$. Moreover, there exist $\tau', C' > 0$ such that*

$$\max\{|\delta g|_s, |\delta u|_s, |\delta b|\} \le \frac{C'}{\sigma^{\tau'}} |\delta v|_{g,s+\sigma},$$

where the constant C' depends only on τ, n, $|g - \mathrm{id}|_{s+\sigma}$ and $|u - (\alpha, -\eta r)|_{s+\sigma}$.

Proof. Following the calculations made to prove Lemma 9 in the demonstration of Theorem 4.2, we get to the following system of homological equations:

$$\begin{gathered} \dot{\varphi}_1' \cdot \bar{\alpha} - Q_{11} \cdot (d\dot{S}_1 + \dot{\xi}_1) = \dot{v}^H_{1,0} \\ d\dot{S}_1' \cdot \bar{\alpha} + \eta(d\dot{S}_1 + \dot{\xi}_1) = \dot{V}^H_{1,0} + \eta\delta\Omega - (\delta b + \partial_{\theta_1} v^1 \delta b) \\ d\dot{S}_2' \cdot \bar{\alpha} + \eta\, d\dot{S}_2 = \dot{V}^H_{2,0} - \partial_{\theta_2} v^1 \delta b. \end{gathered} \tag{6.7}$$

As usual, $\dot{g} = g'^{-1} \cdot \delta g = (\dot{\varphi}, -{}^t\dot{\varphi}' \cdot r + d\dot{S} + \dot{\xi})$; Q_{11} is the element of the torsion matrix $Q = (Q_{i,j})$. The lower indexes of vector fields indicate the component and the order of the corresponding term in r whose the coefficient they are†. Hence, the first one corresponds to the direction of θ and the second two to the zeroth-order term in r in the normal direction.

The tangential equation relative to the time component (which we omitted above) is easily determined: computation gives $\dot{v}_{2,0} = 0$, because of $\delta v \partial_{\theta_2} = \delta 1 = 0$ and the form of g'^{-1}, and $\dot{\varphi}_2 = 0$ as well as $Q(\theta) \cdot d\dot{S} = 0$ in the $\dot{\theta}_2$-direction.

Equations relative to the linear term follow from the Hamiltonian character. Solutions follows from Lemmas 1 and 2, the same kind of estimates as in Lemma 9 hold, hence the required bound. □

† We denote by $v^1 = \varphi_1 - \mathrm{id}$, the vector field coming from the first component of $\varphi - \mathrm{id} = (\varphi_1, \mathrm{id}) - \mathrm{id}$.

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

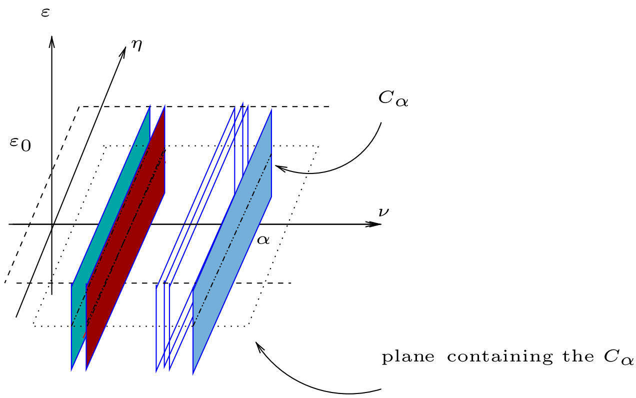

FIGURE 3. The Cantor set of surfaces: transversely cutting with a plane $\varepsilon = const.$, we obtain a Cantor set of curves like the one described in Corollary 6.2.

LEMMA 11. *There exists a constant C'', depending on $|x|_{s+\sigma}$, such that in a neighborhood of* $(\mathrm{id}, u^0, 0) \in \bar{\mathcal{G}}^{\omega}_{s+\sigma} \times \widehat{\mathcal{U}}^{\mathrm{Ham}}_{s+\sigma}(\bar{\alpha}, -\eta) \times \bar{\Lambda}$ *the bilinear map $\phi''(x)$ satisfies the bound*

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

The proof of Corollary 6.1 follows.

6.3. *Surfaces of invariant tori.* The results below will follow from Corollary 6.1 and Theorem 6.1. See Figures 3 and 4 for a representation of the following results in the space of parameters.

THEOREM 6.2. (Cantor set of surfaces) *Let ε_0 be the maximal value that the perturbation can attain. Every Diophantine α identifies a surface $(\varepsilon, \eta) \mapsto \nu(\varepsilon, \eta)$ in the space $(\varepsilon, \eta, \nu) = [0, \varepsilon_0] \times [-\eta_0, \eta_0] \times \mathbb{R}$, which is analytic in ε, smooth in η, for which the following holds: for any parameters $(\varepsilon, \eta, \nu(\varepsilon, \eta))$, $\hat{v}$ admits an invariant α-quasi-periodic torus. This torus is η-normally attractive (respectively, repulsive) if $\eta > 0$ (respectively, $\eta < 0$).*

Fixing an admissible perturbation $\varepsilon > 0$, we obtain the following straightforward corollary.

COROLLARY 6.2. (A curve of normally hyperbolic tori) *Fixing α Diophantine and ε sufficiently small, there exists a unique analytic curve C_α, in the plane $(\eta, \nu) = [0, \varepsilon_0] \times [-\eta_0, \eta_0]$ of the form $\nu = \alpha + O(\varepsilon^2)$, along which the counter term $b(\nu, \alpha, \eta, \varepsilon)$ 'à la Rüssmann' vanishes, so that the perturbed system possesses an invariant α-quasi-periodic torus. This torus is attractive (respectively, repulsive) if $\eta > 0$ (respectively, $\eta < 0$).*

Proof of Theorem 6.2. We just need to observe the following facts.

The existence of the unique local inverse for $\bar{\phi}'$ and the bound on it and $\bar{\phi}''$ allow us to apply Theorem 4.2 and prove the result once we guarantee that

$$|v - u^0|_{s+\sigma} = \max\left(\varepsilon\left|\frac{\partial f}{\partial \theta_1}\right|_{s+\sigma}, \varepsilon\left|\frac{\partial f}{\partial \theta_2}\right|_{s+\sigma}\right) \leq \delta\frac{\sigma^{2\tau}}{2^{8\tau}C^2}$$

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

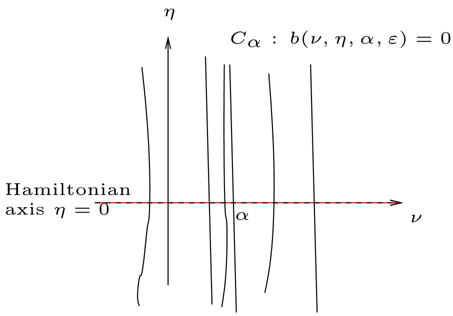

FIGURE 4. The corresponding Cantor set of curves on the plane $\varepsilon = const.$, whose points correspond to an attractive/repulsive invariant torus.

(here we have replaced the constant η appearing in the abstract function theorem with δ, in order not to generate confusion with the dissipation term). This ensures that the inverse mapping theorem can be applied, as well as Propositions (A.1–A.3). Note that the constant C appearing in the bound contains a factor $1/\gamma^2$ coming from the Diophantine condition (6.6), independent of η, since Remark 4.1 still holds here.

For every $\eta \in [-\eta_0, \eta_0]$, apply Theorem 6.1 and find the unique ν such that

$$b(\nu, \eta, \alpha, \varepsilon) = 0,$$

(as in the multidimensional case of Theorem 6.1 b is of the form $b = \eta(\nu - \alpha - \sum_k \delta\xi_k)$, smooth with respect to ν and η and analytic in ε).

In particular, the value of ν that satisfies the equation is of the form

$$\nu(\varepsilon, \eta) = \alpha + O(\varepsilon^2).$$

This follows directly from the very first step of Newton's scheme

$$x_1 = x_0 + \phi'^{-1}(x_0) \cdot (v - \phi(x_0)),$$

where $x_0 = (\mathrm{id}, u^0, \eta(\nu - \alpha))$. Developing the expression, one sees that $\delta\xi_1$ (the term of order ε) is necessarily zero, due to the particular perturbation and the constant torsion.

6.4. *An important dichotomy.* The results obtained for the spin–orbit problem, Theorems 6.1 and 6.2 and Corollary 6.2, are intimately related to the very particular nature of the equations of motion and point to an existing dichotomy between generic dissipative vector fields and the Hamiltonian dissipative case to which the spin–orbit system belongs.

In the case of a general perturbation, if we want to keep the characteristic frequencies $(\alpha, -\eta)$ fixed, even if the system satisfies some torsion property, two counter terms are still necessary: a counter term b to keep track of the average of the equation relocating the torus and $B \in \mathrm{Mat}_n(\mathbb{R})$ (recall condition $[B, \eta] = 0$). Disposing of just n free parameters $\Omega_1, \ldots, \Omega_n$, one possible result is to eliminate b, but it is hopeless to get rid of the obstruction represented by B and have complete control of the normal dynamics of the torus.

In particular, for the spin–orbit problem in one and a half degrees of freedom, by using diffeomorphisms of $\mathbb{T}^2 \times \mathbb{R}$ of the form $g(\theta, r) = (\varphi(\theta), \theta_2, R_0(\theta) + R_1(\theta) \cdot r)$, the

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

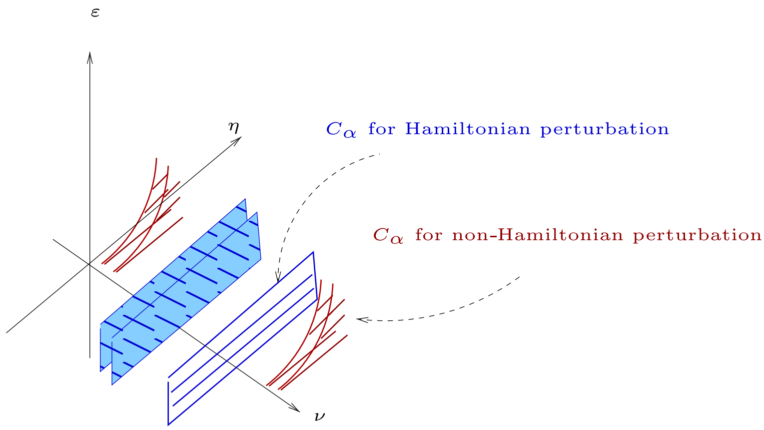

FIGURE 5. The two situations: (1) surfaces $\nu = \nu(\eta, \varepsilon)$ passing through $\eta = 0$ correspond to the case 'Hamiltonian + dissipation' of Theorem 6.2; (2) interrupted surfaces correspond to the case of generic perturbations: they correspond to invariant tori of co-dimension 1 ($B \neq 0$).

system corresponding to (6.7) reads

$$\dot{\varphi}' \cdot \bar{\alpha} - \dot{R}_0 = \dot{v}_0,$$
$$\dot{R}_0' \cdot \bar{\alpha} + \eta \dot{R}_0 = \dot{V}_0 - \dot{b},$$
$$\dot{R}_1' \cdot \bar{\alpha} + \dot{R}_0' = \dot{V}_1 - \dot{B}, \quad \bar{\alpha} = (\alpha, 1),\ \delta b, \delta B \in \mathbb{R}.$$

and disposing of $\nu \in \mathbb{R}$ only, we could at best solve $b = 0$, for any η up to a small neighborhood of the origin, since $b = \eta(\nu - \alpha) + O(\varepsilon)$.

Besides, in the case of no torsion, the direct application of Moser's theorem with counter terms $\beta, B \in \mathbb{R}$ (by considering $\eta(\nu - \alpha)$ part of the perturbation) immediately gives that $\beta = \nu - \alpha + (\int V_0)/\eta + \int v_0$ at first order, and one could solve $\beta = 0$ with respect to ν, for any η in the complement of a small neighborhood of the origin. In fact, the Diophantine condition $|i\, k \cdot \bar{\alpha} + \eta| \geq \gamma/(1 + |k|)^\tau$ for all $k \in \mathbb{Z}^2$ is necessary in order to control the constant part of $\dot{R}_0$ (not necessarily of zero mean), implying that the bound on ε of Theorem A.1 depends on η through γ:

$$\varepsilon < \gamma^4 C' \leq \eta^4 C'.$$

Once ε is fixed, the curves C_α entailed in Corollary 6.2 (obtained by eliminating β) cannot be defined up to the axis $\eta = 0$ in the plane $\varepsilon =$ const. (we grouped in a unique constant C' all the other terms appearing in the bound).

Acknowledgements. During the revision of the present paper, the author has been supported by the ERC starting grant HamPDEs. The author is grateful to A. Chenciner and J. Féjoz for mathematical (and moral) support and advice throughout her Ph.D, at the Astronomie et Systèmes Dynamiques group of the IMCCE, Observatoire de Paris. This work owes much to meaningful conversations of those (and these) times. Thank you also to A. Bounemoura, T. Castan, B. Fayad, J.-P. Marco, P. Mastrolia, L. Niederman, C. and M. Procesi for the enlightening discussions we had and their constant interest in this (and future) work.

A. *Appendix. Inverse function theorem and regularity of ϕ*

We present here the inverse function theorem we used to prove Theorems 2.1, 4.1 and 4.2. This result follows from Féjoz [**11**, **13**]. Remark that we endowed functional spaces with weighted norms and bounds appearing in Proposition 2.1 and Lemma 4 may depend on $|x|_s$ (as opposed to the statements given in [**11**, **13**]); in what follows, we take account of these (slight) differences.

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

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

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. *Under the previous assumptions, assume*

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

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

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

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

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

Define

$$Q : B_{s+2\sigma}^E(\sigma) \times B_{s+2\sigma}^E \to F_s, \quad (x, \hat{x}) \mapsto \phi(\hat{x}) - \phi(x) - \phi'(x)(\hat{x} - x), \tag{A.3}$$

the reminder of the Taylor formula.

LEMMA 12. *For every $x, \hat{x}$ such that $|x - \hat{x}|_s < \sigma$,*

$$|Q(x, \hat{x})|_{x,s} \leq \frac{C''}{2\sigma^2} \, |\hat{x} - x|_{s+\sigma+|\hat{x}-x|_s}^2. \tag{A.4}$$

Proof. Let $x_t = (1 - t)x + t\hat{x}$, $0 \leq t \leq 1$, be the segment joining x to $\hat{x}$. Using Taylor's formula,

$$Q(x, \hat{x}) = \int_0^1 (1 - t)\phi''(x_t)(\hat{x} - x)^2 \, dt,$$

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

hence

$$\begin{aligned}|Q(x,\hat{x})|_{x,s} &\leq \int_0^1 (1-t)|\phi''(x_t)(\hat{x}-x)^2|_{x,s}\,dt\\ &\leq \int_0^1 (1-t)|\phi''(x_t)(\hat{x}-x)^2|_{x_t,s+|x_t-x|_s}\,dt\\ &\leq \int_0^1 (1-t)\,\frac{C''}{\sigma^2}|(\hat{x}-x)|^2_{s+\sigma+|x_t-x|_s}\,dt\\ &\leq \frac{C''}{2\sigma^2}\,|\hat{x}-x|^2_{s+\sigma+|\hat{x}-x|_s}.\end{aligned}$$

□

Proof of Theorem A.1. Let $\eta < s$, σ and ε be fixed positive real numbers and let $y \in B^F_{s+\sigma}(\varepsilon)$. We define the map

$$f : B^E_{s+\sigma}(\sigma) \to E_s, \quad x \mapsto x + \phi'^{-1}(x)(y-\phi(x)).$$

We want to prove that, if ε is sufficiently small, there exists a sequence defined by induction by

$$\begin{cases} x_0 = 0,\\ x_{n+1} = f(x_n),\end{cases}$$

converging towards some point $x \in B^E_s(\eta)$, the pre-image of y by ϕ.

Let us introduce two sequences: a sequence of positive real numbers $(\sigma_n)_{n\geq 0}$ such that $3\sum_n \sigma_n = \sigma$ is the total width of analyticity we will have lost at the end of the algorithm; the decreasing sequence $(s_n)_{n\geq 0}$ defined inductively by $s_0 = s+\sigma$ (the starting width of analyticity), $s_{n+1} = s_n - 3\sigma_n$. Of course, $s_n \to s$ when $n \to +\infty$. Let $C_k = 2C\sigma_k^{-\tau} \geq 1$ for all $k \in \mathbb{N}$ and define $\zeta_n = \prod_{0\leq k\leq n} C_k^{2^{-k}}$ and $\zeta = \prod_{k\geq 0} C_k^{2^{-k}}$. We start to prove that for every $n \geq 1$, there exist $x_0, \ldots, x_n$ and that

$$|x_n - x_{n-1}|_{s_n} \leq (\varepsilon\zeta_{n-1})^{2^{n-1}} \quad \text{and} \quad |x_n|_{s_n} \leq \sum_{k=0}^{n-1} (\varepsilon\zeta)^{2^k}.$$

From $x_k - x_{k-1} = \phi'^{-1}(x_{k-1})(y-\phi(x_{k-1}))$ we see that $y - \phi(x_k) = -Q(x_{k-1}, x_k)$, which allows us to write $x_{k+1} - x_k = -\phi'^{-1}(x_k)Q(x_{k-1}, x_k)$, for $k = 1, \ldots, n$.

First, remark that

$$|x_1 - x_0|_{s_1} = |x_1|_{s_1} \leq \frac{C'}{(3\sigma_0)^{\tau'}}\,|y-\phi(x_0)|_{s_0} \leq \frac{C}{2\sigma_0^{\tau}}\,|y|_{s+\sigma} \leq C_0\varepsilon$$

and $|x_1|_{s_1} \leq \zeta\varepsilon$; the assertion is thus true for $n = 1$.

Assuming that $|x_k - x_{k-1}|_{s_k} \leq \sigma_k$, for $k = 1, \ldots, n$, from the estimate of the right inverse and the previous lemma we get

$$|x_{n+1} - x_n|_{s_{n+1}} \leq \frac{C}{2\sigma_n^{\tau}}|x_n - x_{n-1}|^2_{s_n} \leq \cdots \leq C_n C_{n-1}^2 \ldots C_1^{2^{n-1}}|x_1 - x_0|^{2^n}_{s_1}.$$

Second, observe that since $C_k \geq 1$ (see remark below),

$$|x_{n+1} - x_n|_{s_{n+1}} \leq C_n(\varepsilon\zeta_{n-1})^{2^n} = (\varepsilon\zeta_n)^{2^n} \leq (\varepsilon\zeta)^{2^n} = \left(\varepsilon\prod_{k\geq 0} C_k^{2^{-k}}\right)^{2^n}$$

and $|x_{n+1}|_{s_{n+1}} \leq |x_n|_{s_n} + |x_{n+1} - x_n|_{s_{n+1}} \leq \sum_0^n (\varepsilon\zeta)^{2^k}$.

Third, note that

$$\sum_{n\geq 0} z^{2^n} = z + z^2 + z^4 + \cdots \leq z \sum_{n\geq 0} z^n \leq 2z,$$

if $z \leq \frac{1}{2}$.

The key point is to choose ε such that $\varepsilon \prod_{k\geq 0} C_k^{2^{-k}} \leq \frac{1}{2}$ (or any positive number less than 1) and $\sum_{n\geq 0} |x_{n+1} - x_n|_{s_{n+1}} < \eta$, in order for the whole sequence (x_k) to exist and converge in $B_s(\eta) \subset E_s$. Hence, using the definition of the C_n and the fact that

$$\left(\frac{C}{2}\right)^{-2^{-k}} = \left(\frac{2}{C}\right)^{(1/2)^k} \Longrightarrow \prod \left(\frac{2}{C}\right)^{(1/2)^k} = \left(\frac{2}{C}\right)^{\sum (1/2^k)} = \left(\frac{2}{C}\right)^2,$$

within $\sum_k (1/2^k) = \sum_k k(1/2^k) = 2$, we obtain as a sufficient value

$$\varepsilon = \eta \frac{2}{C^2} \prod_{k\geq 0} \sigma_k^{\tau\,(1/2)^k}. \tag{A.5}$$

Eventually, the constraint $3\sum_{n\geq 0} \sigma_n = \sigma$ gives $\sigma_k = (\sigma/6)\,(\frac{1}{2})^k$, which, plugged into (A.5), gives

$$\varepsilon = \eta\,\frac{2}{C^2}\left(\frac{\sigma}{12}\right)^{2\tau} > \frac{\sigma^{2\tau}\eta}{2^{8^\tau} C^2},$$

hence the theorem.

A posteriori, the exponential decay we proved makes straightforward the further assumption $|x_k - x_{k-1}|_{s_k} < \sigma_k$ to apply Lemma 12.

Concerning the bounds over the constant C, as $\sum_k |x_{k+1} - x_k|_{s_{k+1}} \leq \eta$, we see that all the $|x_n|_{s_n}$ are bounded, hence the constants C' and C'' depending on them.

Moreover, to have all the $C_n \geq 1$, as we previously supposed, it suffices to assume $C \geq \sigma/3$.

Remark A.1. If the operator ϕ is defined only on polynomially small balls

$$\phi : B^E_{s+\sigma}(c_0\sigma^\ell) \to F_s,\ c_0 > 0 \quad \text{for all } s,\ \sigma,$$

the statement and the proof of Theorem A.1 still hold, provided that η is chosen small enough ($\eta < 2c_0(\sigma/12)^\ell$ suffices). This is the case of the operators defined in §§4.1 and 4.2, where $\ell = 2$.

A.1. *Local uniqueness and regularity of the normal form.* We want to show the uniqueness and some regularity properties of the right inverse ψ of ϕ, assuming the additional left invertibility of ϕ' (which is the case for the particular operator ϕ' of interest to us).

Definition A.1. We will say that a family of norms $(|\cdot|_s)_{s>0}$ on a grading $(E_s)_{s>0}$ is *log-convex* if, for every $x \in E_s$, the map $s \mapsto \log |x|_s$ is convex.

LEMMA 13. *If $(|\cdot|_s)$ is log-convex, the following inequality holds:*

$$|x|^2_{s+\sigma} \leq |x|_s |x|_{s+\tilde\sigma} \quad \textit{for all } s,\ \sigma,\ \tilde\sigma = \sigma\left(1 + \frac{1}{s}\right).$$

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

Proof. If $f : s \mapsto \log |x|_s$ is convex, then

$$f\left(\frac{s_1+s_2}{2}\right) \le \frac{f(s_1)+f(s_2)}{2}.$$

Let $x \in E_s$. Then

$$\log |x|_{s+\sigma} \le \log |x|_{(2s+\tilde{\sigma})/2} \le \tfrac{1}{2}(\log |x|_s + \log |x|_{s+\tilde{\sigma}}) = \tfrac{1}{2}\log(|x|_s|x|_{s+\tilde{\sigma}}),$$

hence the lemma. □

Let us assume that the family of norms $(|\cdot|_s)_{s>0}$ of the grading $(E_s)_{s>0}$ are log-convex. To prove the uniqueness of ψ we will assume that ϕ' is also left-invertible.

PROPOSITION A.1. (Lipschitz continuity of ψ) *Let $\sigma < s$. If $y, \hat{y} \in B^F_{s+\sigma}(\varepsilon)$ with $\varepsilon = 3^{-4\tau}2^{-16\tau}(\sigma^{6\tau}/4C^3)$, the inequality*

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

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

Proof. In order to get the desired estimate we introduce an intermediate parameter ξ, which will be chosen later, such that $\eta < \xi < \sigma < s < s+\sigma$.

For ease of notation let $\psi(y) =: x$ and $\psi(\hat{y}) =: \hat{x}$. Let also $\varepsilon = \xi^{2\tau}\eta/2^{8\tau}C^2$ so that if $y, \hat{y} \in B^F_{s+\sigma}(\varepsilon)$, then $x, \hat{x} \in B^E_{s+\sigma-\xi}(\eta)$, by Theorem A.1, provided that $\eta < s+\sigma-\xi$ (to be checked later). In particular, we assume that any $x, \hat{x} \in B^E_{s+\sigma-\xi}$ satisfy $|x-\hat{x}|_{s+\sigma-\xi} \le 2\eta$. Writing

$$(x-\hat{x}) = \phi'^{-1}(x)\cdot\phi(x)(x-\hat{x}),$$

and using

$$\phi'(x)(x-\hat{x}) = \phi(\hat{x}) - \phi(\hat{x}) - Q(x,\hat{x}),$$

we get

$$x-\hat{x} = \phi'^{-1}(x)(\phi(\hat{x}) - \phi(x) - Q(x,\hat{x})).$$

Taking norms, we have

$$\begin{aligned}
|x-\hat{x}|_s &\le \frac{C'}{\sigma^{\tau'}}|y-\hat{y}|_{x,s+\sigma} + \frac{C}{2\xi^{\tau}}|x-\hat{x}|^2_{s+2\xi+|x-\hat{x}|_{s+\xi}}\\
&\le \frac{C'}{\sigma^{\tau'}}|y-\hat{y}|_{x,s+\sigma} + \frac{C}{2\xi^{\tau}}|x-\hat{x}|^2_{s+2\xi+2\eta},
\end{aligned}$$

by Lemma 12 and the fact that $|x-\hat{x}|_{s+\xi} \le |x-\hat{x}|_{s+\sigma-\xi}$ (choosing ξ so that $2\xi < \sigma$ too).

Let us define $\tilde{\sigma} = (2\xi+2\eta)(1+1/s)$ and use the interpolation inequality

$$|x-\hat{x}|^2_{s+2\eta+2\xi} \le |x-\hat{x}|_s|x-\hat{x}|_{s+\tilde{\sigma}}$$

to obtain

$$\left(1 - \frac{C}{2\xi^{\tau}}|x-\hat{x}|_{s+\tilde{\sigma}}\right)|x-\hat{x}|_s \le \frac{C'}{\sigma^{\tau'}}|y-\hat{y}|_{x,s+\sigma}.$$

We now choose η sufficiently small that:

– $\tilde{\sigma} \le \sigma - \xi$, which implies $|x - \hat{x}|_{s+\tilde{\sigma}} \le 2\eta$. It suffices to have $\eta \le \sigma/(2(1+1/s)) - \frac{3}{2}\xi$.
– $\eta \le \xi^\tau/2C$ in order to have $(C/2\xi^\tau)|x - \hat{x}|_{s+\sigma} \le \frac{1}{2}$.

A possible choice is $\xi = \sigma^2/12$ and $\eta = (\sigma/12)^{2\tau}(1/4C)$, hence our choice of ε. □

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

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

and the map $\psi' : B^F_{s+\sigma}(\varepsilon) \to L(F_{s+\sigma}, E_s)$ defined locally by $\psi'(y) = \phi'^{-1}(\psi(y))$ is continuous. In particular, if $\phi : B^E_{s+\sigma}(\sigma) \to F_s$ is C^k, $2 \le k \le \infty$, so is $\psi : B^F_{s+\sigma}(\varepsilon) \to E_s$.

Proof. Let us introduce some terms:

– $\Delta := \psi(\hat{y}) - \psi(y) - \phi'^{-1}(x)(\hat{y} - y)$;
– $\delta := \hat{y} - y$, the increment;
– $\xi := \psi(y+\delta) - \psi(y)$;
– $\Xi := \phi(x+\xi) - \phi(x)$.

With this new notation we can see Δ as

$$\begin{aligned}\Delta &= \xi - \phi'^{-1}(x)\cdot\Xi \\ &= \phi'^{-1}(x)(\phi'(x)\cdot\xi - \Xi) \\ &= \phi'^{-1}(x)(\phi'(x)\xi - \phi(x+\xi) + \phi(x)) \\ &= -\phi'^{-1}(x)Q(x, x+\xi).\end{aligned}$$

Taking norms, we have

$$|\Delta|_s \le K|\hat{y} - y|^2_{x,s+\bar{\sigma}}$$

by Proposition A.1 and Lemma 12, for some $\bar{\sigma}$ which goes to zero when σ does, and some constant $K > 0$ depending on σ. Up to substituting σ for $\bar{\sigma}$, we have proved the statement.

In addition

$$\psi'(y) = \phi^{-1}(y)' = \phi'^{-1} \circ \phi^{-1}(y) = \phi'^{-1}(\psi(y)),$$

the inversion of linear operators between Banach spaces being analytic, the map $y \mapsto \phi'^{-1}(\psi(y))$ 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. Suppose that $\phi(x) = \phi_\nu(x)$ depends on some parameter $\nu \in B^d$ (the unit ball of $\mathbb{R}^d$), that it is C^1 with respect to ν and that estimates of ϕ'^{-1}_ν and ϕ''_ν are uniform with respect to ν over some closed subset D of $\mathbb{R}^d$.

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 $\psi : \mathbb{R}^d \times B^F_{s+\sigma}(\varepsilon) \to B^E_s(\eta)$ of class C^1. If ϕ is C^k, $1 \le k \le \infty$, with respect to ν, this extension is C^k.*

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

Proof. Let $y \in B^F_{s+\sigma}(\varepsilon)$. For $\nu, \nu+\mu \in D$, let $x_\nu = \psi_\nu(y)$ and $x_{\nu+\mu} = \psi_{\nu+\mu}(y)$, implying

$$\phi_{\nu+\mu}(x_{\nu+\mu}) - \phi_{\nu+\mu}(x_\nu) = \phi_\nu(x_\nu) - \phi_{\nu+\mu}(x_\nu).$$

It then follows, since $y \mapsto \psi_{\nu+\mu}(y)$ is Lipschitz, that

$$|x_{\nu+\mu} - x_\nu|_s \le L|\phi_\nu(x_\nu) - \phi_{\nu+\mu}(x_\nu)|_{x_\nu, s+\sigma},$$

taking $y = \phi_{\nu+\mu}(x_\nu)$, $\hat{y} = \phi_{\nu+\mu}(x_{\nu+\mu})$. In particular, since $\nu \mapsto \phi_\nu(x_\nu)$ is Lipschitz, the same holds for $\nu \mapsto x_\nu$. Let us now expand $\phi_{\nu+\mu}(x_{\nu+\mu}) = \phi(\nu+\mu, x_{\nu+\mu})$ as a Taylor series at (ν, x_ν). We have

$$\phi(\nu+\mu, x_{\nu+\mu}) = \phi(\nu, x_\nu) + D\phi(\nu, x_\nu)\cdot(\mu, x_{\nu+\mu} - x_\nu) + O(\mu^2, |x_{\nu+\mu} - x_\nu|_s^2),$$

hence, formally defining the derivative $\partial_\nu x_\nu := -\phi_\nu'^{-1}(x_\nu)\cdot\partial_\nu\phi_\nu(x_\nu)$, we obtain

$$x_{\nu+\mu} - x_\nu - \partial_\nu x_\nu \cdot \mu = \phi_\nu'^{-1}(x_\nu)\cdot O(\mu^2),$$

hence

$$|x_{\nu+\mu} - x_\nu - \partial_\nu x_\nu \cdot \mu|_s = O(\mu^2)$$

by the Lipschitz property of $\nu \mapsto x_\nu$, when $\mu \mapsto 0$, locally uniformly with respect to ν. Hence $\nu \mapsto x_\nu$ is C^1-Whitney smooth and, by the Whitney extension theorem, the claimed extension exists. Similarly, if ϕ is C^k with respect to ν, $\nu \mapsto x_\nu$ is C^k-Whitney smooth. See [**1**] for the straightforward generalization of Whitney's theorem to the case of interest to us: ψ takes values in a Banach space instead of a finite-dimensional vector space; but note that the extension direction is of finite dimension. □

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

We present here a classical result on the inversion of a holomorphism on the complex torus $\mathbb{T}^n_s$ used to guarantee that normal-form operators ϕ were well defined.

All complex extensions of manifolds are defined with the help of the ℓ^∞-norm,

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

Let us 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 $\varphi \circ \psi = \mathrm{id}_{\mathbb{T}^n_{s-2\sigma}}$.*

Furthermore, the following inequalities hold:

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

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

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

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

Proof. Let $\hat{\varphi} := \mathrm{id} + v \circ p : \mathbb{R}^n_s \to \mathbb{R}^n_{s+\sigma}$ be the lift of φ to $\mathbb{R}^n_s$.

Let us start by proving the injectivity and surjectivity of $\hat{\varphi}$; the same properties for φ follow from these.

We first prove that $\hat{\varphi}$ is injective as a map from $\mathbb{R}^n_{s-\sigma} \to \mathbb{R}^n_s$. Let $\hat{\varphi}(x) = \hat{\varphi}(x')$. From the definition of $\hat{\varphi}$ we have

$$|x - x'| = |v \circ p(x') - v \circ p(x)| \le \int_0^1 \sum_{k=1}^n |\partial_{x_k} \hat{v}|_{s-\sigma} |x'_k - x_k| \, dt \le \frac{n}{\sigma} |v|_s |x - x'| < |x - x'|,$$

hence $x' = x$.

Now we prove that $\hat{\varphi} : \mathbb{R}^n_{s-\sigma} \to \mathbb{R}^n_{s-2\sigma} \subset \hat{\varphi}(\mathbb{R}^n_{s-\sigma})$ is surjective. Define, for every $y \in \mathbb{R}^n_{s-2\sigma}$ the map

$$f : \mathbb{R}^n_{s-\sigma} \to \mathbb{R}^n_{s-\sigma}, \ x \mapsto y - v \circ p(x),$$

which is a contraction (see the last but one inequality of the previous step). Hence there exists a unique fixed point such that $\hat{\varphi}(x) = x + v \circ p(x) = y$.

For every $k \in 2\pi\mathbb{Z}^n$, the function $\mathbb{R}^n_s \to \mathbb{R}^n_s$, $x \mapsto \hat{\varphi}(x + k) - \hat{\varphi}(x)$ is continuous and $2\pi\mathbb{Z}^n$-valued. In particular, there exists $A \in \mathrm{GL}_n(\mathbb{Z})$ such that $\hat{\varphi}(x + k) = \hat{\varphi}(x) + Ak$.

We prove that $\varphi : \mathbb{T}^n_{s-\sigma} \to \mathbb{T}^n_s$ is injective. Let $\varphi(p(x)) = \varphi(p(x'))$, with $p(x), p(x') \in \mathbb{T}^n_{s-\sigma}$. Hence $\hat{\varphi}(x') = \hat{\varphi}(x) + k'$, for some $k' \in 2\pi\mathbb{Z}^n$. Hence $\hat{\varphi}(x' - A^{-1}k') = \hat{\varphi}(x)$, and for the injectivity of $\hat{\varphi}$, $p(x) = p(x')$. In particular, φ is biholomorphic.

LEMMA 14. [**14**] *If $G \subset \mathbb{C}^n$ is a domain and $f : G \to \mathbb{C}^n$ injective and holomorphic, then $f(G)$ is a domain and $f : G \to f(G)$ is biholomorphic.*

That $\varphi : \mathbb{T}^n_{s-\sigma} \to \mathbb{T}^n_{s-2\sigma} \subset \varphi(\mathbb{T}^n_{s-\sigma})$ is surjective follows from the surjectivity of $\hat{\varphi}$.

To estimate for $\psi : \mathbb{T}^n_{s-2\sigma} \to \mathbb{T}^n_{s-\sigma}$ the inverse of φ, let $\hat{\psi} : \mathbb{R}^n_{s-2\sigma} \to \mathbb{R}^n_{s-\sigma}$ be the inverse of $\hat{\varphi}$, and $y \in \mathbb{R}^n_{s-2\sigma}$. From the definition of $\hat{\varphi}$, $v \circ p(\hat{\psi}(y)) = y - p(\hat{\psi}(y)) = y - \hat{\psi}(y)$. Hence

$$|\hat{\psi}(y) - y|_{s-2\sigma} = |v \circ p(\hat{\psi}(y))|_{s-2\sigma} \le |v|_{s-2\sigma} \le |v|_{s-\sigma}.$$

Finally, an estimate for $\psi' = \varphi'^{-1} \circ \varphi^{-1}$: we have

$$|\psi' - \mathrm{id}|_{s-2\sigma} \le |\varphi'^{-1} - \mathrm{id}|_{s-\sigma} \le \frac{|\varphi' - \mathrm{id}|_{s-\sigma}}{1 - |\varphi' - \mathrm{id}|_{s-\sigma}} \le \frac{2n}{2n-1} \frac{|v|_s}{\sigma} \le 2 \frac{|v|_s}{\sigma},$$

by triangular and Cauchy inequalities. □

COROLLARY B.1. (Well definition of the operators ϕ) *For all s, σ: if $g \in \mathcal{G}^{\sigma/n}_{s+\sigma}$, then $g^{-1} \in \mathcal{A}(\mathrm{T}^n_s, \mathrm{T}^n_{s+\sigma})$; if $g \in \mathcal{G}^{\omega,\sigma^2/2n}_{s+\sigma}$, then $g^{-1} \in \mathcal{A}(\mathrm{T}^n_s, \mathrm{T}^n_{s+\sigma})$. As a consequence, the operators ϕ in (2.2), (4.1) and (4.10) are well defined.*

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

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

Furthermore,

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

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

Up to rescaling norms by a factor $1/2$ as $\|x\|_s := \frac{1}{2}|x|$, the first statement is straightforward from Theorem B.1. By an abuse of notation, we denote $\|x\|_s$ by $|x|_s$.

Concerning those $g \in \mathcal{G}^{\omega,\sigma^2/2n}_{s+\sigma}$, we recall that g^{-1} is given by

$$g^{-1}(\theta, r) = (\varphi^{-1}(\theta),\ {}^t\varphi' \circ \varphi^{-1}(\theta) \cdot r - \rho \circ \varphi^{-1}(\theta));$$

if $|\varphi^{-1} - \mathrm{id}|_s < \sigma$ and $|\rho|_{s+\sigma} < \sigma/2$ with $|r \cdot \varphi' \circ \varphi^{-1}(\theta)|_s < \sigma/2$, we get the desired thesis. Just note that

$$|\,{}^t(\varphi' - \mathrm{id}) \cdot r|_s \le \frac{n|r|}{\sigma}|\varphi - \mathrm{id}|_{s+\sigma} \le \sigma/2,$$

the factor n coming from the transposition. □

C. *Appendix. Estimates on the Lie brackets of vector fields*

This is just an adaptation to vector fields on $\mathrm{T}^n_{s+\sigma}$ of the analogous lemma for vector fields on the torus $\mathbb{T}^n_s$ in [**22**].

LEMMA 15. *Let f and g be two real analytic vector fields on $\mathrm{T}^n_{s+\sigma}$. The following inequality holds:*

$$|\,[f, g]\,|_s \le \frac{2}{\sigma}\left(1 + \frac{1}{e}\right)|f|_{s+\sigma}|g|_{s+\sigma}.$$

Proof. Consider $f = (f^\theta, f^r) = \sum_{j=1}^n f^{\theta_j}(\partial/\partial\theta_j) + f^{r_j}(\partial/\partial r_j)$ and $g = (g^\theta, g^r) = \sum_{j=1}^n g^{\theta_j}(\partial/\partial\theta_j) + g^{r_j}(\partial/\partial r_j)$. From the definition of the Lie brackets we have $[f, g] = \sum_k f(g^k) - g(f^k)$, where, for every component k,

$$\begin{aligned}[f, g]^k &= \sum_{j=1}^n \left(f^{\theta_j}\frac{\partial g^k}{\partial\theta_j} + f^{r_j}\frac{\partial g^k}{\partial r_j}\right) - \left(g^{\theta_j}\frac{\partial f^k}{\partial\theta_j} + g^{r_j}\frac{\partial f^k}{\partial r_j}\right)\\ &= (Dg \cdot f - Df \cdot g)^k.\end{aligned}$$

We observe that for a holomorphic function $h : \mathrm{T}^n_{s+\sigma} \to \mathbb{C}$,

$$\left|\frac{\partial h}{\partial r_j}\right|_s = \sum_k \left|\frac{\partial h_k(r)}{\partial r_j}\right|_s e^{|k|s} \le \sum_k \frac{1}{\sigma}|h_k(r)|_{s+\sigma}e^{|k|s} \le \frac{1}{\sigma}|h|_{s+\sigma}$$

and

$$\begin{aligned}\left|\frac{\partial h}{\partial\theta_j}\right|_s &= \sum_k |k_j||h_k(r)|_s e^{|k|s} \le \sum_k |k||h_k(r)|_s e^{|k|(s+\sigma)}\, e^{-|k|\sigma}\\ &\le \frac{1}{e\sigma}\sum_k |h_k(r)|_{s+\sigma}e^{|k|(s+\sigma)} = \frac{1}{e\sigma}|h|_{s+\sigma},\end{aligned}$$

where we bound $|k|e^{-|k|\sigma}$ with the maximum attained by $xe^{-x\sigma}$, $x > 0$, in $1/\sigma$, that is, $1/e\sigma$.

Therefore, consider f and g in their Fourier's expansion. We have

$$\begin{aligned}Dg \cdot f &= \sum_{k,\ell} ik \cdot f^\theta_\ell g_k e^{i(k+\ell)\theta} + D_r g_k \cdot f^r_\ell\ e^{i(k+\ell)\cdot\theta}\\ &= \sum_{k,\ell} i\,k \cdot f^\theta_{\ell-k} g_k\, e^{i\ell\cdot\theta} + D_r g_k \cdot f^r_{\ell-k} e^{i\ell\cdot\theta}.\end{aligned}$$

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

Passing to norms, we have the inequality

$$
\begin{aligned}
|Dg \cdot f|_s &\leq \sum_{k,\ell} |k||f^{\theta}_{\ell-k}||g_k|e^{|k|s}e^{|\ell-k|s} + |D_r g_k||f^{r}_{\ell-k}|e^{|k|s}e^{|\ell-k|s} \\
&\leq \sum_{k,\ell} |k||g_k|e^{-|k|\sigma}e^{|k|(s+\sigma)}|f^{\theta}_{\ell-k}|e^{|\ell-k|s} + |D_r g_k|e^{|k|s}\,|f^{r}_{\ell-k}|e^{|\ell-k|s} \\
&\leq \frac{1}{e\sigma}|g|_{s+\sigma}|f|_{s+\sigma} + \frac{1}{\sigma}|g|_{s+\sigma}|f|_{s+\sigma},
\end{aligned}
$$

which follows from the previous remark. Hence the lemma. □

References

[1] R. Abraham and J. Robbin. *Transversal Mappings and Flows*. W. A. Benjamin, New York, 1967, with an appendix by Al Kelley.

[2] V. I. Arnol'd. Small denominators. I. Mapping the circle onto itself. *Izv. Akad. Nauk SSSR Ser. Mat.* **25** (1961), 21–86.

[3] J.-B. Bost. Tores invariants des systèmes dynamiques Hamiltoniens [d'après Kolmogorov, Arnold, Moser, Rüssmann, Zender, Herman, Pöschel, . . .]. Séminaires Bourbaki, Vol. 1984/85, Exp. No. 639. *Astérisque* **133–134** (1986), 113–157.

[4] H. W. Broer, G. B. Huitema, F. Takens and B. L. J. Braaksma. Unfoldings and bifurcations of quasi-periodic tori. *Mem. Amer. Math. Soc.* **83**(421) (1990).

[5] A. Celletti and L. Chierchia. Quasi-periodic attractors in celestial mechanics. *Arch. Ration. Mech. Anal.* **191** (2009), 311–345.

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

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

[8] L. Chierchia. KAM lectures. *Dynamical Systems. Part I (Publ. Centro di Ricerca Matematica Ennio De Giorgi)*. Scuola Normale Superiore, Pisa, 2003, pp. 1–55.

[9] A. C. M. Correia, J. Laskar and R. Dotson. Tidal evolution of exoplanets. *Exoplanets*. University of Arizona Press, Tucson, AZ, 2010, pp. 239–266.

[10] J. Féjoz. Démonstration du 'théorème d'Arnol'd sur la stabilité du système planétaire (d'après Michael Herman). *Ergod. Th. & Dynam. Syst.* **24**(5) (2004), 1521–1582.

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

[12] J. Féjoz. A proof of the invariant torus theorem of Kolmogorov. *Regul. Chaotic Dyn.* **17**(1) (2012), 1–5.

[13] J. Féjoz. Introduction to KAM theory with a view to celestial mechanics. *Variational Methods (Radon Series on Computatinal and Applied Mathematics, 18)*. De Gruyter, Berlin, 2017, pp. 387–433.

[14] K. Fritzsche and H. Grauert. *From Holomorphic Functions to Complex Manifolds (Graduate Texts in Mathematics, 213)*. Springer, New York, 2002.

[15] R. Hamilton. The implicit function theorem of Nash and Moser. *Bull. Amer. Math. Soc. (N.S.)* **7**(1) (1982), 65–222.

[16] M. R. Herman and F. Sergeraert. Sur un théorème d'Arnold et Kolmogorov. *C. R. Acad. Sci. Paris Sér. A* **273** (1971), 409–411.

[17] J. E. Massetti. A normal form à la Moser for diffeomorphisms and a generalization of Rüssmann's translated curve theorem to higher dimensions. *Anal. PDE* **11**(1) (2018), 149–170.

[18] K. R. Meyer. The implicit function theorem and analytic differential equations. *Dynamical Systems—Warwick 1974 (Proc. Symp. Appl. Topology and Dynamical Systems, University of Warwick, Coventry, 1973/1974; presented to E. C. Zeeman on his fiftieth birthday) (Lecture Notes in Mathematics, 468)*. Springer, Berlin, 1975, pp. 191–208.

[19] J. Moser. Convergent series expansions for quasi-periodic motions. *Math. Annal.* **169** (1967), 136–176.

[20] J. Pöschel. On elliptic lower-dimensional tori in Hamiltonian systems. *Math. Z.* **202**(4) (1989), 559–608.

https://doi.org/10.1017/etds.2017.116 Published online by Cambridge University Press

[**21**] J. Pöschel. A lecture on the classical KAM theorem. *Smooth Ergodic Theory and its Applications (Seattle, WA, 1999) (Proceedings of Symposia in Pure Mathematics, 69)*. American Mathematical Society, Providence, RI, 2001, pp. 707–732.

[**22**] J. Pöschel. KAM à la R. *Regul. Chaotic Dyn.* **16**(1–2) (2011), 17–23.

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

[**24**] H. Rüssmann. On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus. *Dynamical Systems, Theory and Applications (Rencontres, Battelle Res. Inst., Seattle, Wash., 1974) (Lecture Notes in Physics, 38)*. Springer, Berlin, 1975, pp. 598–624.

[**25**] M. B. Sevryuk. The lack-of-parameters problem in the KAM theory revisited. *Hamiltonian Systems with Three or More Degrees of Freedom (S'Agaró, 1995) (NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., 533)*. Kluwer, Dordrecht, 1999, pp. 568–572.

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

[**27**] L. Stefanelli and U. Locatelli. Kolmogorov's normal form for equations of motion with dissipative effects. *Discrete Contin. Dyn. Syst. Ser. B* **17**(7) (2012), 2561–2593.

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

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

[**30**] E. Zehnder. Generalized implicit function theorem with applications to some small divisor problems, I. *Comm. Pure Appl. Math.* **28** (1975), 91–140.

[**31**] E. Zehnder. Generalized implicit function theorem with applications to some small divisor problems, II. *Comm. Pure Appl. Math.* **29** (1976), 49–111.