ISSN 1560-3547, Regular and Chaotic Dynamics, 2021, Vol. 26, No. 1, pp. 61–88. © Pleiades Publishing, Ltd., 2021.

V. I. Arnold's "Global" KAM Theorem and Geometric Measure Estimates

Luigi Chierchia[1*] and Comlan E. Koudjinan[2]**

[1]*Dipartimento di Matematica e Fisica, Università "Roma Tre", Largo San Leonardo Murialdo 1, I-00146 Roma, Italy*

[2]*Institute of Science and Technology Austria (IST Austria), Am Campus 1, 3400 Klosterneuburg, Austria*

Received October 26, 2020; revised December 29, 2020; accepted January 04, 2021

Abstract—This paper continues the discussion started in [10] concerning Arnold's legacy on classical KAM theory and (some of) its modern developments. We prove a detailed and explicit "global" Arnold's KAM theorem, which yields, in particular, the Whitney conjugacy of a non-degenerate, real-analytic, nearly-integrable Hamiltonian system to an integrable system on a closed, nowhere dense, positive measure subset of the phase space. Detailed measure estimates on the Kolmogorov set are provided in case the phase space is: (A) a uniform neighbourhood of an arbitrary (bounded) set times the d-torus and (B) a domain with C^2 boundary times the d-torus. All constants are explicitly given.

MSC2010 numbers: 37J40, 37J05, 37J25, 70H08

DOI: 10.1134/S1560354721010044

Keywords: nearly-integrable Hamiltonian systems, perturbation theory, KAM theory, Arnold's scheme, Kolmogorov set, primary invariant tori, Lagrangian tori, measure estimates, small divisors, integrability on nowhere dense sets, Diophantine frequencies

1. INTRODUCTION

a. In [10], we revised Arnold's original analytic "KAM scheme" [2] and showed, in particular, how to implement it so as to get the optimal relation between the size of the perturbation ε and the Diophantine constant α associated to a persistent integrable torus (for generalities, we refer to the Introduction in [10]).

In the present paper we show how Arnold's "pointwise theorem" (Theorem A in [10]) leads naturally to a "global theorem", unifying and improving various previous versions of such a result: compare, in particular, with [9, 14–16]. The term "global" refers here to the simultaneous (and "smooth") construction, in phase space, of all persistent KAM tori having a prefixed Diophantine constant. The main theorem (Theorem 1 below) is formulated in terms of a (Whitney) symplectic transformation conjugating a given (Kolmogorov non-degenerate) analytic, nearly-integrable Hamiltonian system to a Hamiltonian system integrable on a closed, nowhere dense set[1)]. All constants involved in Theorem 1 are explicitly computed, and, in particular, the optimal relation between ε and α is retained.

b. An immediate corollary of "Arnold's global theorem" is that *measure estimates* of the (complement of the) Kolmogorov set (i. e., the set of all persistent integrable tori of a nearly-integrable Hamiltonian system) become essentially trivial (since symplectic transformations preserve Liouville measure on phase space). The problem of finding *explicit measure estimates*

[*]E-mail: luigi@mat.uniroma3.it

[**]E-mail: edmond.koudjinan@ist.ac.at

[1)]Indeed, closed sets of uniform Diophantine numbers may have, in general, isolated points; compare [1].

of the Kolmogorov set in terms of the structure of the phase space is, therefore, reduced to a purely geometrical problem. In particular, as in [5], we are interested in analyzing how such measure estimates depend upon general geometric properties of the action domain, an issue which is particularly relevant in developing KAM theory for *secondary* tori (i. e., those invariant Lagrangian tori which arise because of the perturbation and are not a continuation of integrable tori); compare [3, 4, 6].

In this paper, we shall discuss detailed measure estimates in two different cases, namely:

(A) (*General case*) The Hamiltonian is "uniformly real-analytic" on $\mathcal{D}\times\mathbb{T}^d$, with action domain $\mathcal{D}\subseteq\mathbb{R}^d$ being a *completely arbitrary bounded set*, and the unperturbed frequency map is a local diffeomorphism; "uniformly analytic" means that the Hamiltonian is real-analytic on the union of complex balls with centers in $\mathcal{D}$ and *fixed* radius $\mathtt{R}>0$. In this case the phase space will be $\mathscr{D}\times\mathbb{T}^d$, where $\mathscr{D}$ is a suitable ("minimal") open cover of $\mathcal{D}$. This set-up is similar to that considered in [5].

(B) (*Smooth case*) The Hamiltonian is real-analytic on a phase space $\mathscr{D}\times\mathbb{T}^d$ with $\mathscr{D}$ being a bounded, connected, open set with C^2 boundary and the unperturbed frequency map is a *global* diffeomorphism on $\mathscr{D}$.

c. Let us briefly describe the type of measure estimates we get.

Case (A): As usual in classical KAM theory, we consider real-analytic Hamiltonians

$$H:(y,x)\in\mathscr{D}\times\mathbb{T}^d\mapsto H(y,x):=K(y)+\varepsilon P(y,x)\in\mathbb{R}\,, \qquad (*)$$

where $(y,x)\in\mathbb{R}^d\times\mathbb{T}^d$ are standard action-angle variables (i. e., the phase space is endowed with the standard symplectic form $dy\wedge dx$), ε is a small parameter, and H is real-analytic on the union of $\mathtt{R}$-balls with centers in some bounded set $\mathcal{D}\subseteq\mathbb{R}^d$, while $\mathscr{D}$ is a suitable neighbourhood of $\mathcal{D}$ (see below). The integrable Hamiltonian K is assumed to be Kolmogorov non-degenerate on $\mathscr{D}$ (i. e., the frequency map $y\in\mathscr{D}\mapsto\omega=\partial_y K(y)$ is a real-analytic local diffeomorphism). Let us denote by $\mathscr{K}_{\mathscr{D}}(\alpha,\tau)$ the set of Lagrangian graphs over $\mathbb{T}^d$ in $\mathscr{D}\times\mathbb{T}^d$, which are invariant under the flow governed by H and on which the flow is analytically conjugated to the Kronecker flow $x\in\mathbb{T}^d\mapsto x+\omega t$, with $\omega\in\mathbb{R}^d$ (α,τ)-Diophantine[2)], for some $\tau>d-1$. Then there exist positive numbers C_*, α_*, ε_* and $\mathtt{r}\leqslant\mathtt{R}/9$, depending only on d, τ, K and P (and explicitly given in Theorem 4 below), such that, if $0<\varepsilon<\varepsilon_*$, then

$$\operatorname{meas}\Big((\mathscr{D}\times\mathbb{T}^d)\backslash\mathscr{K}_{\mathscr{D}}(\alpha_*\sqrt{\varepsilon},\tau)\Big)\leqslant C_*\,N_{\mathtt{r}}^{\rm int}(\mathcal{D})\,\sqrt{\varepsilon},$$

where $N_{\mathtt{r}}^{\rm int}(\mathcal{D})$ is the so-called $\mathtt{r}$-internal covering number of $\mathcal{D}$ and $\mathscr{D}$ is a $\mathtt{R}$-neighbourhood of a minimal $\mathtt{R}$-internal cover of $\mathcal{D}$ (compare Section 3.1 for precise definitions).

Case (B): Here H is as above, but $\mathscr{D}$ is assumed to be an open, bounded, connected set with C^2 boundary; H is $\mathtt{R}$-uniformly real-analytic on $\mathscr{D}$ and the unperturbed frequency map is assumed to be a global diffeomorphism on $\mathscr{D}$. Let

$$\mathtt{r}:=\min\{\mathtt{R},\ \operatorname{minfoc}(\partial\mathscr{D})\,,\,1/\kappa\}/\sqrt{d},$$

where "minfoc" denotes the so-called minimal focal distance, and κ is the maximum modulus of the principal curvatures of $\partial\mathscr{D}$. Then there exist positive numbers $\bar{C}_*$, α_*, and ε_* depending only on d, τ, K and P (and explicitly given in Theorem 5 below) such that, if $0<\varepsilon<\varepsilon_*$, then

$$\operatorname{meas}\Big((\mathscr{D}\times\mathbb{T}^d)\backslash\mathscr{K}_{\mathscr{D}}(\alpha_*\sqrt{\varepsilon},\tau)\Big)\leqslant\bar{C}_*\,\max\big\{\sec_{d-1}(\mathscr{D})\,,\,\mathcal{H}^{d-1}(\partial\mathscr{D})\big\}\,\sqrt{\varepsilon},$$

where $\sec_{d-1}(\mathscr{D})$ is the measure of the maximal $(d-1)$-dimensional section of $\mathscr{D}$ and $\mathcal{H}^{d-1}$ denotes the $(d-1)$-dimensional Hausdorff measure (compare Section 3.2 for precise definitions).

2) i. e., $|\omega\cdot k|\geqslant\alpha/|k|^\tau$, for any $k\in\mathbb{Z}^d\backslash\{0\}$.

d. Remarks

(i) For the optimality of the relation between ε and α (and the reason for choosing $\alpha = \alpha_* \sqrt{\varepsilon}$ in the Kolmogorov set), see item **d** in the Introduction of [10].

(ii) Theorem 4 below extends and generalizes the main result (Theorem 1) in [5].

(iii) In Appendix A (see, in particular, Remark A.5), we correct a small flaw (concerning the choice of some constants) in [10].

(iv) In Remark A.4 (Appendix A) all constants appearing in the proof are explicitly given.

e. The paper is organized as follows.

In Section 2.1 we introduce some of the notation used in the paper and in Section 2.2 we state the "global Arnold theorem" (Theorem 1). The statement of such a theorem is quite detailed; in particular, the introduction of apparently arbitrary sets of parameters (such as $\mathcal{D}_0$ or ρ) allows applications to be made in quite different circumstances (such as cases (A) and (B) mentioned above). On the other hand, the proof of this theorem does not really contain novel ideas and is based on the schemes in [2, 12] and [10]. However, since we put some emphasis in making everything explicit, we felt it necessary to outline the proof, detailing, in particular, the choice of the (many) parameters involved (this is done in Appendix A).

Section 3 is devoted to measure estimates and, in particular, to the statements and proofs of Theorem 4 and 5, which have been briefly explained in item **c** above.

Finally, Appendix B contains some of the technical tools used in the paper, namely:

B.1 Classical estimates (Cauchy, Fourier)
B.2 An inverse function theorem
B.3 Internal coverings
B.4 Extensions of Lipschitz continuous functions
B.5 Lebesgue measure and Lipschitz continuous map
B.6 Lipeomorphisms "close" to identity
B.7 Whitney smoothness
B.8 Measure of tubular neighbourhoods of hypersurfaces
B.9 Kolmogorov non-degenerate normal forms.

2. ARNOLD'S GLOBAL KAM THEOREM

2.1. Notations

- $\mathbb{N} := \{1, 2, 3, \cdots\}$ and $\mathbb{N}_0 := \{0, 1, 2, 3, \cdots\}$.

- For $d \in \mathbb{N}$ and $x, y \in \mathbb{C}^d$, we let $x \cdot y := x_1 \bar{y}_1 + \cdots + x_d \bar{y}_d$ be the standard inner product (the bar denotes complex conjugate). We denote, respectively, the sup-norm, the 1-norm and the Euclidean norm, by:

$$|x| := \max_{1 \leqslant j \leqslant d} |x_j|, \qquad |x|_1 := \sum_{j=1}^{d} |x_j|, \qquad |x|_2 := \sqrt{\sum_{j=1}^{d} |x_j|^2}.$$

- $\mathbb{T}^d := \mathbb{R}^d / 2\pi\mathbb{Z}^d$ is the d-dimensional (flat) torus.

- Given $\alpha > 0$, $\tau \geqslant d - 1 \geqslant 1$, we denote by

$$\mathrm{Dioph}^{\tau}_{\alpha} := \left\{ \omega \in \mathbb{R}^d : |\omega \cdot k| \geqslant \frac{\alpha}{|k|_1^{\tau}}, \quad \forall\, 0 \neq k \in \mathbb{Z}^d \right\} \tag{2.1}$$

the set of (α, τ)-*Diophantine vectors* in $\mathbb{R}^d$.

- For $r, s > 0$, $y_0 \in \mathbb{C}^d$, $\emptyset \neq D \subseteq \mathbb{C}^d$, we denote:

$$\begin{aligned}
\mathrm{B}_r(y_0) &\coloneqq \Big\{ y \in \mathbb{R}^d : |y - y_0| < r \Big\} , && (y_0 \in \mathbb{R}^d) ,\\
\mathrm{B}_r(D) &\coloneqq \bigcup_{y_0 \in D} \mathrm{B}_r(y_0) , && (D \subseteq \mathbb{R}^d) ,\\
\mathbb{B}_r(y_0) &\coloneqq \Big\{ y \in \mathbb{C}^d : |y - y_0| < r \Big\} ,\\
\mathbb{B}_r(D) &\coloneqq \bigcup_{y_0 \in D} \mathbb{B}_r(y_0) ,\\
\mathbb{T}^d_s &\coloneqq \Big\{ x \in \mathbb{C}^d : |\operatorname{Im} x| < s \Big\} / 2\pi\mathbb{Z}^d ,\\
\mathbb{B}_{r,s}(y_0) &\coloneqq \mathbb{B}_r(y_0) \times \mathbb{T}^d_s ,\\
\mathbb{B}_{r,s}(D) &\coloneqq \mathbb{B}_r(D) \times \mathbb{T}^d_s ;
\end{aligned}$$

  we shall also denote, in bold face characters, Euclidean balls:

$$\begin{aligned}
\mathbf{B}_r(y_0) &\coloneqq \Big\{ y \in \mathbb{R}^d : |y - y_0|_2 < r \Big\} , && (y_0 \in \mathbb{R}^d),\\
\mathbf{B}_r(D) &\coloneqq \bigcup_{y_0 \in D} \mathbf{B}_r(y_0), && (D \subseteq \mathbb{R}^d).
\end{aligned}$$

- If $\mathbb{1}_d \coloneqq \operatorname{diag}(1)$ is the unit $(d \times d)$ matrix, we denote the standard symplectic matrix by

$$\mathbb{J} \coloneqq \begin{pmatrix} 0 & -\mathbb{1}_d \\ \mathbb{1}_d & 0 \end{pmatrix} .$$

- For $D \subseteq \mathbb{R}^d$, $r \geqslant 0$ and $s > 0$, $\mathcal{B}_{r,s}(D)$ denotes the Banach space of real-analytic functions

$$f : \mathrm{B}_r(D) \times \mathbb{T}_s \to \mathbb{C}$$

  with bounded holomorphic extensions to $\mathbb{B}_{r,s}(D)$, with uniform norm

$$\|f\|_{r,s} \coloneqq \|f\|_{r,s,D} \coloneqq \sup_{\mathbb{B}_{r,s}(D)} |f| < \infty.$$

  Analogously, $\mathcal{B}_r(D)$ denotes the Banach space of real-analytic functions

$$f : \mathrm{B}_r(D) \to \mathbb{C}$$

  with bounded holomorphic extensions to $\mathbb{B}_r(D)$, with

$$\|f\|_r \coloneqq \|f\|_{r,D} \coloneqq \sup_{\mathbb{B}_r(D)} |f| < \infty.$$

- For a differentiable function $f : A \subseteq \mathbb{C}^d \times \mathbb{C}^d \ni (y, x) \mapsto f(y, x) \in \mathbb{C}$, its gradient/Jacobian is denoted by ∇f or by f'.

- We equip $\mathbb{C}^d \times \mathbb{C}^d$ (and its subsets) with the canonical symplectic form

$$\varpi \coloneqq dy \wedge dx = dy_1 \wedge dx_1 + \cdots + dy_d \wedge dx_d ,$$

  and denote by ϕ^t_H the associated Hamiltonian flow governed by the Hamiltonian $H(y, x)$, $y, x \in \mathbb{C}^d$, i. e., $z(t) \coloneqq \phi^t_H(z)$ is the unique solution of

$$\dot{z} = \mathbb{J}\nabla H , \qquad z(0) = z.$$

- Given a linear operator L from the normed space $(V_{\rm a}, \|\cdot\|_{\rm a})$ into the normed space $(V_{\rm b}, \|\cdot\|_{\rm b})$, its "operator norm" is given by
$$\|L\| := \sup_{x\in V_{\rm a}\setminus\{0\}} \frac{\|Lx\|_{\rm b}}{\|x\|_{\rm a}}, \quad \text{so that} \quad \|Lx\|_{\rm b} \leqslant \|L\|\,\|x\|_{\rm a} \quad \text{for any} \quad x \in V_{\rm a}.$$

- Given $\omega \in \mathbb{R}^d$, the directional derivative of a C^1 function f with respect to ω is given by
$$D_\omega f := \omega \cdot f_x = \sum_{j=1}^{d} \omega_j f_{x_j}.$$

- If f is a (smooth or analytic) function on $\mathbb{T}^d$, its Fourier expansion is given by
$$f = \sum_{k\in\mathbb{Z}^d} f_k\, e^{ik\cdot x}, \qquad f_k := \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} f(x)\, e^{-ik\cdot x}\, dx,$$
(where, as usual, $e := \exp(1)$ denotes the Neper number, and i, the imaginary unit). We also set:
$$\langle f\rangle := f_0 = \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} f(x)\, dx\,, \qquad T_N f := \sum_{|k|_1\leqslant N} f_k\, e^{ik\cdot x},\ N > 0.$$

- For a function $f\colon (\mathscr{M}_1, {\rm d}_1) \to (\mathscr{M}_2, {\rm d}_2)$, where $(\mathscr{M}_j, {\rm d}_j)$, $j = 1, 2$ are metric spaces, we denote
$${\rm Lip}_{\mathscr{M}_1}(f) := \|f\|_{L,\mathscr{M}_1} := \sup_{x\neq x'\in\mathscr{M}_1} \frac{{\rm d}_2(f(x), f(x'))}{{\rm d}_1(x, x')} \leqslant \infty,$$
and f is said to be Lipschitz continuous on $\mathscr{M}_1$ if ${\rm Lip}_{\mathscr{M}_1}(f) < \infty$.
If $\mathscr{M}_1 = \mathbb{R}^d$, we usually denote ${\rm Lip}_{\mathbb{R}^d}(f) = {\rm Lip}(f)$.

- $C^k_W(D)$ denotes the set of functions which are C^k in the sense of Whitney on the set D. A C^1_W map $\phi : D\times\mathbb{T}^d \to \mathbb{R}^d\times\mathbb{T}^d$ is symplectic if the Whitney gradient $\nabla\phi = (\partial_y\phi, \partial_x\phi)$ satisfies $(\nabla\phi)\mathbb{J}(\nabla\phi)^T = \mathbb{J}$ on $D\times\mathbb{T}^d$. For more details, see Appendix B.7.

- The s-dimensional Hausdorff measure on $\mathbb{R}^d$ will be denoted by $\mathcal{H}^s$; in particular, $\mathcal{H}^d$, which coincides with the d-dimensional outer Lebesgue measure, will be denoted by "meas".

2.2. KAM Theorem

Given an open set $\mathscr{D} \subseteq \mathbb{R}^d$ and a real-analytic Hamiltonian $H : \mathscr{D}\times\mathbb{T}^d \to \mathbb{R}$, we say that $\mathcal{T} \subseteq \mathscr{D}\times\mathbb{T}^d$ *is a (primary*[3]*) Kolmogorov (or "KAM") torus for* H if $\mathcal{T}$ is a real-analytic Lagrangian embedded torus $\mathcal{T} = \phi(\mathbb{T}^d)$, which is a graph over $\mathbb{T}^d$, and such that
$$\phi^t_H\big(\phi(\theta)\big) = \phi(\theta + \omega t)\,, \qquad \forall\,\theta\in\mathbb{T}^d, t\in\mathbb{R},$$
for a given Diophantine "frequency vector" $\omega \in {\rm Dioph}^\tau_\alpha$ (for some $\alpha, \tau > 0$).

Theorem 1. *Let* $d \geqslant 2$*;* $\mathtt{R} > 0$*;* $0 < s \leqslant 1$*;* $\emptyset \neq \mathcal{D} \subseteq \mathbb{R}^d$*;* $\varepsilon, \alpha > 0$*. Let the "integrable Hamiltonian"* $K \in \mathcal{B}_{\mathtt{R}}(\mathcal{D})$ *be a uniformly (Kolmogorov) non-degenerate (i.e.,* $\det K_{yy} \neq 0$ *on* $\mathbb{B}_{\mathtt{R}}(\mathcal{D})$*) and let the "perturbation"* P *belong to* $\mathcal{B}_{\mathtt{R},s}(\mathcal{D})$*. Define*
$$\mathsf{M} := \|K_{yy}\|_{\mathtt{R},\mathcal{D}}\,, \quad \mathsf{L} := \|K_{yy}^{-1}\|_{\mathtt{R},\mathcal{D}}, \quad \mathsf{P} := \|P\|_{\mathtt{R},s,\mathcal{D}}\,, \quad \theta := \mathsf{ML}, \quad \epsilon := \varepsilon\,\frac{\mathsf{MP}}{\alpha^2}. \tag{2.2}$$

[3] As opposed to *secondary tori* (the same definition, but removing the graph assumption); for a KAM theory for secondary tori, see [3]. In this paper, we shall only consider primary KAM tori.

Choose $0 < \rho < \mathsf{r} \leqslant \mathtt{R}$, $\mathcal{D}_0 \subseteq \mathcal{D}$, $\tau \geqslant d-1$; define the following "action domains":

$$\mathscr{D} := \mathrm{B}_{\mathsf{r}}(\mathcal{D}_0), \qquad \widehat{\mathscr{D}} := \mathrm{B}_{\mathsf{r}-\rho}(\mathcal{D}_0), \qquad \mathscr{D}^* := \{y \in \widehat{\mathscr{D}} : \ K_y(y) \in \mathrm{Dioph}^\tau_\alpha\}, \tag{2.3}$$

and consider the "nearly-integrable", non-degenerate Hamiltonian given by

$$H : (y,x) \in \mathscr{D} \times \mathbb{T}^d \mapsto H(y,x) := K(y) + \varepsilon P(y,x) \in \mathbb{R}\,;$$

the "phase space" $\mathscr{D} \times \mathbb{T}^d$ being endowed with the standard symplectic form ϖ. Fix $0 < s_* < s$. There exist constants $c_*, c_0, c_1, c_2, c_3, c_4 > 1$, depending only on d and τ, such that, if

$$\alpha \leqslant c_0\, \frac{\rho}{\mathsf{L}}; \qquad\qquad \epsilon \leqslant \epsilon_* := \frac{(s-s_*)^a}{c_*\, \theta^6}, \tag{2.4}$$

with $a := 7\nu + 4d + 2$ and $\nu := \tau + 1$, then the following statements hold.
There exists a nowhere dense set $\mathscr{D}_* \subseteq \mathrm{B}_{\mathsf{r}-\frac{\rho}{2}}(\mathcal{D}_0) \subseteq \mathscr{D}$, a lipeomorphism

$$Y^* \colon \mathscr{D}^* \overset{onto}{\longrightarrow} \mathscr{D}_*,$$

a function $K_* \in C^\infty_W(\mathscr{D}_*)$ and a C^∞_W-symplectic transformation

$$\phi_* := \mathtt{id} + (v_*, u_*) \colon \mathscr{D}_* \times \mathbb{T}^d \to \mathscr{K} := \phi_*(\mathscr{D}_* \times \mathbb{T}^d) \subseteq \mathscr{D} \times \mathbb{T}^d\,, \tag{2.5}$$

real-analytic in $x \in \mathbb{T}^d_{s_*}$, such that[4]

$$\partial_{y_*} K_* \circ Y^* = \partial_y K\,, \qquad\qquad on \quad \mathscr{D}^*\,, \tag{2.6}$$

$$\partial^\beta_{y_*}(H \circ \phi_*)(y_*, x) = \partial^\beta_{y_*} K_*(y_*), \qquad \forall\, (y_*, x) \in \mathscr{D}_* \times \mathbb{T}^d, \quad \forall\, \beta \in \mathbb{N}^d_0. \tag{2.7}$$

Furthermore, the following estimates hold:

$$\|Y^* - \mathtt{id}\|_{\mathscr{D}^*} \leqslant c_1\, (s-s_*)^\nu\, \theta^2\, \frac{\varepsilon \mathsf{P}}{\alpha}, \tag{2.8}$$

$$\mathrm{Lip}_{\mathscr{D}^*}(Y^* - \mathtt{id}) \leqslant c_2\, \theta^3\, \ (s-s_*)^{-1}\, \frac{\mathsf{M}\varepsilon\mathsf{P}}{\alpha^2}\, \left(\log \frac{\alpha^2}{\mathsf{M}\varepsilon\mathsf{P}}\right)^\nu \leqslant \frac{1}{4d}, \tag{2.9}$$

$$\max\left\{\|u_*\|_*,\ 2d\sqrt{2}\,\frac{\mathsf{M}\ell^\nu}{\alpha}\|v_*\|_*\right\} \leqslant c_3\, \ell^\nu\, \theta^2\, \frac{\mathsf{M}\varepsilon\mathsf{P}}{\alpha^2}, \tag{2.10}$$

$$\|\partial_x u_*\|_* \leqslant c_4\, \theta_0^2\, \ell^\nu\, \frac{\mathsf{M}\varepsilon\mathsf{P}}{\alpha^2} \leqslant \frac{1}{4(18d^3+70)\theta}\,, \tag{2.11}$$

where

$$\|\cdot\|_* := \sup_{\mathscr{D}_* \times \mathbb{T}^d_{s_*}} |\cdot|, \qquad \ell := 8(s-s_*)^{-1} \log \epsilon^{-1}\,.$$

The "Kolmogorov set" $\mathscr{K}$ defined in (2.5) is foliated, as $y_* \in \mathscr{D}_*$, by Kolmogorov tori $\mathcal{T}_* := \phi_*(\{y_*\} \times \mathbb{T}^d)$, which are Kolmogorov non-degenerate[5].

The proof of this theorem is based upon Arnold's original KAM scheme, revised and improved in [10], where, in particular, all constants are computed and optimal smallness conditions concerning the relation between small divisors and smallness of the perturbation are given. Since essentially no new ideas are needed, details are deferred to Appendix A.
However, let us make here a few observations.

Remark 1. (i) The hypotheses on H can be rephrased by saying that H is $\mathtt{R}$-uniformly real-analytic on $\mathcal{D}$. Notice that $\mathcal{D}$ can be a completely arbitrary subset of $\mathbb{R}^d$, but $\mathscr{D}$ and $\widehat{\mathscr{D}}$ are open sets.

(ii) The introduction of $\mathcal{D}_0$ and ρ is made in order to be able to apply the theorem in quite different contexts; compare, e. g., the next section on measure estimates.

[4] y_*-derivatives are Whitney derivatives.
[5] For a precise definition, see Appendices A and B.9.

(iii) Even if $\mathcal{D}_0$ is a single point, the theorem guarantees, in general, a set of positive measure of Kolmogorov tori for H, since the set $\mathscr{D}^*$ is a set of positive measure, provided $\tau > d-1$ and α is small enough. Precise measure estimates are one of the objectives of this paper and will be given in the next section.

(iv) The parameter θ defined in (2.2) measures the "torsion" of the unperturbed system and is always greater than or equal to 1; indeed, for any $y_0 \in \mathscr{D}$, denoting $T(y) := K_{yy}(y)^{-1}$, one has

$$\theta := \mathsf{LM} \geqslant \|T(y_0)\| \|K_{yy}(y_0)\| = \|T(y_0)\| \|T(y_0)^{-1}\| \geqslant 1. \tag{2.12}$$

(v) The constants c_i appearing in the theorem are explicitly given in Appendix A; compare, in particular, Eq. (A.38).

3. MEASURE ESTIMATES

The fact that the Kolmogorov set $\mathscr{K}$ in Theorem 1 is the image of a (Whitney) symplectic map leads to straightforward measure estimates of its complement:

Theorem 2. *Under the same notations and assumptions of Theorem* 1*, let*

$$\begin{aligned}
&\beta := \big(1 + 2\,\mathrm{Lip}_{\mathscr{D}^*}(Y^* - \mathtt{id})\big)^d (2\pi)^d,\\
&\mathscr{T}_\rho := \mathrm{B}_{\mathsf{r}+\rho}(\mathcal{D}_0) \setminus \mathrm{B}_{\mathsf{r}-\rho}(\mathcal{D}_0),\\
&\mathcal{R}_\alpha := \big\{y \in \mathscr{D} : K_y(y) \notin \mathrm{Dioph}^\tau_\alpha\big\}.
\end{aligned} \tag{3.1}$$

Then one has

$$\begin{aligned}
\mathrm{meas}\,(\mathscr{D} \times \mathbb{T}^d \setminus \mathscr{K}) &\leqslant \beta\ \mathrm{meas}\,\big(\mathrm{B}_{\frac{\rho}{2}}(\mathscr{D}) \setminus \mathscr{D}^*\big)\\
&\leqslant \beta\big(\,\mathrm{meas}\,(\mathscr{T}_\rho) + \mathrm{meas}\,(\mathcal{R}_\alpha)\big).
\end{aligned} \tag{3.2}$$

Proof. By Theorem B.2, we can extend $Y^* - \mathtt{id}$ componentwise to obtain a global Lipschitz continuous function $f\colon \mathbb{R}^d \circlearrowleft$ satisfying $f|_{\mathscr{D}^*} = Y^* - \mathtt{id}$ and

$$\sup_{\mathbb{R}^d} |f| = \sup_{\mathscr{D}^*} |Y^* - \mathtt{id}| \overset{(2.8),(2.4)}{\leqslant} \frac{\rho}{2}\,, \qquad \mathrm{Lip}_{\mathbb{R}^d}(f) = \mathrm{Lip}_{\mathscr{D}^*}(Y^* - \mathtt{id}) < \frac{1}{4d}. \tag{3.3}$$

Set $g := f + \mathtt{id}$. Then, by Lemma 7 and (3.3), one has[6]

$$\mathscr{D} \subseteq g\big(\overline{B_{\frac{\rho}{2}}(\mathscr{D})}\big). \tag{3.4}$$

Notice also that, by (3.3) and Lemma 7, g is a lipeomorphism of $\mathbb{R}^d$. Consequently,

$$\begin{aligned}
\mathrm{meas}\,(\mathscr{D} \times \mathbb{T}^d \setminus \mathscr{K}) &= \mathrm{meas}\,(\mathscr{D} \times \mathbb{T}^d) - \mathrm{meas}\,\big(\phi_*(\mathscr{D}_* \times \mathbb{T}^d)\big)\\
&= \mathrm{meas}\,(\mathscr{D} \times \mathbb{T}^d) - \mathrm{meas}\,(\mathscr{D}_* \times \mathbb{T}^d)\\
&= (2\pi)^d\big(\,\mathrm{meas}\,(\mathscr{D}) - \mathrm{meas}\,(\mathscr{D}_*)\big)\\
&\overset{(3.4)}{\leqslant} (2\pi)^d\big(\,\mathrm{meas}\,\big(g(\overline{B_{\frac{\rho}{2}}(\mathscr{D})})\big) - \mathrm{meas}\,(\mathscr{D}_*)\big)\\
&= (2\pi)^d\,\mathrm{meas}\,\big(g(\overline{B_{\frac{\rho}{2}}(\mathscr{D})}) \setminus g(\mathscr{D}^*)\big)\\
&= (2\pi)^d\,\mathrm{meas}\,\big(g\big(\overline{B_{\frac{\rho}{2}}(\mathscr{D})} \setminus \mathscr{D}^*\big)\big) \qquad \text{(because } g \text{ is injective)}\\
&\overset{(B.11)}{\leqslant} (2\pi)^d(\mathrm{Lip}\,g)^d\,\mathrm{meas}\,\big(\overline{B_{\frac{\rho}{2}}(\mathscr{D})} \setminus \mathscr{D}^*\big)\\
&\overset{(3.3)}{\leqslant} (2\pi)^d(1 + 2\,\mathrm{Lip}(Y^* - \mathtt{id}))^d\,\mathrm{meas}\,\big(B_{\frac{\rho}{2}}(\mathscr{D}) \setminus \mathscr{D}^*\big).
\end{aligned}$$

[6] The bar on sets denotes closure.

Finally, recalling that $\mathscr{D} = \mathrm{B}_{\mathsf{r}}(\mathcal{D}_0)$ and (2.3), one sees that

$$\begin{aligned} \mathrm{B}_{\frac{\varrho}{2}}(\mathscr{D}) \setminus \mathscr{D}^* &= \mathrm{B}_{\frac{\varrho}{2}}(\mathscr{D}) \setminus \widehat{\mathscr{D}} \,\dot\cup\, \widehat{\mathscr{D}} \setminus \mathscr{D}^* \\ &= \mathrm{B}_{\mathsf{r}+\frac{\varrho}{2}}(\mathcal{D}_0) \setminus \mathrm{B}_{\mathsf{r}-\rho}(\mathcal{D}_0) \,\dot\cup\, \big\{ y \in \mathrm{B}_{\mathsf{r}-\rho}(\mathcal{D}_0) : K_y(y) \notin \mathrm{Dioph}^\tau_\alpha \big\} \\ &\subseteq \mathscr{T}_\rho \cup \mathcal{R}_\alpha, \end{aligned}$$

from which the second inequality in (3.2) follows at once. □

Theorem 2 reduces the problem of estimating the measure of the complement of the Kolmogorov set $\mathscr{K}$ to the estimate on the measure of the complement of Diophantine numbers in a given set and to the purely geometrical problem of estimating the measure of the *tubular neighbourhood* $\mathscr{T}_\rho$ of the boundary of $\mathscr{D} = \mathrm{B}_{\mathsf{r}}(\mathcal{D}_0)$. Therefore, concrete measure estimates will depend upon the structure of the action domain $\mathscr{D}$ and of the (unperturbed) frequency map

$$y \in \mathscr{D} \mapsto \omega_0(y) \coloneqq K_y(y) \in \mathbb{R}^d. \tag{3.5}$$

We shall discuss in detail two different cases:

(A) (*General case*) $\mathcal{D}$ is an *arbitrary bounded* set, H uniformly real-analytic on $\mathcal{D} \times \mathbb{T}^d$, and ω_0 is a *local* diffeomorphism on $\mathcal{D}$ (which is always the case if the unperturbed Hamiltonian is assumed to be Kolmogorov non-degenerate). In this case, as phase space we shall consider a "minimal" (in a suitable sense) open cover of $\mathcal{D}$ times $\mathbb{T}^d$. This set-up is analogous to that considered in [5].

(B) (*Smooth case*) $\mathcal{D}$ is a bounded, connected, open set with C^2 boundary and ω_0 is a *global* diffeomorphism on $\mathcal{D}$. In this case the phase space is just $\mathscr{D} \times \mathbb{T}^d \coloneqq \mathcal{D} \times \mathbb{T}^d$.

3.1. General Case

In order to state the result for case (A), let us give two definitions.

- Given a bounded non-empty set $\mathcal{D} \subseteq \mathbb{R}^d$, and given $r > 0$, an **r-internal covering of $\mathcal{D}$** is a subset $\mathcal{D}_0$ of $\mathcal{D}$ such that

$$\mathcal{D} \subseteq \mathrm{B}_r(\mathcal{D}_0) = \bigcup_{y \in \mathcal{D}_0} \mathrm{B}_r(y); \tag{3.6}$$

  $N_r^{\mathrm{int}}(\mathcal{D})$, the **$r$–internal covering number of $\mathcal{D}$**, is defined as[7)]

$$N_r^{\mathrm{int}} \coloneqq \min \big\{ n \in \mathbb{N} : \ \{y_1, \ldots, y_n\} \text{ is an } r-\text{internal covering of } \mathcal{D} \big\}; \tag{3.7}$$

  an r-internal cover $\mathcal{D}_0$ of $\mathcal{D}$ with cardinality equal to the r-internal covering number will be called **a minimal r-internal covering of $\mathcal{D}$**.

- Given a real-analytic Hamiltonian $H : \mathscr{D} \times \mathbb{T}^d \to \mathbb{R}$, we denote the set of KAM tori for H with frequency vector in $\mathrm{Dioph}^\tau_\alpha$ by

$$\mathscr{K}_{\mathscr{D}}(\alpha, \tau) \coloneqq \big\{ \mathcal{T} \subseteq \mathscr{D} \times \mathbb{T}^d | \, \mathcal{T} \text{ is a KAM torus for } H \text{ with frequency } \omega \in \mathrm{Dioph}^\tau_\alpha \big\}. \tag{3.8}$$

[7)] $N_r^{\mathrm{int}}(\mathcal{D})$ is finite if and only if $\mathcal{D}$ is bounded. A simple upper bound on $N_r^{\mathrm{int}}(\mathcal{D})$ for bounded domains $\mathcal{D}$ is: $N_r^{\mathrm{int}}(\mathcal{D}) \leqslant ([\mathrm{diam}(\mathcal{D})/r] + 1)^d$; compare [5] or Appendix B, Section B.3.

Theorem 3. *Let $\mathcal{D}$ be an arbitrary bounded non-empty set in $\mathbb{R}^d$, $\tau > d-1 \geqslant 1$, $\mathtt{R}, s > 0$. Let $K \in \mathcal{B}_{\mathtt{R}}(\mathcal{D})$ be uniformly (Kolmogorov) non-degenerate, $P \in \mathcal{B}_{\mathtt{R},s}(\mathcal{D})$ and let M, L, P, θ as in (2.2). Let c_0 and c_* be as in Theorem 1. Fix $0 < s_* < s$, let ϵ_* be as in (2.4) and define:*

$$\mathtt{r} := \frac{\mathtt{R}}{1+2d^2\theta}, \qquad \alpha_* := \sqrt{\frac{\mathsf{MP}}{\epsilon_*}}, \qquad \varepsilon_* := \Big(\frac{c_0 \mathtt{r}}{\mathsf{L}\alpha_*}\Big)^2,$$
$$\delta_0 := \inf_{\mathrm{B}_{\mathtt{r}}(\mathcal{D})} |\det K_{yy}|, \quad \theta_0 := \max\Big\{\frac{\mathsf{M}^d}{\delta_0}, \theta\Big\}, \quad \rho := \frac{\alpha_* \mathsf{L}}{c_0}\sqrt{\varepsilon}. \tag{3.9}$$

Let $\mathcal{D}_0 \subseteq \mathcal{D}$ be a minimal $\mathtt{r}$-internal covering of $\mathcal{D}$, $\mathscr{D} := \mathrm{B}_{\mathtt{r}}(\mathcal{D}_0)$ and let $\mathscr{K}_{\mathscr{D}}(\alpha_*\sqrt{\varepsilon}, \tau)$ be as in (3.8) with $H = K + \varepsilon P$. Then, if $0 < \varepsilon < \varepsilon_*$, one has

$$\mathrm{meas}\Big((\mathscr{D} \times \mathbb{T}^d) \setminus \mathscr{K}_{\mathscr{D}}(\alpha_*\sqrt{\varepsilon}, \tau)\Big) \leqslant \bar{c}_*\, \theta_0\; N^{\mathrm{int}}_{\mathtt{r}}(\mathcal{D})\; \mathsf{M}^{-1}\; \mathtt{r}^{d-1}\; \alpha_*\; \sqrt{\varepsilon}, \tag{3.10}$$

with

$$\bar{c}_* := \frac{5}{4}\,(2\pi)^d \Big(\frac{d 2^{2d}}{c_0} + 2^d d^{\frac{d-1}{2}} \sum_{k \in \mathbb{Z}^d \setminus \{0\}} \frac{1}{|k|_1^\tau\, |k|_2}\Big). \tag{3.11}$$

Proof. Let $\alpha := \alpha_*\sqrt{\varepsilon}$. Then $\rho = \mathsf{L}\alpha/c_0$, so that the first inequality in (2.4) is satisfied (with the equal sign). Furthermore, with the above positions, ϵ in (2.2) is given by

$$\epsilon = \frac{\mathsf{MP}}{\alpha_*^2},$$

so that the second inequality in (2.4) is also satisfied (with the equal sign). Finally, the relation $\rho < \mathtt{r}$ is equivalent to $\varepsilon < \varepsilon_*$, which is satisfied by hypothesis. Hence, all the assumptions of Theorem 1 are satisfied and therefore the measure estimate (3.2) holds with $\mathscr{K}$ as in (2.5).

We proceed to estimate the two terms on the right-hand side of (3.2) separately. Let us first discuss the measure of $\mathcal{R}_\alpha$.

We claim that *the map $y \in \mathrm{B}_{\mathtt{r}}(y_0) \mapsto \omega_0(y)$ is a diffeomorphism for every $y_0 \in \mathcal{D}$*. To see this, we shall apply the quantitative inverse function theorem B.1 to $f(y) = K_y(y)$. In such a case, we can take $T = K_{yy}(y_0)^{-1}$ (using Cauchy estimates, see Lemma 4),

$$\begin{aligned}\|\mathbb{1}_d - T K_{yy}(y)\| &\leqslant \|T\| \|K_{yy}(y_0) - K_{yy}(y)\| \\ &\leqslant d^2 \mathsf{L} \|\partial_y K_{yy}\|_{\mathtt{r}}\, \mathtt{r} \leqslant d^2 \mathsf{L}\, \frac{\|K_{yy}\|_{\mathtt{R}}}{\mathtt{R}-\mathtt{r}}\, \mathtt{r} \\ &\leqslant d^2\, \mathsf{LM}\, \frac{\mathtt{r}}{\mathtt{R}-\mathtt{r}} = \frac{1}{2},\end{aligned}$$

where $\|\partial_y K_{yy}\|_{\mathtt{r}} := \sup_{\mathrm{B}_{\mathtt{r}}(y_0)} \max\{|\partial^3_{y_i y_j y_k} K| : i,j,k = 1,\cdots,d\}$. Hence, by Theorem B.1, ω_0 is invertible on any ball $\mathrm{B}_{\mathtt{r}}(y_0)$ with $y_0 \in \mathcal{D}$, as claimed.

Now let $\mathcal{D}_0 = \{y_1, \ldots, y_{n_0}\}$ with $n_0 := N^{\mathrm{int}}_{\mathtt{r}}(\mathcal{D})$. Then

$$\begin{aligned}\mathrm{meas}\,(\mathcal{R}_\alpha) &\leqslant \sum_{j=1}^{n_0} \mathrm{meas}\,(\{y \in \mathrm{B}_{\mathtt{r}}(y_j) :\ \omega_0(y) \notin \mathrm{Dioph}^\tau_\alpha\}) \\ &\leqslant \sum_{j=1}^{n_0} \sum_{k \in \mathbb{Z}^d\setminus\{0\}} \mathrm{meas}\Big(\Big\{y \in \mathrm{B}_{\mathtt{r}}(y_j) :\ |\omega_0(y)\cdot e_k| \leqslant \frac{\alpha}{|k|_1^\tau |k|_2}\Big\}\Big),\end{aligned}$$

where $e_k := \frac{k}{|k|_2}$. Since on $\mathrm{B}_{\mathtt{r}}(y_j)$, $y \to \omega_0(y)$ is a diffeomorphism, by the change of variables $y = \omega_0^{-1}(\omega)$, we find

$$\mathrm{meas}\Big(\Big\{y \in \mathrm{B}_{\mathtt{r}}(y_j) :\ |\omega_0(y)\cdot e_k| \leqslant \frac{\alpha}{|k|_1^\tau |k|_2}\Big\}\Big)$$

$$\leqslant \delta_0^{-1}\,\mathrm{meas}\Big(\Big\{\omega\in\omega_0\big(\mathrm{B}_{\mathsf{r}}(y_j)\big):\ |\omega\cdot e_k|\leqslant\frac{\alpha}{|k|_1^\tau|k|_2}\Big\}\Big)$$
$$\leqslant \delta_0^{-1}\Big(\operatorname{diam}\omega_0\big(\mathrm{B}_{\mathsf{r}}(y_j)\big)\Big)^{d-1}\frac{2\alpha}{|k|_1^\tau|k|_2}$$
$$\leqslant \delta_0^{-1}\big(\mathsf{M}2\sqrt{d}\mathsf{r}\big)^{d-1}\frac{2\alpha}{|k|_1^\tau|k|_2}.$$

Summing up over j and k, one gets

$$\mathrm{meas}\,(\mathcal{R}_\alpha)\leqslant\Big(2^d d^{\frac{d-1}{2}}\sum_{k\in\mathbb{Z}^d\setminus\{0\}}\frac{1}{|k|_1^\tau\,|k|_2}\Big)\,n_0\,\delta_0^{-1}\,\mathsf{M}^{d-1}\,\mathsf{r}^{d-1}\,\alpha. \tag{3.12}$$

Let us turn to the estimate of $\mathrm{meas}\,(\mathscr{T}_\rho)$. Observing that

$$\mathscr{T}_\rho:=\mathrm{B}_{\mathsf{r}+\rho}(\mathcal{D}_0)\setminus\mathrm{B}_{\mathsf{r}-\rho}(\mathcal{D}_0)\subseteq\bigcup_{j=1}^{n_0}\mathrm{B}_{\mathsf{r}+\rho}(y_j)\setminus\mathrm{B}_{\mathsf{r}-\rho}(y_j),$$

one finds

$$\mathrm{meas}\,(\mathscr{T}_\rho)\leqslant\sum_{j=1}^{n_0}\mathrm{meas}\Big(\mathrm{B}_{\mathsf{r}+\rho}(y_j)\setminus\mathrm{B}_{\mathsf{r}-\rho}(y_j)\Big)$$
$$=n_0 2^d\big((\mathsf{r}+\rho)^d-(\mathsf{r}-\rho)^d\big)$$
$$\leqslant d2^{2d}\,n_0\,\rho\,\mathsf{r}^{d-1}=\frac{d2^{2d}}{c_0}\,n_0\,\frac{\theta}{\mathsf{M}}\,\mathsf{r}^{d-1}\alpha. \tag{3.13}$$

Observing that $\mathscr{K}\subseteq\mathscr{K}_{\mathscr{D}}(\alpha_*\sqrt{\varepsilon},\tau)$ and that, by (2.9), β in (3.1) satisfies $\beta<\frac{5}{4}(2\pi)^d$, one sees that (3.12) and (3.13) imply (3.10) with $\bar{c}_*$ as in (3.11). □

3.2. Smooth Case

In order to state the result for case (B), we need the following definitions.

- Let S be a compact and connected C^2-hypersurface of $\mathbb{R}^d$. The **minimal focal distance of** S is defined as
$$\mathrm{minfoc}\,(S):=\min\big\{\inf\{e_c(u,\nu^+(u)):u\in S\},\inf\{e_c(u,\nu^-(u)):u\in S\}\big\},$$
where $\nu^\pm(u)$ denotes the outwards/inwards normal to S at u and
$$e_c(u,v):=\sup\{t>0:\mathrm{dist}_2(u+tv,S)=t\},$$
dist_2 being the Euclidean distance.

- Given any bounded set D in $\mathbb{R}^d$, we define the (measure of the) **maximal** $(d-1)$**-dimensional section of** D as
$$\mathrm{sec}_{d-1}(D):=\sup_{\lambda\in\Lambda^{d-1}}\mathcal{H}^{d-1}(\lambda\cap D),$$
where Λ^{d-1} denotes the set of all hyperplanes in $\mathbb{R}^d$ and $\mathcal{H}^{d-1}$ the $(d-1)$-dimensional Hausdorff measure.

- Given a set $D\subseteq\mathbb{R}^d$ and $\rho>0$, we define ρ**-inner domains of** D (which depend upon the choice of the metric) as[8)]
$$D'_\rho:=\big\{y\in D:\ \mathrm{B}_\rho(y)\subseteq D\big\},\qquad D''_\rho:=\big\{y\in D:\ \mathbf{B}_\rho(y)\subseteq D\big\}. \tag{3.14}$$

[8)] Recall that B_ρ denotes a ball with respect to the sup-norm $|\cdot|=|\cdot|_\infty$, while $\mathbf{B}_\rho$ denotes a ball with respect to the Euclidean norm $|\cdot|_2$.

Theorem 4. *Let $\mathscr{D} \subseteq \mathbb{R}^d$ be an open and bounded set with C^2 a compact and connected boundary. Let $\tau > d-1 \geqslant 1$, $s > 0$. Let $K \in \mathcal{B}_{\mathsf{R}}(\mathscr{D})$ be uniformly (Kolmogorov) non-degenerate and so that the unperturbed frequency map $y \in \mathscr{D} \mapsto \omega_0(y) := K_y(y) \in \mathbb{C}^d$ is a global diffeomorphism. Let $P \in \mathcal{B}_{\mathsf{R},s}(\mathscr{D})$ and let M, L, P, θ as in (2.2) and define*

$$\mathsf{r} := \min\{\mathsf{R},\ \mathrm{minfoc}\,(\partial\mathscr{D}),\ 1/\kappa\}/\sqrt{d}, \tag{3.15}$$

where $\kappa := \sup_{\partial\mathscr{D}} \max_{1\leqslant j\leqslant d-1} |\kappa_j|$, κ_j's being the principal curvatures of $\partial\mathscr{D}$. Let c_0 and c_* be as in Theorem 1; fix $0 < s_* < s$, let ϵ_* be as in (2.4); let α_*, ε_*, δ_0, θ_0 and ρ be as in (3.9). Let $\mathscr{K}_{\mathscr{D}}(\alpha_*\sqrt{\varepsilon}, \tau)$ be as in (3.8) with $H = K + \varepsilon P$. Then, if $0 < \varepsilon < \varepsilon_*$, one has

$$\mathrm{meas}\,\Big((\mathscr{D}\times\mathbb{T}^d)\backslash\mathscr{K}_{\mathscr{D}}(\alpha_*\sqrt{\varepsilon},\tau)\Big) \leqslant \hat{c}_*\,\theta_0\ \mathsf{M}^{-1}\ \max\big\{\sec_{d-1}(\mathscr{D})\,,\ \mathcal{H}^{d-1}(\partial\mathscr{D})\big\}\ \alpha_*\ \sqrt{\varepsilon}, \tag{3.16}$$

with

$$\hat{c}_* := \frac{5}{2}\,(2\pi)^d\,\Big(\frac{2^d}{\sqrt{d}\ c_0} + \sum_{k\in\mathbb{Z}^d\backslash\{0\}} \frac{1}{|k|_1^\tau\,|k|_2}\Big). \tag{3.17}$$

Proof. The idea is again to apply Theorem 1 and Theorem 2.

Let $\mathcal{D}_0 := \mathscr{D}''_{\sqrt{d}\mathsf{r}}$. Since $\sqrt{d}\mathsf{r} \leqslant \mathrm{minfoc}\,(\partial\mathscr{D})$, by Lemma 9,

$$\mathrm{B}_{\mathsf{r}}(\mathcal{D}_0) \subseteq \mathbf{B}_{\sqrt{d}\mathsf{r}}(\mathscr{D}''_{\sqrt{d}\mathsf{r}}) = \mathscr{D}\,, \qquad \text{and} \qquad \widehat{\mathscr{D}} = \mathrm{B}_{\mathsf{r}-\rho}(\mathcal{D}_0) \supseteq \mathbf{B}_{\mathsf{r}-\rho}(\mathcal{D}_0) = \mathscr{D}''_{(\sqrt{d}-1)\mathsf{r}+\rho}. \tag{3.18}$$

As in the proof of Theorem 3, we let $\alpha := \alpha_*\sqrt{\varepsilon}$, so that $\rho = \mathsf{L}\alpha/c_0$ and $\epsilon = \mathsf{MP}/\alpha_*^2$ (cfr. (2.2)). Then the inequalities in (2.4) hold with the equal sign. The relation $\rho < \mathsf{r}$ is equivalent to $\varepsilon < \varepsilon_*$, which is satisfied by hypothesis. Hence, all the assumptions of Theorem 1 are satisfied and the measure estimate (3.2) holds with $\mathscr{K}$ as in (2.5).

By hypothesis the frequency map $y \to \omega_0(y)$ is a diffeomorphism on $\mathscr{D}$, so we can repeat the estimate on the measure of $\mathcal{R}_\alpha$ done in the proof of Theorem 3 without the need of localizing the actions. Letting, as above, $e_k := \frac{k}{|k|_2}$, we find

$$\begin{aligned}
\mathrm{meas}\,(\mathcal{R}_\alpha) &= \mathrm{meas}\,(\{y\in\mathscr{D}:\ \omega_0(y)\notin \mathrm{Dioph}^\tau_\alpha\}) \\
&\leqslant \sum_{k\in\mathbb{Z}^d\backslash\{0\}} \mathrm{meas}\,\Big(\Big\{y\in\mathscr{D}:\ |\omega_0(y)\cdot e_k| \leqslant \frac{\alpha}{|k|_1^\tau|k|_2}\Big\}\Big). \\
&\leqslant \sum_{k\in\mathbb{Z}^d\backslash\{0\}} \delta_0^{-1}\,\mathrm{meas}\,\Big(\Big\{\omega\in\omega_0(\mathscr{D}):\ |\omega\cdot e_k| \leqslant \frac{\alpha}{|k|_1^\tau|k|_2}\Big\}\Big) \\
&\leqslant \delta_0^{-1}\,\mathsf{M}^{d-1}\ \sec_{d-1}(\mathscr{D}) \sum_{k\in\mathbb{Z}^d\backslash\{0\}} \frac{2\alpha}{|k|_1^\tau|k|_2} \\
&\leqslant \theta_0\,\mathsf{M}^{-1}\ \sec_{d-1}(\mathscr{D}) \sum_{k\in\mathbb{Z}^d\backslash\{0\}} \frac{2\alpha}{|k|_1^\tau|k|_2}.
\end{aligned}$$

The estimate on the measure of $\mathscr{T}_\rho$ follows from Lemma 10. Indeed, if we denote $\mathfrak{T}_\rho(S) := \{u\in\mathbb{R}^d:\ \mathrm{dist}_2(u,S) < \rho\}$, we have (compare (B.23))

$$\mathscr{T}_\rho = \mathrm{B}_{\mathsf{r}+\rho}(\mathcal{D}_0)\setminus \mathrm{B}_{\mathsf{r}-\rho}(\mathcal{D}_0) \overset{(3.18)}{\subseteq} \mathbf{B}_{\sqrt{d}\rho}(\mathscr{D})\backslash\mathscr{D}''_{(\sqrt{d}-1)\mathsf{r}+\rho} \subseteq \mathfrak{T}_{\sqrt{d}\mathsf{r}}(\partial\mathscr{D}).$$

Since $\mathsf{r} \leqslant \min\{\,\mathrm{minfoc}\,(\partial\mathscr{D})/\sqrt{d},\ 1/(\sqrt{d}\kappa)\}$, by (B.24), we get

$$\begin{aligned}
\mathrm{meas}\,(\mathscr{T}_\rho) &\leqslant \mathrm{meas}\,(\mathfrak{T}_{\sqrt{d}\mathsf{r}}(\partial\mathscr{D})) \\
&\leqslant \frac{2}{d}\,\frac{(1+\sqrt{d}\mathsf{r}\kappa)^d - 1}{\kappa}\,\mathcal{H}^{d-1}(\partial\mathscr{D})
\end{aligned}$$

$$< \frac{2^{d+1}}{\sqrt{d}}\,\rho\,\mathcal{H}^{d-1}(\partial\mathscr{D})$$
$$= \frac{2^{d+1}}{\sqrt{d}c_0}\,\theta\,\mathsf{M}^{-1}\,\mathcal{H}^{d-1}(\partial\mathscr{D})\,\alpha.$$

Since $\alpha = \alpha_*\sqrt{\varepsilon}$, (3.16) follows, with $\hat{c}_*$ as in (3.17). □

APPENDIX A. PROOF OF THEOREM 1

In this appendix we provide the details needed to prove Arnold's Global KAM Theorem (Theorem 1). The main point is the choice of the various parameters and sequences involved in the Newton-like procedure based on the iteration of a "KAM step" (in turn, based upon the original scheme by Arnold; compare [2] and its revisions in [12] and [10]). Although the main ideas are well known, some details are needed, especially in order to compute explicitly constants and to keep the optimal relation between ε and α. Furthermore, the construction of the "integrating map" also requires a discussion. All this is done in the present appendix.

By following [12, Chap. 6], one gets the following:

General Step of the KAM Scheme

Lemma 1 (KAM step). *Let $r > 0$, $0 < 2\sigma < s \leqslant 1$, $\mathscr{D}_\sharp \subseteq \mathbb{R}^d$ be a non-empty, bounded domain. Consider the Hamiltonian parametrized by $\varepsilon \in \mathbb{R}$*

$$H(y, x; \varepsilon) := K(y) + \varepsilon P(y, x),$$

where $K, P \in \mathrm{B}_{r,s}(\mathscr{D}_\sharp)$. Assume that[9]

$$\begin{aligned} &\det K_{yy}(y) \neq 0\ , && T(y) := K_{yy}(y)^{-1}, \quad \forall\, y \in \mathscr{D}_\sharp\ ,\\ &\|K_{yy}\|_{r,\mathscr{D}_\sharp} \leqslant \mathsf{M}, && \|T\|_{\mathscr{D}_\sharp} \leqslant \mathsf{L}\ ,\\ &\|P\|_{r,s,\mathscr{D}_\sharp} \leqslant \mathsf{P}, && K_y(\mathscr{D}_\sharp) \subseteq \Delta^\tau_\alpha. \end{aligned} \tag{A.1}$$

Fix $\varepsilon \neq 0$ and assume that

$$\lambda \geqslant \log\left(\sigma^{2\nu+d}\frac{\alpha^2}{\varepsilon\mathsf{PM}}\right) \geqslant 1\ . \tag{A.2}$$

Let

$$\begin{aligned} &\ell := 4\sigma^{-1}\lambda, \qquad \check{r} \leqslant \frac{r}{32d\mathsf{LM}}, \quad \bar{r} \leqslant \min\left\{\frac{\alpha}{2d\mathsf{M}\ell^\nu}\ ,\ \check{r}\right\},\\ &\tilde{r} := \frac{\check{r}\sigma}{16d\mathsf{LM}}, \qquad \bar{s} := s - \frac{2}{3}\sigma, \quad s' := s - \sigma, \end{aligned} \tag{A.3}$$

and[10]

$$\mathfrak{p} := \mathsf{P}\max\left\{\frac{16\mathsf{L}}{r\bar{r}}\sigma^{-(\nu+d)},\ \frac{\mathsf{C}_4}{\alpha\bar{r}}\sigma^{-2(\nu+d)}\right\}.$$

Assume:

$$\varepsilon\mathfrak{p} \leqslant \frac{\sigma}{3}. \tag{A.4}$$

Then there exists a diffeomorphism $G\colon \mathbb{B}_{\tilde{r}}(\mathscr{D}_\sharp) \to G(\mathbb{B}_{\tilde{r}}(\mathscr{D}_\sharp))$, a symplectic change of coordinates

$$\phi' = \mathtt{id} + \varepsilon\tilde{\phi} : \mathbb{B}_{\bar{r}/2,s'}(\mathscr{D}'_\sharp) \to \mathbb{B}_{2r/3,\bar{s}}(\mathscr{D}_\sharp), \tag{A.5}$$

---

[9] In the sequel, K and P stand for generic real-analytic Hamiltonians which later on will, respectively, play the roles of K_j and P_j, and y_0, r, the roles of y_j, r_j in the iterative step.

[10] Notice that $\mathfrak{p} \geqslant \sigma^{-d}\overline{\mathfrak{p}} \geqslant \overline{\mathfrak{p}}$ since $\sigma \leqslant 1$. Notice also that $\mathsf{LM} \geqslant 1$, so that $\frac{16\mathsf{L}}{r\check{r}}\sigma^{-(\nu+d)} > \frac{16\mathsf{L}}{r^2} \geqslant \frac{4}{\mathsf{M}r^2}$.

such that

$$\begin{cases} H\circ\phi' =: H' =: K' + \varepsilon^2 P' \ , \\ \partial_{y'} K' \circ G = \partial_y K, \quad \det \partial^2_{y'} K' \circ G \neq 0 \quad \text{on } \mathscr{D}_\sharp, \end{cases} \tag{A.6}$$

with $K'(y') := K(y') + \varepsilon \widetilde{K}(y') := K(y') + \varepsilon \langle P(y',\cdot)\rangle$*. Moreover, letting* $\left(\partial^2_{y'} K'(\mathsf{y}')\right)^{-1} =: T(\mathsf{y}') + \varepsilon\, \widetilde{T}(\mathsf{y}')$, $\mathsf{y}' \in G(\mathscr{D}_\sharp)$*, the following estimates hold:*

$$\begin{cases} \|\partial^2_{y'} \widetilde{K}\|_{r/2,\mathscr{D}_\sharp} \leqslant \mathsf{M}\mathfrak{p}, \qquad \|G - \mathtt{id}\|_{\tilde{r},\mathscr{D}_\sharp} \leqslant \sigma^{\nu+d} \bar{r}\varepsilon\mathfrak{p}, \qquad \|\widetilde{T}\|_{\mathscr{D}'_\sharp} \leqslant \mathsf{L}\mathfrak{p}\,, \\ \max\left\{\dfrac{\mathsf{C}_{12}}{\mathsf{C}_4}\|\overline{\mathsf{W}}\nabla\tilde{\phi}\ \overline{\mathsf{W}}^{-1}\|_{\bar{r}/2,s',\mathscr{D}'_\sharp}, \|\mathsf{W}\,\tilde{\phi}\|_{\bar{r}/2,s',\mathscr{D}'_\sharp}\right\} \leqslant \sigma^d \mathfrak{p}\,, \qquad \|P'\|_{\bar{r}/2,s',\mathscr{D}'_\sharp} \leqslant \mathfrak{p}\mathsf{P}, \end{cases} \tag{A.7}$$

where

$$\mathscr{D}'_\sharp := G(\mathscr{D}_\sharp), \quad \left(\partial^2_{y'} K'(\mathsf{y}')\right)^{-1} =: T\circ G^{-1}(\mathsf{y}') + \varepsilon\, \widetilde{T}(\mathsf{y}'),\ \forall\, \mathsf{y}' \in \mathscr{D}'_\sharp,$$

$$\mathsf{W} := \operatorname{diag}(\bar{r}^{-1}\mathbb{1}_d, \mathbb{1}_d)\ , \qquad \overline{\mathsf{W}} := \operatorname{diag}(\sigma^{-\tau}\bar{r}^{-1}\mathbb{1}_d, \mathbb{1}_d)\ .$$

Implementation

As in [10], we shall separate the first step from the others. Let H, K, P, ρ, s, s_*, W, P, M, L, θ, ϵ be as in Section 2. Set

$$\begin{aligned} &\sigma_0 := (s-s_*)/2,\ \epsilon_0 := \epsilon\ ,\ \theta_0 := \theta,\ r_0 := \rho,\ \mathsf{L}_0 := \mathsf{L}\ ,\ \mathsf{M}_0 := \mathsf{M},\ \mathsf{P}_0 := \mathsf{P},\ \mathsf{W}_1 := \mathsf{W}\ , \\ &\lambda_0 := \log \epsilon_0^{-1},\ \lambda_* := \mathsf{C}_7 \sigma_0^{-(4\nu+2d+1)} \theta_0^2 \lambda_0^{2\nu},\ \theta_* := 2^{2\nu+2d+1}\mathsf{C}_5^2\theta_0^2,\ \ell_0 := 4\sigma_0^{-1}\lambda_0\ , \\ &K_0 := K\ ,\ P_0 := P,\ H_0 := H\ ,\quad \mathscr{D}_0 := \mathscr{D}^*. \end{aligned}$$

First step

Let

$$\begin{aligned} &s_1 := s_0 - \sigma_0\ , \quad \check{r}_1 := \frac{r_0}{64 d\theta_0}, \quad \tilde{r}_1 := \frac{\check{r}_1\sigma_0}{32 d\theta_0}, \quad r_1 := \frac{1}{2}\min\left\{\frac{\alpha}{2d\sqrt{2}\mathsf{M}_0\ell_0^\nu}, \check{r}_1\right\}, \\ &\mathsf{M}_1 := \left(1+\frac{\sigma_0}{3}\right)\mathsf{M}_0, \quad \mathsf{L}_1 := \left(1+\frac{\sigma_0}{3}\right)\mathsf{L}_0, \quad \hat{\epsilon}_0 := \mathsf{C}_8\sigma_0^{-(3\nu+2d+1)}\epsilon_0^{1/2}, \quad \mathsf{P}_1 := \frac{\hat{\epsilon}_0\mathsf{P}_0}{\varepsilon}, \\ &\mathfrak{p}_0 := \mathsf{P}_0 \max\left\{\frac{8\mathsf{L}_0}{r_0 r_1}\sigma_0^{-(\nu+d)}, \frac{\mathsf{C}_4}{2\alpha r_1}\sigma_0^{-2(\nu+d)}\right\}. \end{aligned}$$

Lemma 2. *Under the above assumptions and notations, if*

$$\alpha \leqslant \frac{\mathsf{C}_4}{16}\frac{r_0}{\mathsf{L}_0} \qquad \text{and} \qquad \max\{e\,\epsilon_0, \hat{\epsilon}_0\} \leqslant 1, \tag{A.8}$$

then there exist $\mathscr{D}_1 \subseteq \mathscr{D}$*, a real-analytic diffeomorphism*

$$G_1\colon \mathbb{B}_{\tilde{r}_1}(\mathscr{D}^*) \to G_1(\mathbb{B}_{\tilde{r}_1}(\mathscr{D}^*))$$

and a real-analytic symplectomorphism

$$\phi_1 : \mathbb{B}_{r_1,s_1}(\mathscr{D}_1) \to \mathbb{B}_{r_0,s_0}(\mathscr{D}_0) \tag{A.9}$$

such that

$$G_1(\mathscr{D}^*) = \mathscr{D}_1, \tag{A.10}$$

$$\partial_{y_1} K_1 \circ G_1 = \partial_y K_0, \tag{A.11}$$

$$H_1 := H_0 \circ \phi_1 =: K_1 + \varepsilon^2 P_1 \qquad \text{on } \mathbb{B}_{r_1,s_1}(\mathscr{D}_1) \tag{A.12}$$

and[11)]

$$\mathscr{D}_1 \subseteq \mathscr{D}_{r_1}, \tag{A.13}$$

$$\|\partial^2_{y_1} K_1\|_{r_0/4,\mathscr{D}_1} \leqslant \mathsf{M}_1, \qquad \|T_1\|_{\mathscr{D}_1} \leqslant \mathsf{L}_1, \qquad T_1 := (\partial^2_{y_1} K_1)^{-1}, \tag{A.14}$$

$$\|P_1\|_{r_1,s_1,\mathscr{D}_1} \leqslant \mathsf{P}_1, \tag{A.15}$$

$$\|G_1 - \mathtt{id}\|_{\tilde r_1,\mathscr{D}^*} \leqslant 2\sigma_0^{\nu+d}\, r_1\, \varepsilon\mathfrak{p}_0, \tag{A.16}$$

$$\|\partial_z G_1 - \mathbb{1}_d\|_{\tilde r_1/2,\mathscr{D}^*} \leqslant 2^5 d\mathsf{C}_4\sqrt{2}\theta_0\sigma_0^{\tau+d}\ell_0^{-\nu}\varepsilon\mathfrak{p}_0, \tag{A.17}$$

$$\max\{\mathsf{C}_{12}\mathsf{C}_4^{-1}\|\overline{\mathsf{W}}_1\nabla(\phi_1 - id)\overline{\mathsf{W}}_1^{-1}\|_{r_1,s_1,\mathscr{D}_1},\ \|\mathsf{W}_1(\phi_1 - \mathtt{id})\|_{r_1,s_1,\mathscr{D}_1}\} \leqslant \sigma_0^d\, \varepsilon\mathfrak{p}_0. \tag{A.18}$$

Second step, iteration and convergence

For a given $j \geqslant 1$, define[12)]

$$\sigma_j := \frac{\sigma_0}{2^j}, \quad s_{j+1} := s_j - \sigma_j = s_* + \frac{\sigma_0}{2^j}, \quad \bar s_j := s_j - \frac{2\sigma_j}{3}, \quad \ell_j := 4^j\ell_0,$$

$$\mathsf{M}_{j+1} := \mathsf{M}_0\prod_{k=0}^{j}(1+\frac{\sigma_k}{3}) < \mathsf{M}_0\sqrt{2}, \quad \mathsf{L}_{j+1} := \mathsf{L}_0\prod_{k=0}^{j}(1+\frac{\sigma_k}{3}) < \mathsf{L}_0\sqrt{2},$$

$$\epsilon_j := \frac{\mathsf{M}_0\varepsilon^{2^j}\mathsf{P}_j}{\alpha^2}, \quad \check r_{j+1} := \frac{r_j}{64d\theta_0}, \quad \tilde r_{j+1} := \frac{\check r_{j+1}\sigma_j}{32d\theta_0}, \quad r_{j+1} := \frac{1}{2}\min\left\{\frac{\alpha}{2d\sqrt{2}\mathsf{M}_0\ell_j^\nu}, \frac{r_j}{64d\theta_0}\right\},$$

$$\mathsf{P}_{j+1} := \lambda_*\theta_*^{j-1}\frac{\mathsf{M}_0\mathsf{P}_j^2}{\alpha^2}, \quad \hat\epsilon_j := \lambda_*\theta_*^j\epsilon_j, \quad \mathsf{W}_{j+1} := \operatorname{diag}\left((2r_{j+1})^{-1}\mathbb{1}_d\,,\mathbb{1}_d\right),$$

$$\overline{\mathsf{W}}_{j+1} := \operatorname{diag}\left(\sigma_j^{-\tau}(2r_{j+1})^{-1}\mathbb{1}_d, \mathbb{1}_d\right), \quad \mathfrak{p}_j := \mathsf{P}_j\max\left\{\frac{8\mathsf{L}_0\sqrt{2}}{r_jr_{j+1}}\sigma_j^{-(\nu+d)}, \frac{\mathsf{C}_4}{2\alpha r_{j+1}}\sigma_j^{-2(\nu+d)}\right\}.$$

Observe that, for any $j \geqslant 1$,

$$\hat\epsilon_{j+1} = \lambda_*\theta_*^{j+1}\epsilon_{j+1} = \lambda_*\theta_*^{j+1}\frac{\mathsf{M}_0\varepsilon^{2^{j+1}}\mathsf{P}_{j+1}}{\alpha^2} = \lambda_*\theta_*^{j+1}\frac{\mathsf{M}_0\varepsilon^{2^{j+1}}}{\alpha^2}\ \lambda_*\theta_*^{j-1}\frac{\mathsf{M}_0\mathsf{P}_j^2}{\alpha^2}\left(\lambda_*\ \theta_*^j\ \epsilon_j\right)^2 = \hat\epsilon_j^2$$

i.e.

$$\hat\epsilon_j = \hat\epsilon_1^{2^{j-1}}.$$

Lemma 3. *Assume* (A.12) $\div$ (A.15) *with some* $\varepsilon \neq 0$ *and*

$$\max\left\{e\ \epsilon_0\ ,\ 2^{11}d^2\theta_0\sigma_0^{\nu+d}\hat\epsilon/3\ ,\ 2\mathsf{C}_6\theta_0\hat\epsilon_1\right\} \leqslant 1\ . \tag{A.19}$$

Then one can construct a sequence of real-analytic diffeomorphisms

$$G_j\colon\ \mathbb{B}_{\tilde r_j}(\mathscr{D}_{j-1}) \to G_j(\,\mathbb{B}_{\tilde r_j}(\mathscr{D}_{j-1})), \qquad j \geqslant 2,$$

and of real-analytic symplectic transformations

$$\phi_j\,:\,\mathbb{B}_{r_j,s_j}(\mathscr{D}_j) \to \mathbb{B}_{r_{j-1},s_{j-1}}(\mathscr{D}_{j-1}), \tag{A.20}$$

such that

$$G_j(\mathscr{D}_{j-1}) = \mathscr{D}_j \subseteq \mathscr{D}_{r_j},$$
$$\partial_y K_{j+1}\circ G_{j+1} = \partial_y K_j,$$
$$H_j := H_{j-1}\circ\phi_j =: K_j + \varepsilon^{2^j}P_j \qquad \textit{on} \quad \mathbb{B}_{r_j,s_j}(\mathscr{D}_j)$$

converge uniformly. More precisely, we have the following:

[11)] (A.17) follows trivially (A.16) using Cauchy's estimate.

[12)] Notice that $s_j \downarrow s_*$ and $r_j \downarrow 0$.

(*i*) *the sequence* $G^j \coloneqq G_j \circ G_{j-1} \circ \cdots \circ G_2 \circ G_1$ *converges uniformly on* $\mathscr{D}^*$ *to a lipeomorphism* $Y^*\colon \mathscr{D}^* \to \mathscr{D}_* \coloneqq Y^*(\mathscr{D}^*) \subseteq \mathscr{D}$ *and* $Y^* \in C^\infty_W(\mathscr{D}^*)$.

(*ii*) $\varepsilon^{2^j} \partial_y^\beta P_j$ *converges uniformly on* $\mathscr{D}_* \times \mathbb{T}^d_{s_*}$ *to* 0, *for any* $\beta \in \mathbb{N}_0^d$;

(*iii*) $\phi^j \coloneqq \phi_2 \circ \cdots \circ \phi_j$ *converges uniformly on* $\mathscr{D}_* \times \mathbb{T}^d$ *to a symplectic transformation*

$$\phi^*\colon \mathscr{D}_* \times \mathbb{T}^d \overset{into}{\longrightarrow} \mathrm{B}_{r_1}(\mathscr{D}_1) \times \mathbb{T}^d,$$

with $\phi^* \in C^\infty_W(\mathscr{D}_* \times \mathbb{T}^d)$ *and* $\phi^*(y,\cdot)\colon \mathbb{T}^d_{s_*} \ni x \mapsto \phi^*(y,x)$ *holomorphic, for any* $y \in \mathscr{D}_*$;

(*iv*) K_j *converges uniformly on* $\mathscr{D}_*$ *to a function* $K_* \in C^\infty_W(\mathscr{D}_*)$, *with*

$$\partial_{y_*} K_* \circ Y^* = \partial_y K_0 \qquad \text{on} \quad \mathscr{D}^*,$$
$$\partial^\beta_{y_*}(H_1 \circ \phi^*)(y_*,x) = \partial^\beta_{y_*} K_*(y_*), \quad \forall (y_*,x) \in \mathscr{D}_* \times \mathbb{T}^d ,\forall\, \beta \in \mathbb{N}_0^d.$$

Finally, the following estimates hold for any $i \geqslant 2$:[13)]

$$\|G_i - \mathtt{id}\|_{\tilde r_i, \mathscr{D}_{i-1}} \leqslant 2 r_i\, \sigma_{i-1}^{\nu+d}\, \varepsilon^{2^{i-1}} \mathfrak{p}_{i-1}, \tag{A.21}$$

$$\|\partial_z G_i - \mathbb{1}_d\|_{\tilde r_i/2, \mathscr{D}_{i-1}} \leqslant 2^5 d\theta_0\, \sigma_{i-1}^{\tau+d}\, \varepsilon^{2^{i-1}} \mathfrak{p}_{i-1}, \tag{A.22}$$

$$\|P_i\|_{r_i,s_i,\mathscr{D}_i} \leqslant \mathsf{P}_i\ , \tag{A.23}$$

$$\|\mathsf{W}_2(\phi^{i+2} - \phi^{i+1})\|_{r_{i+2},s_{i+2},\mathscr{D}_{i+2}} \leqslant \mathsf{a}_2\ \left(\mathsf{C}_6 \theta_0^{\frac14} \hat\epsilon_1\right)^{2^i}, \tag{A.24}$$

$$|\mathsf{W}_2(\phi^* - \mathtt{id})| \leqslant \frac{2\sigma_0^{d+1}\ \hat\epsilon_1}{3\ \theta_*} \qquad \text{on} \quad \mathscr{D}_* \times \mathbb{T}^d_{s_*}\ , \tag{A.25}$$

where

$$\mathsf{a}_2 \coloneqq \mathbf{a}_1\ \sigma_2^d\ \|\mathsf{W}_2 \phi_2\|_{r_2,s_2,\mathscr{D}_2}.$$

We can now complete the proof of Theorem 1. First of all, observe that

$$(\log t)^a \leqslant \left(\frac{2a}{e}\right)^a \sqrt{t}, \qquad \forall\, t \geqslant e, \quad \forall\, a > \frac12, \tag{A.26}$$

and from the proof we have

$$\varepsilon \mathfrak{p}_0(3\sigma_0^{-1}) \overset{(A.8)}{\leqslant} 6d\mathsf{C}_4\sqrt{2}\sigma_0^{-2(\nu+d)-1}\frac{\mathsf{K}_0 \varepsilon \mathsf{P}_0}{\alpha^2}\ell_0^\nu \tag{A.27}$$

$$\overset{(A.26)}{\leqslant} \hat\epsilon_0 \overset{(A.8)}{\leqslant} 1, \tag{A.28}$$

and, for $j \geqslant 1$,

$$\varepsilon^{2^j}\mathfrak{p}_j(3\sigma_j^{-1}) \leqslant \hat\epsilon_1^{2^{j-1}}/\theta_*. \tag{A.29}$$

Let $\phi_* \coloneqq \phi_1 \circ \phi^*$. Thus, uniformly on $\mathscr{D}_* \times \mathbb{T}^d_{s_*}$,[14)]

$$\begin{aligned}|\mathsf{W}_1(\phi_* - \mathtt{id})| &\leqslant |\mathsf{W}_1(\phi_1 \circ \phi^* - \phi^*)| + |\mathsf{W}_1(\phi^* - \mathtt{id})|\\ &\leqslant \|\mathsf{W}_1(\phi_1 - \mathtt{id})\|_{r_1,s_1,\mathscr{D}_1} + \|\mathsf{W}_1\mathsf{W}_2^{-1}\|\ |\mathsf{W}_2(\phi^* - \mathtt{id})|\\ &\leqslant \sigma_0^d\ \varepsilon\mathfrak{p}_0 + \frac{2\sigma_0^{d+1}\ \hat\epsilon_1}{3\ \theta_*}\end{aligned}$$

13) Observe that (A.22) follows (A.21) using Cauchy's estimate.

14) Observe that $\lambda_0^{2\nu}\epsilon_0 \overset{(A.19)}{\leqslant} (4\nu)^{2\nu}\sqrt{\epsilon_0} \overset{(A.19)}{\leqslant} (4\nu)^{2\nu}(2^{11}d^2\mathsf{C}_5)^{-1/2}\theta_0^{-1}\sigma_0^{(\nu+d+1)/2}$.

$$\overset{(A.26)+(A.27)}{\leqslant} 6d\mathsf{C}_4\sqrt{2}\sigma_0^{-(2\nu+2d+1)}\epsilon_0\ell_0^\nu + \left(\frac{\nu}{2\,e}\right)^\nu \mathsf{C}_7\sigma_0^{-(6\nu+4d+2)}\theta_0^2\ell_0^\nu\epsilon_0$$
$$\leqslant \mathsf{C}_9\theta_0^2\ell_0^\nu\epsilon_0\ ,$$

i. e., (2.10). Moreover, setting $G_0 \coloneqq \mathtt{id}$, we have, for any $i \geqslant 3$,

$$\|G^i - \mathtt{id}\|_{\mathscr{D}^*} \leqslant \sum_{j=0}^{i-1}\|G^{j+1} - G^j\|_{\mathscr{D}^*} = \sum_{j=0}^{i-1}\|G_j - \mathtt{id}\|_{\mathscr{D}_{j-1}} \overset{(A.21)+(A.16)}{\leqslant} 2\sum_{j=0}^{i-1} r_{j+1}\sigma_j^{\nu+d}\varepsilon^{2^j}\mathfrak{p}_j$$
$$\leqslant 2\, r_1\sigma_0^\nu \sum_{j=0}^{\infty}\sigma_j^d\varepsilon^{2^j}\mathfrak{p}_j \overset{(A.28)+(A.29)}{\leqslant} 2r_1\sigma_0^\nu\cdot \mathsf{C}_9\theta_0^2\ell_0^\nu\epsilon_0\ ,$$

and then, passing to the limit, we get

$$\|Y^* - \mathtt{id}\|_{\mathscr{D}^*} \leqslant 2^{2\tau+1/2}d^{-1}\mathsf{C}_9\sigma_0^\nu\theta_0^2\frac{\varepsilon\mathsf{P}_0}{\alpha},$$

i. e., (2.8). Now, observing that, for any $j \geqslant 1$, $\nabla\phi^{j+1} = \nabla\phi^j\nabla\phi_{j+1}$, $\|\overline{\mathsf{W}}_j\overline{\mathsf{W}}_{j+1}^{-1}\| = 1$ and $\|\overline{\mathsf{W}}_{j+1}\overline{\mathsf{W}}_j^{-1}\| \leqslant \mathsf{C}_5\,\theta_0$, we obtain

$$\|\overline{\mathsf{W}}_1(\nabla\phi^{j+1} - \mathbb{1}_{2d})\overline{\mathsf{W}}_{j+1}^{-1}\|_* \leqslant \left(\|\overline{\mathsf{W}}_1(\nabla\phi^j - \mathbb{1}_{2d})\overline{\mathsf{W}}_j^{-1}\|_* + 1\right)\left(\|\overline{\mathsf{W}}_{j+1}(\nabla\phi_{j+1} - \mathbb{1}_{2d})\overline{\mathsf{W}}_{j+1}^{-1}\|_* + 1\right)$$
$$- 1$$
$$\leqslant \left(\|\overline{\mathsf{W}}_1(\nabla\phi^j - \mathbb{1}_{2d})\overline{\mathsf{W}}_j^{-1}\|_* + 1\right)\left(\frac{\mathsf{C}_4}{\mathsf{C}_{12}}\sigma_j^d\varepsilon^{2^j}\mathfrak{p}_j + 1\right) - 1,$$

which iterated yields[15)]

$$\|\overline{\mathsf{W}}_1(\nabla\phi^{j+1} - \mathbb{1}_{2d})\overline{\mathsf{W}}_{j+1}^{-1}\|_* \leqslant \prod_{j=1}^{\infty}\left(\frac{\mathsf{C}_4}{\mathsf{C}_{12}}\sigma_{j-1}^d\varepsilon^{2^{j-1}}\mathfrak{p}_{j-1} + 1\right) - 1$$
$$\leqslant \exp\left(\sum_{j=1}^{\infty}\frac{\mathsf{C}_4}{\mathsf{C}_{12}}\sigma_{j-1}^d\varepsilon^{2^{j-1}}\mathfrak{p}_{j-1}\right) - 1$$
$$\overset{(A.28)+(A.29)}{\leqslant} \exp(\mathsf{C}_4\mathsf{C}_9\mathsf{C}_{12}^{-1}\theta_0^2\ell_0^\nu\epsilon_0) - 1$$
$$\leqslant \exp((4d)^{-1})\mathsf{C}_4\mathsf{C}_9\mathsf{C}_{12}^{-1}\theta_0^2\ell_0^\nu\epsilon_0 \overset{(2.4)+(A.26)}{\leqslant} \frac{1}{4(18d^3+70)\theta},$$

and letting $j \to \infty$, we obtain

$$\|\partial_x u_*\|_* \leqslant \exp((4d)^{-1})\mathsf{C}_4\mathsf{C}_9\mathsf{C}_{12}^{-1}\theta_0^2\ell_0^\nu\epsilon_0 \leqslant \frac{1}{4(18d^3+70)\theta}, \tag{A.30}$$

i. e., (2.11).

Next, we show that $\mathrm{Lip}_{\mathscr{D}^*}(Y^* - \mathtt{id}) < 1$, which will imply that[16)] $Y^*\colon \mathscr{D}^* \overset{onto}{\longrightarrow} \mathscr{D}_*$ is a lipeomorphism. Observe first that, for any $j \geqslant 1$, $0 < r < \tilde{r}_j/2$, $\mathsf{y}_{j-1} \in \mathscr{D}_{j-1}$ and any $y \in \mathbb{B}_r(\mathsf{y}_{j-1})$, we have

$$|G_j(y) - G_j(\mathsf{y}_{j-1})| \leqslant |(G_j(y) - y) - (G_j(\mathsf{y}_{j-1}) - \mathsf{y}_{j-1})| + |y - \mathsf{y}_{j-1}| \overset{(A.22)+(A.19)}{\leqslant} \frac{1}{2}|y - \mathsf{y}_{j-1}| + r < 2r,$$

so that

$$G_j(\mathbb{B}_r(\mathscr{D}_{j-1})) \subseteq \mathbb{B}_{2r}(G_j(\mathscr{D}_{j-1})) = \mathbb{B}_{2r}(\mathscr{D}_j). \tag{A.31}$$

15) Use: $e^t - 1 \leqslant t\,e^t$, for any $t \geqslant 0$.
16) See Proposition II.2 in [20].

Thus, as the sequence $\tilde r_j$ is strictly decreasing, for any $j \geqslant k \geqslant 1$, G^k is well-defined on $\mathbb{B}_{2^{-j-1}\tilde r_{j+1}}(\mathscr{D}_0)$ and we have

$$G^k(\mathbb{B}_{2^{-j-1}\tilde r_{j+1}}(\mathscr{D}_0)) \overset{(A.31)}{\subseteq} G_j \circ \cdots \circ G_2(\mathbb{B}_{2^{-j}\tilde r_{j+1}}(\mathscr{D}_1)) \overset{(A.31)}{\subseteq} \cdots \overset{(A.31)}{\subseteq} \mathbb{B}_{2^{k-j-1}\tilde r_{j+1}}(\mathscr{D}_k) \subseteq \\ \subseteq \mathbb{B}_{2^{-1}\tilde r_{k+1}}(\mathscr{D}_k). \tag{A.32}$$

Therefore, for any $j \geqslant 2$,

$$\begin{aligned}\|G^j - \mathtt{id}\|_{L,\mathbb{B}_{2^{-j-1}\tilde r_{j+1}}(\mathscr{D}^*)} + 1 &= \|(G_j - \mathtt{id}) \circ G^{j-1} + (G^{j-1} - \mathtt{id})\|_{L,\mathbb{B}_{2^{-j-1}\tilde r_{j+1}}(\mathscr{D}_0)} + 1 \\ &\leqslant (\|G_j - \mathtt{id}\|_{L,G^{j-1}(\mathbb{B}_{\frac{\tilde r_{j+1}}{2^{j+1}}}(\mathscr{D}_0))} + 1)(\|G^{j-1} - \mathtt{id}\|_{L,\mathbb{B}_{\frac{\tilde r_{j+1}}{2^{j+1}}}(\mathscr{D}_0)} + 1) \\ &\overset{(A.32)}{\leqslant} (\|G_j - \mathtt{id}\|_{L,\mathbb{B}_{\frac{\tilde r_j}{2}}(\mathscr{D}_{j-1})} + 1)(\|G^{j-1} - \mathtt{id}\|_{L,\mathbb{B}_{\frac{\tilde r_j}{2^j}}(\mathscr{D}_0)} + 1) \\ &= (\|\partial_z G_j - \mathbb{1}_d\|_{\tilde r_j/2,\mathscr{D}_{j-1}} + 1)(\|G^{j-1} - \mathtt{id}\|_{L,\mathbb{B}_{\frac{\tilde r_j}{2^j}}(\mathscr{D}_0)} + 1) \\ &\overset{(A.22)+(A.17)}{\leqslant} (2^5 d\theta_0\, \sigma_{j-1}^{\tau+d}\, \varepsilon^{2^{j-1}} \mathsf{L}_{j-1} + 1)(\|G^{j-1} - \mathtt{id}\|_{L,\mathbb{B}_{\frac{\tilde r_j}{2^j}}(\mathscr{D}^*)} + 1),\end{aligned}$$

which iterated leads to[17)]

$$\begin{aligned}\|G^j - \mathbb{1}_d\|_{L,\mathscr{D}^*} &\leqslant -1 + (1 + (32d)^{-1}) \prod_{i=2}^{\infty} (2^5 d\theta_0 \sigma_{j-1}^{\tau+d} \varepsilon^{2^{i-1}} \mathsf{L}_{i-1} + 1) \\ &\leqslant -1 + \exp((32d)^{-1} + 2^5 d\theta_0 \sum_{i=1}^{\infty} \sigma_i^{\tau+d} \varepsilon^{2^i} \mathsf{L}_i) \\ &\leqslant -1 + \exp((32d)^{-1} + 2^5 d\theta_0 \sigma_1^{\tau} \sum_{i=1}^{\infty} \sigma_i^{d} \varepsilon^{2^i} \mathsf{L}_i) \\ &\leqslant -1 + \exp\left((32d)^{-1} + 2^5 d\theta_0 \sigma_1^{\tau} \frac{2\sigma_0^{d+1}\, \hat\epsilon_1}{3\,\theta_*}\right) \\ &\overset{(A.19)}{<} -1 + \exp\left((32d)^{-1} + (32d)^{-1}\right) \leqslant e^{1/(16d)}/(16d) < \frac{1}{4d}\,.\end{aligned} \tag{A.33}$$

Hence, letting $j \to \infty$, we find that Y^* is Lipschitz continuous, with $\mathrm{Lip}_{\mathscr{D}^*}(Y^* - \mathtt{id})$ satisfying (2.9) as

$$2^5 d\mathsf{C}_4 \sqrt{2}\theta_0 \sigma_0^{\tau+d} \ell_0^{-\nu} \varepsilon \mathsf{L}_0 + \sum_{j\geqslant 2} 2^5 d\theta_0\, \sigma_{j-1}^{\tau+d}\, \varepsilon^{2^{j-1}} \mathsf{L}_{j-1} \overset{(A.19)}{\leqslant} c_2\, \theta^3\ \ (s - s_*)^{-1}\ \frac{\mathsf{M}\varepsilon\mathsf{P}}{\alpha^2}\ \left(\log \frac{\alpha^2}{\mathsf{M}\varepsilon\mathsf{P}}\right)^{\nu}.$$

Next, we show that $\phi_* \in C_W^{\infty}(\mathscr{D}_* \times \mathbb{T}^d)$. For any $n, j \geqslant 1$, we have

$$\begin{aligned}\|G^{n+j} - G^j\|_{\mathscr{D}^*} &\leqslant \sum_{k=j}^{n+j-1} \|G^{k+1} - G^k\|_{\mathscr{D}^*} \\ &\overset{(A.21)}{\leqslant} \sum_{k=j}^{n+j-1} 2r_{k+1} \sigma_k^{\nu+d} \varepsilon^{2^k} \mathsf{L}_k \\ &\leqslant 2r_{j+1} \sigma_j^{\nu} \sum_{k\geqslant 1} \sigma_k^{d} \varepsilon^{2^k} \mathsf{L}_k\end{aligned}$$

---

17) Use, again, $e^t - 1 \leqslant t\,e^t$, $\forall\, t \geqslant 0$, and $2^5 d\mathsf{C}_4\sqrt{2}\theta_0\sigma_0^{\tau+d}\ell_0^{-\nu}\varepsilon\mathsf{L}_0 \overset{(A.27)}{\leqslant} 2^7 d^2 \mathsf{C}_4^2 \theta_0 \sigma_0^{-(\nu+d+1)} \epsilon_0 \overset{(A.19)}{<} (32d)^{-1}$.

$$\leqslant 2r_{j+1}\sigma_j^{\nu}\,\frac{2\sigma_0^{d+1}\,\hat{\epsilon}_1}{3\,\theta_*}$$
$$\overset{(A.19)}{<}\ \sigma_j^{\nu}\,\tilde{r}_{j+1}\,.$$

Now, letting $n \to \infty$, we get

$$\|Y^* - G^j\|_{\mathscr{D}^*} < \sigma_j^{\nu}\,\tilde{r}_{j+1} < \frac{\tilde{r}_{j+1}}{4}. \tag{A.34}$$

Hence[18], for any $j \geqslant 1$,

$$\mathbb{B}_{\frac{\tilde{r}_{j+1}}{4}}\big(G^j(\mathscr{D}^*)\big) \overset{(A.34)}{\subseteq} \mathbb{B}_{\frac{\tilde{r}_{j+1}}{2}}(\mathscr{D}_*) \overset{(A.34)}{\subseteq} \mathbb{B}_{\tilde{r}_{j+1}}(\mathscr{D}_j) \subseteq \mathbb{B}_{r_j}(\mathscr{D}_j)\,. \tag{A.35}$$

Therefore, for any $n \geqslant 1$, we have

$$\begin{aligned}\sum_{j\geqslant 3}\|\mathsf{W}_2(\phi^j-\phi^{j-1})\|_{\tilde{r}_{j+1}/2,s_j,\mathscr{D}_*}\left(\frac{\tilde{r}_{j+1}}{2}\right)^{-n} &\overset{(A.35)}{\leqslant} (2^{12}d^2\theta_0^2)^n\sum_{j\geqslant 3}\|\mathsf{W}_2(\phi^j-\phi^{j-1})\|_{r_j,s_j,\mathscr{D}_j}(r_j\sigma_j)^{-n}\\ &\overset{(A.24)}{\leqslant} (2^{12}d^2\theta_0^2\sigma_1 r_1)^n\mathsf{a}_2\sum_{j\geqslant 3}\left(\mathsf{C}_6\theta_0^{\frac14}\hat{\epsilon}_1\right)^{2^{j-2}}(2\mathsf{a}_1)^{n(j-1)}\\ &< \ +\infty\,,\end{aligned}$$

since, for j sufficiently large,

$$\left(\mathsf{C}_6\theta_0^{\frac14}\hat{\epsilon}_1\right)^{2^{j-1}}(2\mathsf{a}_1)^{nj} < \left(\sqrt{2}\mathsf{C}_6\theta_0^{\frac14}\hat{\epsilon}_1\right)^{2^{j-1}} \overset{(A.19)}{\leqslant} (1/\sqrt{2})^{2^{j-1}}.$$

Thus, letting $\Phi_j \coloneqq \phi_1\circ\phi^j$ and using the mean value theorem, we have

$$\begin{aligned}\sum_{j\geqslant 2}\|\mathsf{W}_2(\Phi_j-\Phi_{j-1})\|_{\tilde{r}_{j+1}/2,s_j,\mathscr{D}_*}\left(\frac{\tilde{r}_{j+1}}{2}\right)^{-n} &\leqslant \|\mathsf{W}_2\nabla\phi_1\mathsf{W}_2^{-1}\|_{r_1,s_1,\mathscr{D}_1}\times\\ &\times\sum_{j\geqslant 3}\|\mathsf{W}_2(\phi^j-\phi^{j-1})\|_{\tilde{r}_{j+1}/2,s_j,\mathscr{D}_*}\left(\frac{\tilde{r}_{j+1}}{2}\right)^{-n}\\ &<\infty.\end{aligned}$$

Consequently, writing

$$\Phi_j = (\Phi_j-\Phi_{j-1})+\cdots+(\Phi_3-\Phi_2)\,,\qquad j\geqslant 2,$$

and invoking Lemma 8 (see Appendix B.7), we conclude that $\phi_* = \lim\Phi_j \in C_W^{\infty}(\mathscr{D}_*\times\mathbb{T}^d)$.

Now we prove $Y^*\in C_W^{\infty}(\mathscr{D}^*)$ analogously. For any $j\geqslant 2$ and $n\geqslant 1$, we have

$$G^j = (G^j-G^{j-1})+\cdots+(G^2-G^1),$$

and, thanks to (A.31), $G^{j+1}-G^j$ is well-defined on $\mathbb{B}_{2^{-j-2}\tilde{r}_{j+2}}(\mathscr{D}^*)$, for any $j\geqslant 1$, so that

$$\begin{aligned}\sum_{j\geqslant 1}\|G^{j+1}-G^j\|_{\frac{\tilde{r}_{j+2}}{2^{j+2}},\mathscr{D}^*}\left(\frac{\tilde{r}_{j+2}}{2^{j+2}}\right)^{-n} &= \sum_{j\geqslant 1}(2^{j+2}\tilde{r}_{j+2}^{-1})^n\|(G_{j+1}-\mathtt{id})\circ G^j\|_{\frac{\tilde{r}_{j+2}}{2^{j+2}},\mathscr{D}^*}\\ &\overset{(A.32)}{\leqslant} \sum_{j\geqslant 2}(2^{j+2}\tilde{r}_{j+2}^{-1})^n\|G_{j+1}-\mathtt{id}\|_{\tilde{r}_{j+1},\mathscr{D}_j}\\ &\leqslant 2\sum_{j\geqslant 1}(2^{j+2}\tilde{r}_{j+2}^{-1})^n r_{j+1}\sigma_j^{\nu+d}\varepsilon^{2^j}\mathsf{L}_j\end{aligned}$$

[18] Recall that, by definition, $G^j(\mathscr{D}^*) = \mathscr{D}_j$ and $Y^*(\mathscr{D}^*) = \mathscr{D}_*$.

$$\overset{(A.19)}{<} \infty ,$$

which proves that $Y^* \in C^\infty_W(\mathscr{D}^*)$.

Finally, we prove Kolmogorov's non-degeneracy[19)] of the Kolmogorov tori $\phi_*(\mathscr{D}_* \times \mathbb{T}^d)$. Fix $y_* \in \mathscr{D}_*$. Let $y_0 := (Y^*)^{-1}(y_*)$ and

$$\hat{\epsilon} := \frac{1}{4(18d^3+70)\theta} .$$

Since $\|\partial_x u_*\|_* \overset{(2.11)}{\leqslant} \hat{\epsilon} < 1/2$, the map $x \longmapsto x + u_*(y_*, x)$ is a diffeomorphism of $\mathbb{T}^d$. Letting

$$(\partial_x(\mathtt{id} + u_*)(y_*, x))^{-1} =: \mathbb{1}_d + A(y_*, x) ,$$

we have

$$\|A\|_* \leqslant 2\|\partial_x u_*\|_* \overset{(2.11)}{\leqslant} 2\hat{\epsilon} < 1 ; \quad \|v_*\|_* \overset{(2.11)}{\leqslant} \frac{\mathsf{C}_9\sqrt{2}}{4d}\theta^2 \frac{\varepsilon\mathsf{P}}{\alpha} \overset{(2.4)}{\leqslant} \frac{\mathsf{C}_4\mathsf{C}_9\sqrt{2}}{2^5 d\mathsf{C}_*}\,\rho < \frac{\rho}{8}. \tag{A.36}$$

Moreover, write $K_{yy}(y_*) = K_{yy}(y_0)(\mathbb{1}_d + K_{yy}(y_0)^{-1}(K_{yy}(y_*) - K_{yy}(y_0)))$ and observe

$$\mathrm{dist}(y_0, \partial\mathscr{D}) \geqslant \rho \quad \text{and} \quad |y_* - y_0| \overset{(2.8)+(2.4)}{\leqslant} \frac{2^{\tau-5}\mathsf{C}_4\mathsf{C}_9\sqrt{2}}{d\mathsf{C}_*}\,\rho < \frac{\rho}{64d},$$

so that

$$\mathrm{dist}(y_*, \partial\mathscr{D}) \geqslant \frac{\rho}{2}. \tag{A.37}$$

Thus, by the mean value theorem, we have

$$\begin{aligned}\|K_{yy}(y_0)^{-1}(K_{yy}(y_*) - K_{yy}(y_0))\| &\overset{(A.37)}{\leqslant} \mathsf{T}\frac{d^2\mathsf{K}}{\rho/2}|y_* - y_0| \\ &\overset{(2.8)}{\leqslant} 2^{\tau+11/2}d^2\mathsf{C}_9\,\theta^3\,\frac{\mathsf{K}\varepsilon\mathsf{P}_0}{\alpha\rho} \overset{(2.4)}{\leqslant} \frac{2^{\tau+15/2}d^2\mathsf{C}_4\mathsf{C}_9}{\mathsf{C}_*} \leqslant \frac{1}{2}.\end{aligned}$$

Hence, $K_{yy}(y_*)$ is invertible and $\|K_{yy}(y_*)^{-1}\| \leqslant 2\|K_{yy}(y_0)^{-1}\| \leqslant 2\mathsf{T}$.

In [17] it is proven that the map

$$\phi^{y_*}(y, x) := (y_* + v_*(y_*, x) + y + A^T y, x + u_*(y_*, x))$$

is symplectic. Then

$$H \circ \phi^{y_*}(y, x) = E^{y_*} + \omega^{y_*} \cdot y + Q^{y_*}(y, x)$$

with

$$\begin{aligned}&E^{y_*} = K(y_*), \quad \omega^{y_*} := K_y(y_0), \quad \langle Q^{y_*}_{yy}(0, \cdot)\rangle = K_{yy}(y_*) + \langle \mathcal{M}\rangle , \\ &\mathcal{M} := \partial_y^2\left(K(y_* + v_* + y + A^T y) - \frac{1}{2}y^T K_{yy}(y_*)y\right)\Big|_{y=0} + \partial_y^2(\varepsilon P \circ \phi)\Big|_{y=0} , \\ &\|K_{yy}(y_*)^{-1}\mathcal{M}\|_* \leqslant 2\mathsf{T}\mathcal{M} \overset{(A.36)}{\leqslant} 2(18d^3 + 70)\hat{\epsilon}\theta = 1/2,\end{aligned}$$

which shows that $\langle Q^{y_*}_{yy}(0, \cdot)\rangle$ is invertible. □

Remark 2. Here we list all the constants, which appear in the above proof and give an explicit expression for the constants c_k's appearing in the statement of Theorem 1.
Recall that $\tau > d - 1 \geqslant 1$ and notice that all the C_i's are greater than 1 and depend only upon d and τ:

$$\nu := \tau + 1, \quad \mathsf{C}_0 := 4\left(\frac{3}{2}\right)^{2\nu+d}\int_{\mathbb{R}^d}\left(|y|_1^\nu + d|y|_1^{2\nu}\right)\,e^{-|y|_1}dy, \quad \mathsf{C}_1 := 2\left(\frac{3}{2}\right)^{\nu+d}\int_{\mathbb{R}^d}|y|_1^\nu\,e^{-|y|_1}dy,$$

19) See Appendix B.9.

$$\mathsf{C}_2 := 2^{3d}d, \quad \mathsf{C}_3 := d^2\mathsf{C}_1^2 + 6d\mathsf{C}_1 + \mathsf{C}_2, \quad \mathsf{C}_4 := \max\left\{(1+d^2)\mathsf{C}_0, \mathsf{C}_3\right\}, \quad \mathsf{C}_5 := \max\left\{2^{2\nu}, 2^7 d\right\},$$

$$\mathsf{C}_6 := \left(2^{-d}\mathsf{C}_5\right)^{\frac{1}{4}}, \quad \mathsf{C}_7 := 3\cdot 2^{4\nu+2d+3} d\sqrt{2}\max\left\{2^{2\nu+6}d, \mathsf{C}_4/2\right\}\mathsf{C}_5, \quad \mathsf{C}_8 := 3\cdot 2^{3\nu+1}\nu^\nu e^{-\nu} d\mathsf{C}_4\sqrt{2},$$

$$\mathsf{C}_9 := 3\cdot 2^{-(4\nu+2d)}d\mathsf{C}_4\sqrt{2} + 2^{-\nu}\nu^\nu e^{-\nu}\mathsf{C}_7\mathsf{C}_8, \quad \mathsf{C}_{10} := 2\left(\frac{3}{2}\right)^{\nu+d+1}\int_{\mathbb{R}^d}|y|_1^{\nu+1} e^{-|y|_1}dy,$$

$$\mathsf{C}_{11} := 8\left(\frac{3}{2}\right)^{3\tau+d+2}\int_{\mathbb{R}^d}(2|y|_1^\tau + 3|y|_1^{2\tau+1} + |y|_1^{3\tau+2})\, e^{-|y|_1}dy\,, \quad \mathsf{C}_{12} := \max\{2\mathsf{C}_{10}\,, 2\mathsf{C}_{11}\,, 12\mathsf{C}_0\}\,,$$

$$\mathsf{C}_* := \max\{2^{11\nu+6d+4}\nu^\nu e^{-\nu}\mathsf{C}_5^2\mathsf{C}_6\mathsf{C}_7\mathsf{C}_8,\ (2^{\nu/2-2d+2}(18d^3+70)\nu^\nu e^{-\nu}\mathsf{C}_4\mathsf{C}_9\mathsf{C}_{12}^{-1})^2,\ 2^{\tau+8}d^2\mathsf{C}_4\mathsf{C}_9\sqrt{2}\}\,.$$

Then

$$\begin{aligned} &c_* := \mathsf{C}_*\,, && c_0 := 2^{-4}\mathsf{C}_4\,, && c_1 := 2^{\tau-1/2}d^{-1}\mathsf{C}_9\,,\\ &c_2 := 2^{2\nu+6}d\mathsf{C}_9\,, && c_3 := \mathsf{C}_9\,, && c_4 := e^{(4d)^{-1}}\mathsf{C}_4\mathsf{C}_9\mathsf{C}_{12}^{-1}\,. \end{aligned} \tag{A.38}$$

Remark 3. There is a small flaw in [10]: The parameter[20] L chosen in [10, Lemma 1] is not big enough to ensure that the new perturbation P' and the symplectic change of coordinates ϕ are well-defined on $D_{\bar r/2,s'}(\mathscr{D}'_\sharp)$. The right choice is the following:

$$\mathsf{L} := \mathsf{P}\max\left\{\frac{40d\mathsf{T}^2\mathsf{K}}{r\bar r\sigma^{\nu+d}}\,, \frac{2\mathsf{C}_4}{\alpha\bar r\sigma^{2(\nu+d)}}\right\}, \quad \mathsf{W} := \mathrm{diag}(\bar r^{-1}\mathbb{1}_d,\ \mathbb{1}_d), \quad \hat\epsilon_0 := \mathsf{C}_9\,\sigma_0^{-2(\nu+d)-1}\epsilon_0\,\theta_0^2\,\lambda_0^\nu\,,$$

$$\mathsf{L}_j := \frac{\mathsf{P}_j}{r_{j+1}}\max\left\{\frac{80d\sqrt{2}\mathsf{T}_0\theta_0}{r_j\sigma_j^{\nu+d}}\,, \frac{\mathsf{C}_4}{\alpha\sigma_j^{2(\nu+d)}}\right\}, \quad \mathsf{W}_j := \mathrm{diag}(2r_{j+1}^{-1}\mathbb{1}_d,\ \mathbb{1}_d), \quad \hat\epsilon_{j+1} := \frac{\mathsf{K}_0\varepsilon^{2^{j+1}}\mathsf{P}_{j+1}}{\alpha^2}\,,$$

$$\mathsf{P}_{j+2} := \lambda_*\theta_*^{j+1}\frac{\mathsf{K}_0\mathsf{P}_{j+1}^2}{\alpha^2}\,, \quad \hat\epsilon_{j+1} := \lambda_*\theta_*^{j+2}\epsilon_{j+1}\,.$$

Of course, one needs then to change accordingly (and in a straightforward way) the constants involved, as follows:

$$\begin{aligned}
\nu &:= \tau+1\\
\mathsf{C}_0 &:= 4\sqrt{2}\left(\frac{3}{2}\right)^{2\nu+d}\int_{\mathbb{R}^d}\left(|y|_1^\nu + |y|_1^{2\nu}\right)\, e^{-|y|_1}dy\,, \quad \mathsf{C}_1 := 2\left(\frac{3}{2}\right)^{\nu+d}\int_{\mathbb{R}^d}|y|_1^\nu\, e^{-|y|_1}dy\,,\\
\mathsf{C}_2 &:= 2^{3d}d\,, \quad \mathsf{C}_3 := \left(d^2\mathsf{C}_1^2 + 6d\mathsf{C}_1 + \mathsf{C}_2\right)\sqrt{2}\,, \quad \mathsf{C}_4 := \max\left\{6d^2\mathsf{C}_0,\, \mathsf{C}_3\right\}\,,\\
\mathsf{C}_5 &:= \frac{3\cdot 2^5 d}{5}\,, \quad \mathsf{C}_6 := \max\left\{2^{2\nu}\,, \mathsf{C}_5\right\}\,, \quad \mathsf{C}_7 := 3d\cdot 2^{4\nu+2}\sqrt{2}\max\left\{640d^2\,, \mathsf{C}_4\right\}\,,\\
\mathsf{C}_8 &:= \left(2^{-d}\mathsf{C}_6\right)^{1/8}\,, \quad \mathsf{C}_9 := 3d\cdot 2^{2\nu+2}\sqrt{2}\max\left\{80d\sqrt{2}\,, \mathsf{C}_4\right\}\,,\\
\mathsf{C}_{10} &:= (4\nu\, e^{-1})^{2\nu}\left(1 + 2^{4\nu+2d+2}(\nu\, e^{-1})^{2\nu}\mathsf{C}_6^2\mathsf{C}_7\right)\mathsf{C}_9/(3d^2)\,, \quad \mathsf{C}_{11} := (5d\cdot 2^{3(\nu+1)})^{-1}\mathsf{C}_{10}\,,\\
\mathsf{C}_{12} &:= 2^{2(5\nu+4d+2)}\,\mathsf{C}_6^2\,\mathsf{C}_7\,\mathsf{C}_8\,\mathsf{C}_9\,, \quad \mathsf{C}_{13} := \mathsf{C}_{10} + \mathsf{C}_{11}\,, \quad \mathsf{C}_{14} := \mathsf{C}_{12}\,,\\
\mathsf{C}_{15} &:= 18d^3 + 70\,, \quad \mathsf{C}_{16} := (6\nu\, e^{-1})^{4\nu}\,, \quad \mathsf{C} := \max\{3\mathsf{C}_{10},\, \mathsf{C}_{13}\}\,,\\
\mathsf{C}_* &:= \max\left\{\mathsf{C}_{16}\mathsf{C}_{14}^{2/3},\ 6\mathsf{C}_{15}\mathsf{C}_{16}\mathsf{C}^2,\ 2^{2(4\nu+2d+1)}\mathsf{C}_{16}\mathsf{C}_9^2,\ \mathsf{C}_{10}^2\right\}\,.
\end{aligned}$$

The smallness condition (14) and the estimate (16) become, respectively,

$$\alpha \leqslant \frac{r}{\mathsf{T}} \qquad \text{and} \qquad \epsilon \leqslant \epsilon_* := \frac{(s-s_*)^a}{\mathsf{C}_*\,\theta^6}\,,$$

[20] In the present remark, we will adopt the notations of [10].

and

$$\max\left\{\|u_*\|_{s_*}\,,\ \|\partial_x u_*\|_{s_*}\,,\ \frac{\mathsf{K}}{\alpha}\,(\log\epsilon^{-1})^\nu\,\|v_*\|_{s_*}\right\}\leqslant\frac{\mathsf{C}\,\theta^3}{(s-s_*)^{a/2}}\,\epsilon\,(\log\epsilon^{-1})^\nu\leqslant\frac{1}{4\,e}\,,$$

where $a := 6\nu + 3d + 2$.

APPENDIX B. TOOLS

B.1. Classical Estimates (Cauchy, Fourier)

Lemma 4 ([7]). *Let $p\in\mathbb{N}$, $r,s>0, y_0\in\mathbb{C}^d$ and let f be a real-analytic function $\mathbb{B}_{r,s}(y_0)$ with $\|f\|_{r,s}:=\sup_{\mathbb{B}_{r,s}(y_0)}|f|<\infty$. Then*

(i) *For any multi–index $(l,k)\in\mathbb{N}^d\times\mathbb{N}^d$ with $|l|_1+|k|_1\leqslant p$ and for any $0<r'<r$, $0<s'<s$,*[21]

$$\|\partial_y^l\partial_x^k f\|_{r',s'}\leqslant p!\,\|f\|_{r,s}(r-r')^{|l|_1}(s-s')^{|k|_1}.$$

(ii) *For any $k\in\mathbb{Z}^d$ and any $y\in\mathbb{B}_r(y_0)$*

$$|f_k(y)|\leqslant\, e^{-|k|_1 s}\|f\|_{r,s}.$$

B.2. An Inverse Function Theorem

Theorem B.1. *Let D be a convex subset of $\mathbb{C}^d$, $y_0\in D$ and let $f\in C^1(D,\mathbb{C}^d)$ such that*[22] $\det f'(y_0)\neq 0$*.Assume*

$$\varrho:=\sup_{y\in D}\|\mathbb{1}-Tf'(y)\|<1\,,\qquad\qquad T:=(f'(y_0))^{-1}\,. \tag{B.1}$$

Then $\det f'(y)\neq 0$*, for each $y\in D$ and*

$$\|(f'(y))^{-1}\|\leqslant\lambda:=\frac{\|T\|}{1-\varrho}\,. \tag{B.2}$$

Moreover, f is injective on D and its inverse function $g: f(D)\overset{\text{onto}}{\to} D$ satisfies

$$\mathrm{Lip}_{f(D)}(g)\leqslant\lambda\,. \tag{B.3}$$

Furthermore, if $D:=B_r(y_0)$, $\rho:=r/\lambda$ and $z_0:=f(y_0)$, then

$$B_\rho(z_0)\subseteq f(D)\,. \tag{B.4}$$

Proof. For every $y\in D$, we have $f'(y)=f'(y_0)(\mathbb{1}-A)$, where $A:=\mathbb{1}-Tf'(y)$ with $\|A\|\leqslant\varrho<1$. Thus, $f'(y)$ is invertible and

$$\|(f'(y))^{-1}\|=\|(\sum_{n\geqslant 0}A^n)T\|\leqslant\frac{\|T\|}{1-\varrho},$$

proving (B.2). Now, consider the auxiliary map $F\colon D\ni y\longmapsto y-Tf(y)$. We have $F\in C^1(D,\mathbb{C}^d)$ and $\sup_D\|F'\|\overset{\text{(B.1)}}{\leqslant}\varrho$. Thus, for every $y,\bar y\in D$ with $y\neq\bar y$, we have

$$\begin{aligned}\|T\|\|f(y)-f(\bar y)\| &\overset{\text{(B.1)}}{\geqslant} \big\|T\big(f(y)-f(\bar y)\big)\big\|\\ &= \|(y-\bar y)+(F(\bar y)-F(y))\|\\ &\geqslant \|y-\bar y\|-\|y-\bar y\|\sup_D\|F'\|\\ &\geqslant \|y-\bar y\|(1-\varrho)\overset{\text{(B.1)}}{>}0\,,\end{aligned} \tag{B.5}$$

which shows that f is injective on D and, hence, that (B.3) holds.

[21] As usual, $\partial_y^l:=\frac{\partial^{|l|_1}}{\partial y_1^{l_1}\cdots\partial y_d^{l_d}}$, $\forall\, y\in\mathbb{R}^d$, $l\in\mathbb{Z}^d$.

[22] f' being the Jacobian matrix of f.

To show (B.4) in the case $D := B_r(y_0)$ and $\rho := r/\lambda$, fix $\eta \in \mathbb{C}^d$ with $\|\eta - z_0\| < \rho$. We have to show that there exists $\bar{y} \in D$ such that $f(\bar{y}) = \eta$. Define the map

$$\Phi : y \in D \mapsto \Phi(y) := y - T\big(f(y) - \eta\big) \in Y\,. \tag{B.6}$$

Then Φ is a contraction on D. Indeed, Φ is C^1, $\Phi'(y) = \mathbb{1} - Tf'(y)$ and

$$\mathrm{Lip}_D\, \Phi = \sup_D \|\Phi'\| = \varrho < 1. \tag{B.7}$$

Furthermore, $\Phi : D \to D$, since, if $y \in D$, then

$$\begin{aligned}\|\Phi(y) - y_0\| &\leqslant \|\Phi(y) - \Phi(y_0)\| + \|\Phi(y_0) - y_0\|\\ &\overset{(B.7)}{\leqslant} \varrho\, r + \|T\|\|\eta - z_0\| < \varrho\, r + \|T\|\,\rho = r\,.\end{aligned}$$

Hence, by the contraction lemma, Φ has a (unique) fixed point $\bar{y} \in D$, but $\Phi(\bar{y}) = \bar{y}$ means $f(\bar{y}) = \eta$. □

B.3. Internal Coverings

Given any non-empty subset D of $\mathbb{R}^d$, and given $r > 0$, an **r-internal covering of** D is a subset P of D such that $D \subseteq \bigcup_{y\in P} \mathrm{B}_r(y)$; the **$r$-internal covering number of** D, denoted $N_r^{\mathrm{int}}(D)$, is the minimal cardinality of any r-internal cover.

In [5] the following simple upper bound (having fixed the sup-norm in $\mathbb{R}^d$) on $N_r^{\mathrm{int}}(D)$ for bounded sets D is given:

Lemma 5. *Let $D \subseteq \mathbb{R}^d$ be a non-empty bounded set. Then, for any $r > 0$, one has*[23)]

$$N_r^{\mathrm{int}}(D) \leqslant \Big(\Big[\frac{\mathrm{diam}\, D}{r}\Big] + 1\Big)^d\,. \tag{B.8}$$

For convenience of the reader, we reproduce here the elementary proof of the lemma.

Proof. It is enough to produce an r-internal cover of D with cardinality N bounded by the right-hand side of (B.8). If D is a singleton, the claim is obvious with $N = 1$. Assume now that $\delta := \mathrm{diam}\, D > 0$, and let $M := [\delta/r] + 1$ and $z_i = \inf\{x_i|\ x \in D\}$. Then $D \subseteq K := z + [0, \delta]^d$ and one can find $0 < r' < r$ close enough to r so that $\lceil \delta/r' \rceil \leqslant [\delta/r] + 1 = M$. Then one can cover K with M^d closed, contiguous cubes K_j, $1 \leqslant j \leqslant M^d$, with edge of length r'. Let j_i be the indices such that $K_{j_i} \cap D \neq \emptyset$ and pick a $y_i \in K_{j_i} \cap E$; let $1 \leqslant N \leqslant M^d$ be the number of such cubes. Observe that, since we have chosen the sup-norm in $\mathbb{R}^d$, we have $K_{j_i} \subseteq \mathrm{B}_r(y_i)$ and (B.8) follows. □

B.4. Extensions of Lipschitz Continuous Functions

Here we recall a theorem due to Minty according to which a Lipschitz continuous function can be extended keeping unchanged both the sup-norm *and* the Lipschitz constant.

Theorem B.2 (G. J. Minty [13]). *Let $(V, \langle \cdot\,, \cdot \rangle)$ be a separable inner product space, $\emptyset \neq A \subseteq V$, $L > 0$, $0 < \alpha \leqslant 1$ and $g \colon A \to \mathbb{R}^d$ a (L, α)-Lipschitz – Hölder continuous function on A, namely, let g satisfy*

$$|g(x_1) - g(x_2)|_2 \leqslant L\, \|x_1 - x_2\|^\alpha, \qquad \forall\, x_1, x_2 \in A\,, \tag{B.9}$$

where $\|\cdot\|$ denotes the norm on V induced by the inner product. Then there exists a global (L, α)-Lipschitz – Hölder continuous function[24)] *$G \colon V \to \mathbb{R}^d$ such that $G|_A = g$. Futhermore, G can be chosen in such a way that $G(V)$ is contained in the closed convex hull of $g(A)$. Hence, in particular,*

$$\sup_{x\in V} |G(x)|_2 = \sup_{x\in A} |g(x)|_2 \quad \textit{and} \quad \sup_{x_1\neq x_2\in V} \frac{|G(x_1) - G(x_2)|_2}{\|x_1 - x_2\|^\alpha} = \sup_{x_1\neq x_2\in A} \frac{|g(x_1) - g(x_2)|_2}{\|x_1 - x_2\|^\alpha}\,. \tag{B.10}$$

[23)] $[x]$ denotes the integer-part (or "floor") function $\max\{n \in \mathbb{Z}|\ n \leqslant x\}$, while $\lceil x \rceil$ denote the "ceiling function" $\min\{n \in \mathbb{Z}|\ n \geqslant x\}$; observe that $\lceil x \rceil \leqslant [x] + 1$.

[24)] I. e., satisfying (B.9) on V.

B.5. Lebesgue Measure and Lipschitz Continuous Map

Lemma 6. *Let $\emptyset \neq A \subseteq \mathbb{R}^d$ be a Lebesgue-measurable set and $f\colon A \to \mathbb{R}^d$ be Lipschitz continuous. Then*

$$\operatorname{meas}\big(f(A)\big) \leqslant \operatorname{Lip}_A(f)^d \operatorname{meas}(A) \tag{B.11}$$

and[25]

$$|\operatorname{meas}(f(A)) - \operatorname{meas}(A)| \leqslant ((1+\delta)^d - 1)\operatorname{meas}(A)\ , \tag{B.12}$$

where

$$\delta := \operatorname{Lip}_A(f - \mathtt{id}). \tag{B.13}$$

Proof. Eq. (B.11) is standard: see, e. g., Theorem 2, Section 2.2 and Theorem 1, Section 2.4 in [11].

Let us prove (B.13). By Theorem B.2, $f - \mathtt{id}$ can be extended to a Lipschitz continuous function $g\colon \mathbb{R}^d \circlearrowleft$ with

$$\operatorname{Lip}(g) = \operatorname{Lip}_A(f - \mathtt{id}) = \delta\ .$$

By Rademacher's theorem, there exists a set $N \subseteq \mathbb{R}^d$ with $\operatorname{meas}(N) = 0$ such that g is differentiable on $\mathbb{R}^d \setminus N$ and

$$\|g_y\|_{\mathbb{R}^d\setminus N} \leqslant \operatorname{Lip}_{\mathbb{R}^d\setminus N}(g) \leqslant \operatorname{Lip}(g) = \delta\ .$$

Now pick $y \in \mathbb{R}^d \setminus N$. Then

$$\begin{aligned}|\det(\mathbb{1}_d + g_y(y)) - 1| &= \left|\int_0^1 \frac{d}{dt}\det(\mathbb{1}_d + tg_y)dt\right| = \left|\int_0^1 \operatorname{tr}\ (\operatorname{Adj}(\mathbb{1}_d + tg_y)g_y)\,dt\right| \\ &\leqslant \int_0^1 d\|\mathbb{1}_d + tg_y\|^{d-1}\|g_y\|dt \leqslant \int_0^1 d\,(1+\delta t)^{d-1}\,\delta dt = (1+\delta)^d - 1.\end{aligned}$$

Thus, by the change of variable (or area) formula[26], we have

$$\begin{aligned}|\operatorname{meas}(f(A)) - \operatorname{meas}(A)| &= |\operatorname{meas}((\mathtt{id}+g)(A)) - \operatorname{meas}(A)| = \left|\int_{(\mathtt{id}+g)(A)} dy - \int_A dy\right| \\ &= \left|\int_{(\mathtt{id}+g)(A\setminus N)} dy - \int_{A\setminus N} dy\right| = \left|\int_{A\setminus N} |\det(\mathbb{1}_d + g_y)|dy - \int_{A\setminus N} dy\right| \\ &\leqslant \int_{A\setminus N} |\det(\mathbb{1}_d + g_y) - 1|dy \leqslant ((1+\delta)^d - 1)\operatorname{meas}(A).\end{aligned}$$

□

B.6. Lipeomorphisms "Close" to Identity

Lemma 7. *Let $g\colon \mathbb{C}^d \to \mathbb{C}^d$ be a Lipschitz continuous function such that*

$$\delta := \sup_{\mathbb{R}^d} |g - \mathtt{id}| < \infty\ , \tag{B.14}$$

$$\theta := \operatorname{Lip}_{\mathbb{R}^d}(g - \mathtt{id}) < 1. \tag{B.15}$$

Then g has a Lipschitz global inverse G satisfying

$$\sup_{\mathbb{R}^d} |G - \mathtt{id}| \leqslant \delta\ , \tag{B.16}$$

[25] Inequality (B.12) is sharp as shown by the example $f = (1+\delta)\,\mathtt{id}$.

[26] See [11], §3.3.

$$\mathrm{Lip}_{\mathbb{R}^d}(G - \mathtt{id})\| < \frac{1}{1-\theta}\,. \tag{B.17}$$

Furthermore, for any $\emptyset \neq A \subseteq \mathbb{C}^d$,

$$A \subseteq g\left(\overline{\mathbb{B}_\delta(A)}\right)\,. \tag{B.18}$$

Proof. Let $f := g - \mathtt{id}$, then, for any $x_i \in \mathbb{R}^d$, one has

$$\begin{aligned}|g(x_1) - g(x_2)| &= \big|x_1 - x_2 + \big(f(x_1) - f(x_2)\big)\big| \overset{(B.15)}{>} |x_1 - x_2| - \theta|x_1 - x_2| \\ &= (1-\theta)|x_1 - x_2|\,,\end{aligned}$$

which proves injectivity of g and that

$$\inf_{x_1 \neq x_2} \frac{|g(x_1) - g(x_2)|}{|x_1 - x_2|} \geqslant 1 - \theta > 0\,. \tag{B.19}$$

Let us now prove (B.18). Let $\bar{y} \in A$. It is enough to show that there exists $|y| \leqslant \delta$ such that $\bar{y} = g(y + \bar{y})$, i.e., $y = -f(y + \bar{y})$, i.e., y is a fixed point of the map

$$h\colon \overline{\mathbb{B}_\delta(0)} \ni y \mapsto -f(y + \bar{y}).$$

But, for any $y \in \overline{\mathbb{B}_\delta(0)}$,

$$|h(y)| = |f(y + \bar{y})| \leqslant \|f\|_{\mathbb{R}^d} \overset{(B.14)}{\leqslant} \delta\,,$$

i.e., $h\colon \overline{\mathbb{B}_\delta(0)} \to \overline{\mathbb{B}_\delta(0)}$. Moreover, h is a contraction since $\mathrm{Lip}_{\overline{\mathbb{B}_\delta(0)}}(h) \leqslant \mathrm{Lip}_{\mathbb{R}^d}(f) \overset{(B.15)}{<} 1$. Thus, by Banach's fixed point theorem, we see that (B.18) holds.

From (B.18) it follows at once that g is onto $\mathbb{R}^d$.

Now (B.16) and (B.17) follow easily from (B.14) and (B.19), respectively. □

B.7. Whitney Smoothness

Definition 1. Let $A \subseteq \mathbb{R}^d$ be non-empty and $n \in \mathbb{N}_0$, $m \in \mathbb{N}$. A function $f\colon A \to \mathbb{R}^m$ is said to be C^n on A in the Whitney sense, with Whitney derivatives $(f_\nu)_{\nu \in \mathbb{N}_0^d, |\nu|_1 \leqslant n}$, $f_0 = f$, and we write $f \in C_W^n(A, \mathbb{R}^m)$ if, for any $\varepsilon > 0$ and $y_0 \in A$, there exists $\delta > 0$ such that, for any $y, y' \in A \cap \mathrm{B}_\delta(y_0)$ and $\nu \in \mathbb{N}_0^d$, with $|\nu|_1 \leqslant n$,

$$\Bigg| f_\nu(y') - \sum_{\substack{\mu \in \mathbb{N}_0^d \\ |\mu|_1 \leqslant n - |\nu|_1}} \frac{1}{\mu!} f_{\nu+\mu}(y)(y' - y)^\mu \Bigg| \leqslant \varepsilon |y' - y|^{n - |\nu|_1}\,. \tag{B.20}$$

Lemma 8 ([8, 12]). *Let $A \subseteq \mathbb{R}^d$ be non-empty and $n \in \mathbb{N}_0$. For $m \in \mathbb{N}$, let f_m be a real-analytic function with holomorphic extension to $D_{r_m}(A)$, with $r_m \downarrow 0$ as $m \to \infty$. Assume that*

$$a := \sum_{m=1}^{\infty} \|f_m\|_{r_m, A}\, r_m^{-n} < \infty, \qquad \|f_m\|_{r_m, A} := \sup_{\mathrm{B}^d_{r_m}(A)} |f_m|\,. \tag{B.21}$$

Then $f := \sum_{m=1}^{\infty} f_m \in C_W^n(A, \mathbb{R})$ *with Whitney derivatives* $f_\nu := \sum_{m=1}^{\infty} \partial_y^\nu f_m$.

For completeness, we recall the beautiful Whitney extension theorem.

Theorem B.3 ([19]). *Let $A \subseteq \mathbb{R}^d$ be a closed set and $f \in C_W^n(A, \mathbb{R})$, $n \in \mathbb{N}_0$. Then there exists $\bar{f} \in C^n(\mathbb{R}^d, \mathbb{R})$, real-analytic on $\mathbb{R}^d \setminus A$ and such that $D^\nu \bar{f} = f_\nu$ on A, for any $\nu \in \mathbb{N}_0^d$, with $|\nu|_1 \leqslant n$.*

B.8. Measure of Tubular Neighbourhoods of Hypersurfaces

Recall the definitions of minimal focal distance and of inner domains given in Section 3.2.

The first elementary remark is that, for smooth domains, taking ρ–inner domains is the inverse operation of taking ρ-neighbourhood:

Lemma 9. *Let $\mathscr{D} \subseteq \mathbb{R}^d$ be an open and bounded set with C^2 boundary $\partial\mathscr{D} = S$ compact and connected. Then, for any $0 < \rho' < \rho \leqslant \operatorname{minfoc}(S)$, one has*

$$\mathbf{B}_\rho\big(\mathscr{D}''_\rho\big) = \mathscr{D}\,, \qquad \textit{and} \qquad \mathbf{B}_{\rho-\rho'}\big(\mathscr{D}''_\rho\big) = \mathscr{D}''_{\rho'}\,. \tag{B.22}$$

Proof. We start proving the first part of (B.22). By definition, $\mathbf{B}_\rho\big(\mathscr{D}''_\rho\big) \subseteq \mathscr{D}$. Thus, it remains only to show that $\mathscr{D} \setminus \mathscr{D}''_\rho \subseteq \mathbf{B}_\rho\big(\mathscr{D}''_\rho\big)$.

Let then $y_0 \in \mathscr{D} \setminus \mathscr{D}''_\rho$. As S is compact and dist_2 is continuous, there exists $\bar{y}_0 \in S$ such that $\operatorname{dist}_2(y_0, \mathbb{R}^d \setminus \mathscr{D}) = \operatorname{dist}_2(y_0, S) = |y_0 - \bar{y}_0|_2$. The vector $\nu := (y_0 - \bar{y}_0)/|y_0 - \bar{y}_0|_2$ is the inward unit normal to $\partial\mathscr{D} = S$ at $\bar{y}_0$. Indeed, for any smooth curve $\gamma\colon [0,1] \to S$ with $\gamma(0) = \bar{y}_0$, 0 is a minimum of the smooth map $f(t) := |\gamma(t) - y_0|_2^2$. Thus,

$$0 = f'(0) = 2\dot\gamma(0) \cdot (\bar{y}_0 - y_0).$$

which, by the arbitrariness of γ, implies that the line $(\bar{y}_0 y_0)$ is perpendicular to the tangent space to S at $\bar{y}_0$ and, therefore ν is the inward unit normal to $\partial\mathscr{D}$ at $\bar{y}_0$. Let $y_1 := \bar{y}_0 + \rho\nu$. By assumption, we have $\operatorname{dist}_2(y_1, S) = \rho$, and, therefore, $y_1 \in \mathscr{D}$. In addition, $y_1 \in \mathscr{D}''_\rho$. Indeed, for any $y \in \mathbf{B}_\rho(y_1)$, $\operatorname{dist}_2\big(y, \mathbb{R}^d \setminus \mathscr{D}\big) \geqslant \operatorname{dist}_2\big(y_1, \mathbb{R}^d \setminus \mathscr{D}\big) - |y_1 - y|_2 = \operatorname{dist}_2\big(y_1, S\big) - |y_1 - y|_2 = \rho - |y_1 - y|_2 > 0$. Thus, as $\mathbb{R}^d \setminus \mathscr{D}$ is a closed set, $y \notin \mathbb{R}^d \setminus \mathscr{D}$, i.e., $y \in \mathscr{D}$. Hence, $\mathbf{B}_\rho(y_1) \subseteq \mathscr{D}$, i.e., $y_1 \in \mathscr{D}''_\rho$. In particular, the argument above shows that:[27] for any $y \in \mathbb{R}^d$, $\operatorname{dist}_2\big(y, \mathbb{R}^d \setminus \mathscr{D}\big) \geqslant \rho$ implies that $y \in \mathscr{D}''_\rho$. Thus, as $y_0 \in \mathscr{D} \setminus \mathscr{D}''_\rho$, we have $\operatorname{dist}_2\big(y_0, \mathbb{R}^d \setminus \mathscr{D}\big) < \rho$, which means y_0 is in the open segment $(\bar{y}_0, y_1)$. Therefore, $|y_0 - y_1|_2 < |\bar{y}_0 - y_1|_2 = \rho$, i.e., $y_0 \in \mathbf{B}_\rho(y_1) \subseteq \mathbf{B}_\rho(\mathscr{D}''_\rho)$.

We now prove the second part of (B.22). We have $\mathbf{B}_{\rho-\rho'}\big(\mathscr{D}''_\rho\big) \subseteq \mathscr{D}''_{\rho'}$. Indeed, for any $y_0 \in \mathscr{D}''_\rho$, $y_1 \in \mathbf{B}_{\rho-\rho'}(y_0)$ and $y \in \mathbf{B}_{\rho'}(y_1)$,

$$|y - y_0| \leqslant |y - y_1| + |y_1 - y_0| < \rho' + (\rho - \rho') = \rho \quad \textit{i.e.,}\ \ y \in \mathbf{B}_\rho(y_0),$$

which implies $\mathbf{B}_{\rho-\rho'}\big(\mathscr{D}''_\rho\big) \subseteq \mathscr{D}''_{\rho'}$. It remains to show that $\mathscr{D}''_{\rho'} \backslash \mathscr{D}''_\rho \subseteq \mathbf{B}_{\rho-\rho'}\big(\mathscr{D}''_\rho\big)$. The proof follows analogously to the previous one. Let $y_0 \in \mathscr{D}''_{\rho'} \backslash \mathscr{D}''_\rho$ and $\bar{y}_0 \in S$ such that $\operatorname{dist}_2\big(y_0, \mathbb{R}^d \setminus \mathscr{D}\big) = \operatorname{dist}_2\big(y_0, S\big) = |y_0 - \bar{y}_0|_2$. Then $\rho' \leqslant |y_0 - \bar{y}_0|_2 < \rho$, and the vector $\nu := (y_0 - \bar{y}_0)/|y_0 - \bar{y}_0|_2$ is the inward unit normal to $\partial\mathscr{D} = S$ at $\bar{y}_0$. Set $y_1' := \bar{y}_0 + \rho'\nu$. Thus, $|y_1' - \bar{y}_0|_2 = \rho' \leqslant |y_0 - \bar{y}_0|_2$ and, hence, $y_1' \in \mathscr{D}''_{\rho'}$ and y_1' is in the semi-open segment $(\bar{y}_0, y_0]$. Therefore, $|y_1' - y_0|_2 = |y_0 - \bar{y}_0|_2 - |y_1' - \bar{y}_0|_2 < \rho - \rho'$. Hence, $y_0 \in \mathbf{B}_{\rho-\rho'}(y_1') \subseteq \mathbf{B}_{\rho-\rho'}(\mathscr{D}''_{\rho'})$, i.e., $\mathscr{D}''_{\rho'} \backslash \mathscr{D}''_\rho \subseteq \mathbf{B}_{\rho-\rho'}\big(\mathscr{D}''_\rho\big)$. □

The next result gives a precise evaluation of tubular domains *in the case where the metric is the Euclidean one*. Define

$$\mathfrak{T}_\rho(S) := \{u \in \mathbb{R}^d :\ \operatorname{dist}_2(u, S) < \rho\}\,. \tag{B.23}$$

Lemma 10. *Let $\mathscr{D} \subseteq \mathbb{R}^d$ be a bounded set with C^2 boundary $\partial\mathscr{D} = S$ compact and connected. Then, for any $0 < \rho \leqslant \operatorname{minfoc}(S)$, then,*

$$\operatorname{meas}\left(\mathfrak{T}_\rho(S)\right) \leqslant \frac{2}{d}\,\frac{(1+\rho\kappa)^d - 1}{\kappa}\,\mathcal{H}^{d-1}(S), \tag{B.24}$$

where $\kappa := \sup_S \max_{1\leqslant j\leqslant d-1} |\kappa_j|$ with κ_j the principal curvatures of S, while $\mathcal{H}^{d-1}$ denotes the $(d-1)$-dimensional Hausdorff measure ("surface area").

[27] Actually, one checks easily that $\partial\mathscr{D}''_\rho = \{y \in \mathbb{R}^d\ :\ \operatorname{dist}_2\big(y, \mathbb{R}^d \setminus \mathscr{D}\big) = \rho\}$ and $\operatorname{int}(\mathscr{D}''_\rho) = \{y \in \mathbb{R}^d\ :\ \operatorname{dist}_2\big(y, \mathbb{R}^d \setminus \mathscr{D}\big) > \rho\}$, $\operatorname{int}(\mathscr{D}''_\rho)$ being the interior of $\mathscr{D}''_\rho$.

Proof. [28] We will estimate the "inner tubular neighbourhoods"

$$\mathfrak{T}'_\rho(S) := \{y \in \mathscr{D} : \mathrm{dist}_2(y, S) < \rho\},$$

as the argument for "outer tubular neighbourhood" $\{y \notin \mathscr{D} : \mathrm{dist}_2(y, S) < \rho\}$ is completely analogous.

Since S is compact and connected, we may assume that $S = f^{-1}(\{0\})$ with $f \in C^2(\mathbb{R}^d, \mathbb{R})$ and 0 a regular value for f. Set

$$\nu(x) = \frac{\nabla f}{|\nabla f|_2} , \qquad |\cdot|_2 := \mathrm{dist}_2(\cdot, 0) ,$$

and replacing eventually f by $-f$, we can assume that ν is the inwards unit normal vector field of S. Let $\{\phi_j : U_j \to \mathbb{R}^m\}_{j=1}^p$ be an atlas of S,

$$\Psi_j(u, t) := \phi_j(u) + t\nu(\phi_j(u)), \qquad O_j := \Psi_j(U_j \times [0, \rho)),$$

and observe that[29]

$$\mathfrak{T}'_\rho(S) = \bigcup_{j=1}^p O_j.$$

Let $\{\psi_j\}_{j=1}^p$ be a partition of unity subordinated to the open covering of $\{O_j\}_{j=1}^p$ of $\mathfrak{T}'_\rho(S)$, i. e.,

- **(i)** $\psi_j \in C_c^\infty(\mathfrak{T}'_\rho(S))$;
- **(ii)** $0 \leqslant \psi_j \leqslant 1$;
- **(iii)** $\mathrm{supp}\, \psi_j \subseteq O_j$;
- **(iv)** $\sum_{j=1}^p \psi_j \equiv 1$ on $\mathfrak{T}'_\rho(S)$.

Given $1 \leqslant j \leqslant p$, define $n_j : U_j \longrightarrow \mathbb{S}^d = \{x \in \mathbb{R}^d \ : \ |x|_2 = x_1^2 + \cdots + x_d^2 = 1\} \subseteq \mathbb{R}^d$ as

$$n_j := \nu \circ \phi_j,$$

and $K_j : U_j \longrightarrow T^*S$ such that[30]

$$K_j(u) := -\nu'(\phi_j(u)).$$

Then K_j is symmetric[31] and therefore diagonalizable, with eigenvalues $\kappa_i \circ \phi_j^{-1}$, $1 \leqslant i \leqslant d-1$, and satisfies

$$\frac{\partial n_j}{\partial u} = -K_j \frac{\partial \phi_j}{\partial u} . \tag{B.25}$$

Thus, recalling that $0 = \partial_x \nu^2 = 2\nu' \cdot \nu$, we have

$$\begin{aligned} \mathrm{meas}\,(\mathfrak{T}'_\rho(S)) &= \sum_{j=1}^p \int_{O_j} \psi_j \; dudt \\ &= \sum_{j=1}^p \int_{\Psi_j(U_j \times [0,\rho))} \psi_j \; dudt \end{aligned}$$

[28] Compare [18], Ch. 1.

[29] As $S = \bigcup_{j=1}^p \phi_j(U_j)$, we have $\mathscr{T}_\rho(S) = \bigcup_{j=1}^p O_j$ for any $0 < \rho \leqslant \mathrm{minfoc}\,(S)$.

[30] T^*S being the cotangent bundle of S.

[31] K_j is actually the Weingarten map $\mathcal{W}_x = -\nu'(x)$ "written in the local chart" (U_j, ϕ_j).

$$
\begin{aligned}
&= \sum_{j=1}^{p} \int_{U_j \times [0,\rho)} \Psi_j^*(\psi_j dudt) \\
&= \sum_{j=1}^{p} \int_{U_j \times [0,\delta)} \psi_j \circ \Psi_j \, \left| \det \left(\frac{\partial \Psi_j}{\partial (u,t)} \right) \right| dudt \\
&\overset{(B.25)}{=} \sum_{j=1}^{p} \int_{U_j \times [0,\rho)} \psi_j \circ \Psi_j \, \left| \det \left[\frac{\partial \phi_j}{\partial u} - tK_j \cdot \frac{\partial \phi_j}{\partial u}, \ \nu(\phi_j(u)) \right] \right| dudt \\
&= \sum_{j=1}^{p} \int_{U_j \times [0,\rho)} \psi_j \circ \Psi_j \, \left| \det \left(\left(\mathbb{1}_{d-1} - tK_j \right) \left[\frac{\partial \phi_j}{\partial u}, \ \nu(\phi_j(u)) \right] \right) \right| dudt \\
&= \sum_{j=1}^{p} \int_{U_j \times [0,\rho)} \psi_j \circ \Psi_j \, |\det(\mathbb{1}_{d-1} - tK_j)| \left| \det \left[\frac{\partial \phi_j}{\partial u}, \ \nu(\phi_j(u)) \right] \right| dudt \\
&\leqslant \int_0^\rho \sum_{j=1}^{p} \int_{U_j} \psi_j \big(\phi_j(u) + t\nu(\phi_j(u)) \big) \left| \det \left[\frac{\partial \phi_j}{\partial u}, \ \nu(\phi_j(u)) \right] \right| du \ (1+t\kappa)^{d-1} dt \\
&= \int_0^\rho \sum_{j=1}^{p} \int_{U_j} \psi_j \big(\phi_j(u) + t\nu(\phi_j(u)) \big) \left(\det \left(\frac{\partial \phi_j}{\partial u} \right)^T \frac{\partial \phi_j}{\partial u} \right)^{1/2} du \ (1+t\kappa)^{d-1} dt \\
&= \int_0^\rho \sum_{j=1}^{p} \int_{\phi_j(U_j)} \psi_j \big(x + t\nu(x) \big) \, d\mathcal{H}^{d-1}(x) \ (1+t\kappa)^{d-1} dt \quad \text{(see [11, Theorem 2, pg. 99])} \\
&\overset{\textbf{(ii)}}{\leqslant} \int_0^\rho \sum_{j=1}^{p} \int_{\bigcup_{i=1}^{p} \phi_i(U_i)} \psi_j \big(x + t\nu(x) \big) \, d\mathcal{H}^{d-1}(x) \ (1+t\kappa)^{d-1} dt \\
&= \int_0^\rho \int_S \sum_{j=1}^{p} \psi_j \big(x + t\nu(x) \big) \, d\mathcal{H}^{d-1}(x) \ (1+t\kappa)^{d-1} dt \\
&\overset{\textbf{(iv)}}{=} \int_0^\rho \int_S d\mathcal{H}^{d-1}(x) \ (1+t\kappa)^{d-1} dt \\
&= \frac{(1+\rho\kappa)^d - 1}{d\,\kappa} \, \mathcal{H}^{d-1}(S) \, .
\end{aligned}
$$

□

B.9. Kolmogorov Non-degenerate Normal Forms

Let $H\colon \mathcal{M} := \mathbb{R}^d \times \mathbb{T}^d \to \mathbb{R}$ be a C^2-Hamiltonian. An embedded torus $\mathcal{T}$ in $\mathcal{M}$ is said to be H-Kolmogorov non-degenerate if there exists a neighbourhood $\mathcal{M}_0$ of $\{0\} \times \mathbb{T}^d$ in $\mathcal{M}$, a symplectic change of coordinates $\phi\colon \mathcal{M}_0 \to \mathcal{M}$ with $\phi(\{0\} \times \mathbb{T}^d) = \mathcal{T}$, a constant $E \in \mathbb{R}$, a vector $\omega \in \mathbb{R}^d$ and a function $Q\colon \mathcal{M}_0 \to \mathbb{R}$ of class C^2 such that

$$
H \circ \phi(y,x) = E + \omega \cdot y + Q(y,x) \qquad \text{and} \qquad \partial_y^\mu Q(0,\cdot) \equiv 0, \ \forall \, \mu \in \mathbb{N}_0^d, \ |\mu|_1 \leqslant 1, \tag{B.26}
$$

and

$$
\det \langle \partial_{yy} Q(0,\cdot) \rangle \neq 0. \tag{B.27}
$$

A Hamiltonian H in the form (B.26) is said to be in Kolmogorov normal form. The Kolmogorov normal form is said to be non-degenerate if, in addition, the quadratic (in y) part Q satisfies (B.27).

ACKNOWLEDGMENTS

We are grateful to Carlangelo Liverani for useful discussions.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

1. Argentieri, F., Isolated Points of Diophantine Sets, arXiv:2011.10267 (2020).
2. Arnol'd, V. I., Proof of a Theorem of A. N. Kolmogorov on the Invariance of Quasi-Periodic Motions under Small Perturbations of the Hamiltonian, *Russian Math. Surveys,* 1963, vol. 18, no. 5, pp. 9–36; see also: *Uspekhi Mat. Nauk,* 1963, vol. 18, no. 5, pp. 13–40.
3. Biasco, L. and Chierchia, L., On the Measure of Lagrangian Invariant Tori in Nearly-Integrable Mechanical Systems, *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.,* 2015, vol. 26, no. 4, pp. 423–432.
4. Biasco, L. and Chierchia, L., KAM Theory for Secondary Tori, arXiv:1702.06480 (2017).
5. Biasco, L. and Chierchia, L., Explicit Estimates on the Measure of Primary KAM Tori, *Ann. Mat. Pura Appl. (4),* 2018, vol. 197, no. 1, pp. 261–281.
6. Biasco, L. and Chierchia, L., On the Topology of Nearly-Integrable Hamiltonians at Simple Resonances, *Nonlinearity,* 2020, vol. 33, no. 7, pp. 3526–3567.
7. Celletti, A. and Chierchia, L., A Constructive Theory of Lagrangian Tori and Computer-Assisted Applications, in *Dynamics Reported,* Dynam. Report. Expositions Dynam. Systems (N. S.), vol. 4, Berlin: Springer, 1995, pp. 60–129.
8. Chierchia, L., Quasi-Periodic Schrödinger Operators in One Dimension, Absolutely Continuous Spectra, Bloch Waves, and Integrable Hamiltonian Systems, *Tech. rep.,* New York: New York Univ., 1986.
9. Chierchia, L. and Gallavotti, G., Smooth Prime Integrals for Quasi-Integrable Hamiltonian Systems, *Nuovo Cimento B (11),* 1982, vol. 67, no. 2, pp. 277–295.
10. Chierchia, L. and Koudjinan, C., V. I. Arnold's "Pointwise" KAM Theorem, *Regul. Chaotic Dyn.,* 2019, vol. 24, no. 6, pp. 583–606.
11. Evans, L. C. and Gariepy, R. F., *Measure Theory and Fine Properties of Functions,* rev. ed., Boca Raton, Fla.: CRC, 2015.
12. Koudjinan, C., Quantitative KAM Normal Forms and Sharp Measure Estimates, *PhD Thesis,* Roma, Università degli Studi Roma Tre, 2019, 222 pp.
13. Minty, G. J., On the Extension of Lipschitz, Lipschitz – Hölder Continuous, and Monotone Functions, *Bull. Amer. Math. Soc.,* 1970, vol. 76, no. 2, pp. 334–339.
14. Neishtadt, A. I., Estimates in the Kolmogorov Theorem on Conservation of Conditionally Periodic Motions, *J. Appl. Math. Mech.,* 1981, vol. 45, no. 6, pp. 766–772; see also: *Prikl. Mat. Mekh.,* 1981, vol. 45, no. 6, pp. 1016–1025.
15. Pöschel, J., Integrability of Hamiltonian Systems on Cantor Sets, *Comm. Pure Appl. Math.,* 1982, vol. 35, no. 5, pp. 653–696.
16. Pöschel, J., A Lecture on the Classical KAM Theorem, in *Smooth Ergodic Theory and Its Applications (Seattle, Wash., 1999),* Proc. Sympos. Pure Math., vol. 69, Providence, R.I.: AMS, 2001, pp. 707–732.
17. Salamon, D. A., The Kolmogorov – Arnold – Moser Theorem, *Math. Phys. Electron. J.,* 2004, vol. 10, Paper 3, 37 pp.
18. Sternberg, Sh., *Curvature in Mathematics and Physics,* Mineola, N.Y.: Dover, 2012.
19. Whitney, H., Analytic Extensions of Differentiable Functions Defined in Closed Sets, *Trans. Amer. Math. Soc.,* 1934, vol. 36, no. 1, pp. 63–89.
20. Zehnder, E., *Lectures on Dynamical Systems: Hamiltonian Vector Fields and Symplectic Capacities,* Berlin: EMS, 2010.