Annali di Matematica
DOI 10.1007/s10231-017-0678-8

Explicit estimates on the measure of primary KAM tori

L. Biasco[1] · **L. Chierchia**[1]

Received: 6 December 2016 / Accepted: 5 July 2017

Abstract From KAM theory it follows that the measure of phase points which do not lie on Diophantine, Lagrangian, "primary" tori in a nearly integrable, real-analytic Hamiltonian system is $O(\sqrt{\varepsilon})$, if ε is the size of the perturbation. In this paper we discuss how the constant in front of $\sqrt{\varepsilon}$ depends on the unperturbed system and in particular on the phase-space domain.

Keywords KAM theory · Invariant tori · Nearly integrable systems · Hamiltonian dynamics

Mathematics Subject Classification 37J40 · 70H08

Contents

This research was partially supported by the European Research Council under FP7 "Hamiltonian PDEs and small divisor problems: a dynamical systems approach" and by MIUR (Italy): PRIN—"Variational and perturbative aspects of nonlinear differential problems".

✉ L. Chierchia
luigi@mat.uniroma3.it

L. Biasco
biasco@mat.uniroma3.it

1 Dipartimento di Matematica e Fisica, Università degli Studi Roma Tre, Largo San L. Murialdo 1, 00146 Rome, Italy

Published online: 11 July 2017

1 Introduction

According to classical KAM theory, the majority of the Lagrangian, nonresonant invariant tori of a "general" completely integrable Hamiltonian system persists under the effect of a small enough perturbations ([1,17,20]; see, also, [2, § 6.3] for a review and [15] for a divulgative exposition).

Indeed, in bounded regions of the phase space $\mathbb{R}^n \times \mathbb{T}^n$ (action-angle) such tori—which are also called "primari" tori—form a set of positive Liouville (Lebesgue) measure, whose complement has a measure proportional to $\sqrt{\varepsilon}$, if ε measures the size of the perturbing function [21,22].

The square root behavior, in such measure estimates, is optimal in the sense that, in general, at simple resonances, for $\varepsilon \neq 0$, there appear regions of size proportional to $\sqrt{\varepsilon}$ free of primary invariant tori as trivially shows the example of the simple pendulum with gravity[1] ε.

It is therefore natural to look for explicit evaluations of the constant in front of $\sqrt{\varepsilon}$ in the KAM measure estimates of the complement of invariant primary tori.

In [21,22] such constant, which depends on analytic properties of the integrable limit, is left implicit, and, somewhat surprisingly, to the best of our knowledge, there are no explicit evaluations of it in the vast literature on classical KAM theory. On the other hand, KAM is a constructive technique and discussions about "KAM constants" are clearly relevant, as also testified by the large literature on them; compare, e.g., [5–14,16,19].

We also point out that an explicit dependence upon the domain in the above measure estimate is crucial in investigating the more complicate problem of the existence and abundance of *secondary tori*, i.e., of those tori which arise by effect of the perturbation around simple resonances. In [2] it is conjectured that "in a generic system with three or more degrees of freedom the measure of the nontorus set has order ε," while in [4] it is given a sketchy proof[2] that *the union of primary and secondary tori leave out (for general mechanical systems) a region of measure* $\varepsilon|\log\varepsilon|^a$ *(for a suitable* $a > 0$*)*. To achieve such result one needs to control simultaneously a large number of regions around simple resonances and to apply KAM measure estimates taking into account different (local) phase-space domains, including neighborhoods of separatrices: To carry out such strategy simple explicit measure estimates—such as the above—are necessary.

In this paper we compute explicitly the constant in front of $\sqrt{\varepsilon}$ up to a constant depending only on n (the number of degrees of freedom) and $\tau > n - 1$ (the uniform "Diophantine exponent").

More precisely, we consider a real-analytic, nearly integrable Hamiltonian

$$\mathtt{H} : (p, q) \in \mathtt{D} \times \mathbb{T}^n \mapsto \mathtt{H}(p, q) = h(p) + f(p, q) \in \mathbb{R}$$

[1] Just look at the phase portrait of the simple pendulum $\frac{1}{2}p^2 + \varepsilon\cos q$, $(p, q) \in \mathbb{R} \times \mathbb{T}$, and observe that the region enclosed by the separatrix $\frac{1}{2}p^2 + \varepsilon\cos q = \varepsilon$ has measure $4\sqrt{2}\cdot\sqrt{\varepsilon}$.

[2] A complete proof will appear elsewhere.

where $\mathtt{D}$ is an *arbitrary* bounded domain in $\mathbb{R}^n$ and $\mathbb{T}^n$ is the standard flat n-torus (with periods 2π); f is a small perturbation function and the integrable limit h is Kolmogorov nondegenerate on $\mathtt{D}$, i.e., its hessian is invertible on $\mathtt{D}$.

The main result—Theorem 1—will be formulated in terms of a few (five) parameters, which we now describe briefly (precise definitions will be given in § 2):

- The Hamiltonian $\mathtt{H}$ is assumed to be real-analytic on $\mathtt{D} \times \mathbb{T}^n$: Therefore there exists $r_0 > 0$ and $0 < s \le 1$ such that $\mathtt{H}$ is holomorphic on a complex r_0-neighborhood of $\mathtt{D}$ and a s-complex neighborhood of $\mathbb{T}^n$.
- The smallness of the perturbation f will be measured by $\epsilon := \dfrac{\varepsilon}{\mathtt{M} r_0^2}$ where: $\varepsilon = \|f\|_{r_0,s}$ denotes the sup-norm on the above complex neighborhood of f and $\mathtt{M} := \|h_{pp}\|_{r_0}$ the sup-norm of the Hessian matrix of h.
- The "torsion" associated to h will be measured by $\mu := \dfrac{\inf_{\mathtt{D}} |\det h_{pp}|}{\mathtt{M}^n}$; note that $0 < \mu \le 1$ (compare (5)).
- Last "independent parameter" will be the number $\lambda := \mathtt{L}\,\mathtt{M}$, where $\mathtt{L}$ denotes a suitable uniform Lipschitz constant of the local complex inverse of the "frequency map" $p \mapsto \omega = h_p(p)$ (compare (9)); indeed one can show that $1 \le \lambda \le 2 \cdot n!\, \mu^{-1}$ (see (14)).

Notice that the parameters ϵ, μ and λ are *dimensionless parameters* (i.e., do not have physical dimensions).

Then, fixed $\nu := \tau + 2 > n + 1$, we will show that there exist a positive constants $c < 1$ depending only on n and ν such that if the perturbation is so small that

$$\epsilon \le c\, \frac{\mu^6}{\lambda^2}\, s^{4\nu},$$

then one can construct a family $\mathcal{T}_\alpha$ of $\mathtt{H}$-invariant primary tori. Such tori live in $\mathtt{D}_{r_0} \times \mathbb{T}^n$ (where $\mathtt{D}_{r_0}$ is a real r_0-neighborhood of $\mathtt{D}$), and the $\mathtt{H}$-flow on them is analytically conjugated to the Kronecker flow $x \in \mathbb{T}^n \mapsto x + \omega t$ for a frequency $\omega \in \mathbb{R}^n$ which is (α, τ)-Diophantine[3] with $\tau = \nu - 2$ and α proportional to $\sqrt{\epsilon}$:

$$\alpha := \frac{\lambda}{\hat{c}\, \mu\, s^{3\nu}}\, (\mathtt{M} r_0)\, \sqrt{\epsilon},$$

where $\hat{c} < 1$ is a suitable constant depending only on n and ν.

The upshot is, then, the following measure estimate (where "meas" denotes outer Lebesgue measure):

$$\text{meas}\,\big((\mathtt{D} \times \mathbb{T}^n) \backslash \mathcal{T}_\alpha\big) \le C\,\sqrt{\epsilon}$$

where the constant C is given by

$$C := \kappa\, \big(\max\big\{\mu^2 r_0\,,\ \text{diam}\ \mathtt{D}\big\}\big)^n \cdot \frac{\lambda^{n+2}}{\mu^3\, s^{3\nu}},$$

where $\kappa > 0$ is a suitable constant depending only on n and ν.

Remarks (i) In fact, we shall prove a stronger statement, which is non trivial even in case of $\mathtt{D}$ of measure zero or even finite (compare (20)).

(ii) Of course, more refined estimates are possible if one adds extra hypotheses on the domain $\mathtt{D}$ (e.g., smooth boundary) and it would be interesting to give bounds which take into account geometrical properties of $\mathtt{D}$.

[3] I.e., $|\omega \cdot k| \ge \alpha / |k|_1^\tau$ for all $k \in \mathbb{Z}^n \backslash \{0\}$; compare (1).

(iii) We do not compute explicitly the dependence upon n (and ν): Indeed it is well known that in such generality explicit bounds on c tend to be quite "pessimistic," however, in concrete example, such as a forced pendulum, the standard map or particular three body problems computer-assisted (rigorous) upper bounds on ε are in excellent agreement with experimental data (see, e.g., [5,6,8,13] and references therein).

2 Notations and setup

Given $r > 0$, p_0 a point of $\mathbb{R}^n$ or $\mathbb{C}^n$ and D a subset of $\mathbb{R}^n$ or $\mathbb{C}^n$, we denote:

$$\begin{array}{ll} \mathrm{B}_r(p_0) := \{p \in \mathbb{R}^n \big| \, |p - p_0| < r\}, & (p_0 \in \mathbb{R}^n), \\ \mathbb{B}_r(p_0) := \{p \in \mathbb{C}^n \big| \, |p - p_0| < r\}, & (p_0 \in \mathbb{C}^n), \\ \mathrm{B}_r(D) := \bigcup_{p_0 \in D} \mathrm{B}_r(p_0), & (D \subseteq \mathbb{R}^n), \\ \mathbb{B}_r(D) := \bigcup_{p_0 \in D} \mathbb{B}_r(p_0), & (D \subseteq \mathbb{C}^n), \end{array}$$

where in $\mathbb{R}^n$ and $\mathbb{C}^n$, $|x| = |(x_1, \ldots, x_n)|$ will denote the sup-norm $\max_i |x_i|$.

For a matrix (or a tensor) A, $\|A\|$ denotes the standard operator norm $\sup_{|x|=1} |Ax|$.

The standard flat n-torus $\mathbb{R}^n/(2\pi\mathbb{Z}^n)$ is denoted by $\mathbb{T}^n$ and, for $s > 0$, $\mathbb{T}^n_s$ denotes its complex neighborhood of points q with norm of the imaginary part $|\operatorname{Im} q| = |(\operatorname{Im} q_1, \ldots, \operatorname{Im} q_n)| < s$:

$$\mathbb{T}^n_s := \{y \in \mathbb{C}^n \big| \, |\operatorname{Im} q| < s\}/(2\pi\mathbb{Z}^n).$$

If D is an arbitrary bounded set in $\mathbb{R}^n$ and h, respectively, f, a real-analytic function (with values in $\mathbb{R}^m$ or in matrix spaces) with bounded holomorphic extension on $\mathbb{B}_r(D)$ for some $r > 0$, respectively, on $\mathbb{B}_r(D) \times \mathbb{T}^n_s$ for some $r, s > 0$, we define its analytic sup-norm as, respectively,

$$\|h\|_{D,r} := \sup_{y \in \mathbb{B}_r(D)} |h(p)|, \qquad \|f\|_{D,r,s} := \sup_{(p,q) \in \mathbb{B}_r(D) \times \mathbb{T}^n_s} |f(p,q)|.$$

The Lipschitz semi-norm of a function $f : \Omega \to \mathbb{R}^m$, will be denoted by

$$|f|_{\mathrm{Lip},\Omega} := \sup_{\omega_1, \omega_2 \in \Omega, \ \omega_1 \neq \omega_2} \frac{|f(\omega_1) - f(\omega_2)|}{|\omega_1 - \omega_2|}.$$

If D is an open set and $H : D \times \mathbb{T}^n \to \mathbb{R}$ is a C^2 function, ϕ^t_H denotes its Hamiltonian flow, namely, $\big(p(t), q(t)\big) = \phi^t_H(p, q)$ solves the standard Hamilton equations[4]

$$\begin{cases} \dot p(t) := \dfrac{\mathrm{d}p}{\mathrm{d}t}(t) = -\partial_q H(p(t), q(t)) \\ \\ \dot q(t) := \dfrac{\mathrm{d}q}{\mathrm{d}t}(t) = \partial_p H(p(t), q(t)) \end{cases}, \qquad (p(0), q(0)) = (p, q).$$

For example, if $H(p, q) = h(p)$, then the flow ϕ^t_h is linear with frequency $\omega := \partial_p h(p)$, namely, $\phi^t_h(p, q) = (p, q + \omega t)$.

[4] Equivalently, ϕ^t_H denotes the Hamiltonian associated to the standard symplectic form $dp \wedge dq = \sum_{i=1}^n dp_i \wedge dq_i$.

Given $\alpha, \tau > 0$, a vector $\omega \in \mathbb{R}^n$ is said to be (α, τ)-Diophantine if

$$|\omega \cdot k| := \left|\sum_{j=1}^{n} \omega_j k_j\right| \geq \frac{\alpha}{|k|_1^\tau}, \qquad \forall\, k \in \mathbb{Z}^n \backslash \{0\}, \tag{1}$$

where $|k|_1 := \sum |k_j|$ denotes the 1-norm. It is well known that, fixed $\tau > n-1$, almost all (in the sense of Lebesgue measure) $\omega \in \mathbb{R}^n$ are (α, τ)-Diophantine for some $\alpha > 0$. Indeed such statement follows immediately observing that[5]

$$\text{meas}\left\{\omega \in \mathrm{B}_R(0) \middle| \, \omega \text{ is not } (\alpha, \tau)\text{-Diophantine}\right\} \leq c\, R^{n-1}\alpha,$$

with a constant c depending only on n and τ.

Finally, given a $2n$-vector (y, x), π_1 and π_2 denote, respectively, the projections on the first and second n components:

$$\pi_1(y, x) = y \qquad \text{and} \qquad \pi_2(y, x) = x. \tag{2}$$

3 Assumptions

Fix $n \geqslant 2$ and $\tau > n - 1$. Let $\mathtt{D}$ be any nonempty, bounded subset of $\mathbb{R}^n$. Let

$$\mathtt{H} := h + f$$

with h and f real-analytic functions with holomorphic extensions on, respectively, $\mathbb{B}_{r_0}(\mathtt{D})$ and $\mathbb{B}_{r_0}(\mathtt{D}) \times \mathbb{T}^n_s$ for some $r_0 > 0$ and $0 < s \leq 1$, and having finite norms:

$$\mathtt{M} := \|h_{pp}\|_{\mathtt{D}, r_0}, \qquad \varepsilon := \|f\|_{\mathtt{D}, r_0, s}. \tag{3}$$

Assume that the frequency map $p \in \mathtt{D} \to \omega = h_p$ is a local diffeomorphism, namely, assume:

$$d := \inf_{\mathtt{D}} |\det h_{pp}| > 0. \tag{4}$$

4 The local frequency map

Under assumption (4) the frequency map is a *local* real-analytic diffeomorphism in the neighborhood of any point of $\mathtt{D}$. More precisely, the following lemma holds. Define[6]

$$\mu := \frac{d}{\mathtt{M}^n} \leq 1. \tag{5}$$

Lemma 1 *Let*

$$c_0 = \frac{1}{8n \cdot n!^2}, \qquad \hat{c}_0 = \frac{1}{4n \cdot n!}. \tag{6}$$

and define

$$r_\star := \hat{c}_0 \mu\, r_0, \qquad \rho_\star := c_0 \mu^2\, \mathtt{M}\, r_0. \tag{7}$$

[5] "meas" stands for Lebesgue measure, or, in general, for outer Lebesgue measure.

[6] Since any eigenvalue of h_{pp} is bounded in absolute value by $\|h_{pp}\| \leq \mathtt{M}$, $d \leq \sup_{\mathtt{D}} |\det h_{pp}| \leq \mathtt{M}^n$.

Then, for every $p_0 \in \mathtt{D}$ *the frequency map* $p \to \omega = h_p$ *has a real-analytic inverse map,* $\omega \to \mathtt{p}(\omega; p_0)$*, defined in a neighborhood of* $\omega_0 := h_p(p_0)$

$$\mathtt{p} = h_p^{-1} : \omega \in \mathbb{B}_{\rho_\star}(\omega_0) \mapsto \mathtt{p}(\omega; p_0) \in \mathbb{B}_{r_\star}(p_0), \tag{8}$$

with uniform Lipschitz constant[7]

$$\mathtt{L} := \sup_{p_0 \in \mathtt{D}} |\mathtt{p}(\cdot; p_0)|_{\mathrm{Lip}, \mathbb{B}_{\rho_\star}(\omega_0)} = \sup_{p_0 \in \mathtt{D}} \sup_{\mathbb{B}_{\rho_\star}(\omega_0)} \| \mathtt{p}_\omega(\cdot; p_0)\| \tag{9}$$

satisfying

$$\mathtt{L} \le \frac{\mathtt{M}^{-1}}{2n\, \hat{c}_0\, \mu}, \tag{10}$$

and

$$\sup_{p_0 \in \mathtt{D}} \sup_{\mathbb{B}_{\frac{3}{4}\rho_\star}(\omega_0)} \| \mathtt{p}_{\omega\omega}(\cdot; p_0)\| \le \frac{\mathtt{L}}{\frac{c_0}{4}\mu^2\, \mathtt{M} r_0}. \tag{11}$$

Proof Writing out the inverse of the matrix h_{pp} (Cramer's rule), by Leibniz formula for the ji-minor of h_{pp}, one has, uniformly on[8] D:

$$|(h_{pp}^{-1})_{ij}| \le \frac{1}{d}\,(n-1)!\, \mathtt{M}^{n-1},$$

which implies

$$\sup_{\mathtt{D}} \|(h_{pp})^{-1}\| \le \frac{n!}{\mu}\, \mathtt{M}^{-1}. \tag{12}$$

Let $T := h_{pp}^{-1}(p_0)$. Then, by standard Cauchy estimates[9], it follows that for any $p \in \mathbb{C}^n$ such that $|p - p_0| \le r_\star$ one has

$$\|I - T h_{pp}(p)\| \le \|T\|\, \|h_{pp}(p) - h_{pp}(p_0)\| \le \|T\|\, n\, \frac{\mathtt{M}}{r_0 - r_\star}\, r_\star \overset{(12)}{\le} \frac{n \cdot n!}{\mu}\, \frac{r_\star}{r_0 - r_\star} \overset{(6)}{\le} \frac{1}{2}.$$

Thus, by the standard inverse function theorem (see "Appendix A," Eqs. (75), (76), (77)) and Cauchy estimates, relations (8), (9), (10) and (11) follow immediately with the constants in (6). □

For later use, we point out that[10]

$$\lambda := \mathtt{L}\,\mathtt{M} \ge 1, \tag{13}$$

and that, by (10),

$$\lambda \le \frac{1}{2n\hat{c}_0}\, \frac{1}{\mu} = 2 \cdot n!\, \frac{1}{\mu}. \tag{14}$$

[7] Notice that on convex domains the Lipschitz semi-norm of a differentiable function coincides with the sup-norm of its Jacobian.

[8] Note that $\sup_{i,j} \sup_D |h_{p_i p_j}| \le \mathtt{M} := \sup_D \sup_i \sum_j |h_{p_i p_j}|$.

[9] *If* $f : \mathbb{B}_r(D) \to \mathbb{C}^m$ *is holomorphic,* ∂^α *is a partial derivative of order* $k = \alpha_1 + \cdots + \alpha_n$ *and* $0 < r' < r$*, then*

$$\sup_{\mathbb{B}_{r'}(D)} |\partial^\alpha f| \le \frac{\sup_{\mathbb{B}_r(D)} |f|}{(r - r')^k}.$$

[10] Indeed: $1 = \|I\| = \|h_{pp}(p) h_{pp}^{-1}(p)\| = \|h_{pp}(p)\, \mathtt{p}_\omega(h_p(p))\| \le \|h_{pp}(p)\|\, \|\mathtt{p}_\omega(h_p(p))\| \le \mathtt{M}\mathtt{L}$.

5 The classical analytic KAM theorem

Theorem 1 *Let the assumptions in Sect. 3 hold and let λ, μ and c_0 be as in Sect. 4 and let $\nu = \tau + 2$.*

There exist positive constants $c_\star < 1/(8 \cdot n!)$ and κ, depending only on n and τ, such that, if

$$c := \frac{c_\star^2}{2^{17} \cdot n^2\,(n!)^6}, \qquad \hat{c} := \frac{c_\star}{2^4}, \tag{15}$$

and if ε is such that

$$\epsilon := \frac{\varepsilon}{\mathtt{M} r_0^2} \le c\,\frac{\mu^6}{\lambda^2}\,s^{4\nu}, \tag{16}$$

then the following holds. Define

$$\alpha := \frac{\lambda}{\hat{c}\,\mu\,s^{3\nu}}\,(\mathtt{M} r_0)\,\sqrt{\epsilon}, \qquad \hat{r} := \frac{c_0}{2}\,\mu^2 r_0, \qquad r_\epsilon := \frac{\lambda}{c_\star}\,\sqrt{\epsilon}\,r_0. \tag{17}$$

Then, there exists a positive measure set $\mathcal{T}_\alpha \subseteq \mathrm{B}_{2\hat{r}}(\mathtt{D}) \times \mathbb{T}^n$ formed by "primary" Kolmogorov's tori; more precisely, for any point $(p, q) \in \mathcal{T}_\alpha$, $\phi^t_{\mathtt{H}}(p, q)$ covers densely an $\mathtt{H}$-invariant, analytic, Lagrangian torus, with $\mathtt{H}$-flow analytically conjugated to a linear flow with (α, τ)-Diophantine frequencies $\omega = h_p(p_0)$, for a suitable $p_0 \in \mathtt{D}$; each of such tori is a graph over $\mathbb{T}^n$ r_ϵ-close to the unperturbed trivial graph $\{(p, \theta) = (p_0, \theta) \,|\, \theta \in \mathbb{T}^n\}$.

Finally, the Lebesgue outer measure of $(\mathtt{D} \times \mathbb{T}^n)\backslash\mathcal{T}_\alpha$ is bounded by:

$$\mathrm{meas}\,\big((\mathtt{D} \times \mathbb{T}^n)\backslash\mathcal{T}_\alpha\big) \le C\,\sqrt{\epsilon} \tag{18}$$

with

$$C := \kappa\,\big(\max\big\{\mu^2 r_0\,,\ \mathrm{diam}\ \mathtt{D}\big\}\big)^n \cdot \frac{\lambda^{n+2}}{\mu^3\,s^{3\nu}}; \tag{19}$$

indeed, there exist N and, for $1 \le i \le N$, $p_i \in \mathtt{D}$, such that $\mathtt{D} \subseteq \bigcup_{i=1}^N \mathrm{B}_{\hat{r}}(p_i)$ and

$$\mathrm{meas}\,\Big(\big(\bigcup_{i=1}^N \mathrm{B}_{\hat{r}}(p_i) \times \mathbb{T}^n\big)\backslash\mathcal{T}_\alpha\Big) \le C\,\sqrt{\epsilon}. \tag{20}$$

6 Remarks

(i) The constant $c_\star$ is, essentially, the "smallness" constant appearing in a local KAM normal form (see Theorem 2). The constant κ is given in (74).

(ii) Notice that the set $\mathcal{T}_\alpha$ of persistent primary Kolmogorov's tori leaves in $\mathrm{B}_{2\hat{r}}(\mathtt{D}) \times \mathbb{T}^n$, where $\mathrm{B}_{2\hat{r}}(\mathtt{D})$ is the $2\hat{r}$-neighborhood of $\mathtt{D}$, which contains the finite $\hat{r}$-covering appearing in (20). Thus, estimate (20) implies at once (18).
Observe that (20) is meaningful also in the case of sets $\mathtt{D}$ of measure zero (such as a singleton).

(iii) To obtain (20), we first prove a covering lemma for $\mathtt{D}$ through N balls $\mathrm{B}_{\hat{r}}(p_i)$ (with suitable $p_i \in \mathtt{D}$) giving an explicit bound on N in terms of the ratio between the diameter of $\mathtt{D}$ and $\hat{r}$ (compare § 6.1); then, we prove the theorem for the special case of a ball. With this strategy the geometric properties of $\mathtt{D}$ enter only in a very primitive (but general) way through its diameter (needed to estimate N) and explains the term $\big(\max\big\{\mu^2 r_0\,,\ \mathrm{diam}\ \mathtt{D}\big\}\big)^n$ appearing in the constant C in (19).
Such strategy reflects our main purpose, which is to give a simple expression for C valid for *completely arbitrary* domain $\mathtt{D}$.

Clearly, sharper estimates can be obtained taking into account the geometry of $\mathtt{D}$ under suitable regularity assumptions. For example, if $\mathtt{D}$ has a piecewise C^1 boundary (and $\hat{r}$ is small enough), it would not be difficult—adapting the covering lemma—to get an estimate in terms of the measure of $\mathtt{D}$ and the $(n-1)$-dimensional measure of the boundary of $\mathtt{D}$ times $\hat{r}$.

6 Proof of KAM theorem 1

The proof of Theorem 1 is divided into six steps.

6.1 Local reduction

The first step consists in covering $\mathtt{D}$ with N balls centered at points of $\mathtt{D}$ (with an explicit upper bound on N), thus reducing the theorem to the special case in which the domain is a ball. Indeed, the following simple result holds.

Lemma 2 (Covering lemma) *Let $E \subseteq \mathbb{R}^n$ be a nonempty set of finite diameter. Then, for any $r > 0$ there exists an integer N, with*[11]

$$1 \le N \le \left(\left[\frac{\operatorname{diam} E}{r} \right] + 1 \right)^n, \tag{21}$$

and N points $p_i \in E$ such that

$$E \subseteq \bigcup_{i=1}^{N} \mathrm{B}_r(p_i). \tag{22}$$

Proof Let $\delta := \operatorname{diam} E$ and let $z_i = \inf\{x_i |\ x \in E\}$. Then $E \subseteq K := z + [0, \delta]^n$. Let $0 < r' < r$ close enough to r so that $\lceil \delta/r' \rceil = [\delta/r] + 1 =: M$. Then, one can cover K with M^n closed, contiguous cubes K_j, $1 \le j \le M^n$, with edge of length r'. Let j_i be the indices such that $K_{j_i} \cap E \neq \emptyset$ and pick a $p_i \in K_{j_i} \cap E$; let $1 \le N \le M^n$ be the number of such cubes. Observe that, since we have chosen the sup-norm in $\mathbb{R}^n$, one has $K_{j_i} \subseteq \mathrm{B}_r(p_i)$ and, therefore, (22) follows with N as in (21). □

We now apply the lemma with $E = \mathtt{D}$ and $r = \hat{r}$ defined by[12]

$$\hat{r} := \frac{\rho_\star}{2\,\mathtt{M}} = \frac{c_0}{2}\,\mu^2 r_0 \le \frac{r_0}{128} < \frac{r_0}{2}\,. \tag{23}$$

Thus, Lemma 2 yields that:

For suitable N points $p_i \in \mathtt{D}$, one has

$$\mathtt{D} \subseteq \bigcup_{i=1}^{N} \mathrm{B}_{\hat{r}}(p_i), \qquad 1 \le N \le \left(\left[\frac{\operatorname{diam} \mathtt{D}}{\frac{c_0}{2}\mu^2 r_0} \right] + 1 \right)^n. \tag{24}$$

Notice that, by (23), $\mathtt{H}$ is holomorphic and bounded on $\mathbb{B}_{r_0/2}\big(\mathrm{B}_{\hat{r}}(p_i)\big) \times \mathbb{T}^n_s$, for every $i \le N$.

Next we shall prove a "local" version of Theorem 1

[11] $[x]$ denotes the integer part (or "floor") function $\max\{n \in \mathbb{Z}|\ n \le x\}$, while $\lceil x \rceil$ denotes the "ceiling function" $\min\{n \in \mathbb{Z}|\ n \ge x\}$.

[12] Recall (5), (6) and (7).

6.2 A KAM local normal form après [23]

Fix one of the balls $\mathrm{B}_{\hat{r}}(p_i)$ in the covering (24). We first prove Theorem 1 with $\mathtt{D}$ *and* r_0 *replaced, respectively, by*

$$D_i := \mathrm{B}_{\hat{r}}(p_i) \qquad \text{and} \qquad \frac{r_0}{2}.$$

We shall use a "KAM normal form with parameters"; more specifically, we shall use Theorem B of [23], whose statement we recall here for convenience of the reader.

Let $r, \alpha, \mathtt{h} > 0$, $0 < s \le 1$, $\tau > n-1$ and let $\Omega \subset \mathbb{R}^n$ be a bounded open set with piecewise smooth boundary; let

$$\Omega_\alpha := \Big\{\omega \in \Omega, \ \ \text{s.t.}\ \ \mathrm{dist}(\omega, \partial\Omega) \ge \alpha \ \ \text{and} \ \ \omega \text{ is } (\alpha,\tau) - \text{Diophantine}\Big\}\,;$$

let $\omega \to e(\omega)$ and $(I,\theta,\omega) \to P(I,\theta,\omega)$ be real-analytic functions with holomorphic extension on, respectively, $\mathbb{B}_{\mathtt{h}}(\Omega_\alpha)$ and $\mathbb{B}_r(0)\times\mathbb{T}^n_s\times\mathbb{B}_{\mathtt{h}}(\Omega_\alpha)$. Finally, if $a > 0$, we define the "action rescaling map":

$$R_a(I,\theta) := (I/a, \theta). \tag{25}$$

Consider the Hamiltonian function, parameterized by ω,

$$H(I,\theta,\omega) := N(I,\omega) + P(I,\theta,\omega)\,, \qquad \text{where} \quad N := e(\omega) + \omega\cdot I\,,$$

with respect to the standard symplectic form $dI\wedge d\theta$; in particular, the integrable flow ϕ^t_N is given by $\phi^t_N(I,\theta) = (I, \theta+\omega t)$.

In [23] the following result is proven.[13]

Theorem 2 *Under the above definitions and assumptions, there exist constants*

$$0 < c_\star < \frac{\hat{c}_\star}{4\cdot n!} < \frac{1}{8\cdot n!}, \tag{26}$$

depending only on n and τ, such that if

$$|P|_{r,s,\mathtt{h}} := \sup_{\mathbb{B}_r(0)\times\mathbb{T}^n_s\times\mathbb{B}_{\mathtt{h}}(\Omega_\alpha)} |P| \le c_\star \alpha r s^\nu, \qquad \alpha s^\nu \le \mathtt{h}, \qquad (\nu := \tau+2), \tag{27}$$

then, there exist a Lipschitz homeomorphism $\varphi : \Omega \circlearrowleft$ and a family of torus embeddings

$$\Phi : \mathbb{T}^n\times\Omega \to \mathrm{B}_r(0)\times\mathbb{T}^n \subseteq \mathbb{R}^n\times\mathbb{T}^n$$

such the following holds. For every $\omega\in\Omega_\alpha$, $\Phi(\mathbb{T}^n,\omega)$ is an invariant torus for $H|_{\varphi(\omega)} := H(I,\theta,\varphi(\omega))$ and

$$(\phi^t_{H|_{\varphi(\omega)}}\circ\Phi)(\theta,\omega) = \Phi(\theta+\omega t, \omega).$$

Moreover, for each $\omega\in\Omega$, $\theta\to\Phi(\theta,\omega)$ is real-analytic on $\mathbb{T}^n_{s/2}$ and if

$$(\theta,\omega) \to \Phi_0(\theta,\omega) := (0,\theta)$$

denotes the trivial torus embedding, one has, uniformly on, respectively, $\mathbb{T}^n_{s/2}\times\Omega$ and Ω, the following estimates:

$$|R_r(\Phi-\Phi_0)|\,, \ \ \alpha s^\nu |R_r(\Phi-\Phi_0)|_{\mathrm{Lip},\Omega} \le \frac{|P|}{\hat{c}_\star \alpha r s^\nu}\,, \tag{28}$$

---

[13] For a detailed discussion and proof of such KAM normal form theorem see, also, [18].

$$|\varphi - \mathtt{id}|\,,\quad \alpha s^{\nu}|\varphi - \mathtt{id}|_{\mathrm{Lip},\Omega} \le \frac{|P|}{\hat c_\star r}\,. \tag{29}$$

Remarks (i) The constants $c_\star$ and $\hat c_\star$ can be taken equal to, respectively, $\gamma/2$ and $1/c$ where γ and c are the constants appearing in Theorem A of [23], where are not explicitly evaluated.
In [3] an infinite dimensional KAM Theorem[14] (implying Theorem 2) is proved, substituting (27) with the stronger (for s small) condition $|P|_{r,s,\mathtt{h}} \le \kappa\alpha r$, where $\kappa = \kappa(s,n,\tau) := \kappa_*^{-c\kappa_*}$ with $\kappa_* = (n+\tau)\ln((n+\tau)/s)$ and where $c > 0$ is an absolute constant.
The numerical relations between $c_\star$ and $\hat c_\star$ in (26) are assumed for later convenience (and, obviously, are compatible with [23]).

(ii) For simplicity—and because it would play no rôle—in (28) we reported slightly weaker estimates with respect to those appearing in Theorem A, where in place of R_r there appears the rescaling $W(I,\theta) := (I/r, \theta/s)$ (which means that the estimates on the angle components in [23] are better by a factor $s < 1$ than those in (28)).

(iii) Actually, the above Theorem is a synthesis of Theorem A and Theorem B in [23]. In particular, the final measure estimate in Theorem B is not reported since the constant (and its dependence upon Ω) is left implicit.

(iv) We point out that from the estimates (28) *does not follows that* Φ *is Lipschitz close to the trivial embedding* Φ_0. Indeed, if π_2 denotes the projection over the θ-component, taking into account that $\alpha = O(\sqrt{\epsilon})$ (compare (17)) , from (27) and (28) it follows that

$$|\pi_2\circ\Phi - \pi_2\circ\Phi_0|_{\mathrm{Lip}} \le \frac{c_\star}{\hat c_\star}\cdot\frac{1}{\alpha s^{\nu}} = O\Big(\frac{1}{\sqrt{\epsilon}}\Big).$$

To overcome this fact one needs suitable asymmetric rescalings of action and angle variables.

(v) As well known [12,22], maps constructed via KAM methods are smooth in the sense of Whitney, or, what is the same, have C^∞ extensions: Indeed, φ and Φ are $C^\infty(\Omega)$.

6.3 Applying the KAM normal form to $\mathtt{H}|_{\mathrm{B}_{\hat r}(p_i)\times\mathbb{T}^n}$

We now apply Theorem 2 to $\mathtt{H}$ restricted to

$$D_i\times\mathbb{T}^n := \mathrm{B}_{\hat r}(p_i)\times\mathbb{T}^n$$

where $\hat r$ is defined in (23) and p_i is one of the points introduced in Lemma 2. Recall that, by (23), $\hat r < r_0/2$ so that $\mathtt{H}$ has holomorphic extension to $\mathbb{B}_{r_0/2}(D_i)\times\mathbb{T}^n_s$.

Let

$$\Omega^{(i)} := h_p(D_i) = h_p\big(\mathrm{B}_{\hat r}(p_i)\big),\qquad \mathtt{h} := \frac{\rho_\star}{4} = \frac{\mathtt{M}\hat r}{2}, \tag{30}$$

and notice that, by (23),

$$\Omega^{(i)}\subseteq\mathrm{B}_{\mathtt{M}\hat r}(h_p(p_i))\subseteq\mathrm{B}_{\rho_\star/2}(h_p(p_i)) = \mathrm{B}_{2\mathtt{h}}(h_p(p_i)) \quad\Longrightarrow\quad \mathbb{B}_{\mathtt{h}}(\Omega^{(i)})\subseteq\mathbb{B}_{\frac34\rho_\star}(h_p(p_i)),$$

which, by Lemma 1, shows that h_p has an inverse[15] $\mathtt{p} = h_p^{-1}$ with holomorphic extension

$$\mathtt{p} = \mathtt{p}(\cdot\,; p_i):\quad \mathbb{B}_{\mathtt{h}}(\Omega^{(i)}) \to \mathbb{B}_{r_\star}(p_i)\,. \tag{31}$$

[14] More precisely, compare Theorem 5.1 in [3], case (H3), p. 755 and Remark 5.3, p. 758.

[15] Recall that $\mathtt{p} = \mathtt{p}(\cdot\,; p_i)$ and therefore depends upon i; however for ease of notation we do not indicate explicitly the dependence upon i.

Following [23], we introduce $\omega \in \mathbb{B}_{\mathtt{h}}(\Omega^{(i)})$ as parameter, and let[16]

$$p = \mathtt{p}(\omega) + I, \qquad \text{with} \qquad \omega \in \mathbb{B}_{\mathtt{h}}(\Omega^{(i)}), \qquad |I| < r := \sqrt{\frac{\varepsilon}{\mathtt{M}}} = \sqrt{\epsilon} \cdot r_0 < \frac{r_0}{4}, \tag{32}$$

and define

$$\begin{cases} N(I,\omega) := e(\omega) + \omega \cdot I := h(\mathtt{p}(\omega)) + \omega \cdot I \\ P(I,\theta,\omega) := \displaystyle\int_0^1 (1-t) h_{pp}(\mathtt{p}(\omega) + tI) I \cdot I \mathrm{d}t \ + f\big(\mathtt{p}(\omega) + I, \theta\big), \end{cases} \tag{33}$$

so that

$$\mathtt{H}(\mathtt{p}(\omega)+I,\theta) = h(\mathtt{p}(\omega) + I) + f(\mathtt{p}(\omega) + I, \theta) = N(I,\omega) + P(I,\theta,\omega) =: H(I,\theta,\omega).$$

By (6), (7), (31), (32), one has that if $\omega \in \mathbb{B}_{\mathtt{h}}(\Omega^{(i)})$ and $I \in \mathbb{B}_r(0)$, then

$$|\mathtt{p}(\omega) + I - p_i| \le |\mathtt{p}(\omega) - p_i| + |I| < r_\star + r = \hat{c}_0 \mu r_0 + \sqrt{\epsilon} r_0 < \frac{r_0}{4} + \frac{r_0}{4} = \frac{r_0}{2}, \tag{34}$$

so that $\mathtt{p}(\omega)+I \in \mathbb{B}_{r_0/2}(p_i)$. Thus, by (3), (32) and (33), H is real-analytic with holomorphic extension to $\mathbb{B}_r(0) \times \mathbb{T}^n_s \times \mathbb{B}_{\mathtt{h}}(\Omega^{(i)})$ with

$$|P|_{r,s,\mathtt{h}} \le 2\varepsilon\,. \tag{35}$$

Thus, if α is as in (17) and r as in (32), then,

$$\frac{|P|_{r,s,\mathtt{h}}}{c_\star \alpha r s^\nu} \le \frac{2\varepsilon}{c_\star \alpha r s^\nu} = \frac{2\hat{c}}{c_\star} \frac{\mu s^{2\nu}}{\lambda} \overset{(15)}{=} \frac{1}{8} \frac{\mu s^{2\nu}}{\lambda} \le \frac{1}{8}, \tag{36}$$

and the first condition in (27) is satisfied. Observe that,

$$\alpha s^\nu = \frac{\lambda}{\hat{c}\mu s^{2\nu}} (\mathtt{M} r_0)\sqrt{\epsilon} \overset{(16)}{\le} \frac{\sqrt{c}}{\hat{c}} \mu^2 \mathtt{M} r_0,$$

and, if $\mathtt{h}$ is as in (30), one has, in view of (15),

$$\frac{\alpha s^\nu}{\mathtt{h}} \overset{(16)}{\le} \frac{4\sqrt{c}}{\hat{c}\, c_0} \le 1.$$

Thus, also the second condition in (27) is satisfied and we can apply Theorem 2, obtaining the family of torus embedding[17]

$$\Phi : \mathbb{T}^n \times \Omega^{(i)} \to \mathrm{B}_r(0) \times \mathbb{T}^n \qquad (r = \sqrt{\epsilon}\, r_0)$$

as described in Theorem 2.

6.4 Kolmogorov's tori: the sets $\mathcal{T}_\alpha^{(i)}$ and $\mathcal{T}_\alpha$

The tori we obtained in the preceding section live in the "local" phase space $\{(I,\theta)|\ (I,\theta) \in \mathrm{B}_r(0) \times \mathbb{T}^n\}$. To translate the invariant tori into the original phase space $\{(p,q)|\ (p,q) \in \mathrm{B}_{r_0/2}(D_i) \times \mathbb{T}^n\}$, we define the ω-family of torus embeddings[18]

$$\theta \mapsto \Psi(\theta,\omega) := \big(\mathtt{p} \circ \varphi(\omega), 0\big) + \Phi(\theta,\omega), \qquad \omega \in \Omega^{(i)}, \tag{37}$$

16 In general, I is complex. Notice that by (5), (13), (16) (and the assumption $s \le 1$), $\epsilon < 1/16$ since $c < 1/16$.

17 Obviously, also Φ depends on i but, as above (compare footnote 15), for ease of notation we do not indicate explicitly the dependence upon i.

18 Recall (34) and footnote 17.

which, as function of θ, is real-analytic on $\mathbb{T}^n_{s/2}$. Then, from Sect. 6.3 it follows that *for* $\omega \in \Omega^{(i)}_\alpha$ *the torus* $\Psi(\mathbb{T}^n, \omega)$ *is invariant for the flow of* $\mathtt{H}$ *and, furthermore*:

$$(\phi^t_{\mathtt{H}} \circ \Psi)(\theta, \omega) = \Psi(\theta + \omega t, \omega).$$

We therefore obtain the following family of "Kolmogorov's tori" (recall, (23), that $\hat{r} < r_0/2$):

$$\mathcal{T}^{(i)}_\alpha := \Psi(\mathbb{T}^n \times \Omega^{(i)}_\alpha) \subseteq \mathrm{B}_{r_0/2}(D_i) \times \mathbb{T}^n, \qquad \mathcal{T}_\alpha := \bigcup_{i=1}^N \mathcal{T}^{(i)}_\alpha \subseteq \mathrm{B}_{r_0}(\mathtt{D}) \times \mathbb{T}^n,$$

Below, we shall show that actually $\mathcal{T}_\alpha$ lives in a smaller neighborhood of $\mathtt{D}$.

Analytic quantitative properties of the torus embedding Ψ, and hence of the family of Kolmogorov's tori, will be described in detailed in the following section.

6.5 Properties of the torus embedding

Lemma 3 *Let* Ψ *be defined as in* (37) *and let* Ψ_0 *denote the "trivial embedding"*

$$\Psi_0 \,:\, (\theta, \omega) \in \mathbb{T}^n_s \times \mathbb{B}_{\mathtt{h}}(\Omega^{(i)}) \;\mapsto\; (\mathtt{p}(\omega), \theta) \in \mathbb{B}_{r_\star}(p_i) \times \mathbb{T}^n_s\,. \tag{38}$$

Then[19]*,*

$$\sup_{\Omega^{(i)}} \sup_{\mathbb{T}^n} |\pi_1\big(\Psi - \Psi_0\big)| \le r_\epsilon \overset{(17)}{:=} \frac{\lambda\sqrt{\epsilon}}{c_\star} r_0\,, \tag{39}$$

and, hence,

$$\mathcal{T}^{(i)}_\alpha \subseteq \mathrm{B}_{r_\epsilon}(D_i) \times \mathbb{T}^n = \mathrm{B}_{\hat{r}+r_\epsilon}(p_i) \times \mathbb{T}^n, \tag{40}$$

$$\mathcal{T}_\alpha \subseteq \bigcup_{i=1}^N \mathrm{B}_{\hat{r}+r_\epsilon}(p_i) \times \mathbb{T}^n \subseteq \mathrm{B}_{2\hat{r}}(\mathtt{D}) \times \mathbb{T}^n. \tag{41}$$

Proof By definitions (38) and (37), one has

$$\Psi = \Psi_0 + \big(\mathtt{p} \circ \varphi - \mathtt{p}, 0\big) + \Phi - \Phi_0\,. \tag{42}$$

Notice that, since $\mathtt{h} = \rho_\star/4$, from (9) and (11), one has

$$\sup_{\mathbb{B}_{\mathtt{h}}(\Omega^{(i)})} \|\mathtt{p}_\omega\| \le \mathtt{L}\,, \qquad \sup_{\mathbb{B}_{\mathtt{h}}(\Omega^{(i)})} \|\mathtt{p}_{\omega\omega}\| \le \frac{4}{c_0}\,\frac{\mathtt{L}}{\mu^2\,\mathtt{M} r_0}. \tag{43}$$

Thus, (29), (35), (43), one gets

$$\sup_{\Omega^{(i)}} |\mathtt{p} \circ \varphi - \mathtt{p}| \le \mathtt{L}\,\frac{2\varepsilon}{\hat{c}_\star r} \overset{(32)}{=} \frac{2}{\hat{c}_\star}\,\lambda r. \tag{44}$$

Then, by (15), (26), (28) and (36), one gets (39).

Since (recall (38))

$$\Psi_0(\mathbb{T}^n \times \Omega^{(i)}_\alpha) = \mathtt{p}(\Omega^{(i)}_\alpha) \times \mathbb{T}^n \subseteq D_i \times \mathbb{T}^n = \mathrm{B}_{\hat{r}}(p_i) \times \mathbb{T}^n,$$

(40) follows from (39); (41) follows since $r_\epsilon \le \hat{r}$. □

[19] Recall, (2), that π_1 denotes projection onto the first n components.

To control Lipschitz norm we introduce suitable partial rescalings. Let

$$1 < \beta := \frac{\lambda}{\check{c}\mu s^{\nu}}, \qquad \text{with} \qquad \check{c} := \frac{n\hat{c}_0\hat{c}_\star}{16} = \frac{\hat{c}_\star}{2^6 \cdot n!}. \tag{45}$$

Define, for any $a > 0$, the following rescaling

$$S_a : (\theta, x) \in \mathbb{T}^n \times \mathbb{R}^n \mapsto (\theta, \omega) := (\theta, ax).$$

Now, recall (25) and let

$$\begin{cases} \tilde{\Phi} := R_{\beta r} \circ \Phi \circ S_\alpha \\ \tilde{\Phi}_0 := R_{\beta r} \circ \Phi_0 \circ S_\alpha \end{cases}, \qquad \begin{cases} \tilde{\Psi} := R_{\beta r} \circ \Psi \circ S_\alpha \\ \tilde{\Psi}_0 := R_{\beta r} \circ \Psi_0 \circ S_\alpha \end{cases} \tag{46}$$

which are defined on the domain $\mathbb{T}^n_{s/2} \times \frac{1}{\alpha}\Omega^{(i)}$. The rescaled version of (42) then becomes:

$$\tilde{\Psi} - \tilde{\Psi}_0 = \Big(\frac{1}{\beta r}(\mathrm{p} \circ \varphi - \mathrm{p}) \circ S_\alpha, 0\Big) + \tilde{\Phi} - \tilde{\Phi}_0. \tag{47}$$

Finally, let

$$\Psi_* := (\tilde{\Psi} - \tilde{\Psi}_0) \circ \tilde{\Psi}_0^{-1}, \tag{48}$$

which is defined on $\frac{1}{\beta r} D_i \times \mathbb{T}^n_{s/2}$.

The above rescaled embeddings may, now, be shown to be close, *in Lipschitz norm*, to the unperturbed rescaled embeddings:

Lemma 4 *The following bounds hold:*

$$\sup_{\mathbb{T}^n_{s/2} \times \frac{1}{\alpha}\Omega^{(i)}} |\tilde{\Phi} - \tilde{\Phi}_0| \le \frac{s^\nu}{8}, \qquad \sup_{\mathbb{T}^n_{s/2}} |\tilde{\Phi} - \tilde{\Phi}_0|_{\mathrm{Lip}, \frac{1}{\alpha}\Omega^{(i)}} \le \frac{s^\nu}{8}, \tag{49}$$

$$\sup_{\mathbb{T}^n_{s/2} \times \frac{1}{\alpha}\Omega^{(i)}} |\tilde{\Psi} - \tilde{\Psi}_0| \le \frac{s^\nu}{4}, \qquad \sup_{\mathbb{T}^n_{s/2}} |\tilde{\Psi} - \tilde{\Psi}_0|_{\mathrm{Lip}, \frac{1}{\alpha}\Omega^{(i)}} \le \frac{1}{4}, \tag{50}$$

$$\sup_{\frac{1}{\beta r} D_i \times \mathbb{T}^n_{s/2}} |\Psi_*| \le \frac{s^\nu}{4}, \qquad \sup_{\mathbb{T}^n_{s/2}} |\Psi_*|_{\mathrm{Lip}, \frac{1}{\beta r} D_i} \le \frac{1}{4}, \tag{51}$$

$$\sup_{\frac{1}{\beta r} D_i \times \mathbb{T}^n} \|\partial_\theta \Psi_*\| \le \frac{s^{\nu-1}}{2}. \tag{52}$$

Proof Since $\beta > 1$, by (28) and (36) we have that

$$\sup_{\mathbb{T}^n_{s/2} \times \frac{1}{\alpha}\Omega^{(i)}} |\tilde{\Phi} - \tilde{\Phi}_0| \le \sup_{\mathbb{T}^n_{s/2} \times \frac{1}{\alpha}\Omega^{(i)}} |R_\beta^{-1}(\tilde{\Phi} - \tilde{\Phi}_0)| = \sup_{\mathbb{T}^n_{s/2} \times \Omega^{(i)}} |R_r(\Phi - \Phi_0)| \le \frac{2\hat{c}}{\hat{c}_\star} \frac{\mu s^{2\nu}}{\lambda} \le \frac{s^\nu}{8}, \tag{53}$$

last inequality holding because of the definition of $\hat{c}$ in (15).

Analogously, (by (28), (36) and the definition of $\hat{c}$) we have that[20]

[20] If f is a Lipschitz map defined on $\Omega^{(i)}$, then $f \circ S_a$ is Lipchitz on $\frac{1}{a}\Omega^{(i)}$ and $|f \circ S_a|_{\mathrm{Lip}, \frac{1}{a}\Omega^{(i)}} = a|f|_{\mathrm{Lip}, \Omega^{(i)}}$.

$$\sup_{\mathbb{T}^n_{s/2}} |\tilde\Phi - \tilde\Phi_0|_{\mathrm{Lip},\frac{1}{\alpha}\Omega^{(i)}} \le \sup_{\mathbb{T}^n_{s/2}} |R_\beta^{-1}(\tilde\Phi - \tilde\Phi_0)|_{\mathrm{Lip},\frac{1}{\alpha}\Omega^{(i)}} = \alpha \sup_{\mathbb{T}^n_{s/2}} |R_r(\Phi - \Phi_0)|_{\mathrm{Lip},\Omega^{(i)}}$$
$$\le \frac{2\hat c}{\hat c_\star}\frac{\mu s^\nu}{\lambda} \le \frac{s^\nu}{8},$$

which, together with (53), proves (49).

Now, by (32), (44), (45), we get

$$\sup_{\mathbb{T}^n_{s/2}\times\frac{1}{\alpha}\Omega^{(i)}} \left|\left(\frac{1}{\beta r}(\mathtt{p}\circ\varphi - \mathtt{p})\circ S_\alpha, 0\right)\right| = \frac{1}{\beta r}\sup_{\Omega}^{(i)} |\mathtt{p}\circ\varphi - \mathtt{p}| \le \frac{2\lambda}{\hat c_\star \beta} \overset{(45)}{=} \frac{\mu s^\nu}{2^5 n!} < \frac{s^\nu}{8}. \quad (54)$$

The first estimate in (50) now follows at once in view of (47), the first inequality in (49) and (54).

In order to prove the second estimate in (50), in view of (47), we need to estimate the Lipschitz semi-norm of $(\mathtt{p}\circ\varphi - \mathtt{p})$. Fix $\omega, \omega' \in \Omega^{(i)}$ and set $p = \mathtt{p}(\omega)$ and $p' = \mathtt{p}(\omega')$. Let also $\gamma(t) := (p' - p)t + p$, for $t \in [0, 1]$ and[21] $\tilde\gamma := h_p \circ \gamma$. Then,[22]

$$\begin{aligned}
&\Big|\mathtt{p}\big(\varphi(\omega')\big) - \mathtt{p}(\omega') - \mathtt{p}\big(\varphi(\omega)\big) + \mathtt{p}(\omega)\Big| \\
&= \left|\left[\int_0^1 \Big(\mathtt{p}_\omega\big(\varphi(\tilde\gamma(t))\big)\varphi_\omega\big(\tilde\gamma(t)\big) - \mathtt{p}_\omega\big(\tilde\gamma(t)\big)\Big) h_{pp}(\gamma(t))\,\mathrm{d}t\right](p' - p)\right| \\
&\le \mathtt{M}|p' - p| \int_0^1 \Big|\mathtt{p}_\omega\big(\varphi(\tilde\gamma(t))\big)\varphi_\omega\big(\tilde\gamma(t)\big) - \mathtt{p}_\omega\big(\tilde\gamma(t)\big)\Big|\,\mathrm{d}t \\
&= \mathtt{M}|p' - p| \int_0^1 \Big|\mathtt{p}_\omega\big(\varphi(\tilde\gamma(t))\big)\big(\partial_\omega(\varphi - \mathtt{id})\big|_{\tilde\gamma(t)}\big) + \mathtt{p}_\omega\big(\varphi(\tilde\gamma(t))\big) - \mathtt{p}_\omega\big(\tilde\gamma(t)\big)\Big|\,\mathrm{d}t \\
&\overset{(43)}{\le} \lambda|\omega' - \omega| \left(\mathtt{L}|\varphi - \mathtt{id}|_{\mathrm{Lip}} + \frac{4\,\mathtt{L}}{c_0\mu^2\,\mathtt{M}r_0}|\varphi - \mathtt{id}|\right) \\
&\overset{(29),(35)}{\le} \lambda|\omega' - \omega| \left(\mathtt{L}\frac{2\varepsilon}{\hat c_\star \alpha r s^\nu} + \frac{4\,\mathtt{L}}{c_0\mu^2\,\mathtt{M}r_0}\frac{2\varepsilon}{\hat c_\star r}\right) \\
&\overset{(16),(32)}{=} |\omega' - \omega|\frac{2\lambda^2 r}{\hat c_\star}\left(\frac{1}{\alpha s^\nu} + \frac{4}{c_0\mu^2\,\mathtt{M}r_0}\right).
\end{aligned}$$

Now, observe that, by (6), (15), (16) and (17) (which implies that $c < \hat c^2 c_0^2/4$), one has

$$\frac{4}{c_0\mu^2\,\mathtt{M}r_0} < \frac{1}{\alpha s^\nu}.$$

Thus,

$$|\mathtt{p}\circ\varphi - \mathtt{p}|_{\mathrm{Lip},\Omega^{(i)}} < \frac{4\lambda^2 r}{\hat c_\star \alpha s^\nu},$$

and, therefore (recalling footnote 20),

$$\frac{1}{\beta r}|(\mathtt{p}\circ\varphi - \mathtt{p})\circ S_\alpha|_{\mathrm{Lip},\frac{1}{\alpha}\Omega^{(i)}} = \frac{\alpha}{\beta r}|\mathtt{p}\circ\varphi - \mathtt{p}|_{\mathrm{Lip},\Omega^{(i)}}$$

[21] The introduction of the lifted curve $\tilde\gamma \subseteq \Omega^{(i)}$ to join ω and ω' is due to the fact that, in general, $\Omega^{(i)}$ is not convex.

[22] By Remark (v) in **7.2**, φ is differentiable; the differentiability of φ almost everywhere also follows, independently, from Rademacher's theorem. Notice also that, if f is a function differentiable (a.e.) on $\Omega^{(i)}$, then $\sup_\Omega^{(i)} |\nabla f| \le |f|_{\mathrm{Lip},\Omega^{(i)}}$ (a.e.), the equality holding if $\Omega^{(i)}$ is convex.

$$
\begin{aligned}
&< \frac{4\lambda^2}{\beta \hat{c}_\star s^\nu}\\
&\overset{(14)}{\le} \frac{2\lambda}{n\hat{c}_0\mu\, \hat{c}_\star s^\nu}\, \frac{1}{\beta}\\
&\overset{(45),(15)}{=} \frac{1}{8},
\end{aligned}
$$

which, together with the second estimate in (49), in view of (47), yields also the second bound in (50).

To estimate Ψ_* (defined in (48)), observe that

$$
\tilde{\Psi}_0^{-1}(y,\theta) = \big(\theta, \alpha^{-1} h_p(\beta r y)\big), \qquad \tilde{\Psi}_0^{-1} : \frac{1}{\beta r} D_i \times \mathbb{T}^n_{s/2} \overset{\text{onto}}{\to} \mathbb{T}^n_{s/2} \times \frac{1}{\alpha}\Omega^{(i)}.
$$

Thus, the first estimate in (51) follows immediately from the first bound in (50). As for the Lipschitz semi-norm of Ψ_*, by (15), (17), (26), (32) and (45) we have, for all $\theta \in \mathbb{T}^n_{s/2}$, that[23]

$$
|\tilde{\Psi}_0^{-1}|_{\mathrm{Lip}, \frac{1}{\beta r} D_i} = \frac{1}{\alpha} |h_p(\beta r \cdot)|_{\mathrm{Lip}, \frac{1}{\beta r} D_i} = \frac{\beta r}{\alpha} |h_p|_{\mathrm{Lip}, D_i} \le \frac{\beta r\, \mathtt{M}}{\alpha} = \frac{\hat{c} s^{2\nu}}{\check{c}} = \frac{c_\star}{\hat{c}_\star} 4 \cdot n! s^{2\nu} < s^{2\nu}.
$$

Thus, in view of the second estimate in (50), we have, for all $\theta \in \mathbb{T}^n_{s/2}$,

$$
|\Psi_*|_{\mathrm{Lip}, \frac{1}{\beta r} D_i} \le |\tilde{\Psi} - \tilde{\Psi}_0|_{\mathrm{Lip}, \frac{1}{\alpha}\Omega^{(i)}}\, |\tilde{\Psi}_0^{-1}|_{\mathrm{Lip}, \frac{1}{\beta r} D_i} < \frac{s^{3\nu}}{4}.
$$

By (51) and Cauchy estimates we get (52). □

We shall also need the following

Lemma 5 *Let*[24]

$$
\rho := \frac{\beta s^\nu}{4} r. \tag{55}
$$

Then,

$$
\tilde{\Psi} \circ \tilde{\Psi}_0^{-1}\Big(\frac{1}{\beta r} D_i \times \mathbb{T}^n\Big) := \tilde{\Psi} \circ \tilde{\Psi}_0^{-1}\Big(\frac{1}{\beta r} \mathrm{B}_{\hat{r}}(p_i) \times \mathbb{T}^n\Big) \supseteq \frac{1}{\beta r} \mathrm{B}_{\hat{r}-\rho}(p_i) \times \mathbb{T}^n. \tag{56}
$$

Proof Since[25]

$$
\frac{1}{\beta r} \mathrm{B}_{\hat{r}-\rho}(p_i) = \mathrm{B}_{\frac{\hat{r}-\rho}{\beta r}}\Big(\frac{p_i}{\beta r}\Big) \overset{(55)}{=} \mathrm{B}_{\frac{\hat{r}}{\beta r} - \frac{s^\nu}{4}}\Big(\frac{p_i}{\beta r}\Big),
$$

one sees that (56) will hold if, for any given $(y_0, \theta_0) \in \frac{1}{\beta r} \mathrm{B}_{\hat{r}-\rho}(p_i) \times \mathbb{T}^n$, there exists a point[26]

$$
(y_1, \theta_1) \in \overline{\mathrm{B}_{s^\nu/4}(0)} \times \mathbb{T}^n
$$

23 $|h_p(\beta r \cdot)|_{\mathrm{Lip}, \frac{1}{\beta r} D_i}$ denotes the Lipschitz norm of the function $y \to h_p(\beta r y)$ on the rescaled domain $(\beta r)^{-1} D_i$.

24 Recall that $r = \sqrt{\epsilon}\, r_0$ is defined in (32).

25 Notice that it is $\hat{r} - \rho > 0$: Indeed, by (32) and (45), we see that $\rho = \lambda \sqrt{\epsilon} r_0/(4\check{c}\mu)$, so that (recalling the definition of $\hat{r}$ in (23)) $\rho < \hat{r}$ is seen to be equivalent to $\epsilon < 4\, \check{c}^2\, c_0^2\, \mu^6/\lambda^2 \overset{(46),(6)}{=} \hat{c}_\star^2 \mu^6/(2^{16} n^2 n!^6 \lambda^2)$, which is guaranteed by (16), observing that, by (15), $c < \hat{c}_\star^2/(2^{16} n^2 n!^6)$.

26 As standard, the overline denotes closure and observe that $y_0 + y_1 \in \frac{1}{\beta r} \mathrm{B}_{\hat{r}}(p_i)$.

such that

$$(y_0, \theta_0) = \tilde{\Psi} \circ \tilde{\Psi}_0^{-1}(y_0 + y_1, \theta_0 + \theta_1) \overset{(48)}{=} (y_0 + y_1, \theta_0 + \theta_1) + \Psi_*(y_0 + y_1, \theta_0 + \theta_1).$$

Such relation is, in turn, equivalent to the fixed point equation

$$(y_1, \theta_1) = -\Psi_*(y_0 + y_1, \theta_0 + \theta_1). \tag{57}$$

We shall solve (57) in two steps: (i), We prove that there exists a unique function $y_*(\theta)$ such that[27]

$$y_*(\theta) = -\pi_1 \Psi_*(y_0 + y_*(\theta), \theta_0 + \theta), \qquad \forall\, \theta \in \mathbb{T}^n, \tag{58}$$

and, (ii), we show that the map

$$\theta \in \mathbb{T}^n \mapsto \theta + \pi_2 \Psi_*(y_0 + y_*(\theta), \theta_0 + \theta) \in \mathbb{T}^n, \tag{59}$$

is onto, guaranteeing that there exists a $\theta_1 \in \mathbb{T}^n$ so that $\theta_1 + \pi_2\Psi_*(y_0 + y_*(\theta_1), \theta_0 + \theta_1) = 0$. These two facts will show that $y_1 := y_*(\theta_1)$ and θ_1 are solutions of (57), proving the claim.

Proof of (i): Let

$$X := \{\theta \mapsto y(\theta) \in C(\mathbb{T}^n, \mathbb{R}^n) |\ \sup_{\mathbb{T}^n} |y| \le s^\nu/4\}, \qquad F(y)(\theta) := -\pi_1 \Psi_*\big(y_0 + y(\theta), \theta_0 + \theta\big).$$

Then, by the first inequality in (51), $F : X \to X$, and the second inequality in (51) shows that F is a contraction from X into X. Hence, there exists a unique fixed point $y_* \in X$ satisfying (58). Furthermore, since Ψ_* is real-analytic, so is y_* and, in particular, its Jacobian $\partial_\theta y_*$ satisfies the equation

$$\Big(\mathtt{id} + \pi_1 \partial_y \Psi_*\big(y_0 + y(\theta), \theta_0 + \theta\big)\Big)\partial_\theta y_* = -\pi_1 \partial_\theta \Psi_*\big(y_0 + y(\theta), \theta_0 + \theta\big),$$

which, by Neumann series, by the second inequality[28] in (51) and by (52), yields

$$\sup_{\mathbb{T}^n} \|\partial_\theta y_*\| \le \frac{1}{1 - \frac{1}{4}} \frac{s^{\nu-1}}{2} = \frac{2}{3} s^{\nu-1}. \tag{60}$$

Proof of (ii): Observe that from the standard contraction lemma it follows easily that[29]:

If g is a $C^1(\mathbb{T}^n, \mathbb{T}^n)$ map such that $\lambda := \sup_{\mathbb{T}^n} \|\partial_\theta g\| < 1$, then, the map $G : \theta \in \mathbb{T}^n \mapsto \theta + g(\theta) \in \mathbb{T}^n$ has a unique inverse $\tilde{G} : \theta \in \mathbb{T}^n \mapsto \theta + \tilde{g}(\theta) \in \mathbb{T}^n$ with $\tilde{g} \in C^1(\mathbb{T}^n, \mathbb{T}^n)$ and $\sup_{\mathbb{T}^n} \|\partial_\theta \tilde{g}\| \le \lambda/(1 - \lambda)$.

Now, recalling (59), to finish the proof is enough to check that the Jacobian of the map

$$\theta \mapsto \pi_2 \Psi_*(y_0 + y_*(\theta), \theta_0 + \theta)$$

has (operator) norm strictly smaller than one. But, by the second inequality in (51), (52) and (60), one has, for any $\theta \in \mathbb{T}^n$,

---

[27] Recall, (2), that π_i denotes projection: $\pi_1(y, \theta) = y$ and $\pi_2(y, \theta) = \theta$.

[28] Recall footnote 7.

[29] Indeed, $G \circ \tilde{G} = \mathtt{id}$ is equivalent to the fixed point equation $\tilde{g} = -g \circ (\mathtt{id} + \tilde{g})$ and if we let X denote $C(\mathbb{T}^n, \mathbb{T}^n)$ endowed with the standard metric $d(h_1, h_2) := \sup_{\mathbb{T}^n} d_{\mathbb{T}^n}\big(h_1(\theta), h_2(\theta)\big)$ (where $d_{\mathbb{T}^n}$ denotes the standard flat metric on $\mathbb{T}^n$), one sees immediately that the assumption implies that the map $h \in X \mapsto -g \circ (\mathtt{id} + h) \in X$ is a contraction from X to X, whose unique fixed point yields $\tilde{g}$. Furthermore, since g is C^1, so is $\tilde{g}$ and the inequality on the Jacobian of $\tilde{g}$ follows by Neumann series after having differentiated the identity $\tilde{g} = -g \circ (\mathtt{id} + \tilde{g})$.

$$\begin{aligned}\big\|\partial_\theta \pi_2 \Psi_*\big(y_0+y_*(\theta),\theta_0+\theta\big)\big\| &\le \big\|\partial_y \Psi_*\big(y_0+y_*(\theta),\theta_0+\theta\big)\big\|\,\big\|\partial_\theta y_*(\theta)\big\|\\ &\quad+\big\|\partial_\theta \Psi_*\big(y_0+y_*(\theta),\theta_0+\theta\big)\big\|\\ &\le \frac14\cdot\frac23 s^{\nu-1}+\frac{s^{\nu-1}}{2}=\frac23 s^{\nu-1}<1.\end{aligned}$$

□

6.6 Measure estimates

We first provide measure estimates on $(D_i\times\mathbb{T}^n)\backslash\mathcal{T}_\alpha^{(i)}=(D_i\times\mathbb{T}^n)\backslash\Psi(\mathbb{T}^n\times\Omega_\alpha^{(i)})$.

Clearly,[30]

$$(D_i\times\mathbb{T}^n)\backslash\mathcal{T}_\alpha^{(i)}\subseteq\big((D_i\times\mathbb{T}^n)\backslash\Psi(\mathbb{T}^n\times\Omega^{(i)})\big)\,\dot\cup\,\Psi\big(\mathbb{T}^n\times(\Omega^{(i)}\backslash\Omega_\alpha^{(i)})\big). \tag{61}$$

Now, the following estimates hold.

Lemma 6 (Measure estimates) *Recall* (6)*,* (15) *and* (45) *and define the following constants:*

$$\kappa_1:=(2\pi)^n\,\frac{nc_0^{n-1}}{2\check c},\qquad \kappa_2:=\Big(\frac{5\pi}{2}\Big)^n; \tag{62}$$

$$\kappa_3':=2\,n^{\frac{n-1}{2}}\Big(\sum_{k\ne0}\frac{1}{|k|_1^{\tau+1}}\Big)\frac{c_0^{n-1}}{\hat c},\qquad \kappa_3'':=\frac{2nc_0^{n-1}}{\hat c},\qquad \kappa_3:=\kappa_3'+\kappa_3.'' \tag{63}$$

Then, one has

$$\operatorname{meas}\big((D_i\times\mathbb{T}^n)\backslash\Psi(\mathbb{T}^n\times\Omega^{(i)})\big)\le\kappa_1\,\mu^{2n-3}r_0^n\sqrt{\epsilon}\,, \tag{64}$$

$$\operatorname{meas}\big(\Psi(\mathbb{T}^n\times\Omega^{(i)}\backslash\Omega_\alpha^{(i)})\big)\le\kappa_2\,L^n\operatorname{meas}(\Omega^{(i)}\backslash\Omega_\alpha^{(i)})\,, \tag{65}$$

$$\operatorname{meas}(\Omega^{(i)}\backslash\Omega_\alpha^{(i)})\le\kappa_3\,\frac{\mu^{2n-3}\lambda^2}{s^{3\nu}}\,(\mathtt{M}r_0)^n\sqrt{\epsilon}. \tag{66}$$

Proof Observe that by (46)

$$R_{\beta r}\circ\Psi=\tilde\Psi\circ\tilde\Psi_0^{-1}\circ R_{\beta r}\circ\Psi_0. \tag{67}$$

Thus, since $\Psi_0(\mathbb{T}^n\times\Omega^{(i)})=D_i\times\mathbb{T}^n$, we have

$$R_{\beta r}\circ\Psi(\mathbb{T}^n\times\Omega^{(i)})=\tilde\Psi\circ\tilde\Psi_0^{-1}\Big(\frac{1}{\beta r}D_i\times\mathbb{T}^n\Big). \tag{68}$$

Therefore,[31]

$$\begin{aligned}\operatorname{meas}\Big((D_i\times\mathbb{T}^n)\backslash\Psi(\mathbb{T}^n\times\Omega^{(i)})\Big)&=(\beta r)^n\operatorname{meas}\Big(R_{\beta r}\big((D_i\times\mathbb{T}^n)\big)\backslash R_{\beta r}\circ\Psi(\mathbb{T}^n\times\Omega^{(i)})\Big)\\ &\overset{(68)}{=}(\beta r)^n\operatorname{meas}\Big(\big(\frac{1}{\beta r}D_i\times\mathbb{T}^n\big)\backslash\tilde\Psi\circ\tilde\Psi_0^{-1}\big(\frac{1}{\beta r}D_i\times\mathbb{T}^n\big)\Big)\\ &\overset{(56)}{\le}(\beta r)^n\operatorname{meas}\Big(\big(\frac{1}{\beta r}D_i\times\mathbb{T}^n\big)\backslash\big(\frac{1}{\beta r}\mathrm{B}_{\hat r-\rho}(p_i)\times\mathbb{T}^n\big)\Big)\\ &=\operatorname{meas}\Big(\big(D_i\backslash\mathrm{B}_{\hat r-\rho}(p_i)\big)\times\mathbb{T}^n\Big)\end{aligned}$$

---

30 The dot over union denotes "disjoint union."

31 Recall that $D_i=\mathrm{B}_{\hat r}(p_i)$; in the last inequality use that for every $0<x<1$ and for every integer $n>1$, one has $1-(1-x)^n<nx$ and for the last equality recall (23), (32), (45), (55).

$$
\begin{aligned}
&= (2\pi)^n\Big((2\hat{r})^n - \big(2(\hat{r}-\rho)\big)^n\Big)\\
&\le (2\pi)^n n 2^n \hat{r}^{n-1}\rho\\
&\stackrel{(62)}{=} \kappa_1\lambda\mu^{2n-3} r_0^n\sqrt{\epsilon},
\end{aligned}
$$

proving (64).

To prove (65), observe that if $A \subseteq \Omega^{(i)}$, from (67) and the identity (recall (38))

$$R_{\beta r}\circ\Psi_0(\mathbb{T}^n\times A) = \frac{1}{\beta r}\,\mathtt{p}(A)\times\mathbb{T}^n,$$

there follows, by (46),

$$\Psi(\mathbb{T}^n\times A) = R_{\beta r}^{-1}\circ\tilde{\Psi}\circ\tilde{\Psi}_0^{-1}\Big(\frac{1}{\beta r}\,\mathtt{p}(A)\times\mathbb{T}^n\Big). \tag{69}$$

Observe also that, since[32]

$$\tilde{\Psi}\circ\tilde{\Psi}_0^{-1} = \mathtt{id} + \Psi_*,$$

from (51) there follows

$$|\tilde{\Psi}\circ\tilde{\Psi}_0^{-1}|_{\mathrm{Lip},\,\frac{1}{\beta r}D_i\times\mathbb{T}^n} \le 5/4.$$

Now, for every measurable set $A \subseteq \Omega^{(i)}$, one has[33]

$$
\begin{aligned}
\mathrm{meas}\big(\Psi(\mathbb{T}^n\times A)\big) &\stackrel{(69)}{=} (\beta r)^n \mathrm{meas}\Big(\tilde{\Psi}\circ\tilde{\Psi}_0^{-1}\big(\frac{1}{\beta r}\,\mathtt{p}(A)\times\mathbb{T}^n\big)\Big)\\
&\le (2\pi)^n\big(|\tilde{\Psi}\circ\tilde{\Psi}_0^{-1}|_{\mathrm{Lip},\,\frac{1}{\beta r}D_i\times\mathbb{T}^n}\big)^n\ \mathtt{L}^n\ \mathrm{meas}(A)\\
&\le \kappa_2\,\mathtt{L}^n\ \mathrm{meas}\,(A),
\end{aligned}
$$

and (65) follows.

To prove (66), observe that

$$\Omega^{(i)}\backslash\Omega^{(i)}_\alpha \subseteq \{\omega\in\Omega^{(i)}|\ \omega \text{ is not } (\alpha,\tau)\text{-Diophantine}\}\cup\Omega^{(i)}(\alpha) \tag{70}$$

where

$$\Omega^{(i)}(\alpha) := \Big\{\omega\in\Omega^{(i)},\ \ \text{s.t. } \mathrm{dist}(\omega,\partial\Omega^{(i)}) < \alpha\Big\}.$$

Let us begin with estimating the measure of the first set in the r.h.s. of (70) keeping track of constants. Notice that, if $\hat{\Omega}^{(i)}$ denotes the Euclidean ball of center $h_p(p_i)$ and radius $\sqrt{n}\,\mathtt{M}\hat{r}$, then

$$\Omega^{(i)} = h_p\big(\mathrm{B}_{\hat{r}}(p_i)\big) \subseteq \hat{\Omega}^{(i)}.$$

Thus, denoting by $\|\cdot\|$ the Euclidean norm in $\mathbb{R}^n$, we have

$$
\begin{aligned}
&\mathrm{meas}\,\{\omega\in\Omega^{(i)}|\ \omega \text{ not } (\alpha,\tau)-\text{Dioph.}\}\\
&\le\ \mathrm{meas}\left\{\omega\in\hat{\Omega}^{(i)}|\ \exists\, k\in\mathbb{Z}^n,\ k\neq 0:\ |\omega\cdot k| < \frac{\alpha}{|k|_1^\tau}\right\}
\end{aligned}
$$

[32] Recall the definition of Ψ_* in (48).

[33] In the first inequality, we use (twice) the following fact: *If $A \subseteq$ is a measurable set and $f : A \to \mathbb{R}^n$ is a Lipschitz map, then* $\mathrm{meas}\, f(A) \le |f|^n_{\mathrm{Lip},A}\,\mathrm{meas}\, A$.

$$
\begin{aligned}
&\le 2^n \sum_{k\neq 0} (\sqrt{n}\,\mathtt{M}\hat{r})^{n-1} \frac{\alpha}{|k|_1^{\tau+1}} \\
&\overset{(23),(17)}{=} \kappa_3' \, \frac{\mathtt{M}^n \lambda \mu^{(2n-3)}}{s^{3\nu}} \, r_0^n \sqrt{\epsilon}.
\end{aligned}
\tag{71}
$$

As for the measure of the second set in (70), we observe that either $\hat{r} \le \mathtt{L}\alpha$ or $\hat{r} > \mathtt{L}\alpha$. In the first case we have

$$
\mathrm{meas}(\Omega^{(i)}(\alpha)) \le \mathrm{meas}(\Omega^{(i)}) \le \mathtt{M}^n \mathrm{meas}(\mathrm{B}_{\hat{r}}(p_i)) = 2^n\,\mathtt{M}^n \hat{r}^n \le 2^n\,\mathtt{M}^n\,\mathtt{L}\hat{r}^{n-1}\alpha\,.
$$

In the second case, let

$$
\check{r} := \hat{r} - \mathtt{L}\alpha > 0\,.
$$

We claim that

$$
h_p\big(\mathrm{B}_{\hat{r}}(p_i)\backslash \mathrm{B}_{\check{r}}(p_i)\big) \supseteq \Omega^{(i)}(\alpha)\,. \tag{72}
$$

Indeed, by contradiction, assume that there exist $\omega = h_p(p) \in \Omega^{(i)}(\alpha)$, $\omega_* = h_p(p_*) \in \partial\Omega^{(i)}$ (namely $|p_* - p_i| = \hat{r}$) with $|p - p_i| < \check{r}$, and $|\omega - \omega_*| < \alpha$. Then

$$
\mathtt{L}\alpha = \hat{r} - \check{r} < |p_* - p| \le \mathtt{L}|\omega_* - \omega| < \mathtt{L}\alpha,
$$

proving (72). Thus,[34]

$$
\begin{aligned}
\mathrm{meas}(\Omega^{(i)}(\alpha)) &\le \mathrm{meas}\Big(h_p\big(\mathrm{B}_{\hat{r}}(p_i)\backslash \mathrm{B}_{\check{r}}(p_i)\big)\Big) \le \mathtt{M}^n \mathrm{meas}\big(\mathrm{B}_{\hat{r}}(p_i)\backslash \mathrm{B}_{\check{r}}(p_i)\big) = 2^n\,\mathtt{M}^n(\hat{r}^n - \check{r}^n) \\
&< (2\,\mathtt{M})^n\; n\hat{r}^{n-1}\,\mathtt{L}\alpha
\end{aligned}
$$

Thus, in either case, by (17) and (23), we have

$$
\mathrm{meas}\,\Omega^{(i)}(\alpha) \le (2\,\mathtt{M})^n\; n\hat{r}^{n-1}\,\mathtt{L}\alpha = \kappa_3''\,\frac{\mathtt{M}^n \lambda^2 \mu^{(2n-3)}}{s^{3\nu}}\, r_0^n \sqrt{\epsilon} \tag{73}
$$

By (70), (71) and (73), we have

$$
\mathrm{meas}\,(\Omega^{(i)}\backslash\Omega^{(i)}_\alpha) \le \kappa_3 \frac{\mathtt{M}^n \lambda^2 \mu^{(2n-3)}}{s^{3\nu}}\, r_0^n \sqrt{\epsilon}.
$$

Lemma 6 is proved. □

From (61) and Lemma 6 there follows

$$
\mathrm{meas}\left((D_i \times \mathbb{T}^n)\backslash \mathcal{T}^{(i)}_\alpha\right) \le (\kappa_1 + \kappa_2\kappa_3)\,\mu^{2n}\, r_0^n \,\frac{\lambda^{n+2}}{\mu^3 s^{3\nu}}\,\sqrt{\epsilon}.
$$

Now, since it is

$$
\begin{aligned}
(\mathtt{D} \times \mathbb{T}^n)\backslash \mathcal{T}_\alpha &= (\mathtt{D} \times \mathbb{T}^n)\backslash \bigcup_{i=1}^N \mathcal{T}^{(i)}_\alpha \\
&\subseteq \bigcup_{i=1}^N (D_i \times \mathbb{T}^n)\backslash \bigcup_{i=1}^N \mathcal{T}^{(i)}_\alpha
\end{aligned}
$$

[34] Recall footnote 31.

$$\subseteq \bigcup_{i=1}^{N} (D_i \times \mathbb{T}^n) \backslash \mathcal{T}_\alpha^{(i)}$$
$$= \bigcup_{i=1}^{N} (\mathrm{B}_{\hat{r}}(p_i) \times \mathbb{T}^n) \backslash \mathcal{T}_\alpha^{(i)}.$$

Now, in view of (24), one obtains (20) with

$$\kappa := \frac{2^{2n}}{c_0^n} (\kappa_1 + \kappa_2 \kappa_3), \tag{74}$$

and (18) follows at once. □

A The standard quantitative inverse function theorem

The following is a standard inverse function theorem in $\mathbb{C}^n$; the bar over sets denotes closure.

Proposition *Let $f : \overline{\mathbb{B}_r(p_0)} \to \mathbb{C}^n$ be a holomorphic function with invertible Jacobian $f_p(p_0)$ and with r such that*

$$\sup_{\overline{\mathbb{B}_r(p_0)}} \|I - f_p^{-1}(p_0) f_p(p)\| \leq \delta < 1. \tag{75}$$

Then, there exists a unique holomorphic inverse g of f such that

$$g : \overline{\mathbb{B}_\rho(\omega_0)} \to \overline{\mathbb{B}_r(p_0)}, \qquad \text{with} \qquad \rho := (1-\delta) \frac{r}{\|f_p^{-1}(p_0)\|}, \qquad \omega_0 := f(p_0). \tag{76}$$

Furthermore,

$$\sup_{\overline{\mathbb{B}_\rho(\omega_0)}} \|g_\omega\| \leq \frac{1}{1-\delta} \|f_p^{-1}(p_0)\|. \tag{77}$$

If f is real-analytic, so is g.

The elementary proof follows by checking that the map $h \mapsto \Phi(h) := h + f_p^{-1}(p_0)\big(f \circ h - \mathtt{id}\big)$ is a contraction on the space of continuous functions from $\overline{\mathbb{B}_\rho(\omega_0)}$ in $\overline{\mathbb{B}_r(p_0)}$. Then, by the contraction lemma, $g = \lim \Phi^n(p_0)$, which also shows, by Weierstrass theorem on the uniform limit of holomorphic functions, that g is holomorphic (and real-analytic, if so is f). Bound (77) is a general fact following from Neumann series: Indeed, if A and B are $(n \times n)$ matrices and $\|I - AB\| \leq \delta < 1$, then, by Neumann series, AB is invertible and so are A and B, furthermore $\|B^{-1}\| \leq \|(AB)^{-1}\| \, \|A\| \leq (1-\delta)^{-1} \|A\|$. □

References

1. Arnold, V.I.: Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian (Russian). Uspehi Mat. Nauk. **18**(5(113)), 13–40 (1963)
2. Arnold, V.I., Kozlov, V.V., Neishtadt, A.I.: Mathematical Aspects of Classical and Celestial Mechanics, Volume 3 of Encyclopaedia of Mathematical Sciences, 3rd edn. Springer, Berlin (2006). [Dynamical systems. III], Translated from the Russian original by E. Khukhro
3. Berti, M., Biasco, L.: Branching of cantor manifolds of elliptic tori and applications to PDEs. Commun. Math. Phys. **305**, 741–796 (2011)

4. Biasco, L., Chierchia, L.: On the measure of Lagrangian invariant tori in nearly-integrable mechanical systems. Rend. Lincei Mat. Appl. **26**, 1–10 (2015)
5. Celletti, A., Chierchia, L.: Construction of analytic KAM surfaces and effective stability bounds. Commun. Math. Phys. **118**(1), 119–161 (1988)
6. Celletti, A., Chierchia, L.: A constructive theory of Lagrangian tori and computer-assisted applications. In: Dynamics Reported, pp. 60–129, Dynam. Report. Expositions Dynam. Systems (N.S.), 4. Springer, Berlin (1995)
7. Celletti, A., Chierchia, L.: On the stability of realistic three-body problems. Commun. Math. Phys. **186**(2), 413–449 (1997)
8. Celletti, A., Chierchia, L.: KAM stability and celestial mechanics. Mem. Am. Math. Soc. **187**(878), 1–134 (2007)
9. Celletti, A., Giorgilli, A.: On the numerical optimization of KAM estimates by classical perturbation theory. Z. Angew. Math. Phys. **39**(5), 743–747 (1988)
10. Celletti, A., Giorgilli, A.: Improved estimates on the existence of invariant tori for Hamiltonian systems (English summary). Nonlinearity **13**(2), 397–412 (2000)
11. Celletti, A., Giorgilli, A., Locatelli, U.: Improved estimates on the existence of invariant tori for Hamiltonian systems. Nonlinearity **13**(2), 397–412 (2000)
12. Chierchia, L., Gallavotti, G.: Smooth prime integrals for quasi-integrable Hamiltonian systems. Nuovo Cimento B (11) **67**(2), 277–295 (1982)
13. de la Llave, R., Rana, D.: Accurate strategies for small divisor problems. Bull. Am. Math. Soc. (N.S.) **22**(1), 85–90 (1990)
14. Delshams, A., Gutiérrez, P.: Effective stability and KAM theory. J. Differ. Equ. **128**(2), 415–490 (1996)
15. Dumas, H.S.: The KAM Story. World Scientific, Singapore (2014)
16. Giorgilli, A., Galgani, L.: Rigorous estimates for the series expansions of Hamiltonian perturbation theory. Celest. Mech. **37**(2), 95–112 (1985)
17. Kolmogorov, A.N.: On conservation of conditionally periodic motions for a small change in Hamilton's function (Russian). Dokl. Akad. Nauk SSSR. (N.S.) **98**, 527–530 (1954)
18. Koudjinan, C.E.: Quantitative KAM normal forms. Preprint (2017)
19. Morbidelli, A., Giorgilli, A.: Quantitative perturbation theory by successive elimination of harmonics (English summary). Celest. Mech. Dyn. Astron. **55**(2), 131–159 (1993)
20. Moser, J.K.: On invariant curves of area-preserving mappings of an annulus. Nachr. Akad. Wiss. Göttingen Math. Phys. Kl. **II**, 1–20 (1962)
21. Neishtadt, A.I.: Estimates in the Kolmogorov theorem on conservation of conditionally periodic motions. J. Appl. Math. Mech. **45**(6), 766–772 (1981)
22. Pöschel, J.: Integrability of Hamiltonian systems on Cantor sets. Commun. Pure Appl. Math. **35**(1), 653–695 (1982)
23. Pöschel, J.: A lecture on the classical KAM theorem. In: Smooth Ergodic Theory and Its Applications (Seattle, WA, 1999), Proc. Sympos. Pure Math., vol. 69, pp. 707–732. Amer. Math. Soc., Providence (2001)