Comlan Edmond Koudjinan

Quantitative KAM normal forms and sharp measure estimates

Dottorato in Matematica, Roma Tre (2019).

<http://www.bdim.eu/item?id=tesi_2019_KoudjinanComlanEdmond_1>

Università degli Studi Roma Tre

Dipartimento di Matematica e Fisica

XXXI Cycle

Doctoral Thesis in Mathematics

Quantitative KAM normal forms and sharp measure estimates

Candidate

Comlan Edmond Koudjinan

Supervisor

Prof. Luigi Chierchia

Coordinator

Prof. Angelo Felice Lopez

May 1, 2019

Abstract

It is widespread since the beginning of KAM Theory that, under "sufficiently small" perturbation, of size ϵ, apart a set of measure $O(\sqrt{\epsilon})$, all the KAM Tori of a non–degenerate integrable Hamiltonian system persist up to a small deformation. However, no explicit, self–contained proof of this fact exists so far. In the present Thesis, we give a detailed proof of how to get rid of a logarithmic correction (due to a Fourier cut–off) in Arnold's scheme and then use it to prove an explicit and "sharp" Theorem of integrability on Cantor–type set. In particular, we give an explicit proof of the above–mentioned measure estimate on the measure of persistent primary KAM tori. We also prove three quantitative KAM normal forms following closely the original ideas of the pioneers Kolmogorov, Arnold and Moser, computing explicitly all the KAM constants involved and fix some "physical dimension" issues by means of appropriate rescalings. Finally, we compare those three quantitative KAM normal forms on a simple mechanical system.

Contents

Acknowledgements

I would like to express my deep gratitude to Prof. Luigi Chierchia for introducing to the absolutely fascinating field of *KAM theory*, for his guidance, availability, patience, support which goes much much beyond mathematics and this thesis, and his precious help all along the trip leading to the accomplishment of this thesis.

I would like aslo to express my sincere gratitude to Prof. Francesco Pappalardi for his inestimable support. I am in Roma thanks to him mainly and without him I would not have the chance to meet Prof. Luigi and this thesis would not be possible simply.

I am very grateful to Prof. Luca Biasco and Prof. Michela Procesi for their availability, kindness, the very helpful discussions i had with them.

I am also deeply grateful to Prof. Antonio Siconolfi, with whom I spent almost the half of the three years of my PhD, for his help, generosity, patience and precious advises. I could have completed my PhD with him, but at some some point, due to time constrained, I had to choose and focus on a single subject.

I would like to address a special thanks to Prof. Wilfrid Gangbo whose impact on my career is considerable and this, since 2015 that I met him. It was through him I met Prof. Siconolfi.

I would like to address my sincere gratitude to Prof. Théophile Olory for his precious advises and assistance.

I am very grateful to Prof. Bernadin Kpamingan for his help all those years along.

I would like to express my gratitude to Prof. Rafael De La Llave for his interest in my work, for offering the opportunity to attend a workshop and for drawing my attention on some related papers in the literature.

I would like to express my gratitude to Prof. Jean–Pierre Marco and Prof. Tere M–Seara for their interest in my thesis and their availability to help.

I am also very grateful to Prof. Stefano Luzzatto and Prof. Carlangelo Liverani for their support and precious time.

A big big thanks to my mom, my sisters and brothers, especially Zizi, Igor, Marcelin, Adèle, Jeannette, Justine, Eulalie, ..., for their permanent support.

Finally, I want to thank my queen Cica and my princess Merquela, for their patience and constant encouragement, I am infinitely grateful to you.

This list is of course far from being exhaustive and detailed; I tried to be as short as possible. Otherwise, I would somehow fall into writing my biography and here is certainly not the place for!!!!

To the memory of my father

Notational conventions

- e denotes the Neper's number *i.e.* $\exp(1)$
- $\mathbb{N} = \{1, 2, 3, \cdots\}$ and $\mathbb{N}_0 = \{1, 2, 3, \cdots\}$
- $\mathbb{R}$ and $\mathbb{C}$ are respectively the set of real and complex numbers
- $\nu! = \nu_1! \cdots \nu_d!$ and $|\nu|_1 = \nu_1 + \cdots + \nu_d$, for any $\nu = (\nu_1, \cdots, \nu_d) \in \mathbb{N}_0^d$
- $y^\beta = y_1^{\beta_1} \cdots y_d^{\beta_d}$ for any $y, \beta \in \mathbb{R}^d$
- $|(y_1, \cdots, y_d)| = \max\{|y_1|, \cdots, |y_d|\}$ and $|(y_1, \cdots, y_d)|_2 = \sqrt{y_1^2 + \cdots + y_d^2}$
- $(x_1, \cdots, x_d) \cdot (y_1, \cdots, y_d) = \langle (x_1, \cdots, x_d), (y_1, \cdots, y_d) \rangle = x_1 y_1 + \cdots + x_d y_d$
- dist denotes the distance function
- $\overline{A}$ denotes the closure of A
- ∂A denotes the "boundary" of A
- $\operatorname{conv}(A)$ denotes the convex–hall of A
- $C^n(A, B)$ (*resp.* $C_c^n(A, B)$) denote respectively the set of functions of class C^n (*resp.* C^n with compact supports) from A into B
- meas_d denotes the d–dimensional Lebesgue–measure
- $\operatorname{dom}(f)$ denotes the domain of f
- $\operatorname{supp}(f)$ denotes the support of f
- $B_r(p)$ (*resp.* $D_r(p)$) denotes the ball centered at p with radius r in $\mathbb{R}^d$ (*resp.* in $\mathbb{C}^d$)
- $B_r(A)$ (*resp.* $D_r(A)$) denotes the r–neighborhood of A in $\mathbb{R}^d$ (*resp.* in $\mathbb{C}^d$)
- Δ_α^τ denotes the set of (α, τ)–Diophantine vectors
- $\mathbb{T}_s^d$ denotes the strip of width s around $\mathbb{T}^d$ in $\mathbb{C}^d$
- $\langle f \rangle$ denotes the average of f on $\mathbb{T}^d$
- $f_\nu = \partial_y^\nu f = \frac{\partial^\nu f}{\partial y^\nu} = \frac{\partial^{|\nu|_1} f}{\partial y_1^{\nu_1} \cdots \partial y_d^{\nu_d}}$ denotes derivative of order ν of f

- Iso(V) denotes the set of isomorphisms from V onto itself
- $\mathcal{M}_{n,m}(\mathbb{C}^d)$ the set of n–by–m matrices with entries in $\mathbb{C}^d$ and $\mathcal{S}_n(\mathbb{C}^d) \subset \mathcal{M}_n(\mathbb{C}^d) := \mathcal{M}_{n,n}(\mathbb{C}^d)$ the set of symmetric square matrices of order n
- $\mathrm{Adj}\,(A)$ denotes the adjoint of A
- $\det(A)$ denotes the determinant of A
- A^T denotes the transposed of A
- NM and TM are respectively the normal and tangent bundle of the manifold M
- $\Gamma(M)$ denotes the space of smooth vector field on M
- $\mathfrak{F}(M)$ denotes the space of smooth functions on M
- $\mathrm{minfoc}\,(M)$ denotes the minimal focal distance of the manifold M

1 | Introduction

In the solar system framework, Celestical Mechanics, a branch of astronomy, consists ultimately in the study of the n–body problem. The n–body problem is the dynamical system that governs the motion of n planets interacting according to Newton's gravitation law. A holy–grail question in Celestical Mechanics was and remains the stability of the solar system, *i.e.* whether the current configuration of the planets will stay unchanged forever under their interaction, or whether some planets will be kicked out of the system or have their trajectories be drastically affected to eventually collapse and give rise to unpredictable behaviors. Across the history of Mathematics, most of the great figures devoted some part of their works to this question. Laplace (1773), Lagrange (1776), Poisson (1809), and Dirichlet (1858) used series expansion techniques to study the question of stability of the solar system and claimed all to have proved it. Then Bruns (1887) proved that, from quantitative point of view, the only method which could solve the n–body problem is the series expansions. But, the works of Haretu (1878) and Poincaré (1892) (see [Poi99]) show that all those series expansion techniques fail as the series expansions they use diverge (see [Dug57, Mou02, Mos73, Dum14, AMM78] for more historical details).

A new viewpoint is thus undeniably needed to overcome this embarrassing fact. The change of paradigm was made by Poincaré. Indeed, Poincaré introduced a completely revolutionary qualitative approach to Mechanics (see [Poi90, Poi99]). The point is that, for question such as stability, one needs to study the entire phase portrait, and in particular the asymptotic time behavior of the solutions.
The phase portrait is the family of solutions curves, which fill up the entire phase space. The phase space is a symplectic manifold (a differentiable manifold together with a symplectic structure). Dynamical system is then just given by a Hamiltonian vector field; this is the Mathematical model for the global study in Mechanics that Poincaré gave in his qualitative theory. With his new geometrical methods, Poincaré discovered the non-integrability of the three body problem. In fact, the small divisor problem was well–known to Poincaré, who was aware that, because of this problem, nearly–integrable Hamiltonian systems are, in general, not integrable (analytically). Poincaré and his successors then

speculate that most of the classical systems were chaotic, and ergodic. There were even a gaped proof of the ergocity of a generic Hamiltonian system by E. Fermi in the 1920' (see [Fer23a, Fer23b, Fer23c, Fer24]). This ergodic hypothesis was accepted by many, including some of the brightest mind of those times, till the discoveries by Kolmogorov and his followers.

At the 1954 International Congress of Mathematician in Amsterdam, against any expectation, Kolmogorov (see [Kol54a, Kol54b]) presented a four–pages note where he sketched the proof of the persistence of the majority of tori for a nearly–integrable Hamiltonian system. Then, his former student Arnol'd (see [Arn63a, Arn63c]) completed the proof in the analytic category, and Moser (see [Mos62, Mos66b, Mos66a]) in the smooth category, whence the acronym KAM Theory.

The object of KAM Theory is the construction of quasiperiodic trajectories, which are sets of perpetual stability, in Hamiltonian dynamics. A KAM scheme is essentially based on the Newcomb idea of successive constructions of change of variable through a Newton–like method. Those successive changes of variables are carried out to eliminate, in a super–exponentianlly increasing manner, the fast phase variables. A KAM scheme of course encounters the small divisor problem that Poincaré faced. Netherless, the super–exponentianlly decay make the whole scheme converge.

Formally, one is given a symplectic manifold (M, ϖ), a (smooth) Hamiltonian $H\colon T^*M \to \mathbb{R}$. To the Hamiltonian H, is associated a (unique!) smooth vector field, the Hamiltonian vector field, say X_H, given by the equation

$$\varpi(X_H, \cdot) = -dH \ .$$

The smooth vector field X_H then generates a flow, say ϕ_H^t, by the relation

$$\frac{d}{dt}\phi_H^t = X_H \circ \phi_H^t \ . \tag{1.0.1}$$

In particular, if $M = \mathbb{R}^d \times \mathbb{T}^d$ and $\varpi = dy_1 \wedge dx_1 + \cdots + dy_d \wedge dx_d$, then $X_H = (-\partial_x H, \partial_y H)$ and, therefore, the equation (1.0.1) reads

$$\begin{cases} \dot{y} = -\partial_x H \\ \dot{x} = \partial_y H \end{cases}, \qquad \phi_H^t = (y(t),\, x(t))\,. \tag{1.0.2}$$

Then, to construct quasi–periodic trajectories for the Hamiltonian system (1.0.1), one looks for a change of variable $\phi'\colon (y', x') \mapsto (y, x) = \phi'(y', x')$, with the following properties

(a) ϕ preserves the Hamiltonian structure of (1.0.1). More precisely, ϕ preserves the symplectic form *i.e.* $\phi^*\varpi = \varpi$;

(b) ϕ conjugates ϕ_H^t to a linear flow:

$$\phi^{-1} \circ \phi_H^t \circ \phi(y,x) = (y, \omega t + x) \ , \qquad \text{with} \qquad \omega := \partial_y H(y,x) \ . \tag{1.0.3}$$

However, one does not solve directly (1.0.3). Instead, one conjugates the Hamiltonian itself *i.e.* construct ϕ in such away that

$$H \circ \phi(y', x') = H_*(y') \ , \tag{1.0.4}$$

as the latter is much easier to carry out than the former. Once (1.0.4) holds, by the property (a), (1.0.3) follows (see for instance [MZ05] for details).
The numerical property of the frequency or winding number ω plays a crucial role in the construction of the invariant tori. The most common assumption is the Diophantine property. A vector $\omega \in \mathbb{R}^d$ is said (α, τ)–Diophantine if

$$|\omega \cdot k| \geqslant \frac{\alpha}{|k|_1^\tau} \ , \qquad \forall \ k \in \mathbb{Z}^d \backslash \{0\}, \tag{1.0.5}$$

where $|k|_1 = |k_1| + \cdots + |k_d|$.
Facts Let

$$\Delta_\alpha^\tau := \left\{ \omega \in \mathbb{R}^d \ : |\omega \cdot k| \geqslant \frac{\alpha}{|k|_1^\tau} \ , \ \forall k \in \mathbb{Z}^d \backslash \{0\} \right\} \ ,$$

and

$$\Delta^\tau := \bigcup_{\alpha > 0} \Delta_\alpha^\tau \ ,$$

be the set of all (α, τ)–Diophantine vectors. Thus

(i) If $\tau < d-1$, then $\Delta^\tau = \varnothing$ (see [Cas57]);

(ii) If $\tau = d-1$, then the set Δ^τ has zero Lebesgue–measure, but is of Hausdorff dimension d. In particular, the intersection of Δ^τ with any open set has the cardinality of $\mathbb{R}$ (see [Sch66, Sch69]);

(iii) If $\tau > d-1$, then the Lebesgue–measure of $\mathbb{R}^d \backslash \Delta^\tau$ is zero. In fact,

$$\text{meas } (B_R(0) \backslash \Delta_\alpha^\tau) \leqslant \left(\sum_{k \in \mathbb{Z}^d \backslash \{0\}} \frac{2^d \sqrt{d}}{|k|_1^{\tau+1}} \right) R^{d-1} \ \alpha \ , \qquad \forall \ R, \alpha > 0. \tag{1.0.6}$$

As soon as the existence of invariant tori is established, one can speak of Kolmogorov sets, which turn out to be very big. A Kolmogorov set associated to a Hamiltonian H

is an union of its invariant maximal KAM tori. A maximal KAM torus for H is an embedded, Lagrangian, Kronecker torus with Diophantine frequency ω. A Kronecker torus with frequency ω is an embedded torus on which the H–flow is conjugated to the linear flow

$$\mathbb{R} \times \mathbb{T}^d \ni (t, x) \mapsto x + \omega t \ .$$

In this thesis, we are mainly concerned with "sharp" measure estimates of Kolmogorov sets, with emphasize on the dependence of those measure estimates upon the geometry of the domain.

Moser introduced in [Mos67] the original idea of parametrizing a non–degenerate quasi–integrable Hamiltonian by the frequency vectors and then apply the KAM technics (see also [Pös01, Pös82]). In [Pös01], the author made a short discussion of the measure of the complement of Kolmogorov set. Then, very recently, Biasco and Chierchia [BC18], give a detailed proof of the measure estimate result in [Pös01] and show how this measure estimate depends upon the domain. We revisit the paper [Pös01] in this thesis and our computation in particular fixes a small gap in the statement of [Pös01] (see §4. of Remark 2.1.5 below).

Arnold's scheme [Arn63a, Arn63b] can be summarized as follows. Let K and P be real–analytic in $D_0 := D_{r_0}(y_0) \times \mathbb{T}^d_{s_0}$, with K integrable and such that

$$K_y(y_0) =: \omega \in \Delta^\tau_\alpha \qquad \text{and} \qquad \det K_{yy}(y_0) \neq 0,$$

with $\alpha > 0$ and $\tau > d - 1$. Thus, the torus $\mathcal{T}_{\omega,0} := \{y_0\} \times \mathbb{T}^d$ is a KAM torus for K on which its flow ϕ^t_K is linear:

$$\phi^t_K(y_0, x) = (y_0, \omega t + x).$$

Then, the idea of Arnold is to construct a near–to–identity symplectic change of variables

$$\phi_1 \colon D_1 := D_{r_1}(y_1) \times \mathbb{T}^d_{s_1} \to D_0 \ ,$$

with $D_1 \subset D_0$ such that

$$\begin{cases} H_1 := H \circ \phi_1 = K_1 + \varepsilon^2 P_1 \ , \\ \partial_y K_1(y_1) = \omega, \quad \det \partial_y^2 K_1(y_1) \neq 0 \, . \end{cases} \tag{1.0.7}$$

And for ε small enough, one can iterate the process and build a sequence of symplectic transformations ($j \geqslant 1$)

$$\phi_j \colon D_j := D_{r_j}(y_j) \times \mathbb{T}^d_{s_j} \to D_{j-1} \ , \qquad D_j \subset D_{j-1} \ ,$$

and satisfying

$$\begin{cases} H_j := H \circ \phi_j = K_1 + \varepsilon^{2^j} P_j \ , \\ \partial_y K_j(y_j) = \omega, \quad \det \partial_y^2 K_j(y_j) \neq 0 \, . \end{cases} \tag{1.0.8}$$

In performing this construction, one is first attempt to solve the linear PDE

$$\partial_y K_j(y) \cdot \partial_x g_j + P_j(y, x) = \text{function of } y \text{ exclusively}, \tag{1.0.9}$$

where g_j is a generating function for ϕ_j. But (1.0.9) does not admit solution, because of small divisor problem (see [Chi12] for more discussion). The key idea of Arnold is then to solve only a truncated version of (1.0.9), with the order of truncation large enough so that the error one commits by solving approximately (1.0.9) is of order the square of the size of the perturbation. The truncation is the origin of the logarithmic correction in the smallness condition required in order to iterate infinitely many times the Arnold process. In particular, the Lebesgue–measure estimate of the complementary of the Kolmogorov set one gets from Arnold's scheme is $O(\sqrt{\varepsilon}\ (\log \varepsilon^{-1})^{3(\tau+1)/4})$. This estimate is not the optimal one, which is $O(\sqrt{\varepsilon})$ (see for instance [BC18], where the constant in front of $\sqrt{\varepsilon}$ in the optimal measure estimate is computed explicitly and the proof uses the KAM Theorem à la Moser). The task of getting rid of the logarithmic correction in the Arnold's scheme is not obvious. The first paper in this direction is the sketchy 7–pages paper [Nei81], where Neishtadt outlines how to overcome the logarithmic correction. The approach we adopt here is essentially equivalent to the one in [Nei81] though conceptually different. Indeed, in our scheme we fix the frequencies of the tori we build up from the beginning once for all. Instead, in [Nei81] as well as in the original paper by Arnold [Arn63b], the tori as well as their respective frequencies are constructed iteratively.

Moreover, in our approach, we focus on the smallest possible α i.e. the situations where the square–root of the sizes of the perturbations are proportional to the Diophantine constant α of the frequency of the tori. We then discuss the measure estimate of the Kolmogorov set we build up. The sharpness of the measure of the Kolmogorov set is in fact intimately related to the power of the Diophantine constant α in the smallness condition under which one performs the KAM scheme. Recently, Villanueva [Vil08] revisited the classical Kolmogorov scheme and succeed to cut down the power of α in the smallness condition, from 4 to the optimal which is 2; but with no measure estimate of Kolmogorov set discussion. See also [Vil18] where he got the exponent 2 for α in the smallness condition, in the framework of exact symplectic maps in Euclidean spaces of even dimensions.

1.1 Main results

As a basic rule in this thesis, we compute explicitly all the KAM constants. Investigating the explicit dependence of the "KAM constants" upon the parameters in a quasi–integrable Hamiltonian system is of great interest, not only in view of its applications (for instance to the n–body problem [CC06], to geodesic flows on surfaces, *etc*) but also for the discussion of explicit measure estimates of Kolmogorov sets. The content of this thesis can be described very roughly as follows:

(i) We prove three quantitative KAM normal forms following closely the original ideas of the pioneers Kolmogorov, Arnold and Moser. We compute in particular explicitly all the KAM constants in them and fix physical dimension issues by rescaling conveniently various quantities. Then, we compare those three quantitative KAM normal forms on a simple mechanical system.

(ii) We give detailed proof of how to get rid of the logarithmic correction in the Arnold's scheme and then use it to prove an explicit and "sharp" Theorem of integrability on Cantor–type set.

(iii) *We prove three types of sharp measure estimate of Kolmogorov sets. In the first one, we adopt the global approach which consists in constructing the Kolmogorov set in a given bounded domain and then estimate its measure. In the two others, we slice the domain into relatively small cubes with equilength sides. In each of those cubes, we construct a Kolmogorov set associated to the restriction of the Hamiltonian to such a cube and estimate its measure. Then, we sum up the local Kolmogorov sets constructed.*
One of the local approaches follows the idea in [BC18] *and recover its result.*
In the second local approach, we introduce a geometric integer constant of a set which is the minimal number of cubes one needs to cover the set by cubes with the same side–length, centered on the set and with total "volume" not exceeding some fixed amount. This third approach is somehow more intrinsic.

(**i**) More precisely, we prove in **Theorem 2.1.1** (following Kolmogorov's proof in [Kol54a], scheme to which a complete proof was given in [Chi08, Chi12]) that for any small enough perturbation of a non–degenerate Kolmogorov normal form, there exists a symplectic change of variables such that in the new variables, the Hamiltonian reduced to a Kolmogorov normal form.

We prove in **Theorem 2.1.2** (following Arnold [Arn63a] and basing on [Chi08, Chi12]) that, under a sufficiently small perturbation, with size say ε, of a non–degenerate integrable Hamiltonian system, the majority ($1 - O(\sqrt{\varepsilon}\,(\log \varepsilon^{-1})^{3\nu/4})$ of the total Lebesgue–measure) of the invariant, Lagrangian, Kronecker tori with Diophantine frequencies of the integrable system persist, being only slightly deformed.

In **Theorem 2.1.4**, we prove (following Moser[Mos67] and basing on [Pös01]) that on a bounded domain, the totality of the invariant maximal KAM tori of the linear normal form, whose frequencies are far enough from the boudary persist under any small enough perturbation. These tori are just a little bit deformed and persist as invariant maximal KAM tori, not of the perturbed Hamiltonian itself, but of the perturbed Hamiltonian plus a small shift of the frequency.

In **Chapter 3**, we compare the explicit KAM mormal forms on a simple mechanical Hamiltonian and compute the numerical values of the thresholds within these Theorems in a concrete case.

In **Theorem 4.2.1**, we prove an explicit Theorem of integrability on a Cantor–like set. Namely, for any given sufficiently small perturbation of a non–degenerate integrable Hamiltonian on a bounded domain, we construct a C^∞–symplectomorphism which conjugates the perturbed Hamiltonian to an integrable Hamiltonian on a Cantor–like set. The Cantor–like set is equipotent to the set of phase points which are at some minimal distance from the boundary and such that their image by the Jacobian of the unperturbed part are Diophantine vectors, with fixed Diophantine parameters. Moreover, the ratio of their respective Lebesgue–measures minus 1 is small with the size of the perturbation.

(**ii**) In **Theorem 5.2.1**, we prove a refinement of the Arnold's Theorem by overcoming the logarithmic correction looming from the original scheme. Indeed, we prove that, for any small enough perturbation of a non–degenerate integrable Hamiltonian system, most of $(1 - O(\sqrt{\varepsilon})$ of the total Lebesgue–measure, where ε is the size of the perturbation) of the invariant maximal KAM tori of the integrable system persist, up to a small deformation. To do so, we isolate the smallness parameter ε from the super–exponential parameter so that, and this is the whole point, as soon as ε is chosen conveniently to perform the scheme one time, one can iterate infinitely many times without any other requirement and, in particular, ε "disappears" once for all from the second step on.

In **Theorem 6.2.1**, *we prove an explicit, intrinsic and sharp integrability Theorem on a Cantor–like set.* Namely, given any small enough, real–analytic perturbation of a non–degenerate integrable Hamiltonian on a bounded domain, we build-up a transformation, C^∞ in the Whitney sense and symplectic. Actually, the two Cantor–like sets are lipeomorphic[1]. In those new variables, the nearly–integrable Hamiltonian becomes integrable on a Cantor–like set. The Cantor-like set is equipotent to the set of phase points which are far enough from the boundary and which images through the Jacobian of the unperturbed part are (α, τ)–Diophantine, for some $\alpha > 0$ and τ a number larger than the half–dimension minus one. In particular, we get a family of invariant maximal KAM torus which complement has a Lebesgue–measure bounded from above by a constant proportional to α.

(**iii**) *The novel part of the present thesis consists mainly in* **Part II** *and, in particular and more interestingly, the various "sharp" and geometric measure estimates of the unstable sets within a Hamiltonian system we provide.*

1.1.1 Measure estimate I

Then, we derive in **Theorem 6.2.2** the following. Let $\mathscr{D} \subset \mathbb{R}^d$ be a non–empty bounded domain with smooth boundary $\partial\mathscr{D}$ and a small enough $\alpha > 0$, depending on the geometry of the hypersurface $\partial\mathscr{D}$. Let H be a sufficiently small perturbation, of order $O(\alpha^2)$, of a non–degenerate integrable Hamiltonian K. Then, the set $\mathscr{I}$ left out of the $H-$ invariant maximal KAM tori is bounded in measure by

$$\operatorname{meas}(\mathscr{I}) \leqslant (3\pi)^d \frac{\mathsf{T}_0}{32 d\sigma_0}\Big(2\,\mathcal{H}^{d-1}(\partial\mathscr{D})\,\alpha + C(d,\sigma_0,\mathsf{T}_0,\mathbf{R}^{\partial\mathscr{D}})\,\alpha^2 + \operatorname{meas}(\mathscr{D}_\delta\backslash\mathscr{D}_{\delta,\alpha})\Big) \tag{1.1.1}$$

$$= O(\alpha)\,, \tag{1.1.2}$$

where[2] $\mathcal{H}^{d-1}$ denotes the $(d-1)$–dimensional Hausdorff measure (or equivalently, the $(d-1)$–surface area), $\mathbf{R}^{\partial\mathscr{D}}$ denotes the curvature tensor of $\partial\mathscr{D}$, σ_0 is the loss of analyticity, T_0 is the norm of the inverse of the Hessian K_{yy} of the unperturbed part,

$$\begin{aligned}
&\delta := \frac{\alpha\mathsf{T}_0}{32 d\sigma_0}\,,\\
&\mathscr{D}_\delta := \{y \in \mathscr{D} \ :\ \operatorname{dist}(y,\partial\mathscr{D}) \geqslant \delta\}\,,\\
&\mathscr{D}_{\delta,\alpha} := \{y \in \mathscr{D}_\delta \ :\ K_y(y) \in \Delta_\alpha^\tau\}\,,\\
&C(d,\sigma_0,\mathsf{T}_0,\mathbf{R}^{\partial\mathscr{D}}) := \frac{\mathsf{T}_0}{16 d\sigma_0}\sum_{j=1}^{\left[\frac{d-1}{2}\right]} \frac{\mathbf{k}_{2j}(\mathbf{R}^{\partial\mathscr{D}})}{1\cdot 3\cdots(2j+1)}\,\delta^{2j-1}\,,
\end{aligned}$$

[1] *i.e.* there exists a bijective Lipschitz continuous function from one onto the other.
[2] See Appendix F for the definitions.

with $\mathbf{k}_{2j}(\mathbf{R}^{\partial\mathscr{D}})$, the $(2j)$–th integrated mean curvature of $\partial\mathscr{D}$ in $\mathbb{R}^d$.

Remark 1.1.1 **(i)** The first two terms of the *r.h.s.* of (1.1.1) arise from the estimation of the δ–strip around $\partial\mathscr{D}$, $\mathscr{D}\backslash\mathscr{D}_\delta$, out of which we construct the family of invariant KAM tori. Notice that the last term $\text{meas}\,(\mathscr{D}_\delta\backslash\mathscr{D}_{\delta,\alpha})$ is of order $O(\alpha)$ by (1.0.6), whence (1.1.2) holds.
(ii) The estimate (1.1.1) might be seen as a "sharp" version of the measure estimate of the invariant set in the Two–scale KAM Theorem of [CP10].

The following is proven in **Theorem 6.4.2**. Let $H = K(y) + \varepsilon\, P(y, x)$ be a perturbation of a non–degenerate integrable Hamiltonian K, where $\mathfrak{D}$ a non–empty, bounded subset of $\mathbb{R}^d$ and K, P two real analytic function on $\mathfrak{D} \times \mathbb{T}^d$ with bounded extension to $D_{\mathsf{r}_0}(\mathfrak{D}) \times \mathbb{T}^d_{s_0}$, for some $\mathsf{r}_0 > 0$ and $0 < s \leqslant 1$. We prove that, for a sufficiently small ε (with explicit upper–bound), one can construct by "localization" argument a family of H–invariant maximal KAM torus, say $\mathscr{K}$, which complement has a Lebesgue–measure of order $O(\alpha)$ and estimated in two ways as follows.

1.1.2 Measure estimate II

More specifically, we show one hand that the Kolmogorov set $\mathscr{K}$ is bounded in measure from above (in the spirit of [BC18]) by

$$\text{meas}\,\big((\mathfrak{D} \times \mathbb{T}^d)\backslash\mathscr{K}\big) \leqslant C\, \mathsf{p}_1\, (\,\text{diam}\,\mathfrak{D} + \ell)^d\, \alpha\ ,$$

with C a positive universal constant depending only upon the dimension d and τ,

$$\begin{aligned}
&\mathsf{p}_1 := \frac{\vartheta\, \eta^2\, \mathsf{T}}{\sigma_0\, \mathsf{r}_0}\ ,\\
&\ell := \frac{\mathsf{r}_0}{2^6 d\eta^2}\ ,\\
&\eta := \mathsf{TK} \geqslant 1\ ,\\
&\vartheta := \frac{\mathsf{K}^d}{\varrho} \geqslant 1\ ,\\
&\varrho := \inf_{y\in\mathfrak{D}} |\det K_{yy}(y)| > 0\ ,\\
&0 < \sigma_0 < 2^{5-2\tau}\, d\, s_0\ ,\\
&\mathsf{T} := \sup_{y\in\mathfrak{D}} \|K_{yy}(y)^{-1}\|\ ,\\
&\mathsf{K} := \|K_{yy}\|_{\mathsf{r}_0,\mathfrak{D}}\ .
\end{aligned}$$

This provides an alternative proof to the result in [BC18] (alternative in the sense that the proof in [BC18] is based upon Moser's idea while here, we use Arnold's scheme) and our proof is somehow more complete as we compute explicitly all the constants while [BC18] refers to [Pös01], where the constants are left implicit.

1.1.3 Measure estimate III

On the other hand, in a more intrinsic way, we build up (under the same basic assumptions as above) a family $\mathscr{K}$ of $H-$ invariant maximal KAM torus such that

$$\operatorname{meas}\left((\mathfrak{D}\times\mathbb{T}^d)\backslash\mathscr{K}\right)\leqslant C'\,\frac{\vartheta\,\mathsf{T}}{\sigma_0}\,n_{\mathfrak{D}}^{\frac{1}{d}}\,\operatorname{meas}(\mathfrak{D})^{\frac{d-1}{d}}\,\alpha\ ,$$

with σ_0, ϑ, T as in §1.1.2, C a positive universal constant depending only upon the dimension d and τ, and $n_{\mathfrak{D}}\in\mathbb{N}$ a "covering number" of $\mathfrak{D}$ defined morally as follows. Given $R>0$, define the set $\mathscr{C}_R$ of coverings of $\mathfrak{D}$ by cubes as follows: $F\in\mathscr{C}_R$ if and only if there exists $n_F\in\mathbb{N}$ and n_F cubes, say F_i, $1\leqslant i\leqslant n_F$, of equal side–length $2R$, centered at a point $y_i\in\mathfrak{D}$ and such that

$$F=\left\{F_i : 1\leqslant i\leqslant n_F\right\}\qquad\text{and}\qquad \mathfrak{D}\subset\bigcup_{i=1}^{n_F}F_i\ .$$

Then define

$$\mathscr{R}:=\left\{R>0:\mathscr{C}_R\neq\varnothing\ \text{and}\ \inf_{F\in\mathscr{C}_R}n_F(2R)^d\leqslant 2^d\operatorname{meas}(\mathfrak{D})\right\}$$

and

$$n_{\mathfrak{D}}:=\min_{R\in\mathscr{R}}\min\left\{n_F : F\in\mathscr{C}_R\quad\text{and}\quad n_FR^d\leqslant\ \operatorname{meas}(\mathfrak{D})\right\}\ .$$

Remark 1.1.2 In the above definition of "covering number", one could replace the coefficient 2^d in front of $\operatorname{meas}(\mathfrak{D})$ in the definition of $\mathscr{R}$ by $\kappa>1$, leading to

$$\mathscr{R}_\kappa:=\left\{R>0:\mathscr{C}_R\neq\varnothing\ \text{and}\ \inf_{F\in\mathscr{C}_R}n_FR^d\leqslant 2^{-d}\kappa\cdot\ \operatorname{meas}(\mathfrak{D})\right\}$$

and

$$n_{\mathfrak{D},\kappa}:=\min_{R\in\mathscr{R}}\min\left\{n_F : F\in\mathscr{C}_R\quad\text{and}\quad n_FR^d\leqslant 2^{-d}\kappa\cdot\ \operatorname{meas}(\mathfrak{D})\right\}\ .$$

1.2 Notations and set up

Fix[3] $d \in \mathbb{N}\backslash\{1\}$. Let $\Omega \subset \mathbb{R}^d$ be non-empty and bounded domain with piecewise smooth boundary and $\mathbb{T}^d := \mathbb{R}^d/2\pi\mathbb{Z}^d$, the d–dimensional torus.
Given $h,\ r,\ s,\ \alpha,\ \varepsilon_0,\ \tau > 0,\ n, p \in \mathbb{N},\ y_0 \in \mathbb{C}^d$ and a non-empty $A \subset \mathbb{C}^d$, $A' \subset \mathbb{R}^d$, we define the following. Let[4]

$$\Delta_\alpha^\tau := \left\{\omega \in \mathbb{R}^d : |\omega \cdot k| \geqslant \frac{\alpha}{|k|_1^\tau}, \forall\ 0 \neq k \in \mathbb{Z}^d\right\}, \tag{1.2.1}$$

be the set of (α, τ)–Diophantine numbers, where $|k|_1 := \sum_{j=1}^{d} |k_j|$ is the 1–norm on[5] $\mathbb{C}^d$ and

$$\Omega_\alpha := \left\{\omega \in \Omega \cap \Delta_\alpha^\tau :\ \mathrm{dist}\,(\omega, \partial\Omega) := \min_{\omega_* \in \partial\Omega} |\omega - \omega_*| \geqslant \alpha\right\},$$

where $|\cdot|$ is some norm on $\mathbb{C}^d$; everywhere in this thesis, we shall use

$$|x| := \max_{1 \leqslant j \leqslant d} |x_j|,$$

the sup–norm on $\mathbb{C}^d$, except in §2.1.4 where we shall use $|\cdot| = |\cdot|_1$. Let[6]

$$\begin{aligned}
\mathbb{J} &:= \begin{pmatrix} 0 & -\mathbb{1}_d \\ \mathbb{1}_d & 0 \end{pmatrix}\ ,\\
\widehat{D}_r^d(y_0) &:= \left\{y \in \mathbb{C}^d : |y - y_0|_1 < r\right\}\ ,\\
D_r^d(y_0) &:= \left\{y \in \mathbb{C}^d : |y - y_0| < r\right\}\ ,\\
D_r^d(A) &:= \bigcup_{y_0 \in A} D_r^d(y_0)\ ,\\
B_r^d(A') &:= \mathbb{R}^d \cap D_r^d(A')\ ,\\
\mathbb{T}_s^d &:= \left\{x \in \mathbb{C}^d : |\,\mathrm{Im}\, x| < s\right\}/2\pi\mathbb{Z}^d\ ,\\
D_{r,s}^d(y_0) &:= \left\{y \in \mathbb{C}^d : |y - y_0| < r\right\} \times \mathbb{T}_s^d\ ,\\
D_{r,s}^d &:= D_{r,s}^d(0)\ ,\\
\Omega_{\alpha,h} &:= \bigcup_{\omega \in \Omega_\alpha} D_h(w)\ ,\\
W_{r,s,\varepsilon_0} &:= D_{rs,s}^d \times \{\varepsilon \in \mathbb{C} : |\varepsilon| < \varepsilon_0\}\ ,
\end{aligned}$$

[3]For us, $\mathbb{N} := \{1, \cdots\}$, $\quad \mathbb{N}_0 := \{0, 1, \cdots\}$.
[4]As usual $\omega \cdot k := \omega_1 k_1 + \cdots + \omega_d k_d$.
[5]And in general on $\mathbb{C}^d$ as well as on all its subsets ($\mathbb{R}^n$, $\mathbb{Z}^n$, $\mathbb{N}^n$, etc).
[6]We shall nevertheless drop the dimension d as it is fixed once for all, and write $B_r(y_0)$ instead of $B_r^d(y_0)$, *etc* .

where $\mathbb{1}_d :=$ diag (1) is the unit matrice of order d.
$\mathbb{C}^d \times \mathbb{C}^d$ will be equipped with the canonical symplectic form $\varpi := dy \wedge dx = dy_1 \wedge dx_1 + \cdots + dy_d \wedge dx_d$.
Given a linear operator $\mathcal{L}\colon (V_1, \|\cdot\|) \to (V_2, \|\cdot\|)$, its "operator–norm" is given by

$$\|\mathcal{L}\| := \sup_{x\in V_1\setminus\{0\}} \frac{\|\mathcal{L}x\|}{\|x\|}, \quad \text{so that} \quad \|\mathcal{L}x\| \leqslant \|\mathcal{L}\|\,\|x\| \quad \text{for any} \quad x \in V_1.$$

Let $\mathcal{A}_{r,s}(y_0)$ (resp. $\mathcal{B}_{r,s}(y_0)$, $\mathcal{A}_{r,s,h,d}$, $\mathcal{B}_{r,s,\varepsilon_0}$) be the set of real–analytic functions f on $\widehat{D}_r(y_0)\times\mathbb{T}^d_s$ (resp. $D_{r,s}\times\Omega_{\alpha,h}$, W_{r,s,ε_0}) with finite norm $|||f|||_{r,s,y_0}$ (resp. $\|f\|_{r,s,y_0}$, $\|f\|_{r,s,h,d}$, $\|f\|_{r,s,\varepsilon_0}$), defined below. Let

$$\mathcal{A}^0_{r,s}(y_0) := \left\{ f \in \mathcal{A}_{r,s}(y_0) : \langle f\rangle := \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} f(y,x)\,dx = 0,\ \forall y \in \widehat{D}_r(y_0) \right\},$$

$\mathcal{A}^0_{r,s,h,d}$ and $\mathcal{B}^0_{r,s,\varepsilon_0}$ are defined analogously. Given $\omega \in \mathbb{R}^d$ and $f \in \mathcal{A}_{r,s}(y_0) \bigcup \mathcal{A}_{r,s,h,d} \bigcup \mathcal{B}_{r,s,\varepsilon_0}$, we define

$$D_\omega f := \omega \cdot f_x = \sum_{j=1}^d \omega_j f_{x_j},$$

write[7]

$$f = \sum_{k\in\mathbb{Z}^d} f_k\, e^{ik\cdot x} = \sum_{k\in\mathbb{Z}^d, l\in\mathbb{N}^d} f_{l,k}\, e^{ik\cdot x}(y-y_0)^l,$$

where $\quad f_k := \dfrac{1}{(2\pi)^d}\displaystyle\int_{\mathbb{T}^d} f(x)\, e^{-ik\cdot x}\, dx =: \sum_{l\in\mathbb{N}^d} f_{l,k}\ (y-y_0)^l, \quad k \in \mathbb{Z}^d,$ define

$$T_\kappa f := \sum_{|k|_1\leqslant\kappa} f_k\, e^{ik\cdot x},\ \kappa \in \mathbb{N}$$

and define on $\mathcal{A}_{r,s}(y_0)$ (resp. $\mathcal{B}_{r,s}(y_0)$, $\mathcal{A}_{r,s,h,d}$, $\mathcal{B}_{r,s,\varepsilon_0}$), the norms[8]

$$|||f|||_{r,s,y_0} := \sum_{k,l\in\mathbb{Z}^d} |f_{l,k}|_1\ e^{s|k|_1} r^{|l|_1} \Bigg(\text{resp.}\quad \|\cdot\|_{r,s,y_0} := \sup_{D_{r,s}(y_0)} |f|,$$
$$\|\cdot\|_{r,s,h,d} := \sup_{D_{r,s}\times\Omega_{\alpha,h}} |f|, \qquad \|\cdot\|_{r,s,\varepsilon_0} := \sup_{W_{r,s,\varepsilon_0}} |f| \Bigg).$$

---

[7] As usual, $(y-y_0)^l := \prod_{j=1}^d (y_j - y_{0j})^{l_j}$. Here, and henceforth, $e := \exp(1)$ denotes the Neper number and i a complex–square–root of -1: $i^2 = -1$.

[8] Notice that the above definitions apply also to vector–valued real analytic functions *i.e.* $f := (f_1, \cdots, f_n)$ with $\{f_j\}_{j=1}^n \subset \mathcal{A}_{r,s}(y_0)$, etc.

Moreover, for a matrix–valued periodic functions $\mathcal{L}(y,x)$, we define[9]

$$|||\mathcal{L}|||_{r,s,y_0} := \sup_{|a|_1=1} |||\mathcal{L}(\cdot,\cdot)a|||_{r,s,y_0} .$$

and for a given $f \in \mathcal{A}_{r,s,y_0}$, the Fourier's norm of the 3–tensor $\partial_x^3 f$ is given by

$$\left|\left|\left|\partial_x^3 f\right|\right|\right|_{r,s,y_0} := \sup_{|b|_1=|c|_1=1} \sum_{j=1}^{d} \left|\left|\left| \sum_{k,l=1}^{d} \frac{\partial^3 f}{\partial x_j \partial x_k \partial x_l} b_l c_k \right|\right|\right|_{r,s,y_0} .$$

Given a map $\varphi\colon A \subset \mathbb{C}^n \to \mathbb{C}^p$, its Lipschitz constant is defined by

$$\|\varphi\|_{L,A} := \sup_{x,y\in A, x\neq y} \frac{|\varphi(x)-\varphi(y)|}{|x-y|} \leqslant \infty.$$

1.3 General remarks

1. We have chosen the norms for simplicity but any others[10] “algebra norms” maybe be used.

2. As we are going to compare the four theorems on a concrete Hamiltonian and since we use two different norms, we need a kind of equivalence between them. Indeed, we have, for any $r>0$, $0<\sigma<s$

$$\begin{aligned}
\|f\|_{r,s-\sigma,y_0} \leqslant |||f|||_{r,s-\sigma,y_0} &= \sum_{n,m\in\mathbb{Z}^d} |f_{m,n}|_1\, e^{(s-\sigma)|n|_1} r^{|m|_1} \\
&\leqslant \sum_{n,m\in\mathbb{Z}^d} \frac{d\|f\|_{r,s,y_0}}{r^{|m|_1}}\, e^{-s|n|_1}\, e^{(s-\sigma)|n|_1} r^{|m|_1} \\
&= d\|f\|_{r,s,y_0} \sum_{n\in\mathbb{Z}^d} e^{-\sigma|n|_1} \\
&= d\|f\|_{r,s,y_0} \left(\sum_{k\in\mathbb{Z}} e^{-\sigma|k|}\right)^d \\
&= d\|f\|_{r,s,y_0} \left(1+\frac{2}{e^{-\sigma}-1}\right)^d \\
&= d\|f\|_{r,s,y_0} \tanh^d\left(\frac{\sigma}{2}\right).
\end{aligned}$$

[9]With an analogous definitions with the other norms. For instance, $\|\mathcal{L}\|_{r,s,y_0} := \sup_{|a|=1} \|\mathcal{L}(\cdot,\cdot)a\|_{r,s,y_0}$.

[10]An algebra norm is a norm satisfying $\|x\cdot y\| \leqslant \|x\|\|y\|$, for any x and y.

3. Through the present thesis, we shall denote by C (resp. c), at any place (with index or not), a constant depending (eventually) only on d, τ and $\bar{\nu}$ (see below) and greater (resp. less) or equal than 1 *i.e.* $C = C(d, \tau, \bar{\nu}) \geqslant 1$ (resp. $c = c(d, \tau, \bar{\nu}) \leqslant 1$).

Part I

Classical KAM Theorems and Quantitative normal forms

2 | Quantitative KAM normal forms

2.1 Statement of the explicit KAM normal forms theorems

2.1.1 Kolmogorov's normal form

2.1.1.1 Assumptions

Let $\alpha,\ r,\ \varepsilon_0 > 0, \quad \tau \geqslant d-1, \quad 0 < 2\sigma < s \leqslant 1$ and

$$s_* := s - 2\sigma.$$

Let's consider a hamiltonian $H \in \mathcal{B}_{r,s,\varepsilon_0}$ such that $K(y,x) := H(y,x;0)$ has the form[11]

$$K = \mathrm{K} + \omega \cdot y + Q(y,x) \quad \text{with} \quad Q = O(|y|^2), \quad \mathrm{K} \in \mathbb{R}, \quad \text{and} \tag{2.1.1}$$

$$\omega \in \Delta_\alpha^\tau \quad i.e. \quad \omega \quad \text{is} \quad (\alpha,\tau)\text{–diophantine.} \tag{2.1.2}$$

Furthermore, assume that K in (2.1.1) is non–degenerate in the sense that[12]

$$\det \left\langle \partial_y^2 Q(0,\cdot) \right\rangle \neq 0 \ .$$

Write

$$\boxed{H =: K + \varepsilon P}$$

and set

$$M := \|P\|_{r,s,\varepsilon_0}\ , \qquad T := \left\langle \partial_y^2 Q(0,\cdot) \right\rangle^{-1}.$$

[11] As usual, $\omega \cdot y = \omega_1 y_1 + \cdots + \omega_d y_d$; $Q = O(|y|^2)$ means that $\partial_y^m Q(0,x) = 0$ for all $m \in \mathbb{N}^d$ with $|m|_1 \leqslant 1$, where $\partial_y^m = \frac{\partial^{|m|_1}}{\partial y_1^{m_1} \cdots \partial y_d^{m_d}}$ and $|m|_1 = m_1 + \cdots + m_d$.

[12] $\langle \cdot \rangle$ being the average over $\mathbb{T}^d$.

Finally define

$$
\begin{aligned}
\mathrm{E} &:= 2\widehat{\mathrm{E}} := 2\max\left(r|\omega|, \|Q\|_{r,s,\varepsilon_0}, |\omega|^2\|T\|\right),\\
\mathcal{W} &:= \operatorname{diag}(|\omega|\,\mathbb{1}_d, r\sigma|\omega|\,\mathbb{1}_d),\\
r_* &:= r(s-2\sigma),\\
B_* &:= B_{r_*}(0),\\
\mathrm{C}_0 &= 2^{d+1-2\tau}\sqrt{\Gamma(2\tau+1)}\,,\\
\mathrm{C}_1 &:= 2\cdot 3^\tau \mathrm{C}_0,\\
\mathrm{C}_2 &:= 2d\mathrm{C}_1 + 2^{-(\tau+1)},\\
\mathrm{C}_3 &:= d\mathrm{C}_2 + 2^{-(\tau+2)},\\
\mathrm{C}_4 &:= \mathrm{C}_2 + 2^{-2}\mathrm{C}_1,\\
\mathrm{C}_5 &:= 3^\tau d\mathrm{C}_0\left(2d\mathrm{C}_4 + 2^{-(\tau+3)}\right),\\
\mathrm{C}_6 &:= 2^{-(\tau+2)}\mathrm{C}_4 + \mathrm{C}_5,\\
\mathrm{C}_7 &:= \frac{3}{2}d\mathrm{C}_5 + 81\cdot 2^{-(\tau+3)}d^3\mathrm{C}_4 + 9\cdot 2^{-(2\tau+5)}d^2,\\
\mathrm{C}_8 &:= 18\mathrm{C}_7,\\
\mathrm{C}_9 &:= 9d^2\mathrm{C}_6^2 + 3\cdot 2^{-(2\tau+5)}d\mathrm{C}_6,\\
\bar{\nu} &:= 4\tau+10,\\
\nu &:= 4\tau+12,\\
\bar{\mathrm{C}} &:= \max\left(2^{-(2\tau+5)}\mathrm{C}_8, \mathrm{C}_9\right),\\
\widetilde{\mathrm{C}} &:= d\left(3d\bar{\mathrm{C}} + 2^{-(2\tau+6)}\mathrm{C}_7\right),\\
\mathrm{C}_\sharp = \mathrm{C}_\sharp(d,\tau) &:= \frac{9d\cdot 2^{4\tau+23}}{5}\left(3d\bar{\mathrm{C}} + 2^{-(2\tau+6)}\mathrm{C}_7\right),\\
\mathrm{c} = \mathrm{c}(d,\tau) &:= \frac{1}{\mathrm{C}_\sharp}\,,\\
\widehat{\mathrm{C}} &:= \frac{6d}{5}\left(3d\bar{\mathrm{C}} + 2^{-(2\tau+6)}\mathrm{C}_7\right),\\
\mathrm{C} = \mathrm{C}(d,\tau) &:= \frac{\widehat{\mathrm{C}}}{3\bar{\mathrm{C}}}\,,\\
\bar{L} &:= \bar{\mathrm{C}}\mathrm{E}^7\sigma^{-\bar{\nu}}r^{-7}\alpha^{-4}|\omega|^{-3}M,\\
\tilde{L} &:= \widetilde{\mathrm{C}}\mathrm{E}^8\sigma^{-\nu}r^{-8}\alpha^{-4}|\omega|^{-4}M,\\
L &:= \frac{6}{5}r^{-2}|\omega|^{-2}\tilde{L}\mathrm{E}^2 = \frac{\mathrm{C}_\sharp}{3\cdot 2^\nu}\widehat{\mathrm{E}}^{10}\sigma^{-\nu}r^{-10}\alpha^{-4}|\omega|^{-6}M,
\end{aligned}
$$

2.1.1.2 Statement of the KAM Theorem

Theorem 2.1.1 (Komogorov [Kol54a], pg. 52) *Under the assumptions in* §2.1.1.1*, the following hold. There exists a real–analytic symplectomorphism* $\phi_*\colon B_* \times \mathbb{T}^d \xrightarrow{into} B_r(0) \times \mathbb{T}^d$*, depending analytically also on* $\varepsilon \in (-\varepsilon_*, \varepsilon_*)$*, with*

$$\boxed{\varepsilon_* := \min\left(\varepsilon_0, \mathrm{c}\, \widehat{\mathrm{E}}^{-9}\sigma^{4\tau+13}r^{10}\alpha^4|\omega|^6 M^{-1}\right)},$$

such that $\phi_*|_{\varepsilon=0}$ *is the identity map and, for any* $|\varepsilon| < \varepsilon_*$*,*

$$H \circ \phi_*(y', x') = K_*(y', x'; \varepsilon) := \mathrm{K}_*(\varepsilon) + \omega \cdot y' + Q_*(y', x'; \varepsilon), \quad \text{with} \quad Q_* = O(|y'|^2)$$

and

$$\frac{\mathrm{CE}^3}{r^3|\omega|^3\sigma^2}\|\mathcal{W}(\phi_* - \mathrm{id})\|_{r_*, s_*, \varepsilon_*}\,, \quad |\mathrm{K} - \mathrm{K}_*|_{\varepsilon_*}, \quad \|Q - Q_*\|_{r_*, s_*, \varepsilon_*}\,, \quad |\omega|^2\|T - T_*\| \leqslant \frac{|\varepsilon| L}{3\sigma}\,. \tag{2.1.3}$$

2.1.2 KAM Theorem après Arnold

2.1.2.1 Assumptions

Let $\alpha, r_0 > 0$, $\tau \geqslant d-1$, $0 < 2\sigma_0 < s_0 \leqslant 1$, $y_0 \in \mathbb{R}^d$ and consider the Hamiltonian parametrized by $\varepsilon \in \mathbb{R}$

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

with

$$K, P \in \mathcal{B}_{r_0, s_0}(y_0)\,.$$

such that

$$\boxed{\omega := K_y(y_0) \in \Delta_\alpha^\tau\,, \qquad \det K_{yy}(y_0) \neq 0.} \tag{2.1.4}$$

Set

$$T := K_{yy}(y_0)^{-1}, \quad M_0 := \|P\|_{r_0, s_0, y_0}, \quad \mathsf{K}_0 := \|K_{yy}\|_{r_0, y_0}, \quad \mathsf{T}_0 := \|T\|\,.$$

Finally, for a given $\varepsilon \neq 0$, define[13]

$$
\begin{aligned}
\mathsf{W}_0 &:= \operatorname{diag}\left(\max\left\{\frac{\mathsf{K}_0}{\alpha}, \frac{1}{r_0}\right\}\mathbb{1}_d, \mathbb{1}_d\right) ,\\
\eta_0 &:= \mathsf{T}_0\mathsf{K}_0 ,\\
\nu &:= \tau + 1 ,\\
\mathsf{C}_0 &:= 4\sqrt{2}\left(\frac{3}{2}\right)^{2\nu+d}\int_{\mathbb{R}^d}\left(|y|_1^{\nu} + d|y|_1^{2\nu}\right) e^{-|y|_1}dy ,\\
\mathsf{C}_1 &:= 2\left(\frac{3}{2}\right)^{\nu+d}\int_{\mathbb{R}^d}|y|_1^{\nu}\, e^{-|y|_1}dy ,\\
\mathsf{C}_2 &:= d^4 3^{8(d-1)} ,\\
\mathsf{C}_3 &:= d^2\mathsf{C}_1^2 + 6d\mathsf{C}_1 + 1 ,\\
\mathsf{C}_4 &:= \max\left\{\mathsf{C}_0, \mathsf{C}_3\right\} ,\\
\mathsf{C}_5 &:= 2^{2(\nu+d)+11}3^2 5^{-2} d^2 ,\\
\mathsf{C}_6 &:= \max\left\{32d\, , \, 10^{-\nu}\mathsf{C}_7\right\} ,\\
\mathsf{C}_7 &:= \max\{\mathsf{C}_2\ ,\ \mathsf{C}_4\} ,\\
\mathsf{C}_8 &:= 3\cdot 5^{\nu}\ \mathsf{C}_6 ,\\
\mathsf{C}_9 &:= \frac{3\cdot 5^{2\nu+1}\sqrt{2}}{8}\ \mathsf{C}_6 ,\\
\mathsf{C}_{10} &:= \max\left\{1\ ,\ \left(\frac{3d\cdot 2^{5-d}}{5}\right)^{\frac{1}{4}}\right\} ,\\
\mathsf{C}_{11} &:= \frac{\mathsf{C}_5^2\mathsf{C}_9\mathsf{C}_{10}}{3} ,\\
s_* &:= s_0 - 2\sigma_0 ,\\
\mathfrak{p}_1 &:= \mathsf{C}_8\ \eta_0\ \sigma_0^{-(3\nu+2d+1)}\ \max\left\{1, \frac{\alpha}{r_0\mathsf{K}_0}\right\} ,\\
\mathfrak{p}_2 &:= \mathsf{C}_{11}\ \eta_0^{\frac{17}{4}}\ \sigma_0^{-(4\nu+2d)} ,\\
\varepsilon_\sharp &:= \min\left\{e^{-1}\, , \, \exp\left(-\frac{\sigma_0}{5}\left(\frac{12\sqrt{2}}{5}\frac{\alpha\mathsf{T}_0}{r_0}\right)^{\frac{1}{\nu}}\right)\right\} ,\\
\mu_0 &:= \frac{\mathsf{K}_0|\varepsilon| M_0}{\alpha^2} .
\end{aligned}
$$

[13]Notice that $\int_{\mathbb{R}^d}|y|_1^{\tau}\, e^{-|y|_1}dy \geqslant \int_{\{|y_j|\geqslant 1 : j=1,\cdots,d\}}|y|_1^{\tau}\, e^{-|y|_1}dy \geqslant d^{\tau}\left(\int_{\{|y_1|\geqslant 1\}}e^{-|y_1|}dy_1\right)^d =$ $= d^{\tau}\left(2\, e^{-1}\right)^d \geqslant d^{d-1}\left(2\, e^{-1}\right)^d = d^{\frac{d}{2}-1}\left(2\sqrt{d}\, e^{-1}\right)^d > 1$ because $\tau \geqslant d-1 \geqslant 1$. Thus, $\mathsf{C}_0 > 1$ and $\mathsf{C}_1 > 1$.

2.1.2.2 Statement of the KAM Theorem

Theorem 2.1.2 (Arnold [Arn63a]) *Under the assumptions in §2.1.2.1, the following hold. For any given ε satisfying*

$$\boxed{\begin{cases} \mu_0 \leqslant \varepsilon_\sharp\,, \\ \\ \mathfrak{p}_1 \cdot \max\left\{1\,,\, \mathfrak{p}_2\, \mu_0\, \left(\log \mu_0^{-1}\right)^{2\nu}\right\} \cdot \mu_0\, \left(\log \mu_0^{-1}\right)^{\nu} < 1\,, \end{cases}} \tag{2.1.5}$$

there exist $y_* \in B_{r_0}(y_0)$ and an embedding $\phi_* \colon \mathbb{T}^d \to D_{r_0,s_0}(y_0)$, real–analytic on $\mathbb{T}^d_{s_*}$ and close to the trivial embedding

$$\phi_0 \colon x \in \mathbb{T}^d \to (y_*, x) \in D_{r_0,s_0}(y_0),$$

such that the d–torus

$$\mathcal{T}_{\omega,\varepsilon} := \phi_*\left(\mathbb{T}^d\right) \tag{2.1.6}$$

is a non-degenrate invariant Kronecker torus for H i.e.

$$\phi_H^t \circ \phi_*(x) = \phi_*(x + \omega t). \tag{2.1.7}$$

Moreover,

$$|\mathsf{W}_0(\phi_* - \mathrm{id})| \leqslant \sigma_0^{d+1}\,, \tag{2.1.8}$$

uniformly on $\{y_*\} \times \mathbb{T}^d_{s_*}$.

Remark 2.1.3 It is not difficult to see that Theorem 2.1.2 is stronger than Theorem 2.1.1. Indeed, let the assumptions in §2.1.1.1 hold. Then, Taylor's expansion yields $H(y,x) = K^\sharp(y) + P^\sharp(y,x)$, where

$$\begin{aligned} K^\sharp(y) &:= \mathrm{K} + \omega \cdot y + \frac{1}{2}(T^{-1}y)\cdot y\,, \\ P^\sharp(y,x) &:= \frac{1}{2}\left((\partial_y^2 Q(0,x) - T^{-1})y\right)\cdot y + 3\sum_{|\beta|_1=3} y^\beta \int_0^1 \frac{(1-t)^2}{\beta!}\partial_y^\beta Q(ty,x)\,dt + \varepsilon P(y,x)\,. \end{aligned}$$

Thus,

$$K_y^\sharp(0) = \omega \in \Delta_\alpha^\tau\,, \qquad \det K_{yy}^\sharp(0) = \det T^{-1} \neq 0 \qquad \text{and} \qquad \|P^\sharp\|_{r,s,\varepsilon_0} = O(r^2) + O(\varepsilon)\,.$$

Consequently, by choosing r proportional to $\sqrt{\varepsilon}$, one can apply Theorem 2.1.2 and recover Theorem 2.1.1.

2.1.3 KAM Theorem après J. Moser (following J. Pöschel)

2.1.3.1 Assumptions

Let $r, h > 0, \quad 0 < s \leqslant 1$ and consider the hamiltonian parametrized by ω

$H(y,x,\omega) := N(I,\omega) + P(y,x,\omega)$ where $N(y,\omega) := K_0(\omega) + \omega \cdot y,$ with $K_0,\ P \in \mathcal{A}_{r,s,h,d}$,

and X_H and X_N the hamiltonian vector fields associate to H and N respectively with respect to the canonical symplectic form ϖ. Let ϕ_H^t and ϕ_N^t be the hamiltonian flow associate to H and N respectively. We have then, $\phi_N^t(y,x) = (y, x+\omega\cdot t)$. Let[14]

$$\alpha > 0, \quad \nu > \bar{\nu} = \tau + 1 > d, \quad \epsilon := \|P\|_{r,s,h,d}\ , \quad \beta := 1 - \frac{\nu - \bar{\nu}}{\nu\bar{\nu}},$$

$$\bar{c} := \min\left(1, \frac{\nu-\bar{\nu}}{\nu\bar{\nu}}\, e\right)\ , \quad B := B_r(0)\ , \quad W := \ \mathrm{diag}\ \left(r^{-1}\mathbb{1}_d, 20s^{-1}\mathbb{1}_d\right),$$

$$\widetilde{W} := \ \mathrm{diag}\ \left(20^{-\bar{\nu}}\alpha s^{\bar{\nu}} r^{-1}\mathbb{1}_d, 20^{-\tau}\alpha s^{\tau}\mathbb{1}_d, \mathbb{1}_d\right),$$

[14] Notice that one could use any $1 > \beta \geqslant 1 - \frac{1}{\bar{\nu}} + \frac{1}{\nu}$.

and define[15]

$$
\begin{aligned}
C_0 &:= 2^{d+1-2\tau}\sqrt{\Gamma(2\tau+1)}\ ,\\
C_1 &:= \frac{4\,e}{3}\ ,\\
C_2 &:= \sum_{j=0}^{d-1}\frac{(d-1)!}{(d-1-j)!(d-1)^{j+1}},\\
C_3 &:= 4(d+1)C_0\ ,\\
C_4 &:= 16(d+1)C_0\ ,\\
C_5 &:= 16(6C_1+1)C_3\ ,\\
C_6 &:= 2dC_5\sum_{j=0}^{\infty}2^{-2\bar\nu\left(\left(\frac{3}{2}\right)^j-j-1\right)-j},\\
C_7 &:= 36d(d+1)(6C_1+1)\exp\left(\frac{36d(d+1)(6C_1+1)}{C_6}\sum_{j=0}^{\infty}2^{-\bar\nu\left(2\left(\frac{3}{2}\right)^j-j-2\right)}\right),\\
C_8 &:= C_7\sum_{j=0}^{\infty}2^{-\bar\nu\left(2\left(\frac{3}{2}\right)^j-j-2\right)-j},\\
C_9 &:= \max\left(1,\frac{6C_1+1}{2C_0}\right),\\
C_{10} &:= 9\cdot 4^{\bar\nu}\left(\max\left\{48d(d+1)^2C_0, 4^d(d+1)C_2\right\}\right)^2,\\
C_{11} &:= \exp\left(\frac{1}{2}\left(\left(\frac{2\bar\nu}{\bar c}\right)^{1/\beta}+\frac{1}{20}\right)\right),\\
C_{12} &:= \left(\frac{2}{\bar c}\right)^{\nu}\left(2\,e^{\frac{1}{40}}C_6\right)^{\nu/\bar\nu},\\
C_{13} &:= \exp\left(\frac{1}{2}\left(\left(\frac{C_0}{6C_1+1}\right)^{1/\bar\nu}+\frac{1}{20}\right)\right)
\end{aligned}
$$

[15] $\Gamma(z) := \int_0^\infty t^{z-1}\,e^{-t}dt$, $\mathrm{Re}\, z > 0$ is the Euler's gamma function. Notice that if $\tau \geqslant d-1$ then $C_0 \geqslant 2^{d+1-2\tau}\sqrt{4^{[2\tau]-3}\cdot 6} \geqslant \sqrt{3}$, where $[2\tau]$ denotes the integer part of 2τ.

and[16]

$$\begin{aligned} C &= C(d,\tau) := e^{\frac{1}{40}} C_6, \\ C_* &= C_*(d,\tau) := \max\left(C_6 C_9, C_8\right), \\ c &= c(d,\tau,\bar\nu) := 20^{-\nu} \min\left(\frac{1}{4^{\bar\nu} C_{10}}, \frac{20^{\nu}\, e^{\frac{1}{40}}}{C_{11}}, \frac{e^{\frac{1}{40}}}{C_{12}}, \frac{20^{\nu}}{C_{13}}\right). \end{aligned} \tag{2.1.9}$$

Finally, let

$$\Phi_0\colon (x,\omega) \in \mathbb{T}^d \times \mathbb{R}^d \mapsto (0,x) \in B \times \mathbb{T}^d,$$

be the trivial embedding.

2.1.3.2 Statement of the KAM Theorem

Theorem 2.1.4 (Pöschel [Pös01]) *Let H, Φ_0, ϵ, ν, $\bar\nu$, C, C_* and c as in §2.1.3.1 and assume that*

$$\boxed{\varepsilon := \frac{\epsilon}{\alpha r} \leqslant c s^{\nu} \quad \textit{and} \quad C\varepsilon \leqslant \frac{h}{\alpha} \leqslant \frac{1}{2\kappa_0^{\bar\nu}}}, \tag{2.1.10}$$

where

$$\kappa_0 := \left[-\frac{40\log\varepsilon - 1}{s}\right].$$

Then, there exist a Lipeomorphism $\varphi\colon \Omega \to \Omega$ close to the identity and a Lipschitz continuous family of real analytic Lagrangian torus embeddings $\Phi\colon \mathbb{T}^d \times \Omega \to B \times \mathbb{T}^d$ closed to the trivial embedding Φ_0 such that the following hold. For any $\omega \in \Omega_\alpha$, $\Phi(\mathbb{T}^d,\omega)$ is a Lagrangian submanifold and an invariant Kronecker torus for $H_{|\varphi(\omega)}$ with $H_{|\varphi(\omega)}(y,x,\omega) := H(y,x,\varphi(\omega))$, i.e.

$$\phi^t_{H_{|\varphi(\omega)}} \circ \Phi(x,\omega) = \Phi(x+\omega t;\omega), \forall\, x \in \mathbb{T}^d. \tag{2.1.11}$$

Moreover, $\Phi(\mathbb{T}^d,\omega)$ is a Lagrangian submanifold and the maps $x \mapsto \Phi(x,\omega)$ is real analytic on $\mathbb{T}^d_{\frac{s}{2}}$ for each given $\omega \in \Omega$ and one has uniformly on $\mathbb{T}^d_{\frac{s}{2}} \times \Omega$ and Ω respectively, the following estimates[17]

$$\|W(\Phi-\Phi_0)\|, \quad h\,\|W(\Phi-\Phi_0)\|_{L,\mathbb{R}^d} \leqslant C_* \frac{\|P\|_{r,s,h}}{rh}, \tag{2.1.12}$$

$$\|\varphi - \mathrm{id}\|, \quad h\,\|\varphi - \mathrm{id}\|_{L,\Omega} \leqslant C_* \frac{\|P\|_{r,s,h}}{r}. \tag{2.1.13}$$

[16] Notice that $C_{12} > C_6$. Moreover, if $\varepsilon < e^{\frac{1}{40}} \min(20^{\nu} C_{11}^{-1}, C_{12}^{-1})\sigma^{\nu}$ then $\kappa_0^{\bar\nu}\sigma_0^{\bar\nu}\, e^{-\kappa_0\sigma_0} \leqslant \varepsilon < 1/(2C_6\kappa_0^{\bar\nu})$ and therefore $\kappa_0\sigma > d-1$; compare Appendix A, with $\vartheta = \frac{1}{2}$ and ε replaced by $\Theta\, e^{\frac{1}{40}}$.

[17] Here and in §2.3.3 as well, we shall denote by $\|f\|_{L,A}$, the uniform Lipschitz' semi-norm of the function f *w.r.t* the ω–argument (parameter) varying in the set A.

Remark 2.1.5

1. If one chooses $h = \frac{\epsilon C}{r}$, then the assumption (2.1.10) in Theorem 2.1.4 reduced to

$$\epsilon \leqslant cr\alpha s^{\nu}.$$

2. Notice that we have some freedom in the choice of C_6. Indeed, one just needs to chose

$$C_6 > \frac{C_5}{2C_0 \log 2} \sum_{j=0}^{\infty} 2^{-\bar{\nu}(2\mu^j - j - 2)} \; . \tag{2.1.14}$$

3. To be precise, in Theorem 2.1.4,

$$\Phi \colon \{0\} \times \mathbb{T}^d \times \Omega \to B \times \mathbb{T}^d.$$

4. Notice that the ν in Theorem 2.1.4 is larger than the $\nu = \tau + 1$ Pöschel uses in Theorem A and B in [Pös01]. In fact the Theorem A and B are not valid for $\nu = \tau + 1$. Indeed[18], assuming the contrary, then for any $\varepsilon = \frac{\epsilon}{\alpha r} \leqslant \gamma s^{\nu}$, there would exists $\kappa_0 \in \mathbb{N}$ such that

$$\kappa_0^{\nu} \sigma^{\nu} \, e^{-\kappa_0 \sigma} \leqslant \varepsilon \leqslant \frac{1}{2\kappa_0^{\nu}}, \quad \text{with} \quad \sigma = \frac{s}{20}. \tag{2.1.15}$$

But then[19], $e^{-\kappa_0 \sigma} \leqslant \varepsilon$ *i.e.* $\kappa_0 \sigma \geqslant \log \varepsilon^{-1}$. Hence, we would have, if we take $\varepsilon = \gamma s^{\nu}$ in particular, for any $0 < s \leqslant 1$,

$$\frac{1}{2} \frac{s^{\nu}}{20^{\nu}} \overset{(2.1.15)}{\geqslant} \varepsilon \kappa_0^{\nu} \sigma^{\nu} \geqslant \varepsilon \left(\log \varepsilon^{-1}\right)^{\nu} = \gamma s^{\nu} \left(\log(\gamma s^{\nu})^{-1}\right)^{\nu}$$

i.e.

$$1 \geqslant 2 \cdot 20^{\nu} \gamma \left(\log(\gamma s^{\nu})^{-1}\right)^{\nu}, \quad \forall\, 0 < s \leqslant 1; \text{contradiction.}$$

2.1.4 KAM Theorem après Salamon–Zehnder

2.1.4.1 Assumptions

Let

$$0 < \hat{s} < s \leqslant \bar{s} \leqslant 1, \quad \sigma := \frac{s - \hat{s}}{2}, \quad r, \alpha, \mathtt{E}, E_{j,k} \geqslant 0, \quad \tau \geqslant d - 1,$$

where $j, k \in \mathbb{N}$. Let's consider a hamiltonian $H \in \mathcal{A}_{r,\bar{s}}(y_0)$, for some $y_0 \in \mathbb{R}^d$ and a pair of real–analytic functions (u, v) on $\mathbb{T}^d_s$ such that

$$\left\| (H_{yy})^{-1} \right\|_{r,\bar{s},y_0} \leqslant \mathtt{E}, \quad \left\| \partial_x^j \partial_y^k H \right\|_{r,\bar{s},y_0} \leqslant E_{j,k} \; ,$$

[18] We are using here the same notations as in [Pös01].

[19] Because $\kappa_0 \sigma \geqslant 1$.

for any $j,k \in \mathbb{N}$ and

$$\left\|u\right\|_s \leqslant U, \quad \left\|v\right\|_s \leqslant V \quad \text{and} \quad \rho := \left\|v - y_0\right\|_s < r,$$

for some

$$0 \leqslant U \leqslant \bar{s} - s \quad \text{and} \quad V > 0,$$

$$\mathcal{M} := \mathbb{1}_d + u_\theta, \quad \text{and} \quad H^0_{yy}(\theta) := H_{yy}\left(v(\theta), \theta + u(\theta)\right)$$

are invertible for each given $\theta \in \mathbb{T}^d$ and, defining[20],

$$\mathcal{T} := \mathcal{M}^{-1} H^0_{yy} \mathcal{M}^{-T},$$

$\langle \mathcal{T} \rangle$ is invertible. Let $\omega \in \Delta^\tau_\alpha$ and define f and g by

$$\begin{cases} \omega + D_\omega u - H_y(v, \mathrm{id} + u) &= f \\ D_\omega v + H_x(v, \mathrm{id} + u) &= g \end{cases} \tag{2.1.16}$$

Futhermore, assume that

$$\left\|\mathcal{M}\right\|_s \leqslant \mathtt{M}, \quad \left\|\mathcal{M}^{-1}\right\|_s \leqslant \overline{\mathtt{M}}, \quad \left\|v_\theta\right\|_s \leqslant \widetilde{V}, \quad \left\|f\right\|_s \leqslant F, \quad \left\|g\right\|_s \leqslant G, \quad |\langle \mathcal{T} \rangle^{-1}| \leqslant \widetilde{T},$$

for some $\widetilde{V}, F, G \geqslant 0$ and $\mathtt{M}, \overline{\mathtt{M}}, \widetilde{T} > 0$. Finally, define

$$\begin{aligned}
\widehat{V} &:= \max\{\widetilde{V}, r - \rho\}, \\
\mathtt{A}_1 &:= |\omega|, \quad \mathtt{A}_2 := \max\left\{\mathtt{A}_1 \, , \, E_{1,1}\right\}, \\
\mathtt{A}_3 &:= \max\left\{E_{2,1} \, , \, \mathtt{E}E_{1,2}\mathtt{A}_2 \, , \, \mathtt{E}^2 E_{0,3}\mathtt{A}_2^2\right\}, \\
\mathtt{A}_4 &:= \max\left\{E_{3,0} \, , \, \mathtt{E}E_{2,1}\mathtt{A}_2 \, , \, \mathtt{E}^2 E_{1,2}\mathtt{A}_1^2\right\}, \\
\mathtt{A}_5 &:= \max\left\{\mathtt{A}_4 \, , \, \mathtt{E}\mathtt{A}_1\mathtt{A}_2\right\}, \quad \mathtt{A}_6 := \max\left\{\mathtt{A}_5 \, , \, \widehat{V}\mathtt{A}_3\right\}, \\
\mathtt{A}_7 &:= \max\left\{\mathtt{E}E_{0,2} \, , \, E_{0,2}\widetilde{T}\right\} \cdot \max\left\{\alpha^{-2}E_{0,2}\mathtt{A}_6 \, , \, \alpha^{-1}\mathtt{A}_3\right\}, \\
\mathtt{A}_8 &:= (s - \hat{s})^{2\tau} \max\left\{1 \, , \, \frac{\mathtt{E}\mathtt{A}_2}{r - \rho} \, , \, E_{0,2}\widetilde{T} \, , \, E_{1,2}\widetilde{T} \, , \, \mathtt{E}E_{0,3}\mathtt{A}_2\widetilde{T}\right\}, \\
\mathtt{A}_9 &:= \max\left\{\mathtt{A}_7 \, , \, \mathtt{A}_8\right\}, \\
\mathtt{A}_* &:= \max\left\{\mathtt{E}E_{0,2} \, , \, E_{0,2}\widetilde{T}\right\} \cdot \max\left\{\alpha^{-1}F \, , \, \alpha^{-2}E_{0,2}\widehat{V}F \, , \, \alpha^{-2}E_{0,2}G\right\}.
\end{aligned}$$

[20] A^{-T} stands for the transpose of the inverse of A: $A^{-T} := \left(A^{-1}\right)^T$.

2.1.4.2 Statement of the KAM Theorem

Theorem 2.1.6 (Celletti–Chierchia [CC97]) *Under the assumptions in §2.1.4.1, the following holds. There exists a polynomial Ξ in (s,σ) satisfying*

$$\frac{5}{4} \leqslant \Xi(a,b) \leqslant 21+88a \leqslant 109, \quad \forall\, 0<a\leqslant 1, \forall\, 0<b<\frac{1}{2}, \tag{2.1.17}$$

such that, if

$$\boxed{\mathtt{A}_*\mathtt{A}_9\mathtt{M}^7\overline{\mathtt{M}}^9(s-\hat{s})^{-2(2\tau+1)}2^{8\tau+13}\tau!^4\,\Xi(s,\sigma)\leqslant 1}, \tag{2.1.18}$$

then there exists $(\tilde{u},\tilde{v})$, real–analytic on $\mathbb{T}^d_{\hat{s}}$, $\mathtt{A}_*$–close to (u,v) and solving

$$\begin{cases} \omega + D_\omega \tilde{u} - H_y(\tilde{v}, \mathrm{id}+\tilde{u}) &= 0 \\ D_\omega \tilde{v} + H_x(\tilde{v}, \mathrm{id}+\tilde{u}) &= 0 \end{cases} \tag{2.1.19}$$

Futhermore, $\langle \tilde{u}\rangle = \langle u\rangle$ and the solution $(\tilde{u},\tilde{v})$ is uniquely determined in the $\mathtt{A}_*$–neighborhood of (u,v) by the condition $\langle \tilde{u}\rangle = \langle u\rangle$.

For a proof, see [CC97].

Remark 2.1.7 Notice that instead of the bound $\widetilde{V}$ on v_θ used in [CC97] to define the parameters, here we use $\widehat{V}$. The point is that, with this change, one is then allowed to chose $\widetilde{V}=0$ when v is constant.

2.2 Some preliminary facts

As we are going to use the same idea as in [Chi90] to extend maps obtained through the *KAM step* in the proof of *Theorem* 2.1.4, we will need a *cut–off* function.

Lemma 2.2.1 (Cut–Off) *Then, for any $n\in\mathbb{N}$, there exists a constant $\mathcal{C}_n>0$ such that for any given $R>0$ and a non–empty $\Delta\subset\mathbb{R}^d$, there exists $\chi\in C(\mathbb{C}^d)\cap C^\infty(\mathbb{R}^d)$ with $0\leqslant\chi\leqslant 1$, supp$\chi\subset\Delta_R := \bigcup_{\omega\in\Delta} D_R(\omega)$, $\chi\equiv 1$ on $\Delta_{\frac{R}{2}}$ and for any $m\in\mathbb{N}^d$ with $|m|_1\leqslant n$,*

$$\|\partial_\omega^m \chi\|_0 \stackrel{def}{=} \sup_{\mathbb{R}^d}\|\partial_\omega^m\chi\| \leqslant \mathcal{C}_n \frac{(|m|_1+2)!}{R^{|m|_1}}\,. \tag{2.2.1}$$

Proof Let $a,b>0$ such that $0<a\leqslant\frac{1}{4}$ and $\frac{1}{2}+a\leqslant b\leqslant 1-a$. Consider

$$\chi_1\colon t\in\mathbb{R}\mapsto\chi_1(t)=\begin{cases} e^{-\frac{1}{1-t^2}} & \text{if} \quad |t|<1 \\ 0 & \text{if} \quad |t|\geqslant 1 \end{cases}, \quad \chi_d(\omega) := R^{-d}N_a\prod_{j=1}^d \chi_1\left(\frac{\omega_j}{aR}\right), \text{ with}$$

$$N_a := \left(a \int_{\mathbb{R}} \chi_1(t)dt \right)^{-d}.$$

Define[21]

$$\chi(\omega) := \int_{\mathbb{R}^d} \chi_{\Delta_{bR}}(y)\chi_d(\omega - y)dy.$$

Thus, $\chi \in C(\mathbb{C}^d) \cap C^\infty(\mathbb{R}^d)$ and

- $supp\chi \subset \Delta_R$. Indeed, for any $\omega \in \mathbb{R}^d$, $\chi(\omega) > 0$ implies that $\chi_{\Delta_{bR}}(y) > 0$ and $\chi_d(\omega - y) > 0$ for *a.e.* $y \in \mathbb{R}^d$ (with respect to the Lebesgue measure on $\mathbb{R}^d$); which implies in particular that there exist $y_0 \in \mathbb{R}^d$ and $\omega_* \in \Delta$ such that $|y_0 - \omega_*| < bR$ and $|\omega - y_0| < aR$, so that $|\omega - \omega_*| < (a+b)R \leqslant R$ *i.e.* $supp\chi \subset \Delta_R$.

- $\chi \equiv 1$ on $\Delta_{\frac{R}{2}}$. Indeed, let $\omega \in \Delta_{\frac{R}{2}}$ *i.e.* $|\omega - \omega_*| < \frac{R}{2}$, for some $\omega_* \in \Delta$. Then, for any $y \in \mathbb{R}^d$,

$$|\omega - y| < aR \implies |y - \omega_*| \leqslant |y - \omega| + |\omega - \omega_*| < (a + \frac{1}{2})R \leqslant bR \implies y \in \Delta_{bR}.$$

  Hence,

$$1 \geqslant \chi(\omega) \geqslant \int_{B^d_{aR}(\omega)} \chi_d(\omega - y)dy = N_a \left(a \int_{\mathbb{R}} \chi_1(t)dt \right)^d = 1.$$

Moreover, for any $\omega \in \Delta_R$ and for any $n \in \mathbb{N}$, we have[22]

$$\begin{aligned}
\partial^n_{\omega_1}\chi(\omega) &= \int_{\mathbb{R}^d} \chi_{\Delta_{bR}}(y)\partial^n_{\omega_1}\chi_d(\omega - y)dy = \int_{B^d_{aR}(\omega)} \chi_{\Delta_{bR}}(y)\partial^n_{\omega_1}\chi_d(\omega - y)dy \\
&= \int_{B^d_{aR}(\omega)} \partial^n_{\omega_1}\chi_d(\omega - y)dy = R^{-d}N_a \int_{\mathbb{R}^d} \partial^n_{\omega_1}\chi_1 \left(\frac{\omega_1 - y_1}{aR}\right) \prod_{j=2}^d \chi_1 \left(\frac{\omega_j - y_j}{aR}\right) dy \\
&= R^{-d}N_a \left((aR)^{-n+1} \int_{\mathbb{R}} \frac{d^n\chi_1(t)}{dt^n}dt \right) \left(aR \int_{\mathbb{R}} \chi_1(t)dt \right)^{d-1} \\
&= \left((aR)^{-n} \int_{\mathbb{R}} \frac{d^n\chi_1(t)}{dt^n}dt \right) \left(\int_{\mathbb{R}} \chi_1(t)dt \right)^{-1} \left(aR \int_{\mathbb{R}} \chi_1(t)dt \right)^d R^{-d}N_a \\
&= (aR)^{-n} \left(\int_{\mathbb{R}} \chi_1(t)dt \right)^{-1} \int_{\mathbb{R}} \frac{d^n\chi_1}{dt^n}(t)dt.
\end{aligned}$$

---

[21] $\chi_{\Delta_{bR}}$ denotes the characteristic function of the set Δ_{bR}, *i.e.* $\chi_{\Delta_{bR}} \equiv 1$ on Δ_{bR} and $\chi_{\Delta_{bR}} \equiv 0$ on $\mathbb{R}^d \backslash \Delta_{bR}$.

[22] In fact, one checks easily that for any $m \in \mathbb{N}^d$, $\|\partial^m_\omega \chi\|_0 \leqslant \|\partial^{|m|_1}_{\omega_1}\chi\|_0 = \cdots = \|\partial^{|m|_1}_{\omega_d}\chi\|_0$.

One checks easily that for any $t \in \mathbb{R}$ and $n \in \mathbb{N}$,

$$\frac{d^n \chi_1}{dt^n}(t) = \frac{P_n(t)}{(1-t^2)^{2n}} \chi_1(t),$$

with $\quad P_0 := 1, \quad P_1(t) := -2t \quad$ and for any $n \geqslant 1$

$$P_{n+1} := (-2 + 4n(1-t^2))tP_n(t) + (1-t^2)^2 \frac{dP_n}{dt}(t),$$

and one has, for any $n \geqslant 1$,

$$\deg(P_n) = 3n - 2.$$

The existence of the sequence $(\mathcal{C}_n)_n$ then follows easily and in particular, by choosing $a = \frac{1}{4}$ and, then, $b = \frac{3}{4}$, we can take[23]

$$\mathcal{C}_1 := \frac{2\,e^2}{3!a} \int_{\mathbb{R}} \left| \frac{d\chi_1}{dt} \right| (t)dt = \frac{4\,e}{3} \tag{2.2.2}$$

and

$$\begin{aligned} \frac{2\,e^2}{4!a^2} \int_{\mathbb{R}} \left| \frac{d^2\chi_1}{dt^2} \right| (t)dt &\leqslant \frac{4\,e^2}{3} \cdot 8 \int_0^1 \frac{1}{(1-t)^4}\, e^{-\frac{1}{2(1-t)}} dt \\ &= \frac{32\,e^2}{3} \int_1^{\infty} s^4\, e^{-\frac{s}{2}} ds = \frac{832\,e^{\frac{3}{2}}}{3} =: \mathcal{C}_2. \end{aligned}$$

∎

The following lemma establishes a bound on the Lipschitz constant of a map obtained as the composition of infinetly many Lipschitz maps and will be used to prove the Lipschitz continuity of the map in *Theorem* 2.1.4, obtained through the infinite iterative KAM scheme.

Lemma 2.2.2 *Let $(\mathcal{X}, \|\cdot\|)$ be a real or complex normed vector space, $\mathcal{L}_j \colon \mathcal{X} \to \mathcal{X}$ be a sequence of invertible linear operators and $(g_j)_{j \geqslant 0}$ be a sequence of Lipschitz continuous maps from $\mathcal{X}$ to itself. Define*

$$l_j := \|\mathcal{L}_j(g_j - \mathrm{id})\mathcal{L}_0^{-1}\|_{L,\mathcal{X}}, \quad G_0 := \mathrm{id}, \quad G_{j+1} := g_0 \circ \cdots \circ g_j,$$

$$L_j := \|\mathcal{L}_0(G_j - \mathrm{id})\mathcal{L}_0^{-1}\|_{L,\mathcal{X}} \quad \textit{and assume that}$$

[23] We have $\frac{1}{2\,e^2} \leqslant \int_{\mathbb{R}} \chi_1(t)dt \leqslant \frac{2\sqrt{6}}{7}$.

$$\delta := \sup_{j\geqslant 0} \|\mathcal{L}_j \mathcal{L}_{j+1}^{-1}\| < \infty \quad \text{and} \quad \sum_{j=0}^{\infty} \delta^j l_k \leqslant \frac{1}{2}. \tag{2.2.3}$$

Then $(G_j)_{j\geqslant 0}$ *is a sequence of Lipschitz continuous maps and for any* $j \geqslant 0$*,*

$$\|\mathcal{L}_0 (G_{j+1} - \mathrm{id}) \mathcal{L}_0^{-1}\|_{L,\mathcal{X}} \leqslant 2 \sum_{k=0}^{j} \delta^k l_k. \tag{2.2.4}$$

Proof For any $j \geqslant 0$, we have

$$\begin{aligned} L_{j+1} &= \Big\|\mathcal{L}_0 (G_j - \mathrm{id}) \mathcal{L}_0^{-1} \left\{ \left(\mathcal{L}_0 \mathcal{L}_1^{-1}\right) \cdots \left(\mathcal{L}_{j-1} \mathcal{L}_j^{-1}\right) \mathcal{L}_j (g_j - \mathrm{id}) \mathcal{L}_0^{-1} + \mathrm{id} \right\} + \\ &\quad + \left(\mathcal{L}_0 \mathcal{L}_1^{-1}\right) \cdots \left(\mathcal{L}_{j-1} \mathcal{L}_j^{-1}\right) \mathcal{L}_j (g_j - \mathrm{id}) \mathcal{L}_0^{-1} \Big\|_{L,\mathcal{X}} \\ &\leqslant (1 + \delta^j l_j) L_j + \delta^j l_j. \end{aligned}$$

Hence, we get, inductively, that for any $j \geqslant 0$, $L_{j+1} \leqslant L_j + 2\delta^j l_j \leqslant 2 \sum_{k=0}^{j} \delta^k l_k$. ∎

We recall the very famous *Cauchy estimate* used to control the derivatives of an analytic function.

Lemma 2.2.3 (Cauchy's estimate) *Let* $p \in \mathbb{N}$, $f \in \mathcal{A}_{r,s,h,d}$*. Then, for any multi–index* $(l,k) \in \mathbb{N}^d \times \mathbb{N}^d$ *with* $|l|_1 + |k|_1 \leqslant p$ *and for any* $0 < r' < r$, $0 < s' < s$,[24]

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

In the next lemma, we recall some properties of the Fourier's coefficients of an analytic function.

Lemma 2.2.4 *Let* $\kappa \in \mathbb{N}$, $f \in \mathcal{A}_{r,s,h,d}$, $0 < \sigma < s$ *with* $\kappa > \frac{d-1}{\sigma}$*. Then*

$$\begin{aligned} &(i) \qquad |f_k(y,\omega)| \leqslant e^{-|k|_1 s} \|f\|_{r,s,h,d}\,, \quad \forall\, k \in \mathbb{Z}^d,\, y \in D_r(0),\, \omega \in \Omega_{\alpha,h}, \\ &(ii) \qquad \|f - T_\kappa f\|_{r,s-\sigma,h,d} \leqslant 4^d C_2 \kappa^d\, e^{-\kappa\sigma} \|f\|_{r,s,h,d}. \end{aligned}$$

Proof

(*i*) Let $k \in \mathbb{Z}^d$, $y \in D_r(0)$, $\omega \in \Omega_{\alpha,h}$. Then

$$f_k(y,\omega) = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} f(y,x,\omega)\, e^{-ik\cdot x}\, dx.$$

---

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

But, for any given $\beta \in \mathbb{R}^d$ such that $|\beta| < s$, by periodicity of v in each argument and Cauchy's theorem, we get

$$f_k(y,\omega) = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} f(y, x - i\beta, \omega)\, e^{-ik\cdot(x-i\beta)}\, dx.$$

Now, we choose $\beta := (s-\sigma)\,(\operatorname{sign}(k_1), \cdots, \operatorname{sign}(k_d))$ with $0 < \sigma < s$. Thus, we get

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

and letting $\sigma \to 0^+$, we get the desired inequality.

(ii) We have

$$\begin{aligned}
\|f - T_\kappa f\|_{r,s-\sigma,h,d} &\leqslant \|f\|_{r,s,h,d} \sum_{|k|_1 > \kappa} e^{-|k|_1\sigma} \\
&= \|f\|_{r,s,h,d} \sum_{l=\kappa+1}^{\infty} \sum_{\substack{k\in\mathbb{Z}^d \\ |k|_1 = l}} e^{-|k|_1\sigma} \\
&\leqslant \|f\|_{r,s,h,d} \sum_{l=\kappa+1}^{\infty} 4^d l^{d-1}\, e^{-l\sigma} \\
&\leqslant \|f\|_{r,s,h,d} \int_{\kappa}^{\infty} 4^d t^{d-1}\, e^{-t\sigma}\, dt \\
&\leqslant 4^d \kappa^d\, e^{-\kappa\sigma} \|f\|_{r,s,h,d} \int_0^{\infty} (t+1)^{d-1}\, e^{-t(d-1)}\, dt \\
&= 4^d C_2 \kappa^d\, e^{-\kappa\sigma} \|f\|_{r,s,h,d}.
\end{aligned}$$

∎

In the following lemma, we recall some facts about the homological equation.

Lemma 2.2.5 ([CC95]) *Let $p \in \mathbb{N}$, $\omega \in \Delta_\alpha^\tau$ and $f \in \mathcal{A}^0_{r,s,h,d}$. Then, for any $0 < \sigma < s$, the equation*

$$D_\omega g = f$$

has a unique solution in $\mathcal{A}^0_{r,s-\sigma,h,d}$ and there exist constants $\bar{B}_p = \bar{B}_p(d,\tau) \geqslant 1$ and $k_p = k_p(d,\tau) \geqslant 1$ such that for any multi–index $k \in \mathbb{N}^d$ with $|k|_1 \leqslant p$

$$\|\partial_\omega^k g\|_{r,s-\sigma,h,d} \leqslant \bar{B}_p \frac{\|f\|_{r,s,h,d}}{\alpha} \sigma^{-k_p}.$$

In particular, one can take $\bar{B}_1 = C_0$ (see [Rüs75, CC95]).

Now, we recall the classical implicit function theorem, in a quantitative framework.

Lemma 2.2.6 (Implicit Function Theorem I[Chi12]) *Let $r, s > 0$, $n, m \in \mathbb{N}$, $(y_0, x_0) \in \mathbb{C}^n \times \mathbb{C}^m$ and* [25]

$$F\colon (y, x) \in D^n_r(y_0) \times D^m_s(x_0) \subset \mathbb{C}^{n+m} \mapsto F(y, x) \in \mathbb{C}^n$$

be continuous with continuous Jacobian matrix F_y. Assume that $F_y(y_0, x_0)$ is invertible with inverse $T := F_y(y_0, x_0)^{-1}$ such that

$$\sup_{D^n_r(y_0)\times D^m_s(x_0)} \|\mathbb{1}_n - TF_y(y, x)\| \leqslant c < 1 \quad \text{and} \quad \sup_{D^m_s(x_0)} |F(y_0, \cdot)| \leqslant \frac{(1-c)r}{\|T\|}. \tag{2.2.5}$$

Then, there exists a unique continuous function $g\colon D^m_s(x_0) \to D^n_r(y_0)$ such that the following are equivalent

(i) $(y, x) \in D^n_r(y_0) \times D^m_s(x_0)$ and $F(y, x) = 0$;

(ii) $x \in D^m_s(x_0)$ and $y = g(x)$.

Moreover, g satisfies

$$\sup_{D^m_s(x_0)} |g - y_0| \leqslant \frac{\|T\|}{1-c} \sup_{D^m_s(x_0)} |F(y_0, \cdot)|. \tag{2.2.6}$$

Finally, we recall some inversion function theorems.
Taking $n = m$, $c = c' = \frac{1}{2}$ and $F(y, x) = f(y) - x$ in Lemma 2.2.6, for a given $f \in C^1(D^n_r(y_0), \mathbb{C}^n)$, then the following holds.

Lemma 2.2.7 (Inversion Function Theorem I) *Let $f\colon D^n_r(y_0) \to \mathbb{C}^n$ be a C^1 function with invertible Jacobian $f_y(y_0)$ and assume that*

$$\sup_{D^n_r(y_0)} \|\mathbb{1}_n - Tf_y\| \leqslant \frac{1}{2}, \quad T := f_y(y_0)^{-1}.$$

Then, there exists a unique C^1 function

$$g\colon D^n_s(x_0) \to D^n_r(y_0), \quad x_0 := f(y_0), \quad s := \frac{r}{2\|T\|},$$

such that

$$f \circ g(x) = x, \quad g \circ f(y) = y, \quad \forall\, x \in D^n_s(x_0), \quad \forall\, y \in g\left(D^n_s(x_0)\right).$$

Moreover,

$$\sup_{D^n_s(x_0)} \|g_x\| \leqslant 2\|T\|\,. \tag{2.2.7}$$

[25] Let us point out that any other norm (different!) may be used on $\mathbb{C}^n$, $\mathbb{C}^m$ and $\mathbb{C}^{n+m}$.

Remark 2.2.8 (*i*) Notice that in Lemmata 2.2.6 and 2.2.7 if, in addition, F is periodic in x (resp. analytic, real on reals) then so is g.

(*ii*) Notice that Lemmata 2.2.6 and 2.2.7 still hold if, everywhere therein, open balls are replaced by closed balls or complex–balls by real–balls.

Another consequence of the Implicit Function Theorem is the following version of Inversion Function Theorem.

Lemma 2.2.9 (Inversion Function Theorem II [Pös01]) *Assume that f is a real analytic function from $\Omega_{\alpha,h}$ into $\mathbb{C}^d$ such that*

$$\|f - \mathrm{id}\| \leqslant \delta \leqslant \frac{h}{4}.$$

Then f has a real analytic inverse g defined on $\Omega_{\alpha,\frac{h}{4}}$ and on which it satisfies

$$\|g - \mathrm{id}\|, \quad \frac{h}{4}\|Dg - \mathrm{Id}\| \leqslant \delta.$$

Remark 2.2.10 Notice that, all the Lemmata above are valid if one replace $(\mathcal{A}_{r,s,h,d}, \|\cdot\|_{r,s,h,d})$ by $(\mathcal{B}_{r,s,\rho_0}, \|\cdot\|_{r,s,\rho_0})$ or $(\mathcal{B}_{r,s}(y_0), \|\cdot\|_{r,s,y_0})$.

2.3 Proofs

2.3.1 Proof of Theorem 2.1.1

The proof is essentially the one in [Chi08] though one needs to re–scale various quantities; therefore we shall skip some details. First of all, notice that $K, P \in \mathcal{B}_{s,\varepsilon_0}$. For simplicity, sometimes, the explicit dependence on r or ε_0, ε_* will not be denoted in the norm $\|\cdot\|_{r,s,\varepsilon_0}$ or in the $\mathcal{B}$–spaces, etc, as r, ε_0 and ε_* will not be changed during the iteration. We begin by describing completely one step of the scheme, namely the KAM step, which will be then iterated infinitely many time to compute the symplectic change of variable.

KAM step Kolmogorov's idea is to construct a near–to–the–identity symplectic transformation ϕ_1, such that

$$H_1 := H \circ \phi_1 = K_1 + \varepsilon^2 P_1 \ , \quad K_1 = K_1^* + \omega \cdot y' + Q_1(y', x') \ , Q_1 = O(|y'|^2) \ ; \qquad (2.3.1)$$

if this is achieved, the Hamiltonian K_1 has the same basic properties of K (the linear part in y is the same and, being ϕ_1 close to the identity, K_1 is non–degenerate) and the procedure can be iterated.

For, Kolmogorov considers the generating function of ϕ_1 of the form[26]

$$g(y',x) := y'\cdot x + \varepsilon\big(b\cdot x + \mathrm{s}(x) + y'\cdot a(x)\big)\ , \tag{2.3.2}$$

where, s and a are (respectively, scalar and vector–valued, ε–dependent) real–analytic functions on $\mathbb{T}^d$ with zero average and $b\in\mathbb{R}^d$. Define[27]

$$u_0 = u_0(x) := b + \mathrm{s}_x\ , \qquad A = A(x) := a_x \quad \text{and} \quad u = u(y',x) := u_0 + A\,y'\ .$$

Then ϕ_1 is implicitely defined by

$$\begin{cases} y = y' + \varepsilon u(y',x) := y' + \varepsilon(u_0(x) + A(x)y') \\[2ex] x' = x + \varepsilon a(x)\ . \end{cases}$$

Moreover, for ε small, $x\in\mathbb{T}^d \to x+\varepsilon a(x)\in\mathbb{T}^d$ defines a diffeomorphism of $\mathbb{T}^d$ with inverse

$$x = \varphi(x') := x' + \varepsilon\widetilde{\varphi}(x';\varepsilon)\ ,$$

for a suitable real–analytic function $\widetilde{\varphi}$. Thus ϕ_1 is explicitly given by

$$\phi_1 : (y',x') \to \begin{cases} y = y' + \varepsilon\ u\big(y',\varphi(x')\big) \\[2ex] x = \varphi(x')\ . \end{cases} \tag{2.3.3}$$

To determine b, s and a, observe that by Taylor's formula

$$H(y'+\varepsilon\ u,x) = \mathrm{K} + \omega\cdot y' + Q(y',x) + \varepsilon\Big[\omega\cdot u + Q_y(y',x)\cdot u + P(y',x)\Big] + \varepsilon^2 P'(y',x) \tag{2.3.4}$$

where $P' := P'(y',x;\varepsilon) := P^{(1)} + P^{(2)}$ with

$$\begin{cases} P^{(1)} := \frac{1}{\varepsilon^2}[Q(y'+\varepsilon u,x) - Q(y',x) - \varepsilon Q_y(y',x)\cdot u] = \displaystyle\int_0^1 (1-t)Q_{yy}(y'+t\varepsilon u,x)\,u\cdot u\,dt \\ P^{(2)} := \frac{1}{\varepsilon}[P(y'+\varepsilon u,x) - P(y',x)] = \displaystyle\int_0^1 P_y(y'+t\varepsilon u,x)\cdot u\,dt\ . \end{cases} \tag{2.3.5}$$

[26] Compare [DS01, AKN06] for generalities on symplectic transformations and their generating functions. For simplicity, we do not report in the notation the dependence of various functions on ε, but, in fact, $P = P(y,x;\varepsilon)$, $\mathrm{s} = \mathrm{s}(x;\varepsilon)$, $a = a(x;\varepsilon)$, etc.

[27] As usual, we denote $\mathrm{s}_x = \partial_x \mathrm{s} = (\mathrm{s}_{x_1},\dots,\mathrm{s}_{x_d})$ and a_x denotes the matrix $(a_x)_{ij} := \frac{\partial a_j}{\partial x_i}$; as above, we often do not report in the notation the dependence upon ε (but u_0, A and u do depend also on ε).

Note that

$$Q_y(y',x)\cdot(a_x y') =: Q^{(1)}(y',x) = O(|y'|^2)\ , \tag{2.3.6}$$

and that (again by Taylor's formula)

$$\begin{cases} Q_y(y',x)\cdot u_0 = Q_{yy}(0,x)y'\cdot u_0 + Q^{(2)}(y',x)\ , \quad Q^{(2)} := \displaystyle\int_0^1 (1-t)Q_{yyy}(ty',x)y'\cdot y'\cdot u_0\, dt \\ \\ P(y',x) = P(0,x) + P_y(0,x)\cdot y' + Q^{(3)}(y',x)\ , \quad Q^{(3)} := \displaystyle\int_0^1 (1-t)P_{yy}(ty',x)\, y'\cdot y'\, dt\ . \end{cases} \tag{2.3.7}$$

Thus, since[28] $\omega\cdot u = \omega\cdot b + D_\omega \mathrm{s} + D_\omega a\cdot y'$, we find

$$H(y'+\varepsilon u,x) = \mathrm{K} + \omega\cdot y' + Q(y',x) + \varepsilon Q'(y',x) + \varepsilon F'(y',x) + \varepsilon^2 P'(y',x) \tag{2.3.8}$$

with P' as in (2.3.4)–(2.3.5) and

$$\begin{cases} Q'(y',x) := Q^{(1)} + Q^{(2)} + Q^{(3)} = O(|y'|^2) \\ \\ F'(y',x) := \omega\cdot b + D_\omega \mathrm{s} + P(0,x) + \Big\{D_\omega a + Q_{yy}(0,x)b + Q_{yy}(0,x)\mathrm{s}_x + P_y(0,y')\Big\}\cdot y' \end{cases} \tag{2.3.9}$$

By Lemma 2.2.5, *there exist a unique constant b and unique functions s and a* (with zero average) *such that F' is constant.* In fact, if

$$\begin{cases} \mathrm{s} := -D_\omega^{-1}\Big(P(0,x) - P_0(0)\Big) = -\sum_{\substack{n\in\mathbb{Z}^d \\ n\neq 0}} \dfrac{P_n(0)}{i\omega\cdot n}\, e^{in\cdot x} \\ b := -\langle Q_{yy}(0,\cdot)\rangle^{-1}\Big(\langle Q_{yy}(0,\cdot)\mathrm{s}_x\rangle + \langle P_y(0,\cdot)\rangle\Big) \\ a := -D_\omega^{-1}\Big(Q_{yy}(0,x)(b+\mathrm{s}_x) + P_y(0,x)\Big) \end{cases}$$

then $F' = \omega\cdot b + P_0(0)$. Thus, with this determination of g in (2.3.2), recalling (2.3.3), we find that (2.3.1) holds with

$$\begin{cases} \mathrm{K}_1 := \mathrm{K} + \varepsilon\widetilde{\mathrm{K}}\ , & \widetilde{\mathrm{K}} := \omega\cdot b + P_0(0) \\ Q_1(y',x') := Q(y',x') + \varepsilon\widetilde{Q}(y',x')\ , & \widetilde{Q} := \displaystyle\int_0^1 Q_x(y',x'+t\varepsilon\alpha(x'))\cdot\alpha\, dt\ + Q'(y',\varphi(x')) \\ P_1(y',x') := P'(y',\varphi(x'))\ . & \end{cases}$$

Clearly, for ε small enough $\big\langle \partial_{y'}^2 Q_1(0,\cdot)\big\rangle$ is invertible and, if $T := \langle Q_{yy}(0,\cdot)\rangle^{-1}$, we may write

$$T_1 := \big\langle \partial_{y'}^2 Q_1(0,\cdot)\big\rangle^{-1} =: T + \varepsilon\widetilde{T}\ . \tag{2.3.10}$$

[28] Recall that $\omega\cdot \mathrm{s}_x = D_\omega \mathrm{s}$ and $\omega\cdot(a_x y') = (D_\omega a)\cdot y'$.

Next, we provide the KAM step with carefull estimates. Actually, we shall do the estimates in term of a lower bound of $|\omega|$ instead of $|\omega|$, so that, by taking this lower bound equal to $|\omega|$, we shall get the estimates in Theorem 2.1.1[29]. Thus, we fix, for the remainder of the proof,

$$0 < \underline{\omega} \leqslant |\omega| \ . \tag{2.3.11}$$

Recall the definition in §2.1.1; in particular[30]

$$r|\omega| \ , \ \|Q\|_{s,\varepsilon_0}, |\omega|^2\|T\| < \mathrm{E} \ . \tag{2.3.12}$$

Finally, fix[31]

$$0 < \sigma < \frac{s}{2} \quad \text{and define} \quad \bar{s} := s - \frac{2}{3}\sigma \ , \quad s' := s - \sigma \ .$$

Lemma 2.3.1 *Then*[32]

$$\underline{\omega}\|\mathrm{s}_x\|_{\bar{s}}, \underline{\omega}|b|, |\widetilde{\mathrm{K}}|, r\sigma\underline{\omega}\|a\|_{\bar{s}}, r\underline{\omega}\|a_x\|_{\bar{s}}, \underline{\omega}\|u_0\|_{\bar{s}}, \underline{\omega}\|u\|_{\bar{s}}, \|Q'\|_{\bar{s}}, r^2\sigma^2\|\partial^2_{y'}Q'(0,\cdot)\|_0 \leqslant \bar{L} \tag{2.3.13}$$

Furthermore, if $\varepsilon_* \leqslant \varepsilon_0$ satisfies

$$\varepsilon_* r^{-1}\sigma^{-1}\underline{\omega}^{-1}\bar{L} \leqslant \frac{\sigma}{3} \ , \tag{2.3.14}$$

then

$$\|P'\|_{\bar{s}} \leqslant r^{-1}\sigma^{-1}\underline{\omega}^{-1}\bar{L}M \ . \tag{2.3.15}$$

and the following hold. For $|\varepsilon| < \varepsilon_*$, the map $\psi_\varepsilon(x) := x + \varepsilon a(x)$ has an analytic inverse $\varphi(x') = x' + \varepsilon\widetilde{\varphi}(x';\varepsilon)$ such that, for all $|\varepsilon| < \varepsilon_*$,

$$\|\widetilde{\varphi}\|_{s',\varepsilon_*} \leqslant r^{-1}\sigma^{-1}\underline{\omega}^{-1}\bar{L} \qquad \text{and} \ , \forall \ |\varepsilon| < \varepsilon_* \ , \quad \varphi = \mathrm{id} + \varepsilon\widetilde{\varphi} : \mathbb{T}^d_{s'} \to \mathbb{T}^d_{\bar{s}} \ ; \tag{2.3.16}$$

for any $(y', x, \varepsilon) \in W_{\bar{s},\varepsilon_*}$, $|y' + \varepsilon u(y', x)| < rs$; the map ϕ_1 is a symplectic diffeomorphism and

$$\phi_1 = \big(y' + \varepsilon u(y', \varphi(x')), \varphi(x')\big) : W_{s',\varepsilon_*} \to D_{rs} \times \mathbb{T}^d_s, \quad \text{and} \quad \|\mathcal{W}\,\tilde{\phi}\|_{s',\varepsilon_*} \leqslant \bar{L} \ , \tag{2.3.17}$$

---

[29] The point is that, if we replace $|\omega|$ by $\underline{\omega}$ everywhere in §2.1.1, except in the expressions of E and $\widehat{\mathrm{E}}$, then, Theorem 2.1.1 holds for any ω such that $|\omega| \geqslant \underline{\omega}$.

[30] The notation in Eq. (2.3.12) means that each term on the l.h.s. is bounded by the r.h.s.

[31] The parameter s' will be the size of the domain of analyticity of the new symplectic variables (y', x'), domain on which we shall bound the Hamiltonian $H_1 = H \circ \phi_1$, while $\bar{s}$ is an intermediate domain where we shall bound various functions of y' and x. Note that $\sigma < \frac{1}{2}$.

[32] Here $\|\cdot\|_{\bar{s}} = \|\cdot\|_{\bar{s},\varepsilon_0}$.

where $\tilde\phi$ *is defined by the relation* $\phi_1 =: \mathrm{id} + \varepsilon\tilde\phi$.
Finally, if [33]

$$\varepsilon_* \frac{L}{\mathrm{E}} \leqslant \frac{\sigma}{3} \ , \tag{2.3.18}$$

then

$$\begin{cases} |\widetilde{\mathrm{K}}| \ , \ \|\widetilde{Q}\|_{s',\varepsilon_*} \ , \ |\omega|^2\|\widetilde{T}\| \ , \ \|\mathcal{W}\,\tilde\phi\|_{s',\varepsilon_*} \leqslant L \\ \\ \|P_1\|_{s',\varepsilon_*} \leqslant \frac{LM}{\mathrm{E}} \ . \end{cases} \tag{2.3.19}$$

Proof We begin by estimating $\|\mathrm{s}_x\|_{\bar s}$. Actually these estimate will be given on a larger intermediate domain, namely, $W_{s-\frac{\sigma}{3},\varepsilon_0}$, allowing to give the remaining bounds on the smaller domain $W_{\bar s,\varepsilon_0}$. Let $f(x) := P(0,x) - \langle P(0,\cdot)\rangle$. By definition of $\|\cdot\|_s$ and M, it follows that $\|f\|_s \leqslant \|P(0,x)\|_s + \|\langle P(0,\cdot)\rangle\| \leqslant 2M$. By Lemma 2.2.5 with $p=1$ and $s' = s - \frac{\sigma}{3}$, one gets

$$\|\mathrm{s}_x\|_{s-\frac{\sigma}{3}} \leqslant C_0\frac{2M}{\alpha}\,3^\tau\sigma^{-\tau} = \mathrm{C}_1\sigma^{-\tau}\alpha^{-1}M \leqslant \bar L,$$

so that

$$\underline{\omega}\|\mathrm{s}_x\| \leqslant 2^{-(3\tau+10)}\mathrm{C}_1\mathrm{E}^7\sigma^{-\bar\nu}r^{-7}\alpha^{-4}\underline{\omega}^{-3}M \leqslant \bar L$$

Next, we estimate b. By definitions and Lemma 2.2.3, we have[34]

$$\begin{aligned} |b| &\leqslant \|T\|\left(\max_{1\leqslant l\leqslant d}\sum_{j=1}^d \|Q_{y_ly_j}\|_{s-\sigma}\|\mathrm{s}_{x_j}\|_{s-\frac{\sigma}{3}} + \max_{1\leqslant j\leqslant d}\|P_{y_j}\|_{s-\sigma}\right) \\ &\leqslant \underline{\omega}^{-2}\mathrm{E}\left(2d\mathrm{E}r^{-2}\sigma^{-2}\mathrm{C}_1\sigma^{-\tau}\alpha^{-1}M + Mr^{-1}\sigma^{-1}\right) \\ &\leqslant \left(2d\mathrm{C}_1\mathrm{E} + \sigma^{\tau+1}r\alpha\right)\underline{\omega}^{-2}\mathrm{E}\sigma^{-(\tau+2)}r^{-2}\alpha^{-1}M \\ &\leqslant \left(2d\mathrm{C}_1 + 2^{-(\tau+1)}\right)\underline{\omega}^{-2}\mathrm{E}^2\sigma^{-(\tau+2)}r^{-2}\alpha^{-1}M \\ &= \mathrm{C}_2\mathrm{E}^2\sigma^{-(\tau+2)}r^{-2}\alpha^{-1}\underline{\omega}^{-2}M, \end{aligned}$$

so that

$$\underline{\omega}|b| \leqslant 2^{-(3\tau+8)}\mathrm{C}_2\mathrm{E}^7\sigma^{-\bar\nu}r^{-7}\alpha^{-4}\underline{\omega}^{-3}M \leqslant \bar L$$

---

[33] Notice that $L \geqslant (r^{-1}|\omega|^{-1}\mathrm{E})^2\tilde L \geqslant \tilde L \geqslant r^{-1}\sigma^{-1}|\omega|^{-1}\bar L C \geqslant \bar L$ since $r|\omega| \leqslant \mathrm{E}$, so that (2.3.18) implies (2.3.14).

[34] Remember that $\sigma < 1/2$, $\|T\| \leqslant |\omega|^{-2}\mathrm{E} \leqslant \underline{\omega}^{-2}\mathrm{E}$ and $r\alpha \leqslant r|\omega| \leqslant \mathrm{E}$.

Next, we estimate $\widetilde{\mathrm{K}}$. We have

$$\begin{aligned}
|\widetilde{\mathrm{K}}| &\leqslant d|\omega|\cdot|b| + \|P\|_s \\
&\leqslant d\underline{\omega}\mathrm{C}_2\mathrm{E}^2\sigma^{-(\tau+2)}r^{-2}\alpha^{-1}\underline{\omega}^{-2}M + M \\
&\leqslant (d\underline{\omega}^{-1}\mathrm{C}_2\mathrm{E}^2 + \sigma^{\tau+2}r^2\alpha)\sigma^{-(\tau+2)}r^{-2}\alpha^{-1}M \\
&\leqslant (d\underline{\omega}^{-1}\mathrm{C}_2\mathrm{E}^2 + 2^{-(\tau+2)}\underline{\omega}^{-1}r^2|\omega|^2)\sigma^{-(\tau+2)}r^{-2}\alpha^{-1}M \\
&= \mathrm{C}_3\mathrm{E}^2\sigma^{-(\tau+2)}r^{-2}\alpha^{-1}\underline{\omega}^{-1}M \\
&\leqslant 2^{-(3\tau+8)}\mathrm{C}_3\mathrm{E}^7\sigma^{-\bar{\nu}}r^{-7}\alpha^{-4}\underline{\omega}^{-2}M \leqslant \bar{L}.
\end{aligned}$$

Next, we estimate u_0. We have

$$\begin{aligned}
\|u_0\|_{\bar{s}} &\leqslant |b| + \|s_x\|_{s-\frac{\sigma}{3}} \leqslant \mathrm{C}_2\mathrm{E}^2\sigma^{-(\tau+2)}r^{-2}\alpha^{-1}\underline{\omega}^{-2}M + \mathrm{C}_1\sigma^{-\tau}\alpha^{-1}M \\
&\leqslant (\mathrm{C}_2\mathrm{E}^2\underline{\omega}^{-2} + \mathrm{C}_1 r^2\sigma^2)\sigma^{-(\tau+2)}r^{-2}\alpha^{-1}M \\
&\leqslant (\mathrm{C}_2\mathrm{E}^2\underline{\omega}^{-2} + 2^{-2}\mathrm{C}_1 r^2|\omega|^2\underline{\omega}^{-2})\sigma^{-(\tau+2)}r^{-2}\alpha^{-1}M \\
&\leqslant (\mathrm{C}_2 + 2^{-2}\mathrm{C}_1)\mathrm{E}^2\sigma^{-(\tau+2)}r^{-2}\alpha^{-1}\underline{\omega}^{-2}M \\
&= \mathrm{C}_4\mathrm{E}^2\sigma^{-(\tau+2)}r^{-2}\alpha^{-1}\underline{\omega}^{-2}M,
\end{aligned}$$

so that

$$\underline{\omega}\|u_0\|_{\bar{s}} \leqslant 2^{-(3\tau+8)}\mathrm{C}_4\mathrm{E}^7\sigma^{-\bar{\nu}}r^{-7}\alpha^{-4}\underline{\omega}^{-3}M \leqslant \bar{L}$$

Next, we estimate a and a_x. Let $f(x) := Q_{yy}(0,x)(b+\mathrm{s}_x)+P_y(0,x)$. Then, by Lemma 2.2.32.2.5, we have

$$\begin{aligned}
\|f\|_{s-\frac{\sigma}{3}} &\leqslant \max_{1\leqslant l\leqslant d}\sum_{j=1}^{d}\|Q_{y_l y_j}\|_{s-\sigma}\left(|b_j| + \|\mathrm{s}_{x_j}\|_{s-\frac{\sigma}{3}}\right) + \max_{1\leqslant j\leqslant d}\|P_{y_j}\|_{s-\sigma} \\
&\leqslant \max_{1\leqslant l\leqslant d}\sum_{j=1}^{d}\|Q_{y_l y_j}\|_{s-\sigma}\left(|b| + \|\mathrm{s}_x\|_{s-\frac{\sigma}{3}}\right) + \max_{1\leqslant j\leqslant d}\|P_{y_j}\|_{s-\sigma} \\
&\leqslant 2d\mathrm{E}\sigma^{-2}r^{-2}\cdot\mathrm{C}_4\mathrm{E}^2\sigma^{-(\tau+2)}r^{-2}\alpha^{-1}\underline{\omega}^{-2}M + \sigma^{-1}r^{-1}M \\
&\leqslant \left(2d\mathrm{C}_4\mathrm{E}^3\underline{\omega}^{-2} + \sigma^{\tau+3}r^3\alpha\right)\sigma^{-(\tau+4)}r^{-4}\alpha^{-1}M \\
&\leqslant \left(2d\mathrm{C}_4\mathrm{E}^3\underline{\omega}^{-2} + \sigma^{\tau+3}r^3|\omega|^3\underline{\omega}^{-2}\right)\sigma^{-(\tau+4)}r^{-4}\alpha^{-1}M \\
&\leqslant \left(2d\mathrm{C}_4 + 2^{-(\tau+3)}\right)\mathrm{E}^3\sigma^{-(\tau+4)}r^{-4}\alpha^{-1}\underline{\omega}^{-2}M
\end{aligned}$$

Thus, by Lemma 2.2.5, we obtain[35]

$$\begin{aligned}\|a\|_{\bar{s}}, \|a_x\|_{\bar{s}} &\leqslant dC_0\frac{\|f\|_{s-\frac{\sigma}{3}}}{\alpha}3^\tau\sigma^{-\tau}\\ &\leqslant 3^\tau dC_0\left(2d\mathrm{C}_4+2^{-(\tau+3)}\right)\mathrm{E}^3\sigma^{-(2\tau+4)}r^{-4}\alpha^{-2}\underline{\omega}^{-2}M\\ &= \mathrm{C}_5\mathrm{E}^3\sigma^{-(2\tau+4)}r^{-4}\alpha^{-2}\underline{\omega}^{-2}M,\end{aligned}$$

so that

$$r\sigma\underline{\omega}\|a\|_{\bar{s}}, r\underline{\omega}\|a_x\|_{\bar{s}} \leqslant 2^{-(2\tau+6)}\mathrm{C}_5\mathrm{E}^7\sigma^{-\bar{\nu}}r^{-7}\alpha^{-4}\underline{\omega}^{-3}M \leqslant \bar{L}$$

Next, we estimate u. We have

$$\begin{aligned}\|u\|_{\bar{s}} &\leqslant \|u_0\|_{\bar{s}}+\|Ay'\|_{\bar{s}} \leqslant \|u_0\|_{\bar{s}}+\max_{1\leqslant l\leqslant d}\sum_{j=1}^{d}\|A_{lj}\|_{\bar{s}}r\bar{s}\\ &\leqslant \mathrm{C}_4\mathrm{E}^2\sigma^{-(\tau+2)}r^{-2}\alpha^{-1}\underline{\omega}^{-2}M+d\frac{\mathrm{C}_5}{d}\mathrm{E}^3\sigma^{-(2\tau+4)}r^{-3}\alpha^{-2}\underline{\omega}^{-2}M\\ &\leqslant (2^{-(\tau+2)}\mathrm{C}_4r\alpha+\mathrm{C}_5\mathrm{E})\mathrm{E}^2\sigma^{-(2\tau+4)}r^{-3}\alpha^{-2}\underline{\omega}^{-2}M\\ &\leqslant (2^{-(\tau+2)}\mathrm{C}_4+\mathrm{C}_5)\mathrm{E}^3\sigma^{-(2\tau+4)}r^{-3}\alpha^{-2}\underline{\omega}^{-2}M\\ &= \mathrm{C}_6\mathrm{E}^3\sigma^{-(2\tau+4)}r^{-3}\alpha^{-2}\underline{\omega}^{-2}M,\end{aligned}$$

so that

$$\underline{\omega}\|u\|_{\bar{s}} \leqslant 2^{-(2\tau+6)}\mathrm{C}_6\mathrm{E}^7\sigma^{-\bar{\nu}}r^{-7}\alpha^{-4}\underline{\omega}^{-3}M \leqslant \bar{L}$$

Next, we estimate Q'. To do this, we need to estimate $Q^{(1)}, Q^{(2)}$ and $Q^{(3)}$. By definitions and Lemma 2.2.3, we have

$$\begin{aligned}\|Q^{(1)}\|_{\bar{s}} &\leqslant \sum_{1\leqslant l,j\leqslant d}\|Q_{y_l}\|_{\bar{s}}\|A_{lj}\|_{\bar{s}}r\bar{s}\\ &\leqslant d^2\mathrm{E}\frac{3}{2}\sigma^{-1}r^{-1}\cdot\frac{\mathrm{C}_5}{d}\mathrm{E}^3\sigma^{-(2\tau+4)}r^{-3}\alpha^{-2}\underline{\omega}^{-2}M\\ &= \frac{3}{2}d\mathrm{C}_5\mathrm{E}^4\sigma^{-(2\tau+5)}r^{-4}\alpha^{-2}\underline{\omega}^{-2}M\end{aligned}$$

[35] The factor d comes from the fact $\|a_x\|_{\bar{s}}=\|A\|_{\bar{s}}\leqslant\max_{1\leqslant l\leqslant d}\sum_{j=1}^{d}\|A_{lj}\|_{\bar{s}}$.

$$\begin{aligned}
\|Q^{(2)}\|_{\bar s} &\leqslant \int_0^1 (1-t) \sum_{1\leqslant j,l,k\leqslant d} \|Q_{y_jy_ly_k}\|_{\bar s}\|y'_j\|_{\bar s}\|y'_l\|_{\bar s}\|(u_0)_k\|_{\bar s}dt\\
&\leqslant \frac{d^3}{2}6\mathrm{E}\frac{27}{8}\sigma^{-3}r^{-3}\cdot r^2\bar s^2\cdot \mathrm{C}_4\mathrm{E}^2\sigma^{-(\tau+2)}r^{-2}\alpha^{-1}\underline{\omega}^{-2}M\\
&= \frac{81d^3\mathrm{C}_4}{8}\mathrm{E}^3\sigma^{-(\tau+5)}r^{-3}\alpha^{-1}\underline{\omega}^{-2}M\\
&\leqslant \frac{81d^3\mathrm{C}_4}{8}\mathrm{E}^4\sigma^{-(\tau+5)}r^{-4}\alpha^{-2}\underline{\omega}^{-2}M\\
&\leqslant 81\cdot 2^{-(\tau+3)}d^3\mathrm{C}_4\mathrm{E}^4\sigma^{-(2\tau+5)}r^{-4}\alpha^{-2}\underline{\omega}^{-2}M
\end{aligned}$$

and

$$\begin{aligned}
\|Q^{(3)}\|_{\bar s} &\leqslant \int_0^1 (1-t) \sum_{1\leqslant j,l\leqslant d} \|P_{y_ly_j}\|_{\bar s}\|y'_j\|_{\bar s}\|y'_l\|_{\bar s}dt\\
&\leqslant \frac{d^2}{2}2M\frac{9}{4}\sigma^{-2}r^{-2}\cdot r^2\bar s^2\\
&\leqslant \frac{9d^2}{4}\sigma^{-2}M
\end{aligned}$$

Thus

$$\|Q'\|_{\bar s}\leqslant \|Q^{(1)}\|_{\bar s}+\|Q^{(2)}\|_{\bar s}+\|Q^{(3)}\|_{\bar s}\leqslant \mathrm{C}_7\mathrm{E}^4\sigma^{-(2\tau+5)}r^{-4}\alpha^{-2}\underline{\omega}^{-2}M,$$

so that

$$\|Q'\|_{\bar s}\leqslant 2^{-(2\tau+5)}\mathrm{C}_7\mathrm{E}^7\sigma^{-\bar\nu}r^{-7}\alpha^{-4}\underline{\omega}^{-3}M\leqslant \bar L.$$

Finally, we estimate $\partial^2_{y'}Q'(0,\cdot)$. We have, once again by Lemma 2.2.3,

$$\|\partial^2_{y'}Q'(0,\cdot)\|_0\leqslant \|\partial^2_{y'}Q'(0,\cdot)\|_{s-\sigma}\leqslant 2\mathrm{C}_7\mathrm{E}^4\sigma^{-(2\tau+5)}r^{-4}\alpha^{-2}\underline{\omega}^{-2}M\cdot 9\sigma^{-2}r^{-2}=\mathrm{C}_8\mathrm{E}^4\sigma^{-(2\tau+7)}r^{-6}\alpha^{-2}M,$$

so that

$$r^2\sigma^2\|\partial^2_{y'}Q'(0,\cdot)\|_0\leqslant 2^{-(2\tau+5)}\mathrm{C}_8\mathrm{E}^7\sigma^{-\bar\nu}r^{-7}\alpha^{-4}\underline{\omega}^{-3}M\leqslant \bar L.$$

Now, under the assumption (2.3.14), we prove (2.3.15). For $(y',x;\varepsilon)\in W_{\bar s,\varepsilon_0}$ and $0\leqslant t\leqslant 1$, by (2.3.13) one has

$$|y'+t\varepsilon u(x)|\leqslant r\bar s+\varepsilon\|u\|_{\bar s}\leqslant r\bar s+\varepsilon_*\underline{\omega}^{-1}\bar L\leqslant r\bar s+r\frac{\sigma}{3}=rs-r\frac{\sigma}{3}<rs, \tag{2.3.20}$$

so that

$$\begin{aligned}
\|P^{(1)}\|_{\bar s} &\leqslant \int_0^1 (1-t) \sum_{1\leqslant j,l\leqslant d} \|Q_{y_ly_j}\|_{s-\frac{\sigma}{3}}\|u_j\|_{\bar s}\|u_l\|_{\bar s}dt\\
&\leqslant \frac{d^2}{2}2\mathrm{E}\cdot 9\sigma^{-2}r^{-2}\cdot\left(\mathrm{C}_6\mathrm{E}^3\sigma^{-(2\tau+4)}r^{-3}\alpha^{-2}\underline{\omega}^{-2}M\right)^2\\
&= 9d^2\mathrm{C}_6^2\mathrm{E}^7\sigma^{-(4\tau+10)}r^{-8}\alpha^{-4}\underline{\omega}^{-4}M^2
\end{aligned}$$

and

$$\begin{aligned}
\|P^{(2)}\|_{\bar{s}} &\leqslant \int_0^1 \sum_{1\leqslant j\leqslant d} \|P_{y_j}\|_{s-\frac{\sigma}{3}} \|u_j\|_{\bar{s}} dt \\
&\leqslant dM \cdot 3\sigma^{-1} r^{-1} \cdot \mathrm{C}_6 \mathrm{E}^3 \sigma^{-(2\tau+4)} r^{-3} \alpha^{-2} \underline{\omega}^{-2} M \\
&= 3d\mathrm{C}_6\mathrm{E}^3 \sigma^{-(2\tau+5)} r^{-4} \alpha^{-2} \underline{\omega}^{-2} M^2 \\
&\leqslant 3 \cdot 2^{-(2\tau+5)} d\mathrm{C}_6 \mathrm{E}^3 \underline{\omega}^{-2} r^4 \alpha^2 |\omega|^2 \sigma^{-(4\tau+10)} r^{-8} \alpha^{-4} \underline{\omega}^{-2} M^2 \\
&\leqslant 3 \cdot 2^{-(2\tau+5)} d\mathrm{C}_6 \mathrm{E}^7 \sigma^{-(4\tau+10)} r^{-8} \alpha^{-4} \underline{\omega}^{-4} M^2.
\end{aligned}$$

Thus

$$\|P'\|_{\bar{s}} \leqslant \|P^{(1)}\|_{\bar{s}} + \|P^{(2)}\|_{\bar{s}} \leqslant \mathrm{C}_9 \mathrm{E}^7 \sigma^{-(4\tau+10)} r^{-8} \alpha^{-4} \underline{\omega}^{-4} M^2 \leqslant r^{-1} \sigma^{-1} \underline{\omega}^{-1} \bar{L} M.$$

Next, we show how (2.3.14) implies the existence of the inverse of ψ_ε satisfying (2.3.16). The defining relation $\psi_\varepsilon \circ \varphi = \mathrm{id}$ implies that $\widetilde{\varphi}(x') = -a(x' + \varepsilon\widetilde{\varphi}(x'))$, where $\widetilde{\varphi}(x')$ is short for $\widetilde{\varphi}(x';\varepsilon)$ and such relation is a fixed point equation for the non–linear operator $f : u \to f(v) := -a(\mathrm{id}+\varepsilon v)$. To find a fixed point for this equation one can use a standard contraction Lemma (see [KF99]). Let Y denote the closed ball (with respect to the sup–norm) of continuos functions $v : \mathbb{T}^d_{s'} \times \{|\varepsilon| < \varepsilon_*\} \to \mathbb{C}^d$ such that $\|v\|_{s',\varepsilon_*} \leqslant r^{-1}\sigma^{-1}\underline{\omega}^{-1}\bar{L}$. By (2.3.14), $|\operatorname{Im}(x' + \varepsilon v(x'))| < s' + \varepsilon_* r^{-1}\sigma^{-1}\underline{\omega}^{-1}\bar{L} < s' + \frac{\sigma}{3} = \bar{s}$, for any $v \in Y$, and any $x' \in \mathbb{T}^d_{s'}$; thus, $\|f(v)\|_{r,s',\varepsilon_*} \leqslant \|a\|_{\bar{s}} \leqslant r^{-1}\sigma^{-1}\underline{\omega}^{-1}\bar{L}$ by (2.3.13), so that $f : Y \to Y$; notice that, in particular, this means that f sends x–periodic functions into x–periodic functions. Moreover, (2.3.14) implies also that f is a contraction: if $v_1, v_2 \in Y$, then, by the mean value theorem and (2.3.13), $|f(v_1) - f(v_2)| \leqslant \|a_x\|_{\bar{s}}\, |\varepsilon|\, |v_1 - v_2| \leqslant r^{-1}\sigma^{-1}\underline{\omega}^{-1}\bar{L}|\varepsilon|\, |v_1 - v_2|$, so that, by taking the sup–norm, one has $\|f(v_1) - f(v_2)\|_{s'} \leqslant \varepsilon_* r^{-1}\sigma^{-1}\underline{\omega}^{-1}\bar{L}\|v_1 - v_2\|_{s'} < \frac{1}{6}\|v_1 - v_2\|_{s'}$ showing that f is a contraction. Thus, there exists a unique $\widetilde{\varphi} \in Y$ such that $f(\widetilde{\varphi}) = \widetilde{\varphi}$. Furthermore, recalling that the fixed point is achieved as the uniform limit $\lim_{n\to\infty} f^n(0)$ $(0 \in Y)$ and since $f(0) = -a$ is analytic, so is $f^n(0)$ for any n and, hence, by Weierstrass Theorem on the uniform limit of analytic function, the limit $\widetilde{\varphi}$ itself is analytic. In conclusion, $\varphi \in \mathcal{B}_{s',\varepsilon_*}$ and (2.3.16) holds. Next, (2.3.16) and (2.3.20) imply (2.3.17) and therefore, ϕ_1 defines a symplectic diffeomorphism[36] satisfying (2.3.17) and the fourth inequality in the first line of (2.3.19). It remains to show the other estimates in (2.3.19). Since $L \geqslant \bar{L}$, the bound on $|\widetilde{E}|$ follows (2.3.13). By (2.3.15), (2.3.17) and (2.3.20), one has $\|P_1\|_{s',\varepsilon_*} \leqslant \|P'\|_{\bar{s},\varepsilon_*} \leqslant r^{-1}\sigma^{-1}\underline{\omega}^{-1}\bar{L}M \leqslant LM/\mathrm{E}$. Now, by Cauchy estimates, (2.3.13),

[36]Notice, in particular that the matrix $\mathbb{1}_d + \varepsilon a_x$ is, for any $x \in \mathbb{T}^d_{\bar{s}}$, invertible with inverse $\mathbb{1}_d + \varepsilon S(x;\varepsilon)$; in fact, since $\|\varepsilon a_x\|_{\bar{s}} < \varepsilon_* L/\mathrm{E} \leqslant \sigma/3 \leqslant 1/6$ the matrix $\mathbb{1}_d + \varepsilon a_x$ is invertible with inverse given by the "Neumann series" $(\mathbb{1}_d + \varepsilon a_x)^{-1} = \mathbb{1}_d + \sum_{k=1}^{\infty}(-1)^k(\varepsilon a_x)^k =: \mathbb{1}_d + \varepsilon S(x;\varepsilon)$, so that $\|S\|_{\bar{s},\varepsilon_*} \leqslant (\|a_x\|_{\bar{s},\varepsilon_*})/(1 - |\varepsilon|\|a_x\|_{\bar{s},\varepsilon_*}) < \frac{6}{5} r^{-1}\sigma^{-1}\underline{\omega}^{-1}\bar{L}$.

(2.3.14) and (2.3.17), it follows that

$$\begin{aligned}
\|\widetilde{Q}\|_{s',\varepsilon_*} &\leqslant \int_0^1 \sum_{j=1}^d \|Q_{x_j}\|_{\bar{s}} \|\widetilde{\varphi}_j\|_{s'} dt + \|Q'\|_{\bar{s}} \\
&\leqslant d\mathrm{E}\frac{3}{2}\sigma^{-1}r^{-1}\sigma^{-1}\underline{\omega}^{-1}\bar{L} + \|Q'\|_{\bar{s}} \\
&\leqslant \frac{3}{2}d\bar{\mathrm{C}}\mathrm{E}^8\sigma^{-(\bar{\nu}+2)}r^{-8}\alpha^{-4}\underline{\omega}^{-4}M + \mathrm{C}_7\mathrm{E}^4\sigma^{-(2\tau+5)}r^{-4}\alpha^{-2}\underline{\omega}^{-2}M \\
&\leqslant \left(\frac{3}{2}d\bar{\mathrm{C}}\mathrm{E}^4\underline{\omega}^{-2} + 2^{-(2\tau+7)}r^4\alpha^2\mathrm{C}_7\right)\mathrm{E}^4\sigma^{-(\bar{\nu}+2)}r^{-8}\alpha^{-4}\underline{\omega}^{-2}M \\
&\leqslant \left(\frac{3}{2}d\bar{\mathrm{C}}\underline{\omega}^{-2} + 2^{-(2\tau+7)}\mathrm{C}_7\underline{\omega}^{-2}\right)\mathrm{E}^8\sigma^{-(\bar{\nu}+2)}r^{-8}\alpha^{-4}\underline{\omega}^{-2}M \\
&= \frac{\widetilde{\mathrm{C}}}{2d}\mathrm{E}^8\sigma^{-(\bar{\nu}+2)}r^{-8}\alpha^{-4}M \leqslant \widetilde{L}
\end{aligned}$$

and[37]

$$\begin{aligned}
\|\partial_{y'}^2\widetilde{Q}(0,\cdot)\|_{0,\varepsilon_*} &\leqslant \max_{1\leqslant l\leqslant d}\sum_{j=1}^d \|\widetilde{Q}_{y'_l y'_j}\|_{s'-\sigma,\varepsilon_*} \\
&\leqslant 2d\frac{\widetilde{\mathrm{C}}}{2d}\mathrm{E}^8\sigma^{-(\bar{\nu}+2)}r^{-8}\alpha^{-4}M\cdot\sigma^{-2}r^{-2} \\
&= \widetilde{\mathrm{C}}\mathrm{E}^8\sigma^{-(\bar{\nu}+2)}r^{-10}\alpha^{-4}M = r^{-2}\widetilde{L}.
\end{aligned}$$

so that[38]

$$\|\widetilde{Q}\|_{s',\varepsilon_*}\ ,\ \frac{6}{5}\underline{\omega}^{-2}\mathrm{E}^2\,\|\partial_{y'}^2\widetilde{Q}(0,\cdot)\|_{0,\varepsilon_*} \leqslant 6r^{-2}\underline{\omega}^{-2}\widetilde{L}\mathrm{E}^2/5 = L\ , \tag{2.3.21}$$

Thus,

$$\begin{aligned}
\left\langle\partial_{y'}^2 Q_1(0,\cdot)\right\rangle &= \left\langle\partial_y^2 Q(0,\cdot)\right\rangle + \varepsilon\left\langle\partial_{y'}^2\widetilde{Q}(0,\cdot)\right\rangle = T^{-1}\left(\mathbb{1}_d + \varepsilon T\left\langle\partial_{y'}^2\widetilde{Q}(0,\cdot)\right\rangle\right) \\
&=: T^{-1}(\mathbb{1}_d + \varepsilon R)\ , \qquad (2.3.22)
\end{aligned}$$

and, in view of (2.3.12) and (2.3.21), we have

$$\begin{aligned}
\|R\| &\leqslant \|T\|\,\left\|\left\langle\partial_{y'}^2\widetilde{Q}(0,\cdot)\right\rangle\right\| \\
&\leqslant \underline{\omega}^{-2}\mathrm{E}\|\partial_{y'}^2\widetilde{Q}(0,\cdot)\|_{0,\varepsilon_*} \leqslant \frac{5L}{6\mathrm{E}}
\end{aligned}$$

[37] Recall that $0 < 2\sigma < s$ so that $s' - \sigma = s - 2\sigma > 0$.

[38] It is only here that a constant $L > r^{-1}\sigma^{-1}|\omega|^{-1}\bar{L}\mathrm{E}$ is needed; the (irrelevant) factor $6\underline{\omega}^{-2}\mathrm{E}^2/5$ has been introduced for later convenience.

Therefore, by (2.3.18), $\varepsilon_*\|R\| \leqslant \sigma/3 \leqslant 1/6 < 1$, implying that $(1+\varepsilon R)$ is invertible and

$$(\mathbb{1}_d + \varepsilon R)^{-1} = \mathbb{1}_d + \sum_{k=1}^{\infty} (-1)^k \varepsilon^k R^k =: 1 + \varepsilon D$$

with $\|D\| \leqslant \|R\|/(1 - |\varepsilon|\,\|R\|) < L/\mathrm{E}$. In conclusion, by (2.3.22), and the estimate on $\|D\|$,

$$T_1 = (1+\varepsilon R)^{-1} T = T + \varepsilon D T =: T + \varepsilon \widetilde{T}\ , \qquad |\omega|^2 \|\widetilde{T}\| \leqslant \|D\|\,|\omega|^2 \|T\| \leqslant \|D\|\mathrm{E} \leqslant \frac{L}{\mathrm{E}}\,\mathrm{E} = L\ ,$$

proving last estimate in (2.3.19) and, hence, Lemma 2.3.1. ∎

Next Lemma shows that, for $|\varepsilon|$ small enough, Kolmogorov's construction can be iterated and convergence proved.

Lemma 2.3.2 *Fix $0 < s_* < s$ and, for $j \geqslant 0$, let*[39]

$$\begin{cases} s_0 := s \\[2ex] \sigma_0 := \dfrac{s - s_*}{2} \end{cases} \qquad \begin{cases} \sigma_j := \dfrac{\sigma_0}{2^j} \\[2ex] s_{j+1} := s_j - \sigma_j = s_* + \frac{\sigma_0}{2^j} \end{cases} .$$

Let also $H_0 := H$, $\mathrm{K}_0 := \mathrm{K}$, $Q_0 := Q$, $K_0 := K$, $P_0 := P$, with $\mathcal{W}$, $\widehat{\mathrm{C}}$, $\widetilde{\mathrm{C}}$, E, L and ν as in §2.1.1 and assume that $\varepsilon_* \leqslant \varepsilon_0$ satisfies

$$\varepsilon_*\, \mathrm{e}_*\, d_*\, \|P\|_{s,\varepsilon_0} \leqslant 1 \quad \text{where} \quad \mathrm{e}_* := 3\widehat{\mathrm{C}}\,\sigma_0^{-(\nu+1)}\,\mathrm{E}^9 r^{-10} \alpha^{-4} \underline{\omega}^{-6}\ , \quad d_* := 2^{\nu+1}\ . \tag{2.3.23}$$

Then, one can construct a sequence of symplectic transformations

$$\phi_j : W_{rs_j, s_j, \varepsilon_*} \to D_{rs_{j-1}} \times \mathbb{T}^d_{s_{j-1}}\ , \tag{2.3.24}$$

so that

$$H_j := H_{j-1} \circ \phi_j =: K_j + \varepsilon^{2^j} P_j\ , \tag{2.3.25}$$

converges uniformly to a Kolmogorov's normal form. More precisely, $\varepsilon^{2^j} P_j$, $\Phi_j := \phi_1 \circ \phi_2 \circ \cdots \circ \phi_j$, K_j, K_j, Q_j converge uniformly on W_{s_*,ε_*} to, respectively, 0, ϕ_*, K_*, K_*, Q_*, which are real–analytic on W_{s_*,ε_*} and $H \circ \phi_* = K_* = \mathrm{K}_* + \omega \cdot y + Q_*$ with $Q_* = O(|y|^2)$. Finally, the following estimates hold for any $|\varepsilon| < \varepsilon_*$ and for any $i \geqslant 0$:

$$|\varepsilon|^{2^i} M_i := |\varepsilon|^{2^i} \|P_i\|_{s_i,\varepsilon_*} \leqslant \frac{(\,|\varepsilon|\mathrm{e}_*\, d_* M)^{2^i}}{\mathrm{e}_*\, {d_*}^{i+1}}\ , \tag{2.3.26}$$

$$\frac{\mathrm{C}\mathrm{E}^3}{r^3 \underline{\omega}^3 \sigma^2} \|\mathcal{W}(\phi_* - \mathrm{id})\|_{s_*},\ |\mathrm{K} - \mathrm{K}_*|,\ \|Q - Q_*\|_{s_*},\ |\omega|^2 \|T - T_*\| \leqslant \frac{|\varepsilon| L}{3\sigma}\ , \tag{2.3.27}$$

where $T_* := \left\langle \partial_y^2 Q_*(0,\cdot) \right\rangle^{-1}$.

[39] Notice that $s_j \downarrow s_*$.

Proof Notice that (2.3.23) implies (2.3.18) (and, hence, (2.3.14)). For $i \geqslant 0$, define

$$\mathcal{W}_i := \operatorname{diag}(\underline{\omega}\,\mathbb{1}_d, r\sigma_i\underline{\omega}\,\mathbb{1}_d) \quad \text{and} \quad \bar{L}_i := \bar{\mathrm{C}}\mathrm{E}^7\sigma_i^{-\bar{\nu}}r^{-7}\alpha^{-4}\underline{\omega}^{-3}M_i\,.$$

Let us assume (*inductive hypothesis*) that we can iterate $j \geqslant 1$ times Kolmogorov transformation obtaining j symplectic transformations $\phi_{i+1} : W_{rs_{i+1},s_{i+1},\varepsilon_*} \to D_{rs_i} \times \mathbb{T}^d_{s_i}$, for $0 \leqslant i \leqslant j-1$, and j Hamiltonians $H_{i+1} = H_i \circ \phi_{i+1} = K_{i+1} + \varepsilon^{2^{i+1}}P_{i+1}$ real–analytic on $W_{s_{i+1},\varepsilon_*}$ such that, for any $0 \leqslant i \leqslant j-1$,

$$\left\{\begin{array}{l} r|\omega|, \|Q_i\|_{s_i}, |\omega|^2\|T_i\| \leqslant \mathrm{E} \\[2ex] |\varepsilon|^{2^i}L_i := |\varepsilon|^{2^i}\widehat{\mathrm{C}}\mathrm{E}^{10}\sigma_0^{-\nu}\,2^{\nu i}r^{-10}\alpha^{-4}M_i \leqslant \mathrm{E}\frac{\sigma_i}{3}\ . \end{array}\right. \tag{2.3.28}$$

Observe that for $j=1$, it is $i=0$ and (2.3.28) is implied by the definition of E and by condition (2.3.23).
Because of (2.3.28), (2.3.18) holds for H_i and Lemma 2.3.1 can be applied to H_i and one has, for $0 \leqslant i \leqslant j-1$ and for any $|\varepsilon| < \varepsilon_*$ (compare (2.3.19)):

$$\begin{aligned} &|\mathrm{K}_{i+1}| \leqslant |\mathrm{K}_i| + |\varepsilon|^{2^i}L_i\ , \quad \|Q_{i+1}\|_{s_{i+1}} \leqslant \|Q_i\|_{s_i} + |\varepsilon|^{2^i}L_i\ , \quad |\omega|^2\|T_{i+1}\| \leqslant |\omega|^2\|T_i\| + |\varepsilon|^{2^i}L_i\ , \\ &\|\mathcal{W}_i(\phi_{i+1} - \mathrm{id})\|_{s_{i+1}} \leqslant |\varepsilon|^{2^i}\bar{L}_i\ , \quad M_{i+1} \leqslant M_iL_i\mathrm{E}^{-1}\ . \end{aligned} \tag{2.3.29}$$

Observe that, by definition of e_*, d_* in (2.3.23) and of L_i in (2.3.28), one has $|\varepsilon|^{2^j}L_j(3\sigma_j^{-1}\mathrm{E}^{-1}) = \mathrm{e}_*\,{d_*}^j|\varepsilon|^{2^j}M_j =: \theta_j/d_*$, so that $L_i\mathrm{E}^{-1} < \mathrm{e}_*\,{d_*}^iM_i$, thus by last relation in (2.3.29), for any $0 \leqslant i \leqslant j-1$, $|\varepsilon|^{2^{i+1}}M_{i+1} < \mathrm{e}_*\,{d_*}^i(M_i|\varepsilon|^{2^i})^2$ *.i.e.* $\theta_{i+1} < \theta_i^2$, which iterated, yields (2.3.26) *.i.e.* $\theta_i \leqslant \theta_0^{2^i}$ for $0 \leqslant i \leqslant j$.
Next, we show that, thanks to (2.3.23), (2.3.28) holds also for $i=j$. In fact, by (2.3.28) and the definition of E in §2.1.1, we have

$$\|Q_i\|_{s_j} \leqslant \|Q\|_s + \sum_{i=0}^{j-1}\varepsilon_*^{2^i}L_i \leqslant \|Q\|_s + \frac{1}{3}\mathrm{E}\sum_{i\geqslant 0}\sigma_i < \|Q\|_s + \frac{1}{3}\mathrm{E}\sum_{i\geqslant 0}2^{-(i+1)} = \|Q\|_s + \frac{1}{3}\mathrm{E} < \mathrm{E}\ .$$

The bound for $\|T_i\|$ is proven in an identical manner. Now, by $(2.3.26)_{i=j}$ and (2.3.23),

$$\theta_j/d_* = |\varepsilon|^{2^j}L_j(3\sigma_j^{-1}\mathrm{E}^{-1}) = \mathrm{e}_*\,{d_*}^j|\varepsilon|^{2^j}M_j \leqslant \mathrm{e}_*\,{d_*}^j(\mathrm{e}_*\,d_*\varepsilon_*M)^{2^j}/(\mathrm{e}_*\,{d_*}^{j+1}) \leqslant 1/d_* < 1\ ,$$

which implies the second inequality in (2.3.28) with $i=j$; the proof of the induction is finished and one can construct an *infinite sequence* of Kolmogorov transformations satisfying (2.3.28), (2.3.29) and (2.3.26) *for all* $i \geqslant 0$.
To check (2.3.27), we observe that

$$|\varepsilon|^{2^i}L_i\mathrm{E}^{-1} = \frac{\sigma_0}{3\cdot 2^i}\ \mathrm{e}_*\,{d_*}^i|\varepsilon|^{2^i}M_i < \frac{1}{2^{i+1}d_*}(|\varepsilon|\mathrm{e}_*\,d_*M)^{2^i} \leqslant \frac{1}{d_*}\Big(\frac{|\varepsilon|\mathrm{e}_*\,d_*M}{2}\Big)^{i+1}$$

and therefore

$$\sum_{i\geqslant 0}|\varepsilon|^{2^i}L_i \leqslant \frac{\mathrm{E}}{d_*}\sum_{i\geqslant 1}\left(\frac{|\varepsilon|\mathrm{e}_*\, d_* M}{2}\right)^i \leqslant |\varepsilon|\mathrm{e}_*\mathrm{E}M = \frac{|\varepsilon|L}{3\sigma_0}\ .$$

Thus,

$$\|Q-Q_*\|_{s_*} \leqslant \sum_{i\geqslant 0}|\varepsilon|^{2^i}\|\tilde{Q}_i\|_{s_i} \leqslant \sum_{i\geqslant 0}|\varepsilon|^{2^i}L_i \leqslant \frac{|\varepsilon|L}{3\sigma_0}\ ;$$

and analogously for $|\mathrm{K}-\mathrm{K}|_*$ and $\|T-T_*\|$.

Next, we prove that Φ_j is convergent by proving that it is Cauchy. For any $j\geqslant 1$, we have[40]

$$\begin{aligned}
\|\mathcal{W}_0(\Phi_j-\Phi_{j-1})\|_{s_*,\varepsilon_*} &= \|\mathcal{W}_0\Phi_{j-1}\circ\phi_j-\mathcal{W}_0\Phi_{j-1}\|_{s_*,\varepsilon_*}\\
&\leqslant \|\mathcal{W}_0 d\Phi_{j-1}\mathcal{W}_j^{-1}\|_{\varepsilon_*\bar{L}_{j-1}}\,\|\mathcal{W}_j(\phi_j-\mathrm{id})\|_{s_*,\varepsilon_*}\\
&\leqslant \|\mathcal{W}_0 d\Phi_{j-1}\|_{s_*+\sigma_j/3,\varepsilon_*}\,\|\mathcal{W}_j(\phi_j-\mathrm{id})\|_{s_*,\varepsilon_*}\\
&\leqslant \|\mathcal{W}_0 d\Phi_{j-1}\|_{s_{j-1}-\frac{11}{6}\sigma_{j-1},\varepsilon_*}\,\varepsilon^{2^{j-1}}\bar{L}_{j-1}\\
&\leqslant \frac{6}{11}\|\mathcal{W}_0\Phi_{j-1}\|_{s_{j-1},\varepsilon_*}\sigma_{j-1}^{-1}\max\left(r^{-1}\underline{\omega}^{-1}, r^{-1}\sigma_{j-1}^{-1}\underline{\omega}^{-1}\right)\varepsilon^{2^{j-1}}\bar{L}_{j-1}\\
&\leqslant \frac{6}{11}\|\mathcal{W}_0\Phi_0\|_{s_0,\varepsilon_*}\cdot\varepsilon^{2^{j-1}}\cdot r^{-1}\sigma_{j-1}^{-2}\underline{\omega}^{-1}\bar{L}_{j-1}\\
&\leqslant \frac{6}{11}\max\left(rs_0\underline{\omega}, r\sigma_0\underline{\omega}s_0\right)\varepsilon^{2^{j-1}}L_{j-1}\mathrm{E}^{-1}\\
&\leqslant \frac{6}{11}\cdot rs_0\underline{\omega}\cdot\frac{1}{3}\sigma_{j-1}.
\end{aligned}$$

Therefore, for any $n\geqslant 0$, $j\geqslant 1$,

$$\|\Phi_{n+j}-\Phi_n\|_{s_{n+j},\varepsilon_*} \leqslant \sum_{i=n}^{n+j}\|\Phi_{i+1}-\Phi_i\|_{s_i,\varepsilon_*} \leqslant \frac{2}{11}rs_0\underline{\omega}\sum_{i=n}^{n+j}\sigma_i.$$

Hence Φ_j converges uniformly on W_{s_*,ε_*} to some ϕ_*, which is then real–analytic function on W_{s_*,ε_*}.

To estimate $\|\mathcal{W}_0(\phi_*-\mathrm{id})\|_{s_*}$, observe that

$$\|\mathcal{W}_0(\Phi_i-\mathrm{id})\|_{s_i} \leqslant \|\mathcal{W}_0(\Phi_{i-1}\circ\phi_i-\phi_i)\|_{s_i}+\|\mathcal{W}_0(\phi_i-\mathrm{id})\|_{s_i} \leqslant \|\mathcal{W}_0(\Phi_{i-1}-\mathrm{id})\|_{s_{i-1}}+|\varepsilon|^{2^i}\bar{L}_i\ ,$$

[40] Notice that $\Phi_0=\mathrm{id}$, for any $j\geqslant 0$, $L_j\geqslant r^{-1}\sigma_j^{-2}\underline{\omega}^{-1}\bar{L}_j\mathrm{E}$ and, by (2.3.13), (2.3.14), (2.3.16) and (2.3.23), we have,

$$\widetilde{\phi}_j := \phi_j-\mathrm{id}\colon W_{s_*,\varepsilon_*}\to W_{s_*+\sigma_j/3,\varepsilon_*}\quad\text{and}\quad s_*+\sigma_j/3 = s_{j-1}-\frac{11}{6}\sigma_{j-1}.$$

which iterated yields

$$\|\mathcal{W}_0(\Phi_i - \mathrm{id})\|_{s_i} \leqslant \sum_{k=0}^{i} |\varepsilon|^{2^k} \bar{L}_k = \frac{\bar{\mathrm{C}}}{\widehat{\mathrm{C}}} \mathrm{E}^{-3} r^3 \underline{\omega}^3 \sum_{k=0}^{i} |\varepsilon|^{2^k} L_k \sigma_k^2 \leqslant \frac{\bar{\mathrm{C}}}{\widehat{\mathrm{C}}} \mathrm{E}^{-3} r^3 \underline{\omega}^3 |\varepsilon| \mathrm{e}_* M \sigma_0^2$$
$$= \frac{\bar{\mathrm{C}}}{\widehat{\mathrm{C}}} \mathrm{E}^{-3} r^3 \underline{\omega}^3 |\varepsilon| L \sigma_0 = \frac{r^3 \underline{\omega}^3}{3\mathrm{C}\mathrm{E}^3} |\varepsilon| L \sigma.$$

Therefore, taking the limit over i completes the proof of (2.3.27), Lemma 2.3.2 and, whence, of Kolmogorov's Theorem. ∎

2.3.2 Proof of Theorem 2.1.2

Lemma 2.3.3 (KAM step) *Let $r > 0$, $0 < 2\sigma < s \leqslant 1$ and consider the hamiltonian parametrized by $\varepsilon \in \mathbb{R}$*

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

with

$$K, P \in \mathcal{B}_{r,s}(\mathsf{y}) .$$

Assume that[41,42]

$$\begin{aligned} &\det K_{yy}(\mathsf{y}) \neq 0 , && T := K_{yy}(\mathsf{y})^{-1} , \\ &\|K_{yy}\|_{r,\mathsf{y}} \leqslant \mathsf{K} , && \|T\| \leqslant \mathsf{T} , \\ &\|P\|_{r,s,\mathsf{y}} \leqslant M , && \omega := K_{yy}(\mathsf{y}) \in \Delta_\alpha^\tau . \end{aligned} \tag{2.3.30}$$

Fix $\varepsilon \neq 0$ and let

$$\begin{aligned} &\lambda \geqslant \frac{4}{5} \log\left(\sigma^{2\nu+d} \frac{\alpha^2}{|\varepsilon| M K}\right), \quad \kappa := 5\sigma^{-1}\lambda, \quad \bar{r} \leqslant \min\left\{\frac{\alpha}{2d\mathsf{K}\kappa^\nu}, \frac{5}{24d} \frac{r}{\mathsf{T}\mathsf{K}}\right\}, \\ &\bar{s} := s - \frac{2}{3}\sigma, \quad s' := s - \sigma , \end{aligned} \tag{2.3.31}$$

Finally, define[43]

$$\begin{aligned} \bar{\mathsf{L}} &:= \frac{\mathsf{C}_0}{\sqrt{2}} \max\left\{1, \frac{\alpha}{r\mathsf{K}}\right\} \frac{M\mathsf{K}}{\alpha^2} \sigma^{-(2\nu+d)} , \\ \mathsf{L} &:= M \max\left\{\frac{8\mathsf{T}}{r\bar{r}} \sigma^{-(\nu+d)}, \frac{\mathsf{C}_7}{\sqrt{2}} \max\left\{1, \frac{\alpha}{r\mathsf{K}}\right\} \frac{\mathsf{K}}{\alpha^2} \sigma^{-2(\nu+d)}\right\} \\ &= M \max\left\{\frac{8\mathsf{T}}{r\bar{r}} \sigma^{-(\nu+d)}, \frac{4}{\mathsf{K}r^2}, \frac{\mathsf{C}_7}{\sqrt{2}} \max\left\{1, \frac{\alpha}{r\mathsf{K}}\right\} \frac{\mathsf{K}}{\alpha^2} \sigma^{-2(\nu+d)}\right\} . \end{aligned}$$

---

[41] 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, r, the roles of y_j, r_j in the iterative step.

[42] Notice that $\mathsf{T}\mathsf{K} \geqslant \mathsf{T}\|K_{yy}(\mathsf{y})\| \geqslant \|T\|\|K_{yy}(\mathsf{y})\| = \|T\|\|T^{-1}\| \geqslant 1$.

[43] Notice that $\mathsf{L} \geqslant \sigma^{-d}\bar{\mathsf{L}} \geqslant \bar{\mathsf{L}}$ since $\sigma \leqslant 1$.

Then, there exists a generating function $g \in \mathcal{B}_{\bar r,\bar s}(\mathsf{y})$ with the following properties:

$$\begin{cases} \|g_x\|_{\bar r,\bar s,\mathsf{y}} \leqslant \mathsf{C}_1 \dfrac{M}{\alpha}\sigma^{-(\nu+d)}\,, \\ \|g_{y'}\|_{\bar r,\bar s,\mathsf{y}},\ \|\partial^2_{y'x} g\|_{\bar r,\bar s,\mathsf{y}} \leqslant \bar{\mathsf{L}}\,, \\ \|\partial^2_{y'}\widetilde K\|_{\bar r,\mathsf{y}} \leqslant \mathsf{KL}\,, \end{cases} \tag{2.3.32}$$

where

$$\widetilde K(y') := \langle P(y',\cdot)\rangle\,.$$

If, in addition,

$$|\varepsilon| \leqslant \varepsilon_\sharp \quad \text{and} \quad |\varepsilon|\mathsf{L} \leqslant \frac{\sigma}{3}\,, \tag{2.3.33}$$

then, there exists $\mathsf{y}' \in \mathbb{R}^d$ such that

$$\begin{cases} \partial_{y'}K'(\mathsf{y}') = \omega\,, & \det \partial^2_{y'}K'(\mathsf{y}') \neq 0\,, \\ |\varepsilon|\|g_x\|_{\bar r,\bar s,\mathsf{y}} \leqslant \dfrac{r}{3}\,, & |\mathsf{y}'-\mathsf{y}| \leqslant \dfrac{8|\varepsilon|\mathsf{T}M}{r}\,, \\ |\varepsilon|\|\widetilde T\| \leqslant \mathsf{T}|\varepsilon|\mathsf{L}\,, & \|P_+\|_{\bar r,\bar s,\mathsf{y}} \leqslant \mathsf{L}M\,, \end{cases} \tag{2.3.34}$$

where

$$K' := K + \varepsilon\widetilde K\,, \qquad \big(\partial^2_{y'}K'(\mathsf{y}')\big)^{-1} =: T + \varepsilon\,\widetilde T\,, \qquad P_+(y',x) := P(y' + \varepsilon g_x(y',x),x)\,.$$

and the following hold. For $y' \in D_{\bar r}(\mathsf{y})$, the map $\psi_\varepsilon(x) := x + \varepsilon g_{y'}(y',x)$ has an analytic inverse $\varphi(x') = x' + \varepsilon\widetilde\varphi(y',x';\varepsilon)$ such that

$$\|\widetilde\varphi\|_{\bar r,s',\mathsf{y}} \leqslant \bar{\mathsf{L}} \qquad \text{and} \quad \varphi = \mathrm{id} + \varepsilon\widetilde\varphi : D_{\bar r/2,s'}(\mathsf{y}') \to \mathbb{T}^d_{\bar s}\,; \tag{2.3.35}$$

for any $(y',x) \in D_{\bar r,\bar s}(\mathsf{y})$, $|y'+\varepsilon g_x(y',x)-\mathsf{y}| < \frac{2}{3}r$; the map ϕ' is a symplectic diffeomorphism and

$$\phi' = \big(y' + \varepsilon g_x(y',\varphi(y',x')),\varphi(y',x')\big) : D_{\bar r/2,s'}(\mathsf{y}') \to D_{2r/3,\bar s}(\mathsf{y}), \tag{2.3.36}$$

with

$$\|\mathsf{W}\,\tilde\phi\|_{\bar r/2,s',\mathsf{y}'} \leqslant \sigma^d\mathsf{L}\,, \tag{2.3.37}$$

where $\tilde\phi$ is defined by the relation $\phi' =: \mathrm{id} + \varepsilon\tilde\phi$,

$$\mathsf{W} := \begin{pmatrix} \max\{\frac{\mathsf{K}}{\alpha}\,,\frac{1}{r}\}\,\mathbb{1}_d & 0 \\ 0 & \mathbb{1}_d \end{pmatrix}$$

and

$$\|P'\|_{\bar r/2,s',\mathsf{y}'} \leqslant \mathsf{L}M\,, \tag{2.3.38}$$

with

$$P'(y',x') := P_+(y',\varphi(x')) = P\circ\phi'(y',x')\,.$$

Proof
Step 1: Construction of the Arnold's transformation We seek for a near–to–the–identity symplectic transformation

$$\phi'\colon D_{r_1,s_1}(\mathsf{y}') \to D_{r,s}(\mathsf{y}),$$

with $D_{r_1,s_1}(\mathsf{y}') \subset D_{r,s}(\mathsf{y})$, generated by a function of the form $y' \cdot x + \varepsilon g(y', x)$, so that

$$\phi'\colon \begin{cases} y = y' + \varepsilon g_x(y', x) \\ x' = x + \varepsilon g_{y'}(y', x)\,, \end{cases} \tag{2.3.39}$$

such that

$$\begin{cases} H' := H \circ \phi' = K' + \varepsilon^2 P'\ , \\ \partial_{y'} K'(\mathsf{y}') = \omega, \quad \det \partial_{y'}^2 K'(\mathsf{y}') \neq 0\,. \end{cases} \tag{2.3.40}$$

By Taylor's formula, we get[44]

$$\begin{aligned} H(y' + \varepsilon g_x(y', x), x) =& K(y') + \varepsilon \widetilde{K}(y') + \varepsilon \Big[K_y(y') \cdot g_x + T_\kappa P(y', \cdot) - \widetilde{K}(y') \Big] + \\ & + \varepsilon^2 \left(P^{(1)} + P^{(2)} + P^{(3)} \right)(y', x) \\ =& K'(y') + \varepsilon \Big[K_y(y') \cdot g_x + T_\kappa P(y', \cdot) - \widetilde{K}(y') \Big] + \varepsilon^2 P_+(y', x), \end{aligned} \tag{2.3.41}$$

with $\kappa \in \mathbb{N}$, which will be chosen large enough so that $P^{(3)} = O(\varepsilon)$ and

$$\begin{cases} P_+ := P^{(1)} + P^{(2)} + P^{(3)} \\ P^{(1)} := \dfrac{1}{\varepsilon^2} \left[K(y' + \varepsilon g_x) - K(y') - \varepsilon K_y(y') \cdot g_x \right] = \displaystyle\int_0^1 (1-t) K_{yy}(\varepsilon t g_x) \cdot g_x \cdot g_x dt \\ P^{(2)} := \dfrac{1}{\varepsilon} \left[P(y' + \varepsilon g_x, x) - P(y', x) \right] = \displaystyle\int_0^1 P_y(y' + \varepsilon t g_x, x) \cdot g_x dt \\ P^{(3)} := \dfrac{1}{\varepsilon} \left[P(y', x) - T_\kappa P(y', \cdot) \right] = \dfrac{1}{\varepsilon} \displaystyle\sum_{|n|_1 > \kappa} P_n(y')\, e^{i n \cdot x}\ . \end{cases} \tag{2.3.42}$$

By the non–degeneracy condition in (2.3.30), for ε small enough (to be made precised below), $\det \partial_{y'}^2 K'(\mathsf{y}) \neq 0$ and, therefore, by Lemma 2.2.6, there exists a unique $\mathsf{y}' \in D_r(\mathsf{y})$ such that the second part of (2.3.40) holds. In view of (2.3.41), in order to get the first

[44] Recall that $\langle \cdot \rangle$ stands for the average over $\mathbb{T}^d$.

part of (2.3.40), we need to find g such that $K_y(y')\cdot g_x + T_\kappa P(y',\cdot) - \widetilde{K}(y')$ vanishes; such a g is indeed given by

$$g := \sum_{0<|n|_1\leqslant\kappa} \frac{-P_n(y')}{iK_y(y')\cdot n}\, e^{in\cdot x}, \tag{2.3.43}$$

provided that

$$K_y(y')\cdot n \neq 0, \quad \forall\, 0<|n|_1\leqslant\kappa, \quad \forall\, y'\in D_{r_1}(\mathsf{y}') \quad (\subset D_r(\mathsf{y}))\,. \tag{2.3.44}$$

But, in fact, since $K_y(\mathsf{y})$ is rationally independent, then, given any $\kappa\in\mathbb{N}$, there exists $\bar r\leqslant r$ such that

$$K_y(y')\cdot n \neq 0, \quad \forall\, 0<|n|_1\leqslant\kappa, \quad \forall\, y'\in D_{\bar r}(\mathsf{y}). \tag{2.3.45}$$

The last step is to invert the function $x\mapsto x+\varepsilon g_{y'}(y',x)$ in order to define P'. But, by Lemma 2.2.6, for ε small enough, the map $x\mapsto x+\varepsilon g_{y'}(y',x)$ admits an real–analytic inverse of the form

$$\varphi(y',x';\varepsilon) := x' + \varepsilon\widetilde{\varphi}(y',x';\varepsilon), \tag{2.3.46}$$

so that the Arnod's symplectic transformation is given by

$$\phi'\colon (y',x') \mapsto \begin{cases} y = y' + \varepsilon g_x(y',\varphi(y',x')) \\ x = \varphi(y',x';\varepsilon) = x' + \varepsilon\widetilde{\varphi}(y',x';\varepsilon). \end{cases} \tag{2.3.47}$$

Hence, (2.3.40) holds with

$$P'(y',x') := P_+(y',\varphi(y',x')). \tag{2.3.48}$$

Step 2: Quantitative estimates

First of all, notice that[45]

$$\bar r \leqslant \frac{5r}{24d\mathsf{TK}} < \frac{r}{2}\,. \tag{2.3.49}$$

We begin by extending the "diophantine condition w.r.t. K_y" uniformly to $D_{\bar r}(\mathsf{y})$ up to the order κ. Indeed, by the Mean Value Inequality and $K_y(\mathsf{y})=\omega\in\Delta_\alpha^\tau$, we get, for any $0<|n|_1\leqslant\kappa$ and any $y'\in D_{\bar r}(\mathsf{y})$,

$$\begin{aligned} |K_y(y')\cdot n| &= |\omega\cdot n + (K_y(y')-K_y(\mathsf{y}))\cdot n| \geqslant |\omega\cdot n|\left(1 - d\frac{\|K_{yy}\|_{\bar r,\mathsf{y}}}{|\omega\cdot n|}|n|_1\bar r\right) \\ &\geqslant \frac{\alpha}{|n|_1^\tau}\left(1-\frac{d\mathsf{K}}{\alpha}|n|_1^{\tau+1}\bar r\right) \geqslant \frac{\alpha}{|n|_1^\tau}\left(1-\frac{d\mathsf{K}}{\alpha}\kappa^{\tau+1}\bar r\right) \geqslant \frac{\alpha}{2|n|_1^\tau}, \end{aligned} \tag{2.3.50}$$

[45] Recall footnote 42.

so that, by Lemma 2.2.4–(i), we have

$$
\begin{aligned}
\|g_x\|_{\bar r,\bar s,\mathsf{y}} &\overset{def}{=} \sup_{D_{\bar r,\bar s}(\mathsf{y})}\left|\sum_{0<|n|_1\leqslant\kappa}\frac{nP_n(y')}{K_y(y')\cdot n}\,e^{in\cdot x}\right| \leqslant \sum_{0<|n|_1\leqslant\kappa}\frac{\|P_n\|_{\bar r,\bar s,\mathsf{y}}}{|K_y(y')\cdot n|}|n|_1\,e^{\left(s-\frac{2}{3}\sigma\right)|n|_1}\\
&\leqslant \sum_{0<|n|_1\leqslant\kappa} M\,e^{-s|n|_1}\frac{2|n|_1^{\nu}}{\alpha}\,e^{\left(s-\frac{2}{3}\sigma\right)|n|_1} \leqslant \frac{2M}{\alpha}\sum_{n\in\mathbb{Z}^d}|n|_1^{\nu}\,e^{-\frac{2}{3}\sigma|n|_1}\\
&\leqslant \frac{2M}{\alpha}\int_{\mathbb{R}^d}|y|_1^{\nu}\,e^{-\frac{2}{3}\sigma|y|_1}dy\\
&= \left(\frac{3}{2\sigma}\right)^{\nu+d}\frac{2M}{\alpha}\int_{\mathbb{R}^d}|y|_1^{\nu}\,e^{-|y|_1}dy\\
&= \mathsf{C}_1\frac{M}{\alpha}\sigma^{-(\nu+d)}\,,
\end{aligned}
$$

$$
\begin{aligned}
\|\partial_{y'}g\|_{\bar r,\bar s,\mathsf{y}} &\overset{def}{=} \sup_{D_{\bar r,\bar s}(\mathsf{y})}\left|\sum_{0<|n|_1\leqslant\kappa}\left(\frac{\partial_y P_n(y')}{K_y(y')\cdot n}-P_n(y')\frac{K_{yy}(y')n}{(K_y(y')\cdot n)^2}\right)e^{in\cdot x}\right|\\
&\leqslant \sum_{0<|n|_1\leqslant\kappa}\sup_{D_{\bar r}(\mathsf{y})}\left(\frac{\|(P_y)_n\|_{\bar r,s,\mathsf{y}}}{|K_y(y')\cdot n|}+d\|P_n\|_{r,s,\mathsf{y}}\frac{\|K_{yy}\|_{r,\mathsf{y}}|n|_1}{|K_y(y')\cdot n|^2}\right)e^{\left(s-\frac{2}{3}\sigma\right)|n|_1}\\
&\overset{(2.3.30)+(2.3.50)}{\leqslant} \sum_{0<|n|_1\leqslant\kappa}\left(\frac{M}{r-\bar r}\,e^{-s|n|_1}\frac{2|n|_1^{\tau}}{\alpha}+dM\,e^{-s|n|_1}\mathsf{K}|n|_1\left(\frac{2|n|_1^{\tau}}{\alpha}\right)^2\right)e^{\left(s-\frac{2}{3}\sigma\right)|n|_1}\\
&\overset{(2.3.49)}{\leqslant} \frac{4M}{\alpha^2 r}\sum_{0<|n|_1\leqslant\kappa}\left(|n|_1^{\tau}\alpha+dr\mathsf{K}|n|_1^{2\tau+1}\right)e^{-\frac{2}{3}\sigma|n|_1}\\
&\leqslant \max\{\alpha,r\mathsf{K}\}\frac{4M}{\alpha^2 r}\sum_{0<|n|_1\leqslant\kappa}\left(|n|_1^{\tau}+d|n|_1^{2\tau+1}\right)e^{-\frac{2}{3}\sigma|n|_1}\\
&\leqslant \max\left\{1,\frac{\alpha}{r\mathsf{K}}\right\}\frac{4M\mathsf{K}}{\alpha^2}\int_{\mathbb{R}^d}\left(|y|_1^{\tau}+d|y|_1^{2\tau+1}\right)e^{-\frac{2}{3}\sigma|y|_1}dy\\
&= \left(\frac{3}{2\sigma}\right)^{2\tau+d+1}\max\left\{1,\frac{\alpha}{r\mathsf{K}}\right\}\frac{4M\mathsf{K}}{\alpha^2}\int_{\mathbb{R}^d}\left(|y|_1^{\tau}+d|y|_1^{2\tau+1}\right)e^{-|y|_1}dy\\
&\leqslant \frac{\mathsf{C}_0}{\sqrt{2}}\max\left\{1,\frac{\alpha}{r\mathsf{K}}\right\}\frac{M\mathsf{K}}{\alpha^2}\sigma^{-(2\tau+d+1)}\\
&\leqslant \bar{\mathsf{L}}\,,
\end{aligned}
$$

and, analogously,

$$\begin{aligned}
\|\partial^2_{y'x} g\|_{\bar r,\bar s,\mathsf{y}} &\overset{def}{=} \sup_{D_{\bar r,\bar s}(\mathsf{y})} \left| \sum_{0<|n|_1\leqslant \kappa} \left(\frac{\partial_y P_n(y')}{K_y(y')\cdot n} - P_n(y') \frac{K_{yy}(y')n}{(K_y(y')\cdot n)^2} \right) \cdot n\, e^{in\cdot x} \right| \\
&\leqslant \sum_{0<|n|_1\leqslant\kappa} \sup_{D_{\bar r}(\mathsf{y})} \left(\frac{\|(P_y)_n\|_{\bar r,s,\mathsf{y}}}{|K_y(y')\cdot n|} + d\|P_n\|_{r,s,\mathsf{y}} \frac{\|K_{yy}\|_{r,\mathsf{y}}|n|_1}{|K_y(y')\cdot n|^2} \right) |n|_1\, e^{\left(s-\frac{2}{3}\sigma\right)|n|_1} \\
&\leqslant \max\{\alpha, r\mathsf{K}\} \frac{4M}{\alpha^2 r} \sum_{0<|n|_1\leqslant\kappa} \left(|n|_1^\tau + d|n|_1^{2\tau+1}\right) |n|_1\, e^{-\frac{2}{3}\sigma|n|_1} \\
&\leqslant \max\left\{1, \frac{\alpha}{r\mathsf{K}}\right\} \frac{4M\mathsf{K}}{\alpha^2} \int_{\mathbb{R}^d} \left(|y|_1^\tau + d|y|_1^{2\tau+1}\right) |y|_1\, e^{-\frac{2}{3}\sigma|y|_1} dy \\
&= \left(\frac{3}{2\sigma}\right)^{2\tau+d+2} \max\left\{1, \frac{\alpha}{r\mathsf{K}}\right\} \frac{4M\mathsf{K}}{\alpha^2} \int_{\mathbb{R}^d} \left(|y|_1^{\tau+1} + d|y|_1^{2\tau+2}\right) e^{-|y|_1} dy \\
&= \frac{\mathsf{C}_0}{\sqrt{2}} \max\left\{1, \frac{\alpha}{r\mathsf{K}}\right\} \frac{M\mathsf{K}}{\alpha^2} \sigma^{-(2\nu+d)} \\
&= \overline{\mathsf{L}}\,,
\end{aligned}$$

and, for $|\varepsilon| < \varepsilon_*$,

$$\|\widetilde{K}_y\|_{r/2,\mathsf{y}} = \|\,[P_y]\,\|_{r/2,\mathsf{y}} \leqslant \|P_y\|_{r/2,\bar s,\mathsf{y}} \leqslant \frac{M}{r-\frac{r}{2}} \leqslant \frac{2M}{r}\,,$$

$$\|\partial^2_{y'}\widetilde{K}\|_{r/2,\mathsf{y}} = \|\,[P_{yy}]\,\|_{r/2,\mathsf{y}} \leqslant \|P_{yy}\|_{r/2,\bar s,\mathsf{y}} \leqslant \frac{M}{(r-\frac{r}{2})^2} \leqslant \frac{4M}{r^2} \leqslant \mathsf{K}\mathsf{L}\,.$$

Next, we prove the existence and uniqueness of y' in (2.3.40). Consider then

$$\begin{aligned}
F\colon D_{\bar r}(\mathsf{y}) \times D^1_{2|\varepsilon|}(0) &\longrightarrow \qquad \mathbb{C}^d \\
(y,\eta) \quad &\longmapsto K_y(y) + \eta \widetilde{K}_{y'}(y) - K_y(\mathsf{y}).
\end{aligned}$$

Then

- $F(\mathsf{y},0) = 0, \quad F_y(\mathsf{y},0)^{-1} = K_{yy}(\mathsf{y})^{-1} = T;$

- For any $(y,\eta)\in D_{\bar r}(\mathsf{y})\times D^1_{2|\varepsilon|}(0)$,

$$\begin{aligned}
\|\mathbb{1}_d - TF_y(y,\eta)\| &\leqslant \|\mathbb{1}_d - TK_{yy}\| + |\eta|\ \|T\|\ \|\partial^2_{y'}\widetilde{K}\|_{r/2,\mathsf{y}}\\
&\leqslant d\|T\|\|K_{yyy}\|_{\bar r,\mathsf{y}}\bar r + 2|\varepsilon|\mathsf{T}\frac{4M}{r^2}\\
&\leqslant d\mathsf{T}\mathsf{K}\frac{\bar r}{r-\bar r} + 8\mathsf{T}\frac{|\varepsilon|M}{r^2}\\
&\overset{(2.3.49)}{\leqslant} d\mathsf{T}\mathsf{K}\frac{2\bar r}{r} + |\varepsilon|\frac{8\mathsf{T}M}{r^2}\\
&\leqslant 2d\mathsf{T}\mathsf{K}\frac{\bar r}{r} + \frac{1}{2}|\varepsilon|\mathsf{L}\\
&\overset{(2.3.49)+(2.3.33)}{\leqslant} \frac{5}{12} + \frac{\sigma}{6}\\
&\leqslant \frac{5}{12} + \frac{1}{12} = \frac{1}{2}\ ;
\end{aligned}$$

- Recalling $\sigma\leqslant\frac{1}{2}$, we have

$$\begin{aligned}
2\|T\|\|F(\mathsf{y},\cdot)\|_{2|\varepsilon|,0} &= 2\|T\|\sup_{B^1_{2|\varepsilon|}(0)}|\eta\widetilde{K}_{y'}(\mathsf{y})|\\
&\leqslant 2\mathsf{T}\frac{4|\varepsilon|M}{r}\\
&< \bar r\sigma^d|\varepsilon|\mathsf{L}\\
&\overset{(2.3.33)}{<} \bar r\ \frac{\sigma}{3}\\
&< \frac{\bar r}{2}\ .
\end{aligned}\tag{2.3.51}$$

Therefore, Lemma 2.2.6 applies. Hence, there exists a function $g\colon D^1_{2|\varepsilon|}(0)\to D_{\bar r}(\mathsf{y})$ such that its graph coincides with $F^{-1}(\{0\})$. In particular, $\mathsf{y}':=g(\varepsilon)$ is the unique $y\in D_{\bar r}(\mathsf{y})$ satisfying $0=F(y,\varepsilon)=\partial_y K'(y)-\omega$ *i.e.* the second part of (2.3.40). Moreover,

$$|\mathsf{y}'-\mathsf{y}|\leqslant 2\|T\|\|F(\mathsf{y},\cdot)\|_{2|\varepsilon|,0}\leqslant\frac{8|\varepsilon|\mathsf{T}M}{r}\overset{(2.3.51)}{\leqslant}\bar r\ \sigma^d|\varepsilon|\mathsf{L}<\frac{\bar r}{2}\ ,\tag{2.3.52}$$

so that

$$D_{\frac{\bar r}{2}}(\mathsf{y}')\subset D_{\bar r}(\mathsf{y}).\tag{2.3.53}$$

Next, we prove that $\partial_y^2 K'(\mathsf{y}')$ is invertible. Indeed, by Taylor' formula, we have

$$\begin{aligned}\partial_y^2 K'(\mathsf{y}') &= K_{yy}(\mathsf{y}) + \int_0^1 K_{yyy}(\mathsf{y}+t\varepsilon\widetilde{y})\cdot\varepsilon\widetilde{y}dt + \varepsilon\widetilde{K}_{yy}(\mathsf{y}')\\ &= T^{-1}\left(\mathbb{1}_d + \varepsilon T\left(\int_0^1 K_{yyy}(\mathsf{y}+t\varepsilon\widetilde{y})\cdot\widetilde{y}dt + \widetilde{K}_{yy}(\mathsf{y}')\right)\right)\\ &=: T^{-1}(\mathbb{1}_d+\varepsilon A),\end{aligned}$$

and, by Cauchy's estimate,

$$\begin{aligned}|\varepsilon|\|A\| &\leqslant \|T\|\left(d\|K_{yyy}\|_{r/2,\mathsf{y}}|\varepsilon||\mathsf{y}'-\mathsf{y}| + |\varepsilon|\|\partial_{y'}^2\widetilde{K}\|_{r/2,\mathsf{y}}\right)\\ &\leqslant \|T\|\left(\frac{d\|K_{yy}\|_{r,\mathsf{y}}}{r-\frac{r}{2}}|\varepsilon||\mathsf{y}'-\mathsf{y}| + |\varepsilon|\|\widetilde{K}_{yy}\|_{r/2,\mathsf{y}}\right)\\ &\overset{(2.3.52)}{\leqslant} \mathsf{T}\left(\frac{2d\mathsf{K}}{r}\frac{8|\varepsilon|\mathsf{T}M}{r} + \frac{4|\varepsilon|M}{r^2}\right)\\ &\leqslant \frac{4|\varepsilon|\mathsf{T}M}{r^2}(4d\mathsf{T}\mathsf{K}+1)\\ &\leqslant \frac{20d|\varepsilon|\mathsf{T}^2\mathsf{K}M}{r^2}\\ &\overset{(2.3.49)}{\leqslant} \frac{25}{6d}|\varepsilon|\frac{\mathsf{T}M}{r\bar{r}}\\ &< \frac{1}{2}|\varepsilon|\mathsf{L}\\ &\overset{(2.3.33)}{\leqslant} \frac{\sigma}{6}\\ &\leqslant \frac{1}{2}.\end{aligned}$$

Hence $\partial_{y'}^2 K'(\mathsf{y}')$ is invertible with

$$\partial_{y'}^2 K'(\mathsf{y}')^{-1} = (\mathbb{1}_d+\varepsilon A)^{-1}T = T + \sum_{k\geqslant 1}(-\varepsilon)^k A^k T =: T + \varepsilon\widetilde{T},$$

and

$$|\varepsilon|\|\widetilde{T}\| \leqslant |\varepsilon|\frac{\|A\|}{1-|\varepsilon|\|A\|}\|T\| \leqslant 2|\varepsilon|\|A\|\|T\| \leqslant |\varepsilon|\mathsf{L}\mathsf{T} \leqslant 2\frac{\sigma}{6}\mathsf{T} = \mathsf{T}\frac{\sigma}{3}\,.$$

Next, we prove estimate on P_+. We have,

$$|\varepsilon|\|g_x\|_{\bar{r},\bar{s},\mathsf{y}} \leqslant |\varepsilon|\mathsf{C}_1\frac{M}{\alpha}\sigma^{-(\tau+d+1)} \leqslant |\varepsilon|\frac{r}{3}\mathsf{L} \overset{(2.3.33)}{\leqslant} \frac{r}{3}\frac{\sigma}{3} \leqslant \frac{r}{3}$$

so that, for any $(y',x)\in D_{\bar r,\bar s}(\mathsf{y})$,

$$|y'+\varepsilon g_x(y',x)-\mathsf{y}|\leqslant \bar r+\frac{r}{3}<\frac{r}{8d}+\frac{r}{3}<\frac{2r}{3}<r\,,$$

and thus

$$\begin{aligned}\|P^{(1)}\|_{\bar r,\bar s,\mathsf{y}} &\leqslant d^2\|K_{yy}\|_{r,\mathsf{y}}\|g_x\|^2_{\bar r,\bar s,\mathsf{y}}\leqslant d^2\mathsf{K}\left(\mathsf{C}_1\frac{M}{\alpha}\sigma^{-(\nu+d)}\right)^2\\ &= d^2\mathsf{C}_1^2\frac{\mathsf{K}M^2}{\alpha^2}\sigma^{-2(\nu+d)},\end{aligned}$$

$$\begin{aligned}\|P^{(2)}\|_{\bar r,\bar s,\mathsf{y}} &\leqslant d\|P_y\|_{\frac{5r}{6},\bar s,\mathsf{y}}\|g_x\|_{\bar r,\bar s,\mathsf{y}}\leqslant d\frac{6M}{r}\mathsf{C}_1\frac{M}{\alpha}\sigma^{-(\nu+d)}\\ &= 6d\mathsf{C}_1\frac{M^2}{\alpha r}\sigma^{-(\nu+d)}\end{aligned}$$

and by Lemma 2.2.4–(i), we have,

$$\begin{aligned}|\varepsilon|\|P^{(3)}\|_{\bar r,s-\frac{\sigma}{2},\mathsf{y}} &\leqslant \sum_{|n|_1>\kappa}\|P_n\|_{\bar r,\mathsf{y}}\,e^{(s-\frac{\sigma}{2})|n|_1}\leqslant M\sum_{|n|_1>\kappa}e^{-\frac{\sigma|n|_1}{2}}\\
&\leqslant M\,e^{-\frac{\kappa\sigma}{4}}\sum_{|n|_1>\kappa}e^{-\frac{\sigma|n|_1}{4}}\leqslant M\,e^{-\frac{\kappa\sigma}{4}}\sum_{|n|_1>0}e^{-\frac{\sigma|n|_1}{4}}\\
&= M\,e^{-\frac{\kappa\sigma}{4}}\left(\left(\sum_{k\in\mathbb{Z}}e^{-\frac{\sigma|k|}{4}}\right)^d-1\right)=M\,e^{-\frac{\kappa\sigma}{4}}\left(\left(1+\frac{2\,e^{-\frac{\sigma}{4}}}{1-\,e^{-\frac{\sigma}{4}}}\right)^d-1\right)\\
&= M\,e^{-\frac{\kappa\sigma}{4}}\left(\left(1+\frac{2}{e^{\frac{\sigma}{4}}-1}\right)^d-1\right)\leqslant M\,e^{-\frac{\kappa\sigma}{4}}\left(\left(1+\frac{2}{\frac{\sigma}{4}}\right)^d-1\right)\\
&\leqslant \sigma^{-d}M\,e^{-\frac{\kappa\sigma}{4}}\left((\sigma+8)^d-\sigma^d\right)\leqslant d8^d\sigma^{-d}M\,e^{-\frac{\kappa\sigma}{4}}\\
&= \mathsf{C}_2\sigma^{-d}M\,e^{-\frac{5}{4}\lambda}\\
&\overset{(2.3.31)}{\leqslant} \mathsf{C}_2\sigma^{-d}M\sigma^{-(2\nu+d)}\frac{|\varepsilon|M\mathsf{K}}{\alpha^2}\\
&= \mathsf{C}_2M\frac{|\varepsilon|M\mathsf{K}}{\alpha^2}\sigma^{-2(\nu+d)}\,.\end{aligned}$$

Hence[46],

$$\begin{aligned}
\|P_+\|_{\bar r,\bar s,\mathsf{y}} &\leqslant \|P^{(1)}\|_{\bar r,\bar s,\mathsf{y}} + \|P^{(2)}\|_{\bar r,\bar s,\mathsf{y}} + \|P^{(3)}\|_{\bar r,\bar s,\mathsf{y}}\\
&\leqslant d^2\mathsf{C}_1^2\frac{\mathsf{K}M^2}{\alpha^2}\sigma^{-2(\nu+d)} + 6d\mathsf{C}_1\frac{M^2}{\alpha r}\sigma^{-(\nu+d)} + \mathsf{C}_2 M\frac{|\varepsilon|M\mathsf{K}}{\alpha^2}\sigma^{-2(\nu+d)}\\
&= \left(d^2\mathsf{C}_1^2 r\mathsf{K} + 6d\mathsf{C}_1\alpha\sigma^{\nu+d} + \mathsf{C}_2 r\mathsf{K}\right)\frac{M^2}{\alpha^2 r}\sigma^{-2(\tau+d+1)}\\
&\leqslant \left(d^2\mathsf{C}_1^2 + 6d\mathsf{C}_1 + \mathsf{C}_2\right)\max\{\alpha, r\mathsf{K}\}\frac{M^2}{\alpha^2 r}\sigma^{-2(\tau+d+1)}\\
&\overset{(2.3.30)}{\leqslant} \frac{\mathsf{C}_3}{\sqrt{2}}\max\left\{1, \frac{\alpha}{r\mathsf{K}}\right\}\frac{M^2\mathsf{K}}{\alpha^2}\sigma^{-2(\nu+d)}\\
&\leqslant \mathsf{L}M\ .
\end{aligned}$$

The proof of the claims on ϕ' and P' are proven in a similar way as in Lemma 2.3.1. ∎

Finally, we prove the convergence of the scheme by mimicking Lemma 2.3.2.

Lemma 2.3.4 *Let $H_0 := H$, $K_0 := K$, $P_0 := P$, $\phi^0 = \phi_0 :=$ id, and r_0, s_0, s_*, σ_0, μ_0,*

[46] Recall that $\sigma < 1$.

W_0, M_0, K_0, T_0 *and* $\varepsilon_\sharp$ *be as in* §2.1.2*. For a given* $\varepsilon \neq 0$*, define*[47]

$$
\begin{aligned}
\sigma_j &:= \frac{\sigma_0}{2^j}\,,\\
s_{j+1} &:= s_j - \sigma_j = s_* + \frac{\sigma_0}{2^j}\,,\\
\bar{s}_j &:= s_j - \frac{2\sigma_i}{3}\,,\\
\mathsf{K}_{j+1} &:= \mathsf{K}_0 \prod_{k=0}^{j} (1+\frac{\sigma_k}{3}) \leqslant \mathsf{K}_0\, e^{\frac{2\sigma_0}{3}} < \mathsf{K}_0\sqrt{2}\,,\\
\mathsf{T}_{j+1} &:= \mathsf{T}_0 \prod_{k=0}^{j} (1+\frac{\sigma_k}{3}) \leqslant \mathsf{T}_0\, e^{\frac{2\sigma_0}{3}} < \mathsf{T}_0\sqrt{2}\,,\\
\lambda_0 &:= \log \mu_0^{-1}\,,\\
\lambda_j &:= 2^j \lambda_0\,,\\
\kappa_j &:= 5\sigma_j^{-1}\lambda_j\,,\\
r_{j+1} &:= \min\left\{ \frac{\alpha}{4d\sqrt{2}\mathsf{K}_0\kappa_j^{\nu}}\,,\, \frac{5}{96d}\frac{r_j}{\eta_0} \right\}\,,\\
\mathsf{d}_* &:= \mathsf{C}_5\, \eta_0^2\,,\\
\mathsf{e}_* &:= \mathsf{C}_9 \frac{\mathsf{K}_0}{\alpha^2} \sigma_0^{-(4\nu+2d+1)}\, \lambda_0^{2\nu}\,,\\
\mathsf{f}_* &:= \mathsf{C}_8\; \max\left\{1, \frac{\alpha}{r_0\mathsf{K}_0}\right\} \eta_0\; \sigma_0^{-(3\nu+2d+1)}\; \mu_0\; \lambda_0^{\nu}\,.
\end{aligned}
$$

Assume that ε *is such that*

$$
\mu_0 \leqslant \varepsilon_\sharp \quad \textit{and} \quad \mathsf{f}_* \max\left\{1\,,\, \frac{\mathsf{C}_{10}}{3}\, \sigma_0\;\; \eta_0^{\frac14}\; \mathsf{e}_*\; d_*^2\; |\varepsilon|\; M_0\right\} < 1\,. \tag{2.3.54}
$$

Then, one can construct a sequence of symplectic transformations

$$
\phi_j : D_{r_j,s_j}(y_j) \to D_{r_{j-1},s_{j-1}}(y_{j-1})\;, \tag{2.3.55}
$$

so that

$$
H_j := H_{j-1}\circ\phi_j =: K_j + \varepsilon^{2^j} P_j\;, \tag{2.3.56}
$$

converges uniformly. More precisely, $\varepsilon^{2^j}P_j$*,* $\phi^j := \phi_0\circ\phi_1\circ\phi_2\circ\cdots\circ\phi_j$*,* K_j*,* y_j *converge uniformly on* $\{y_*\}\times\mathbb{T}^d_{s_*}$ *to, respectively,* 0*,* ϕ_**,* K_**,* y_* *which are real–analytic on* $\mathbb{T}^d_{s_*}$ *and*

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

$H \circ \phi_* = K_*$ *with* $\det \partial_y^2 K_*(y_*) \neq 0$. *Finally, the following estimates hold for any* $i \geqslant 1$*:*

$$|\varepsilon|^{2^i} \|P_i\|_{r_i,s_i,y_i} \leqslant \frac{\left(\frac{1}{3}\sigma_0 \mathsf{f}_* \mathsf{e}_* d_*^2 |\varepsilon| M_0\right)^{2^{i-1}}}{\mathsf{e}_* \, d_*^{\,i+1}} \,, \tag{2.3.57}$$

$$|\mathsf{W}(\phi_* - \mathrm{id})| \leqslant \sigma_0^{d+1} \quad on \quad \{y_*\} \times \mathbb{T}^d_{s_*} \,. \tag{2.3.58}$$

Proof For $i \geqslant 0$, define

$$\begin{aligned}
\mathsf{W}_i &:= \operatorname{diag}\left(\max\left\{\frac{\mathsf{K}_i}{\alpha}, \frac{1}{r_i}\right\} \mathbb{1}_d \,, \mathbb{1}_d\right) , \\
\bar{\mathsf{L}}_i &:= \mathsf{C}_0 \max\left\{1, \frac{\alpha}{r_i \mathsf{K}_i}\right\} \frac{M_i \mathsf{K}_0}{\alpha^2} \sigma_i^{-(2\nu+d)} \,, \\
\mathsf{L}_i &:= M_i \max\left\{\frac{4\sqrt{2}\mathsf{T}_0}{r_i r_{i+1}} \sigma_i^{-(\nu+d)} \,, \mathsf{C}_7 \max\left\{1, \frac{\alpha}{r_i \mathsf{K}_i}\right\} \frac{\mathsf{K}_0}{\alpha^2} \sigma_i^{-2(\nu+d)}\right\} \\
&\geqslant M_i \max\left\{\frac{4\mathsf{T}_i}{r_i r_{i+1}} \sigma_i^{-(\nu+d)} \,, \frac{4}{\mathsf{K}_i r_i^2} \,, \mathsf{C}_7 \max\left\{1, \frac{\alpha}{r_i \mathsf{K}_i}\right\} \frac{\mathsf{K}_0}{\alpha^2} \sigma_i^{-2(\nu+d)}\right\} .
\end{aligned}$$

Let us assume (*inductive hypothesis*) that we can iterate $j \geqslant 1$ times the KAM step obtaining j symplectic transformations[48]

$$\phi_{i+1} : D_{r_{i+1},s_{i+1}}(y_{i+1}) \to D_{2r_i/3,s_i}(y_i), \quad \text{for} \quad 0 \leqslant i \leqslant j-1, \tag{2.3.59}$$

and j Hamiltonians $H_{i+1} = H_i \circ \phi_{i+1} = K_{i+1} + \varepsilon^{2^{i+1}} P_{i+1}$ real–analytic on $D_{r_{i+1},s_{i+1}}(y_{i+1})$ such that, for any $0 \leqslant i \leqslant j-1$,

$$\begin{cases}
\|\partial_y^2 K_i\|_{r_i,y_i} \leqslant \mathsf{K}_i \,, \\
\|T_i\| \leqslant \mathsf{T}_i \,, \\
\|P_i\|_{r_i,s_i,y_i} \leqslant M_i \,, \\
\lambda_i \geqslant \frac{4}{5} \log\left(\sigma_i^{2\nu+d} \frac{\alpha^2}{|\varepsilon|^{2^i} M_i K_i}\right) , \\
|\varepsilon|^{2^i} \mathsf{L}_i \leqslant \frac{\sigma_i}{3} \,.
\end{cases} \tag{2.3.60}$$

Observe that for $j = 1$, it is $i = 0$ and (2.3.60) is implied by the definitions of K_0, T_0, λ_0, M_0 and by condition (2.3.54).

[48] Compare (2.3.36).

Because of (2.3.54) and (2.3.60), (2.3.33) holds for H_i and Lemma 2.3.3 can be applied to H_i and one has, for $0 \leqslant i \leqslant j-1$ (see (2.3.32), (2.3.34), (2.3.37) and (2.3.38)):

$$\begin{cases} |y_{i+1}-y_i| \leqslant 2\sigma_i^{\nu+d} r_{i+1}|\varepsilon|^{2^i}\mathsf{L}_i\ , \\ \|T_{i+1}\| \leqslant \|T_i\| + \mathsf{T}_i|\varepsilon|^{2^i}\mathsf{L}_i\ , \\ \|\mathsf{K}_{i+1}\|_{r_{i+1},y_{i+1}} \leqslant \|\mathsf{K}_i\|_{r_i,y_i} + |\varepsilon|^{2^i} M_i\ , \\ \|\partial_y^2\mathsf{K}_{i+1}\|_{r_{i+1},y_{i+1}} \leqslant \|\partial_y^2\mathsf{K}_i\|_{r_i,y_i} + \mathsf{K}_i|\varepsilon|^{2^i}\mathsf{L}_i\ , \\ \|\mathsf{W}_i(\phi_{i+1}-\mathrm{id})\|_{r_{i+1},s_{i+1},y_{i+1}} \leqslant \sigma_i^d\ |\varepsilon|^{2^i}\mathsf{L}_i\ , \\ \|P_{i+1}\|_{r_{i+1},s_{i+1},y_{i+1}} \leqslant M_{i+1} := M_i\mathsf{L}_i\ . \end{cases} \tag{2.3.61}$$

Let $0 \leqslant i \leqslant j-1$. Since

$$\mu_0 \leqslant \varepsilon_\sharp \implies \frac{\alpha}{4d\sqrt{2}\mathsf{K}_0\kappa_0^\nu} \leqslant \frac{5}{96d}\frac{r_0}{\eta_0}\,,$$

then

$$r_1 = \frac{\alpha}{4d\sqrt{2}\mathsf{K}_0\kappa_0^\nu}\ ,$$

and therefore

$$\begin{aligned} r_{i+1} &= \min\left\{\frac{\alpha}{4d\sqrt{2}\mathsf{K}_0\kappa_i^\nu}\,,\,\frac{5}{96d}\frac{r_i}{\eta_0}\right\} \\ &= \min\left\{\frac{\alpha}{4d\sqrt{2}\mathsf{K}_0\kappa_i^\nu}\,,\,\frac{5}{96d\eta_0}\frac{\alpha}{4d\sqrt{2}\mathsf{K}_0\kappa_{i-1}^\nu}\,,\,\left(\frac{5}{96d\eta_0}\right)^2 r_{i-1}\right\} \\ &\ \vdots \\ &= \min\left\{\frac{\alpha}{4d\sqrt{2}\mathsf{K}_0\kappa_i^\nu}\,,\,\frac{5}{96d\eta_0}\frac{\alpha}{4d\sqrt{2}\mathsf{K}_0\kappa_{i-1}^\nu}\,,\,\cdots\,,\,\left(\frac{5}{96d\eta_0}\right)^i r_1\right\} \\ &= \frac{r_1}{4^i}\min\left\{1\,,\,\frac{5}{24d\eta_0}\,,\,\cdots\,,\,\left(\frac{5}{24d\eta_0}\right)^i\right\} \\ &= \left(\frac{5}{96d\eta_0}\right)^i r_1\,. \end{aligned}$$

Thus, since

$$\mu_0 \leqslant \varepsilon_\sharp \implies \mu_0 \leqslant e^{-1} \implies \kappa_0 \geqslant 5\sigma_0^{-1} \geqslant 10\,, \tag{2.3.62}$$

we have

$$\begin{aligned}
|\varepsilon|\mathsf{L}_0(3\sigma_0^{-1}) &= 3|\varepsilon|M_0 \max\left\{\frac{4\sqrt{2}\mathsf{T}_0}{r_0r_1}\sigma_0^{-(\nu+d)}\,,\ \mathsf{C}_7\max\left\{1,\frac{\alpha}{r_0\mathsf{K}_0}\right\}\frac{\mathsf{K}_0}{\alpha^2}\sigma_0^{-2(\nu+d)}\right\}\\
&\leqslant 3\max\left\{4\sqrt{2}\mathsf{T}_0\frac{\alpha}{r_1}\frac{\alpha}{r_0\mathsf{K}_0}\,,\ \mathsf{C}_7\max\left\{1,\frac{\alpha}{r_0\mathsf{K}_0}\right\}\right\}\sigma_0^{-2(\nu+d)-1}\frac{\mathsf{K}_0|\varepsilon|M_0}{\alpha^2}\\
&= 3\max\left\{32d\eta_0\kappa_0^{\nu}\frac{\alpha}{r_0\mathsf{K}_0}\,,\ \mathsf{C}_7\max\left\{1,\frac{\alpha}{r_0\mathsf{K}_0}\right\}\right\}\sigma_0^{-2(\nu+d)-1}\frac{\mathsf{K}_0|\varepsilon|M_0}{\alpha^2}\\
&\leqslant 3\max\left\{32d\,,\ 10^{-\nu}\mathsf{C}_7\right\}\cdot\max\left\{1,\frac{\alpha}{r_0\mathsf{K}_0}\right\}\eta_0\ \sigma_0^{-2(\nu+d)-1}\frac{\mathsf{K}_0|\varepsilon|M_0}{\alpha^2}\kappa_0^{\nu}\\
&= \mathsf{C}_8\ \max\left\{1,\frac{\alpha}{r_0\mathsf{K}_0}\right\}\eta_0\ \sigma_0^{-(3\nu+2d+1)}\mu_0\lambda_0^{\nu}\\
&= \mathsf{f}_* \overset{(2.3.54)}{\leqslant} 1\ .
\end{aligned}$$

Now, fix $i \geqslant 1$. We have

$$r_i\mathsf{K}_i \leqslant r_1\mathsf{K}_0\sqrt{2} \overset{(2.3.62)}{\leqslant} \frac{\alpha}{4d\cdot 10^{\nu}} < \alpha \tag{2.3.63}$$

so that

$$\begin{aligned}
|\varepsilon|^{2^i}\mathsf{L}_i(3\sigma_i^{-1}) &= 3|\varepsilon|^{2^i}M_i\max\left\{\frac{4\sqrt{2}\mathsf{T}_0}{r_ir_{i+1}}\sigma_i^{-(\nu+d)}\,,\ \mathsf{C}_7\max\left\{1,\frac{\alpha}{r_i\mathsf{K}_i}\right\}\frac{\mathsf{K}_0}{\alpha^2}\sigma_i^{-2(\nu+d)}\right\}\sigma_i^{-1}\\
&= 3|\varepsilon|^{2^i}M_i\max\left\{\frac{4\sqrt{2}\mathsf{T}_0}{r_ir_{i+1}}\sigma_i^{-(\nu+d)}\,,\ \mathsf{C}_7\frac{1}{\alpha r_i}\sigma_i^{-2(\nu+d)}\right\}\sigma_i^{-1}\\
&\leqslant 3\max\left\{\frac{4\sqrt{2}\alpha\mathsf{T}_0}{r_{i+1}}\,,\ \mathsf{C}_7\right\}\sigma_i^{-2(\nu+d)-1}\frac{|\varepsilon|^{2^i}M_i}{\alpha r_i}\\
&= 3\max\left\{32d\eta_0\kappa_0^{\nu}\left(\frac{96d\eta_0}{5}\right)^i\,,\ \mathsf{C}_7\right\}\sigma_i^{-2(\nu+d)-1}\frac{|\varepsilon|^{2^i}M_i}{\alpha r_i}\\
&\leqslant 3\max\left\{32d\,,\ 10^{-\nu}\mathsf{C}_7\right\}\left(\frac{96d\eta_0}{5}\right)^i\eta_0\ \kappa_0^{\nu}\ \sigma_i^{-2(\nu+d)-1}\frac{|\varepsilon|^{2^i}M_i}{\alpha r_i}\\
&\leqslant 12d\sqrt{2}\max\left\{32d\,,\ 10^{-\nu}\mathsf{C}_7\right\}\left(\frac{96d\eta_0}{5}\right)^{2i-1}\eta_0\ \kappa_0^{2\nu}\ \sigma_i^{-2(\nu+d)-1}\frac{\mathsf{K}_0|\varepsilon|^{2^i}M_i}{\alpha^2}\\
&= \mathsf{C}_9\frac{\mathsf{K}_0}{\alpha^2}\sigma_0^{-(4\nu+2d+1)}\ \lambda_0^{2\nu}\ \left(2^{2(\nu+d)+11}3^2 5^{-2}d^2\eta_0^2\right)^i\ |\varepsilon|^{2^i}M_i\\
&= \mathsf{e}_*\mathsf{d}_*^i|\varepsilon|^{2^i}M_i =: \frac{\theta_i}{\mathsf{d}_*}\ ,
\end{aligned}$$

so that

$$\mathsf{L}_i < \mathsf{e}_* \, \mathsf{d}_*^i M_i \ ,$$

thus by last relation in (2.3.61), for any $1 \leqslant i \leqslant j-1$,

$$|\varepsilon|^{2^{i+1}} M_{i+1} < \mathsf{e}_* \mathsf{d}_*^i |\varepsilon|^{2^{i+1}} M_i^2$$

i.e. $\theta_{i+1} < \theta_i^2$, which iterated, yields $\theta_i \leqslant \theta_1^{2^{i-1}}$ for $1 \leqslant i \leqslant j$. Next, we show that, thanks to (2.3.54), (2.3.60) holds also for $i = j$. In fact, by (2.3.60) and (2.3.61), we have

$$\|T_{i+1}\| \leqslant \|T_i\| + \mathsf{T}_i |\varepsilon|^{2^i} \mathsf{L}_i \leqslant \mathsf{T}_i + \mathsf{T}_i \frac{\sigma_i}{3} = \mathsf{T}_{i+1} \, ,$$

and similarly for $\|\partial_y^2 K_{i+1}\|_{r_{i+1},y_{i+1}}$. Now, by $(2.3.57)_{i=j}$,

$$|\varepsilon|^{2^j} \mathsf{L}_j (3\sigma_j^{-1}) \leqslant \frac{\theta_j}{\mathsf{d}_*} \leqslant \frac{1}{\mathsf{d}_*} (\mathsf{e}_* \, d_*^2 \varepsilon^2 M_1)^{2^{j-1}} \leqslant \frac{1}{\mathsf{d}_*} \left(\frac{\sigma_0}{3} \mathsf{f}_* \mathsf{e}_* d_*^2 |\varepsilon| M_0 \right)^{2^{j-1}} \overset{(2.3.54)}{\leqslant} \frac{1}{\mathsf{d}_*} < 1 \ ,$$

which implies the last inequality in (2.3.60) with $i = j$.
Next, we check the fourth inequality in (2.3.60) for $i = j$. We have[49]

$$\begin{aligned} \lambda_j &= 2\lambda_{j-1} \\ &\overset{(2.3.60)_{i=j-1}}{\geqslant} \frac{4}{5} \log \left(\sigma_{j-1}^{2(2\nu+d)} \frac{\alpha^4}{|\varepsilon|^{2^j} M_{j-1}^2 K_{j-1}^2} \right) \\ &\geqslant \frac{4}{5} \log \left(\sigma_{j-1}^{2(2\nu+d)} \frac{\alpha^4}{|\varepsilon|^{2^j} M_{j-1} K_{j-1}^2} \cdot \sigma_{j-1}^{-2(\nu+d)} \frac{\mathsf{C}_7 \mathsf{K}_0}{\alpha^2 \mathsf{L}_{j-1}} \right) \\ &= \frac{4}{5} \log \left(\sigma_{j-1}^{2\nu} \frac{\alpha^2}{|\varepsilon|^{2^j} M_j K_{j-1}} \cdot \frac{\mathsf{C}_7 \mathsf{K}_0}{\mathsf{K}_{j-1}} \right) \\ &> \frac{4}{5} \log \left(\sigma_j^{2\nu+d} \frac{\alpha^2}{|\varepsilon|^{2^j} M_j K_j} \right) \ . \end{aligned}$$

The proof of the induction is then finished and one can construct an *infinite sequence* of Arnold's transformations satisfying (2.3.60), (2.3.61) and (2.3.57) *for all* $i \geqslant 0$.
Next, we prove that ϕ^j is convergent by proving that it is Cauchy. For any $j \geqslant 3$, we have,

[49]Notice that $\mathsf{L}_i \geqslant M_i \mathsf{C}_7 \frac{\mathsf{K}_0}{\alpha^2} \sigma_i^{-2(\nu+d)}$, $\forall \ i \geqslant 0$.

using again Cauchy's estimate,[50]

$$
\begin{aligned}
\|\mathsf{W}_{j-1}(\phi^{j-1}-\phi^{j-2})\|_{r_j,s_j,y_j} &= \|\mathsf{W}_{j-1}\phi^{j-2}\circ\phi_{j-1}-\mathsf{W}_{j-1}\phi^{j-2}\|_{r_j,s_j,y_j}\\
&\overset{(2.3.59)}{\leqslant} \|\mathsf{W}_{j-1}D\phi^{j-2}\mathsf{W}_{j-1}^{-1}\|_{2r_{j-1}/3,s_{j-1},y_{j-1}}\,\|\mathsf{W}_{j-1}(\phi_{j-1}-\mathrm{id})\|_{r_j,s_j,y_j}\\
&\overset{(2.3.61)}{\leqslant} \max\left(r_{j-1}\frac{3}{r_{j-1}},\frac{3}{2\sigma_{j-1}}\right)\|\mathsf{W}_{j-1}\phi^{j-2}\|_{r_{j-1},s_{j-1},y_{j-1}}\times\\
&\qquad\times\|\mathsf{W}_{j-1}(\phi_{j-1}-\mathrm{id})\|_{r_j,s_j,y_j}\\
&= \frac{3}{2\sigma_{j-1}}\|\mathsf{W}_{j-1}\phi^{j-2}\|_{r_{j-1},s_{j-1},y_{j-1}}\,\|\mathsf{W}_{j-1}(\phi_{j-1}-\mathrm{id})\|_{r_j,s_j,y_j}\\
&\leqslant \frac{1}{2}\|\mathsf{W}_{j-1}\phi^{j-2}\|_{r_{j-1},s_{j-1},y_{j-1}}\cdot\sigma_{j-1}^{d}\left(|\varepsilon|^{2^{j-1}}\mathsf{L}_{j-1}3\sigma_{i-1}^{-1}\right)\\
&\leqslant \frac{1}{2}\|\mathsf{W}_{j-1}\phi_1\|_{r_2,s_2,y_2}\cdot\sigma_{j-1}^{d}\;\theta_{j-1}\\
&\leqslant \frac{1}{2}\left(\prod_{i=1}^{j-2}\|\mathsf{W}_{i+1}\mathsf{W}_i^{-1}\|\right)\|\mathsf{W}_1\phi_1\|_{r_2,s_2,y_2}\cdot\sigma_{j-1}^{d}\;\theta_{j-1}\\
&\overset{(2.3.63)}{=} \frac{1}{2}\left(\prod_{i=1}^{j-2}\frac{r_i}{r_{i+1}}\right)\|\mathsf{W}_1\phi_1\|_{r_2,s_2,y_2}\cdot\sigma_{j-1}^{d}\;\theta_{j-1}\\
&= \frac{r_1}{2r_{j-1}}\|\mathsf{W}_1\phi_1\|_{r_2,s_2,y_2}\cdot\sigma_{j-1}^{d}\;\theta_{j-1}\\
&= \frac{48d}{5}\sigma_2^{d}\;\eta_0\;\|\mathsf{W}_1\phi_1\|_{r_2,s_2,y_2}\cdot\left(\frac{3d\cdot 2^{5-d}\eta_0}{5}\right)^{j-3}\cdot\;\theta_1^{2^{j-2}}\\
&\leqslant \frac{48d}{5}\sigma_2^{d}\;\eta_0\;\|\mathsf{W}_1\phi_1\|_{r_2,s_2,y_2}\cdot\left(\frac{3d\cdot 2^{5-d}\eta_0}{5}\right)^{2^{j-4}}\cdot\;\theta_1^{2^{j-2}}\\
&= \frac{48d}{5}\sigma_2^{d}\;\eta_0\;\|\mathsf{W}_1\phi_1\|_{r_2,s_2,y_2}\cdot\left(\left(\frac{3d\cdot 2^{5-d}}{5}\right)^{\frac14}\eta_0^{\frac14}\;\theta_1\right)^{2^{j-2}}\\
&= \frac{48d}{5}\sigma_2^{d}\;\eta_0\;\|\mathsf{W}_1\phi_1\|_{r_2,s_2,y_2}\cdot\left(\mathsf{C}_{10}\;\eta_0^{\frac14}\;\theta_1\right)^{2^{j-2}}\quad.
\end{aligned}
$$

[50]Notice that (2.3.63) $\implies W_i=\mathrm{diag}\left(\frac{1}{r_i}\,\mathbb{1}_d\,,\mathbb{1}_d\right)\,,\ \forall\, i\geqslant 1$ and recall that $2^{i-1}\geqslant i,\ \forall\, i\geqslant 0$.

Therefore, for any $n \geqslant 1$, $j \geqslant 0$,

$$\begin{aligned}
\|\mathsf{W}_1(\phi^{n+j+1}-\phi^n)\|_{r_{n+j+2},s_{n+j+2},y_{n+j+2}} &\leqslant \sum_{i=n}^{n+j} \|\mathsf{W}_1(\phi^{i+1}-\phi^i)\|_{r_{i+2},s_{i+2},y_{i+2}} \\
&\leqslant \sum_{i=n}^{n+j}\left(\prod_{k=1}^{i}\|\mathsf{W}_k\mathsf{W}_{k+1}^{-1}\|\right)\|\mathsf{W}_{i+1}(\phi^{i+1}-\phi^i)\|_{r_{i+2},s_{i+2},y_{i+2}} \\
&\overset{(5.3.46)}{=} \sum_{i=n}^{n+j}\prod_{k=1}^{i}\max\left\{1\,,\,\frac{r_{k+1}}{r_k}\right\}\|\mathsf{W}_{i+1}(\phi^{i+1}-\phi^i)\|_{r_{i+2},s_{i+2},y_{i+2}} \\
&= \sum_{i=n}^{n+j}\|\mathsf{W}_{i+1}(\phi^{i+1}-\phi^i)\|_{r_{i+2},s_{i+2},y_{i+2}} \\
&\leqslant \frac{48d}{5}\sigma_2^d\,\eta_0\,\|\mathsf{W}_1\phi_1\|_{r_2,s_2,y_2}\sum_{i=n}^{n+j}\left(\mathsf{C}_{10}\,\eta_0^{\frac{1}{4}}\,\theta_1\right)^{2^i}
\end{aligned}$$

and

$$\mathsf{C}_{10}\,\eta_0^{\frac{1}{4}}\,\theta_1 \overset{(2.3.54)}{<} 1\,.$$

Hence, ϕ^j converges uniformly on $\{y_*\}\times\mathbb{T}^d_{s_*}$ to some ϕ^*, which is then real–analytic map in $x\in\mathbb{T}^d_{s_*}$.
To estimate $|\mathsf{W}_0(\phi^*-\mathrm{id})|$ on $\{y_*\}\times\mathbb{T}^d_{s_*}$, observe that , for $i\geqslant 1$,[51]

$$\sigma_i^d\,|\varepsilon|^{2^i}\mathsf{L}_i \leqslant \frac{\sigma_0^{d+1}}{3\cdot 2^{i(d+1)}}\,\frac{\theta_1^{2^{i-1}}}{\mathsf{d}_*} \leqslant \frac{\sigma_0^{d+1}}{3\cdot 2^{i(d+1)}\mathsf{d}_*}\theta_1^i = \frac{\sigma_0^{d+1}}{3\mathsf{d}_*}\left(\frac{\theta_1}{2^{d+1}}\right)^i$$

and therefore

$$\sum_{i\geqslant 1}\sigma_i^d\,|\varepsilon|^{2^i}\mathsf{L}_i \leqslant \frac{\sigma_0^{d+1}}{3\mathsf{d}_*}\sum_{i\geqslant 1}\left(\frac{\theta_1}{2^{d+1}}\right)^i \leqslant \frac{\sigma_0^{d+1}\,\theta_1}{3\cdot 2^d\,\mathsf{d}_*} \overset{(2.3.54)}{<} \frac{1}{2}\,\sigma_0^{d+1}\,.$$

Moreover, for any $i\geqslant 1$,

$$\begin{aligned}
\|\mathsf{W}_1(\phi^i-\mathrm{id})\|_{r_{i+1},s_{i+1},y_{i+1}} &\leqslant \|\mathsf{W}_1(\phi^{i-1}\circ\phi_i-\phi_i)\|_{r_{i+1},s_{i+1},y_{i+1}} + \|\mathsf{W}_1(\phi_i-\mathrm{id})\|_{r_{i+1},s_{i+1},y_{i+1}} \\
&\leqslant \|\mathsf{W}_1(\phi^{i-1}-\mathrm{id})\|_{r_i,s_i,y_i} + \left(\prod_{j=0}^{i-1}\|\mathsf{W}_j\mathsf{W}_{j+1}^{-1}\|\right)\|\mathsf{W}_i(\phi_i-\mathrm{id})\|_{r_{i+1},s_{i+1},y_{i+1}} \\
&= \|\mathsf{W}_1(\phi^{i-1}-\mathrm{id})\|_{r_i,s_i,y_i} + \|\mathsf{W}_i(\phi_i-\mathrm{id})\|_{r_{i+1},s_{i+1},y_{i+1}} \\
&= \|\mathsf{W}_1(\phi^{i-1}-\mathrm{id})\|_{r_i,s_i,y_i} + \|\mathsf{W}_i(\phi_i-\mathrm{id})\|_{r_{i+1},s_{i+1},y_{i+1}} \\
&\leqslant \|\mathsf{W}_1(\phi^{i-1}-\mathrm{id})\|_{r_i,s_i,y_i} + \sigma_i^d\,|\varepsilon|^{2^i}\mathsf{L}_i\,,
\end{aligned}$$

[51] Recall that $2^i\geqslant i+1$, $\forall\, i\geqslant 0$.

which iterated yields

$$\begin{aligned}\|\mathsf{W}_1(\phi^i - \mathrm{id})\|_{r_i,s_i,y_i} &\leqslant \sum_{k=1}^{i-1} \sigma_k^d\, |\varepsilon|^{2^k} \mathsf{L}_k \\ &\leqslant \sum_{k\geqslant 1} \sigma_k^d\, |\varepsilon|^{2^k} \mathsf{L}_k \\ &\leqslant \frac{1}{2}\, \sigma_0^{d+1}\,.\end{aligned}$$

Therefore, taking the limit over i completes yields, uniformly on $\{y_*\} \times \mathbb{T}^d_{s_*}$,

$$|\mathsf{W}_1(\phi^* - \mathrm{id})| \leqslant \frac{1}{2}\, \sigma_0^{d+1}\,.$$

Now, to complete the proof of the Lemma and, consequently, of the Theorem, just set $\phi_* := \phi_0 \circ \phi^*$ and observe that, uniformly on $\{y_*\} \times \mathbb{T}^d_{s_*}$,

$$\begin{aligned}|\mathsf{W}_0(\phi_* - \mathrm{id})| &\leqslant |\mathsf{W}_0(\phi_0 \circ \phi^* - \phi^*)| + |\mathsf{W}_0(\phi^* - \mathrm{id})| \\ &\leqslant \|\mathsf{W}_0(\phi_0 - \mathrm{id})\|_{r_1,s_1,y_1} + \|\mathsf{W}_0\mathsf{W}_1^{-1}\|\, |\mathsf{W}_1(\phi^* - \mathrm{id})| \\ &= \|\mathsf{W}_0(\phi_0 - \mathrm{id})\|_{r_1,s_1,y_1} + \max\left\{\frac{r_1\mathsf{K}_0}{\alpha}\,, \frac{r_1}{r_0}\,, 1\right\} |\mathsf{W}_1(\phi^* - \mathrm{id})| \\ &\overset{(2.3.63)}{=} \|\mathsf{W}_0(\phi_0 - \mathrm{id})\|_{r_1,s_1,y_1} + |\mathsf{W}_1(\phi^* - \mathrm{id})| \\ &\leqslant \sigma_0^d\, |\varepsilon| \mathsf{L}_0 + \frac{1}{2}\, \sigma_0^{d+1} \\ &\overset{(2.3.54)}{\leqslant} \frac{1}{3}\, \sigma_0^{d+1} + \frac{1}{2}\, \sigma_0^{d+1} \\ &< \sigma_0^{d+1}\,.\end{aligned}$$

∎

2.3.3 Proof of Theorem 2.1.4

As usual, the proof is inductive: at each step $j \in \mathbb{N}$, a small perturbation of some normal form $N_j = e_j(\omega) + \omega \cdot y$,

$$H_j = N_j + P_j$$

is considered. Then, a coordinates and parameter transformation $\mathcal{F}_j$ is constructed so that

$$H_j \circ \mathcal{F}_j = N_{j+1} + P_{j+1}$$

with another normal form N_{j+1}, some much smaller error term P_{j+1} satisfying

$$\|P_{j+1}\| \leqslant C\|P_j\|^{\frac{3}{2}}$$

for some constant $C > 0$ and the sequence $\mathcal{F}^{j+1} := \mathcal{F}_0 \circ \cdots \circ \mathcal{F}_j$ converges to an embedding of an invariant Kronecker torus.
The first step, called **KAM step**, will be then to describe one cycle of this iterative scheme in which, for readability, we drop the subscribe j and consider a generic hamiltonian $H = N + P$. First of all, instead of H, we consider the hamiltonian $\bar{H}$ obtained from H by first linearizing the perturbation P in y and then truncating its Fourier series in x at some suitable high order κ.
The transformation $\mathcal{F}$ is of the form

$$\mathcal{F} := (\Phi, \varphi) := \left(\widetilde{\Phi} \circ (\pi_1, \pi_2; \varphi \circ \pi_3), \varphi \circ \pi_3\right) : (y, x; \omega) \mapsto \left(\widetilde{\Phi}(y, x; \varphi(\omega)), \varphi(\omega)\right),$$

where $\widetilde{\Phi}$ is obtained as the time–1–map of the flow Φ_F^t of some hamiltonian F and

$$\pi_j \colon \mathbb{C}^d \times \mathbb{C}^d \times \mathbb{C}^d \to \mathbb{C}^d,\ j = 1, 2, 3 : \pi_1(y, x, \omega) = y,\ \pi_2(y, x, \omega) = x,\ \pi_3(y, x, \omega) = \omega\ . \tag{2.3.64}$$

In particular, Φ is then symplectic for ω fixed[52]. Then, we iterate this cycle and prove the convergences.

In all this section, the sup–norm on $D_{r,s} \times \Omega_{\alpha,h}$ will be denoted by $\|\cdot\|_{r,s,h} := \|\cdot\|_{r,s,h,d}$, while on $D_{r,s} \times \mathbb{C}^d$ (resp. $D_{r,s} \times \mathbb{R}^d$), it will be denoted by $\|\cdot\|_{r,s,\infty}$ (resp. $\|\cdot\|_{r,s,0}$).

2.3.3.1 KAM step

Lemma 2.3.5 (KAM step) *Assume that* $\|P\|_{r,s,h} \leqslant \epsilon$ *with*

(a) $\epsilon \leqslant \frac{1}{C_4}\alpha\eta r\sigma^\nu$,

(b) $\epsilon \leqslant \frac{1}{C_6}hr$,

(c) $h \leqslant \frac{\alpha}{2\kappa^\nu}$,

[52] Indeed, denoting the Lie derivative by $\mathcal{L}$ and the contraction operator by ι, we have

$$\frac{d}{dt}\left(\phi_F^t\right)^* \varpi = \left(\phi_F^t\right)^* \mathcal{L}_{X_F}\varpi = \left(\phi_F^t\right)^* \left(\iota_{X_F}\underbrace{d\varpi}_{0} + d\underbrace{\iota_{X_F}\varpi}_{-dF}\right) = 0 \implies \left(\phi_F^t\right)^* \varpi = \left(\phi_F^0\right)^* \varpi = \mathrm{id}^*\varpi = \varpi.$$

for some $0 < \eta < 1/8$, $0 < \sigma < s/10$ *and sufficiently large* $\kappa > \frac{d-1}{\sigma}$. *Then there exist* $\varphi\colon \mathbb{R}^d \to \mathbb{R}^d$ *a* C^∞*–diffeomorphism with* $\varphi(\Omega) = \Omega$ *and* $\varphi \equiv \mathrm{id}$ *on* $\mathbb{R}^d \backslash \Omega$, $\widetilde{\Phi}\colon D_{\eta r, s-5\sigma} \times \mathbb{C}^d \to D_{r,s}$ *a (ω–) family of symplectic transformations parametrized over* $\mathbb{C}^d$*, each being real–analytic with holomorphic extention to* $D_{\eta r,s-5\sigma}$ *and* C^∞ *in* ω *on* $\mathbb{R}^d$ *and such that, if* $\mathcal{F} := (\Phi, \varphi) := \left(\widetilde{\Phi} \circ (\pi_1, \pi_2; \varphi \circ \pi_3), \varphi \circ \pi_3\right)$ *, the following hold: its restriction map*

$$\mathcal{F}_h = (\Phi, \varphi)\colon D_{\eta r, s-5\sigma} \times \Omega_{\alpha, \frac{h}{4}} \to D_{r,s} \times \Omega_{\alpha,h}$$

is well–defined, real–analytic (in all arguments), $H \circ \mathcal{F}_h = N_+ + P_+$ *with another normal form* $N_+ := e_+(\omega) + \omega \cdot y$ *and*

$$\|P_+\|_{\eta r, s-5\sigma, \frac{h}{4}} \leqslant \frac{\sqrt{C_{10}}}{3 \cdot 2^{\bar\nu}} \left(\frac{\epsilon^2}{\alpha r \sigma^{\bar\nu}} + (\eta^2 + \kappa^n e^{-\kappa\sigma})\epsilon\right).$$

Moreover[53]*,*

$$\|\bar W(\mathcal{F} - \mathrm{id})\|_{\frac{r}{4}, s-4\sigma, 0} \leqslant \left(\frac{\sigma_0}{\sigma}\right)^{2\bar\nu - 1} \cdot \max\left(4C_3 \frac{\epsilon}{\alpha r \sigma^{\bar\nu}}, \frac{C_3}{C_0}\frac{\epsilon}{rh}\right) , \tag{2.3.65}$$

$$\left\|\bar W(D\mathcal{F} - \mathrm{Id})\bar W^{-1}\right\|_{\frac{r}{8}, s-5\sigma, 0} \leqslant \left(\frac{\sigma_0}{\sigma}\right)^{2\bar\nu - 1} \cdot \max\left(dC_5 \frac{\epsilon}{\alpha r \sigma^{\bar\nu}}, \frac{C_5}{C_0}\frac{\epsilon}{rh}\right) , \tag{2.3.66}$$

$$\|\varphi - \mathrm{id}\|_\infty , \, h\, \|D\varphi - \mathrm{Id}\|_0 \leqslant \frac{C_5}{2C_0}\frac{\epsilon}{r} , \tag{2.3.67}$$

for a given $\sigma_0 \geqslant \sigma$*, with*

$$\bar W := \ \mathrm{diag}\left(\frac{1}{r}\mathbb{1}_d,\ \frac{1}{\sigma}\left(\frac{\sigma_0}{\sigma}\right)^{2\bar\nu-1}\mathbb{1}_d,\ \frac{1}{h}\mathbb{1}_d\right).$$

Proof For convenience, we will follow the scheme of the proof in [Pös01] and add two more steps allowing us, later, as we said, to estimate the Lipschitz' semi–norm of the symplectic transformation we are going to build–up without invoking the Whitney's extension theorem.

1. Truncation. Let $Q(x, \omega) := P(0, x, \omega) + P_y(0, x, \omega) \cdot y$, the linerization of P and $R := T_\kappa Q$. Then by Cauchy's estimate we get

$$\|Q\|_{r,s,h} \leqslant \|P\|_{r,s,h} + d\frac{\|P\|_{r,s,h}}{r} r \leqslant (d+1)\epsilon,$$

[53] We denote by $D\Phi$ and $D\varphi$, respectively, the jacobian of Φ with respect to (y, x, ω) and of φ with respect to ω.

And

$$\begin{aligned}\|P-Q\|_{2\eta r,s,h} &= \left\|\int_0^1 (1-t)P_{yy}(ty,x,\omega)(y,y)dt\right\|_{2\eta r,s,h}\\ &\leqslant \sum_{1\leqslant j,l\leqslant d}\ \sup_{(y,x,\omega)\in D_{2\eta r,s}\times\Omega_{\alpha,h}}\int_0^1 (1-t)\|\partial^2_{y_jy_l}P(ty,x,\omega)\|\|y_j\|\|y_l\|dt\\ &\leqslant \sum_{1\leqslant j,l\leqslant d}\int_0^1 (1-t)\frac{\|P\|_{r,s,h}}{(r-2\eta r)^2}(2\eta r)^2 dt\\ &= \frac{2\eta^2 d^2}{(1-2\eta)^2}\epsilon\\ &\leqslant \frac{32d^2}{9}\eta^2\epsilon.\end{aligned}$$

By Lemma 2.2.4, we have

$$\|R-Q\|_{r,s-\sigma,h}\leqslant 4^d C_2\kappa^d\, e^{-\kappa\sigma}\|Q\|_{r,s,h}\leqslant 4^d(d+1)C_2\kappa^d\, e^{-\kappa\sigma}\epsilon,$$

and therefore

$$\begin{aligned}\|R\|_{r,s-\sigma,h} &\leqslant \|R-Q\|_{r,s-\sigma,h}+\|Q\|_{r,s-\sigma,h}\\ &\leqslant \left(4^d C_2\kappa^d\, e^{-\kappa\sigma}+1\right)(d+1)\epsilon\leqslant 2(d+1)\epsilon=\frac{C_3}{2C_0}\epsilon,\end{aligned}$$

because $C_2\leqslant C_{10}$ and, later, κ will be chosen so that $\kappa^d\, e^{-\kappa\sigma}\leqslant (4^{\bar\nu}C_{10})^{-1}$.

2. Extending the Diophantine condition. The Diophantine condition (compare (1.2.1)) is assumed to hold only on Ω_α. Nevertheless, given $\omega\in\Omega_{\alpha,h}$, there exits $\omega_*\in\Omega_\alpha$ such that $|\omega-\omega_*|<h$, so that, for any $|k|_1\leqslant\kappa$

$$|k\cdot(\omega-\omega_*)|\leqslant |k|_1\cdot|\omega-\omega_*|\leqslant \kappa h\overset{(c)}{\leqslant}\frac{\alpha}{2\kappa^\tau}\leqslant\frac{\alpha}{2|k|_1^\tau}$$

and thanks to (1.2.1), we get, for any $\omega\in\Omega_{\alpha,h}$

$$|k\cdot\omega|\geqslant\frac{\alpha}{2|k|_1^\tau},\qquad \forall\, 0\neq|k|_1\leqslant\kappa. \tag{2.3.68}$$

3. Finding the hamiltonian F by solving a homological equation. We have

$$H=\bar H+(P-Q)+(Q-R)\text{ with }\bar H=N+R$$

Let's remind that we are looking at for a hamiltonian F such that its flow ϕ_F^t satisfies[54]

$$H \circ \phi_F^1 = N_+^1 + P_+^1,$$

for some hamiltonian N_+^1 closed to a normal form and much smaller error term P_+^1. We have

$$H \circ \phi_F^1 = \bar{H} \circ \phi_F^1 + (P - Q) \circ \phi_F^1 + (Q - R) \circ \phi_F^1.$$

Next we expend $\bar{H} \circ \phi_F^t$ around $t = 0$ to get[55]

$$\begin{aligned}
\bar{H} \circ \phi_F^1 &= N \circ \phi_F^1 + R \circ \phi_F^1 \\
&= N + \frac{d}{ds} N \circ \phi_F^s\big|_{s=0} + \int_0^1 (1-t) \frac{d^2}{ds^2} N \circ \phi_F^s\big|_{s=t}\, dt + \\
&\quad + R + \int_0^1 \frac{d}{ds} R \circ \phi_F^s\big|_{s=t}\, dt \\
&= N + \{N, F\} + \int_0^1 (1-t)\{\{N, F\}, F\} \circ \phi_F^t\, dt + R + \int_0^1 \{R, F\} \circ \phi_F^t\, dt \\
&= \underbrace{N + [R]}_{:= N_+^1} + \{N, F\} + (R - [R]) + \underbrace{\int_0^1 \{(1-t)\{N, F\} + R, F\} \circ \phi_F^t\, dt}_{:= P_+^0}
\end{aligned}$$

Since Q is affine in the variable y , then so is R and a fortiori $[R]$; moreover $[R]$ does not depend on x. Therefore, there exist analytic functions $e_+^0 \colon \omega \to e_+^0(\omega)$ and $v \colon \omega \to v(\omega)$ such that $[R](y, \omega) = e_+^0(\omega) + v(\omega) \cdot y$ (in fact $v = [R_y]$) so that

$$N_+^1 = \underbrace{e(\omega) + e_+^0(\omega)}_{:= e_+^1(\omega)} + \underbrace{(\omega + v(\omega))}_{:= \rho(\omega)} \cdot y \tag{2.3.69}$$

Let

$$P_+^1 := P_+^0 + (P - Q) \circ \phi_F^1 + (Q - R) \circ \phi_F^1 \tag{2.3.70}$$

The main point is then to determine F by solving the homological equation

$$\{N, F\} + (R - [R]) = 0$$

[54] In fact, rigorously, one should write $H \circ (\phi_F^1, \mathrm{id})$ and so on.

[55] Given a function K and denoting the Poisson bracket by $\{\cdot, \cdot\}$, we have

$$\frac{d}{dt} K \circ \phi_F^t = dK(\phi_F^t) \cdot \frac{d}{dt} \phi_F^t = dK(\phi_F^t) \cdot \mathbb{J} dF(\phi_F^t) = \{K, F\} \circ \phi_F^t.$$

$i.e.$ (recall that $N = e(\omega) + \omega \cdot y$)

$$D_\omega F = \{F, N\} = R - [R] \tag{2.3.71}$$

so that we have

$$(1-t)\{N, F\} + R = (1-t)([R] - R) + R = (1-t)[R] + tR,$$

$$P_+^0 = \int_0^1 \{(1-t)[R] + tR, F\} \circ \phi_F^t \, dt, \tag{2.3.72}$$

and

$$H \circ \phi_F^1 = N_+^1 + P_+^1. \tag{2.3.73}$$

Since $[R - [R]] = 0$, $\|[R]\|_{r,h} \leqslant \|R\|_{r,s-\sigma,h}$ and $\|R - [R]\|_{r,s-\sigma,h} \leqslant 2\|R\|_{r,s-\sigma,h} < \infty$, Lemma 2.2.5 applies to (2.3.71) and we find F with

$$\|F\|_{r,s-2\sigma,h} \leqslant \frac{2C_0}{\alpha\sigma^\tau}\|R - [R]\|_{r,s-\sigma,h} \leqslant \frac{4C_0}{\alpha\sigma^\tau}\|R\|_{r,s-\sigma,h} \leqslant 2C_3\frac{\epsilon}{\alpha\sigma^\tau} \tag{2.3.74}$$

Then by Cauchy's estimate we get

$$\|F_x\|_{r,s-3\sigma,h} \leqslant \frac{\|F\|_{r,s-2\sigma,h}}{\sigma} \leqslant 2C_3\frac{\epsilon}{\alpha\sigma^{\bar\nu}},$$

$$\|F_y\|_{\frac{r}{2},s-2\sigma,h} \leqslant \frac{2\|F\|_{r,s-2\sigma,h}}{r} \leqslant 4C_3\frac{\epsilon}{\alpha r\sigma^\tau}$$

and

$$\|F_\omega\|_{r,s-2\sigma,\frac{h}{2}} \leqslant \frac{2\|F\|_{r,s-2\sigma,h}}{h} \leqslant 4C_3\frac{\epsilon}{\alpha\sigma^\tau h} \tag{2.3.75}$$

so that

$$\frac{1}{r}\|F_x\|_{\frac{r}{2},s-3\sigma,h},\ \frac{1}{\sigma}\|F_y\|_{\frac{r}{2},s-3\sigma,h} \leqslant 4C_3\frac{\epsilon}{\alpha r\sigma^{\bar\nu}}$$

and by using assumption (a), we get

$$\|F_x\|_{r,s-3\sigma,h} \leqslant \frac{2C_3}{C_4}\eta r \leqslant \eta r \leqslant \frac{r}{8} \tag{2.3.76}$$

$$\|F_y\|_{\frac{r}{2},s-2\sigma,h} \leqslant \frac{4C_3}{C_4}\eta\sigma \leqslant \sigma \tag{2.3.77}$$

4. Extending the hamiltonian F. Thanks to Lemma 2.2.1, there exists a cut–off $\chi_1 \in C(\mathbb{C}^d) \cap C^\infty(\mathbb{R}^d)$ with $0 \leqslant \chi_1 \leqslant 1$, $supp\chi_1 \subset \Omega_{\alpha,\frac{h}{2}}$ and $\chi_1 \equiv 1$ on $\Omega_{\alpha,\frac{h}{4}}$. Now we extend F, witch we call $\tilde F$, as follows: $\tilde F \equiv 0$ on $D_{r,s-2\sigma} \times \left(\mathbb{C}^d \backslash \Omega_{\alpha,\frac{h}{2}}\right)$ and $\tilde F(y,x,\omega) = \chi_1(\omega)F(y,x,\omega)$ on $D_{r,s-2\sigma} \times \Omega_{\alpha,\frac{h}{2}}$. Thus

(i) $\tilde{F}$ concide with F on $D_{r,s-2\sigma}\times\Omega_{\alpha,\frac{h}{4}}$, is continuous on $D_{r,s-2\sigma}\times\mathbb{C}^d$, C^∞ on $D_{r,s-2\sigma}\times\mathbb{R}^d$ and for any $\omega\in\mathbb{C}^d$ given, the map $(y,x)\mapsto\tilde{F}(y,x,\omega)$ is real–analytic with holomorphic extention to $D_{r,s-2\sigma}$.

(ii)

$$\|\tilde{F}\|_{r,s-2\sigma,\infty}\overset{def}{=}\sup_{D_{r,s-2\sigma}\times\mathbb{C}^d}\|\tilde{F}\|\leqslant\|F\|_{r,s-2\sigma,h}\leqslant 2C_3\frac{\epsilon}{\alpha\sigma^\tau}\,,\tag{2.3.78}$$

$$\|\tilde{F}_y\|_{\frac{r}{2},s-2\sigma,\infty}\leqslant\|F_y\|_{\frac{r}{2},s-2\sigma,h}\leqslant 4C_3\frac{\epsilon}{\alpha r\sigma^\tau}\leqslant\sigma\,,\tag{2.3.79}$$

$$\|\tilde{F}_x\|_{r,s-3\sigma,\infty}\leqslant\|F_x\|_{r,s-3\sigma,h}\leqslant 2C_3\frac{\epsilon}{\alpha\sigma^{\bar\nu}}\leqslant\frac{r}{8}\tag{2.3.80}$$

and by using (2.2.1), (2.3.74) and (2.3.75), we get

$$\|\tilde{F}_\omega\|_{r,s-2\sigma,0}\leqslant\|\partial_\omega\chi_1\|_0\cdot\|F\|_{r,s-2\sigma,h}+\|F_\omega\|_{r,s-\sigma,\frac{h}{2}}\tag{2.3.81}$$

$$\leqslant 24C_1C_3\frac{\epsilon}{\alpha\sigma^\tau h}+4C_3\frac{\epsilon}{\alpha\sigma^\tau h}\leqslant\frac{C_5}{4}\frac{\epsilon}{\alpha\sigma^\tau h}\,.\tag{2.3.82}$$

5. Transforming coordinates. As we said, the coordinates transformation $\widetilde{\Phi}$ is obtained as the time–1–map of the flow $\phi^t_{\tilde{F}}$ of the hamiltonian $\tilde{F}$ with equations of motion

$$\dot{y}=-\tilde{F}_x,\quad\dot{x}=\tilde{F}_y\text{ or equivalently }\frac{d}{dt}\phi^t_{\tilde{F}}=\mathbb{J}d\tilde{F}(\phi^t_{\tilde{F}}).$$

By using (2.3.79) and (2.3.80), we deduce that, given $\omega\in\mathbb{C}^d$, the flow $\phi^t_{\tilde{F}}$ is well–defined, real–analytic with holomorphic extention to $D_{\frac{r}{4},s-4\sigma}$ and C^∞ in ω on $\mathbb{R}^d$ for any $0\leqslant t\leqslant 1$, with

$$\phi^t_{\tilde{F}}\colon D_{\frac{r}{4},s-4\sigma}\to D_{\frac{r}{2},s-3\sigma}\tag{2.3.83}$$

and, setting $\widetilde{\Phi}:=\phi^1_{\tilde{F}}=:(U,V)$, we have

$$\|U-\mathrm{id}\|_{\frac{r}{4},s-4\sigma,0}\leqslant\|\tilde{F}_x\|_{\frac{r}{2},s-3\sigma,\infty}\leqslant 2C_3\frac{\epsilon}{\alpha\sigma^{\bar\nu}}\tag{2.3.84}$$

$$\|V-\mathrm{id}\|_{\frac{r}{4},s-4\sigma,0}\leqslant\|\tilde{F}_y\|_{\frac{r}{2},s-3\sigma,\infty}\leqslant 4C_3\frac{\epsilon}{\alpha r\sigma^\tau}\,.\tag{2.3.85}$$

Moreover, since R is affine in the variable y, then so is F and then $\tilde{F}$ so that $\tilde{F}_y$ and V do not depend on y, therefore the jacobian of Φ is of the form

$$D\widetilde{\Phi}=\begin{pmatrix}U_y & U_x\\ 0 & V_x\end{pmatrix},\tag{2.3.86}$$

with, by using Cauchy's estimate, the following bounds

$$\|U_y - \mathrm{Id}\|_{\frac{r}{8},s-5\sigma,0} \leqslant d\frac{8\|U-\mathrm{id}\|_{\frac{r}{4},s-4\sigma,0}}{r} \leqslant 16dC_3\frac{\epsilon}{\alpha r\sigma^{\bar{\nu}}}\,, \tag{2.3.87}$$

$$\|U_x\|_{\frac{r}{8},s-5\sigma,0} \leqslant d\frac{\|U-\mathrm{id}\|_{\frac{r}{4},s-4\sigma,0}}{\sigma} \leqslant 2dC_3\frac{\epsilon}{\alpha\sigma^{\bar{\nu}+1}}\,, \tag{2.3.88}$$

$$\|V_x - \mathrm{Id}\|_{\frac{r}{8},s-5\sigma,0} \leqslant d\frac{\|V-\mathrm{id}\|_{\frac{r}{4},s-4\sigma,0}}{\sigma} \leqslant 4dC_3\frac{\epsilon}{\alpha r\sigma^{\bar{\nu}}}\,, \tag{2.3.89}$$

$$\|U_\omega\|_{\frac{r}{4},s-4\sigma,0} \leqslant \|\tilde{F}_{\omega\theta}\|_{r,s-3\sigma,0} \leqslant d\frac{\|\tilde{F}_\omega\|_{r,s-2\sigma,0}}{\sigma} \leqslant \frac{dC_5}{4}\frac{\epsilon}{\alpha\sigma^{\bar{\nu}}h}\,, \tag{2.3.90}$$

$$\|V_\omega\|_{\frac{r}{4},s-4\sigma,0} \leqslant \|\tilde{F}_{\omega y}\|_{\frac{r}{2},s-2\sigma,0} \leqslant d\frac{2\|\tilde{F}_\omega\|_{r,s-2\sigma,0}}{r} \leqslant \frac{dC_5}{2}\frac{\epsilon}{\alpha r\sigma^{\tau}h}\,. \tag{2.3.91}$$

6. New error term estimate. To estimate P_+^1 (compare (2.3.70)), we need to estimate $\{R, F\}$. By Cauchy's estimate, we have

$$\begin{aligned}
\|\{R,F\}\|_{\frac{r}{2},s-3\sigma,\frac{h}{2}} &\leqslant \sum_{j=1}^{d}\|R_{x_j}\|_{\frac{r}{2},s-3\sigma,h}\|F_{y_j}\|_{\frac{r}{2},s-3\sigma,\frac{h}{2}} + \|R_{y_j}\|_{\frac{r}{2},s-3\sigma,h}\|F_{x_j}\|_{\frac{r}{2},s-3\sigma,\frac{h}{2}}\\
&\leqslant \sum_{j=1}^{d}\frac{\|R\|_{\frac{r}{2},s-2\sigma,h}}{\sigma}\|F_y\|_{\frac{r}{2},s-3\sigma,\frac{h}{2}} + \frac{2\|R\|_{r,s-3\sigma,h}}{r}\|F_x\|_{\frac{r}{2},s-3\sigma,\frac{h}{2}}\\
&\leqslant \sum_{j=1}^{d}\frac{C_3}{2C_0}\frac{\epsilon}{\sigma}\cdot 4C_3\frac{\epsilon}{\alpha r\sigma^{\tau}} + \frac{2C_3}{2C_0}\frac{\epsilon}{r}\cdot 2C_3\frac{\epsilon}{\alpha\sigma^{\bar{\nu}}}\\
&= \frac{4dC_3^2}{C_0}\frac{\epsilon^2}{\alpha r\sigma^{\bar{\nu}}},
\end{aligned}$$

$$\begin{aligned}
\|\{[R],F\}\|_{\frac{r}{2},s-3\sigma,\frac{h}{2}} &= \|[R]_y\cdot F_x\|_{\frac{r}{2},s-3\sigma,\frac{h}{2}} = \|[R_y]\cdot F_x\|_{\frac{r}{2},s-3\sigma,\frac{h}{2}}\\
&\leqslant d\,\|[R_y]\|_{\frac{r}{2},s-3\sigma,h}\cdot\|F_x\|_{\frac{r}{2},s-3\sigma,\frac{h}{2}}\\
&\leqslant d\|R_y\|_{\frac{r}{2},s-3\sigma,h}\cdot\|F_x\|_{\frac{r}{2},s-3\sigma,\frac{h}{2}}\\
&\leqslant d\frac{2\|R\|_{r,s-3\sigma,h}}{r}\|F_x\|_{\frac{r}{2},s-3\sigma,\frac{h}{2}} \leqslant \frac{2dC_3^2}{C_0}\frac{\epsilon^2}{\alpha r\sigma^{\bar{\nu}}}
\end{aligned}$$

and therefore

$$\begin{aligned}
\|P_+^0\|_{\eta r,s-5\sigma,\frac{h}{2}} &\overset{(2.3.72)}{=} \left\|\int_0^1 \{(1-t)[R]+tR,F\}\circ\phi_F^t\,dt\right\|_{\eta r,s-5\sigma,\frac{h}{2}}\\
&\overset{(2.3.76)+(2.3.77)}{\leqslant} \int_0^1 \|\{(1-t)[R]+tR,F\}\|_{2\eta r,s-4\sigma,\frac{h}{2}}\ dt\\
&\leqslant \int_0^1 \|\{(1-t)[R]+tR,F\}\|_{\frac{r}{2},s-3\sigma,\frac{h}{2}}\ dt\\
&\leqslant \frac{3dC_3^2}{C_0}\frac{\epsilon^2}{\alpha r\sigma^{\bar\nu}}\,.
\end{aligned}$$

Hence

$$\begin{aligned}
\|P_+^1\|_{\eta r,s-5\sigma,\frac{h}{2}} &\overset{(2.3.70)}{=} \|P_+^0+(P-Q)\circ\phi_F^1+(Q-R)\circ\phi_F^1\|_{\eta r,s-5\sigma,\frac{h}{2}}\\
&\overset{(2.3.76)+(2.3.77)}{\leqslant} \|P_+^0\|_{\eta r,s-5\sigma,\frac{h}{2}}+\|P-Q\|_{2\eta r,s-4\sigma,\frac{h}{2}}+\|Q-R\|_{2\eta r,s-4\sigma,\frac{h}{2}}\\
&\leqslant \frac{3dC_3^2}{C_0}\frac{\epsilon^2}{\alpha r\sigma^{\bar\nu}}+\frac{32d^2}{9}\eta^2\epsilon+4^d(d+1)C_2\kappa^d\,e^{-\kappa\sigma}\epsilon\\
&\leqslant \max\left(\frac{3dC_3^2}{C_0},4^d(d+1)C_2\right)\left(\frac{\epsilon^2}{\alpha r\sigma^{\bar\nu}}+\left(\eta^2+\kappa^d\,e^{-\kappa\sigma}\right)\epsilon\right)\\
&= \frac{\sqrt{C_{10}}}{3\cdot 2^{\bar\nu}}\left(\frac{\epsilon^2}{\alpha r\sigma^{\bar\nu}}+\left(\eta^2+\kappa^d\,e^{-\kappa\sigma}\right)\epsilon\right).
\end{aligned}\tag{2.3.92}$$

7. Transforming frequencies. In view of (2.3.69), we need to invert the map

$$\rho\colon \omega\mapsto\omega+v(\omega)=\omega+[R_y].$$

But we have

$$\|\rho-\mathrm{id}\|_h=\|[R_y]\|_h\leqslant\|R_y\|_{\frac{r}{2},s-\sigma,h}\leqslant\frac{C_3}{C_0}\frac{\epsilon}{r}\overset{(b)}{\leqslant}\frac{C_3}{C_0C_6}h\leqslant\frac{h}{64\sqrt{3}}\leqslant\frac{h}{4}$$

Therefore, we apply Lemma 2.2.9 and get a real analytic map $\tilde\varphi\colon \Omega_{\alpha,\frac{h}{4}}\to\Omega_{\alpha,\frac{h}{2}}$, inverse of ρ and satisfies

$$\|\tilde\varphi-\mathrm{id}\|_{\frac{h}{4}},\ \frac{h}{4}\|D\tilde\varphi-\mathrm{Id}\|_{\frac{h}{4}}\leqslant\frac{C_3}{C_0}\frac{\epsilon}{r}$$

Now we extend $\tilde\varphi$: by Lemma 2.2.1, there exists a cut–off $\chi_2\in C(\mathbb{C}^d)\cap C^\infty(\mathbb{R}^d)$ with $0\leqslant\chi_2\leqslant 1$, $supp\chi_2\subset\Omega_{\alpha,\frac{h}{4}}$, $\chi_2\equiv 1$ on $\Omega_{\alpha,\frac{h}{8}}$. Then let $\varphi:=\mathrm{id}+(\tilde\varphi-\mathrm{id})\chi_2$ on $\Omega_{\alpha,\frac{h}{8}}$ and

$\varphi := \mathrm{id}$ on $\mathbb{C}^d \backslash \Omega_{\alpha,\frac{h}{4}}$. Thus, setting

$$\begin{aligned} N_+(y,\omega_+) &:= N_+^1(y,\tilde\varphi(\omega_+)) = e_+^1\circ\tilde\varphi(\omega_+) + \omega_+\cdot y, \\ \mathcal{F} &:= (\Phi,\varphi) := \left(\widetilde\Phi\circ(\pi_1,\pi_2,\varphi\circ\pi_3),\varphi\circ\pi_3\right), \\ P_+ &:= P_+^1\circ\mathcal{F}_{|D_{\frac{r}{4},s-4\sigma}\times\Omega_{\alpha,\frac{h}{8}}}\ , \end{aligned}$$

we have $H\circ\mathcal{F} = N_+ + P_+$ on $D_{\frac{r}{4},s-4\sigma}\times\Omega_{\alpha,\frac{h}{8}}$, with all the required properties. Moreover,

$$\|\varphi-\mathrm{id}\|_\infty \leqslant \|\tilde\varphi-\mathrm{id}\|_{\frac{h}{4}} \leqslant \frac{C_3}{C_0}\frac{\epsilon}{r} \overset{(b)}{\leqslant} \frac{h}{64\sqrt{3}} \tag{2.3.93}$$

and

$$\begin{aligned} \|D\varphi-\mathrm{Id}\|_0 &\leqslant \|D\tilde\varphi-\mathrm{Id}\|_{\frac{h}{4}} + \|\tilde\varphi-\mathrm{id}\|_{\frac{h}{4}}\|D_\omega\chi_2\|_0 \overset{(2.2.1)}{\leqslant} \frac{4C_3}{C_0}\frac{\epsilon}{hr} + \frac{C_3}{C_0}\frac{\epsilon}{r}\cdot\frac{24C_1}{h} \\ &= \frac{C_5}{4C_0}\frac{\epsilon}{hr} \overset{(b)}{\leqslant} \frac{C_5}{4C_0C_6} < \frac{1}{2}\ . \end{aligned} \tag{2.3.94}$$

So, in particular, $\varphi_{|\mathbb{R}^d}$ is C^∞–diffeomorphism from $\mathbb{R}^d$ onto itself and since[56] $\varphi = \mathrm{id}$ outside of Ω then $\varphi_{|\Omega}$ is C^∞–diffeomorphism from Ω onto itself.
8. Estimating Φ. By (2.3.84), (2.3.85) and (2.3.93), we have[57]

$$\begin{aligned} \|\bar W(\mathcal{F}-\mathrm{id})\|_{\frac{r}{4},s-4\sigma,0} &\leqslant \max\left(6C_3\frac{\epsilon}{\alpha r\sigma^{\bar\nu}}\left(\frac{\sigma_0}{\sigma}\right)^{2\bar\nu-1},\ \frac{C_3}{C_0}\frac{\epsilon}{rh}\right) \\ &\leqslant \left(\frac{\sigma_0}{\sigma}\right)^{2\bar\nu-1}\cdot\max\left(6C_3\frac{\epsilon}{\alpha r\sigma^{\bar\nu}},\ \frac{C_3}{C_0}\frac{\epsilon}{rh}\right)\ . \end{aligned}$$

Now[58], since

$$\partial_y\Phi = \partial_y\widetilde\Phi|_{(\pi_1,\pi_2;\varphi\circ\pi_3)},\quad \partial_x\Phi = \partial_x\widetilde\Phi|_{(\pi_1,\pi_2;\varphi\circ\pi_3)},\quad \partial_\omega(\Phi-\mathrm{id}) = \partial_\omega\widetilde\Phi\big|_{(\pi_1,\pi_2;\varphi\circ\pi_3)}\circ D\varphi|_{\pi_3},$$

$$\bar W(D\mathcal{F}-\mathrm{Id})\bar W^{-1} = \begin{pmatrix} \widetilde U_y-\mathrm{Id} & \left(\frac{\sigma}{\sigma_0}\right)^{2\bar\nu-1}\frac{\sigma}{r}\widetilde U_x & \frac{h}{r}\widetilde U_\omega \\ 0 & \widetilde V_x-\mathrm{Id} & \left(\frac{\sigma_0}{\sigma}\right)^{2\bar\nu-1}\frac{h}{\sigma}\widetilde V_\omega \\ 0 & 0 & \varphi_\omega-\mathrm{Id} \end{pmatrix},$$

[56] because $\Omega_\alpha\subset\Omega$, $\operatorname{dist}(\Omega_\alpha,\partial\Omega)\geqslant\alpha$ and h will be chosen (just below) in such away that $h\leqslant\frac{\alpha}{2}$ so that $\Omega_{\alpha,\frac{h}{4}}\cap\mathbb{R}^d\subset\Omega$.
[57] Recall that $\sigma_0\geqslant\sigma$.
[58] Recall that $D\Phi$ denotes the Jacobian of Φ *w.r.t* (y,x), $D\varphi$ the Jacobian of φ *w.r.t* ω and $\Phi-\mathrm{id}$ means $\Phi-(\pi_1,\pi_2)$.

with $\Phi =: (\widetilde{U}, \widetilde{V})$, then by (2.3.87)–(2.3.91) and (2.3.94), we have

$$\begin{aligned}\left\|\bar{W}(D\mathcal{F} - \mathrm{Id})\bar{W}^{-1}\right\|_{\frac{r}{8}, s-5\sigma, 0} &\leqslant \max\left\{\left(16dC_3 + 2dC_3 + \frac{dC_5}{4}\right)\frac{\epsilon}{\alpha r \sigma^{\bar{\nu}}},\right.\\ &\qquad \left.\left(4dC_3 + \frac{dC_5}{2}\right)\frac{\epsilon}{\alpha r \sigma^{\bar{\nu}}}\left(\frac{\sigma_0}{\sigma}\right)^{2\bar{\nu}-1}, \frac{C_5}{C_0}\frac{\epsilon}{rh}\right\}\\ &\leqslant \max\left(dC_5\frac{\epsilon}{\alpha r \sigma^{\bar{\nu}}}\left(\frac{\sigma_0}{\sigma}\right)^{2\bar{\nu}-1}, \frac{C_5}{C_0}\frac{\epsilon}{rh}\right)\\ &\leqslant \left(\frac{\sigma_0}{\sigma}\right)^{2\bar{\nu}-1}\cdot\max\left(dC_5\frac{\epsilon}{\alpha r \sigma^{\bar{\nu}}}, \frac{C_5}{C_0}\frac{\epsilon}{rh}\right).\end{aligned}$$

Since $D_{\frac{r}{8}, s-5\sigma} \supseteq D_{\eta r, s-5\sigma}$, then the estimates on Φ are proven. ∎

2.3.3.2 Iteration of the KAM step

Since we are going to iterate the *KAM step* infinitely many times, we need to choose the sequences r_j, s_j, h_j, κ_j, σ_j, η_j conveniently so that at each step, all the assumptions in the *KAM step* hold. See [Pös01], for details on how those sequences are choosen. First, we set up the sequences, then we prove that at each step they meet all the assumptions in *KAM step* and then we prove the iterative lemma.

Let then $\mu := \frac{3}{2}$ and (recall [16])

$$\begin{cases} 0 < s_0 \leqslant 1 \\ s_{j+1} = s_j - 5\sigma_j \end{cases}, \quad \begin{cases} \sigma_0 = \frac{s_0}{20} \\ \sigma_{j+1} = \frac{\sigma_j}{2} \end{cases}, \quad \begin{cases} E_0 \leqslant 20^{\nu} c \sigma_0^{\nu-\bar{\nu}} \\ E_{j+1} = C_{10}^{\mu-1} E_j^{\mu} \end{cases},$$

$$\begin{cases} \kappa_0 = \left[-\frac{40\log\varepsilon - 1}{s_0}\right] \\ \kappa_{j+1} = 4\kappa_j \end{cases}, \quad \begin{cases} \alpha C\varepsilon \leqslant h_0 \leqslant \frac{\alpha}{2\kappa_0^{\bar{\nu}}} \\ h_{j+1} = \frac{h_j}{4^{\bar{\nu}}} \end{cases} \quad \text{and}$$

$$\begin{cases} 0 < r_0 \leqslant 1 \\ \eta_j^2 = E_j,\ r_{j+1} = \eta_j r_j \text{ and } \epsilon_j = \alpha E_j r_j \sigma_j^{\bar{\nu}} \end{cases}.$$

Thus the following hold

Lemma 2.3.6 *For any $j \in \mathbb{N}$,*

(*i*) $\epsilon_j \leqslant \frac{1}{C_4}\alpha \eta_j r_j \sigma_j^{\bar{\nu}}$

(*ii*) $\epsilon_j \leqslant \frac{1}{C_6} h_j r_j$

(*iii*) $h_j \leqslant \frac{\alpha}{2\kappa_j^{\bar{\nu}}}$

(*iv*) $\epsilon_{j+1} \geqslant \frac{\sqrt{C_{10}}}{3\cdot 2^{\bar{\nu}}}\left(\epsilon_j E_j + (\eta_j^2 + \kappa_j^d e^{-\kappa_j\sigma_j})\epsilon_j\right)$

(*v*) $\kappa_j \sigma_j > d-1$, $\quad 0 < \eta_j < 1/8$, $\quad$ *and* $\quad 0 < \sigma_j < s_j/10$

Proof (i) As E_j is decreasing (super–exponentially) and $(i) \Leftrightarrow C_4^2 E_j \leqslant 1$ then it is enough to check it for $j = 0$. But by definitions, we have

$$E_0 \leqslant \frac{1}{4^{\bar\nu} C_{10}} \leqslant \frac{1}{C_4^2}.$$

$(ii) + (iii)$ By definitions, it follows[59]

$$\kappa_0^{\bar\nu} \sigma_0^{\bar\nu}\, e^{-\kappa_0\sigma_0} \leqslant E_0 \sigma_0^{\bar\nu} = \frac{1}{\alpha}\frac{\epsilon_0}{r_0} \leqslant \frac{h_0}{\alpha C_6} \leqslant \frac{1}{2C_6 \kappa_0^{\bar\nu}}.$$

Now let $j \in \mathbb{N}$ and assume

$$\kappa_j^{\bar\nu} \sigma_j^{\bar\nu}\, e^{-\kappa_j\sigma_j} \leqslant E_j \sigma_j^{\bar\nu} = \frac{1}{\alpha}\frac{\epsilon_j}{r_j} \leqslant \frac{h_j}{\alpha C_6} \leqslant \frac{1}{2C_6 \kappa_j^{\bar\nu}}.$$

Then, by using the above definitions we get

$$\begin{aligned}
\kappa_{j+1}^{\bar\nu} \sigma_{j+1}^{\bar\nu}\, e^{-\kappa_{j+1}\sigma_{j+1}} &= 4^{\bar\nu} \kappa_j^{\bar\nu} \sigma_{j+1}^{\bar\nu}\, e^{-2\kappa_j\sigma_j} \\
&\leqslant \left(\kappa_j^{\bar\nu}\, e^{-\kappa_j\sigma_j}\right)^2 \sigma_{j+1}^{\bar\nu} \leqslant E_j^2 \sigma_{j+1}^{\bar\nu} \\
&\leqslant C_{10}^{\mu-1} E_j^{\mu} \sigma_{j+1}^{\bar\nu} = E_{j+1}\sigma_{j+1}^{\bar\nu} = \frac{1}{\alpha}\frac{\epsilon_{j+1}}{r_{j+1}} \\
&= \sqrt{C_{10} E_j}\, E_j \frac{\sigma_j^{\bar\nu}}{2^{\bar\nu}} \leqslant \frac{1}{4^{\bar\nu}} E_j \sigma_j^{\bar\nu} \\
&\leqslant \frac{h_j}{\alpha 4^{\bar\nu} C_6} = \frac{h_{j+1}}{\alpha C_6} \\
&\leqslant \frac{1}{2C_6 \cdot 4^{\bar\nu} \kappa_j^{\bar\nu}} = \frac{1}{2C_6 \kappa_{j+1}^{\bar\nu}}
\end{aligned}$$

which ends the proof of (ii) and (iii).
(iv) We have

$$\begin{aligned}
\frac{\sqrt{C_{10}}}{3\cdot 2^{\bar\nu}} \left(\epsilon_j E_j + (\eta_j^2 + \kappa_j^d\, e^{-\kappa_j\sigma_j})\epsilon_j\right) &\leqslant \frac{\sqrt{C_{10}}}{3\cdot 2^{\bar\nu}} \left(\epsilon_j E_j + (E_j + E_j)\epsilon_j\right) \\
&= \frac{\sqrt{C_{10}}}{2^{\bar\nu}} \epsilon_j E_j = \frac{\sqrt{C_{10}}}{2^{\bar\nu}} \alpha E_j^2 r_j \sigma_j^{\bar\nu} \\
&= \alpha C_{10}^{\mu-1} E_j^{\mu} \sqrt{E_j} \frac{r_{j+1}}{\eta_j} \left(\frac{\sigma_j}{2}\right)^{\bar\nu} \\
&= \alpha r_{j+1} \sigma_{j+1}^{\bar\nu} E_{j+1} = \epsilon_{j+1}.
\end{aligned}$$

[59]See Appendix A.

(v) We have

$$k_j\sigma_j = 2^j\kappa_0\sigma_0 \geqslant \kappa_0\sigma_0 > d-1,$$
$$0 < \eta_j^2 = E_j = \frac{1}{C_{10}}(C_{10}E_0)^{\mu^j} \leqslant \frac{1}{C_{10}4^{\bar\nu\mu^j}} < \frac{1}{4^3},$$
$$s_j - 10\sigma_j = \frac{s_0}{2} > 0.$$

∎

Now we arrive to the iterative lemma. Given $j \in \mathbb{N}$, let[60]

$$D_j := D_{r_j,s_j}, \quad O_j := \Omega_{\alpha,h_j}, \quad \bar W_j := \operatorname{diag}\left(r_j^{-1}\mathbb{1}_d,\ 2^{j(2\bar\nu-1)}\sigma_j^{-1}\mathbb{1}_d,\ h_j^{-1}\mathbb{1}_d\right),$$
$$u_j := \frac{\epsilon_j}{r_j h_j}, \quad v_j := C_{10}E_j.$$

Then, by the definitions, we have $u_{j+1} = 2^{\bar\nu}u_j\sqrt{v_j}$ and $v_{j+1} = v_j^{\mu}$. Thus,

$$v_j = v_0^{\mu^j}, \quad v_0 = C_{10}E_0 \leqslant 4^{-\bar\nu}, \quad u_j = 2^{j\bar\nu}u_0 v_0^{\mu^j-1}, \quad u_0 \leqslant \frac{1}{C_6} \quad \text{and} \quad \frac{u_j}{h_j} = \frac{u_0}{h_0}2^{3j\bar\nu}v_0^{\mu^j-1}.$$

In particular for any $j \geqslant 0$,

$$u_j \leqslant 2^{j\bar\nu}\cdot u_0 \cdot 4^{-\bar\nu(\mu^j-1)}, \quad \frac{u_j}{h_j} \leqslant 2^{3j\bar\nu}\cdot\frac{u_0}{h_0}\cdot 4^{-\bar\nu(\mu^j-1)}. \tag{2.3.95}$$

Lemma 2.3.7 *Suppose $H_0 := N + P_0$ is real analytic on $D_0 \times O_0$ with*

$$\|P_0\|_{r_0,s_0,h_0} \leqslant \epsilon_0 = \alpha E_0 r_0 \sigma_0^{\bar\nu}.$$

Then for each $j \in \mathbb{N}$, there exist a normal form N_j and a transformation $\mathcal{F}^j = (\Phi^j;\varphi^j) := \mathcal{F}_0\circ\cdots\circ\mathcal{F}_{j-1}\colon D_j\times\mathbb{R}^d \to D_0\times\mathbb{R}^d$ such that

- *(i) $\varphi^j\colon \mathbb{R}^d \to \mathbb{R}^d$ is a C^∞–diffeomorphism with $\varphi^j \equiv \mathrm{id}$ on Ω and $\Phi^j\colon D_j\times\mathbb{R}^d \to D_0$ is a (ω–) family of real–anatylitic, symplectic transformations parametrized over $\mathbb{R}^d$ and C^∞ in ω;*
- *(ii) $\mathcal{F}^j$ is Lipschitz–continuous in ω with*

$$\|\bar W_0\left(\mathcal{F}^j - \mathrm{id}\right)\|_{L,\mathbb{R}^d} \leqslant C_9\frac{\epsilon_0}{r_0 h_0^2}, \tag{2.3.96}$$

uniformly on $D_j\times\mathbb{R}^d$.

[60] Notice that $\sigma_0/\sigma_j = 2^j$.

(iii) The restriction $\tilde{\mathcal{F}}^j := \mathcal{F}^j_{|D_j \times O_j}$ *is real–analytic with holomorphic extension to* $D_j \times O_j$ *for each given* $\omega \in O_j$ *and satisfies* $\tilde{\mathcal{F}}^j \colon D_j \times O_j \to D_0 \times O_0$, $H \circ \tilde{\mathcal{F}}^j = N_j + P_j$ *with*

$$\|P_j\|_{r_j, s_j, h_j} \leqslant \epsilon_j = \alpha E_j r_j \sigma_j^{\bar{\nu}}.$$

Furthermore

$$\|\bar{W}_0(\mathcal{F}^{j+1} - \mathcal{F}^j)\|_{r_{j+1}, s_{j+1}, 0} \leqslant C_7 \frac{\epsilon_0}{r_0 h_0} \cdot 2^{-\bar{\nu}(2\mu^j + 3j - 2) + j} \ . \tag{2.3.97}$$

Proof For $j = 0$ we take $\mathcal{F}^1 = \mathcal{F}_0 = \mathrm{id}$, $N_1 = N$, $P_1 = P_0$ and we are done.
Next we pick $j \geqslant 0$ and we assume that it holds at the step j. Then we have to check it for the step $j+1$. But, thanks to lemma 2.3.6, we can apply the *KAM step* to H_j to get a transformation $\mathcal{F}_j = (\Phi_j; \varphi_j) \colon D_{j+1} \times \mathbb{R}^d \to D_j \times \mathbb{R}^d$ for which every properties in *KAM step* hold. So, its restriction

$$\tilde{\mathcal{F}}_j := \mathcal{F}_{j|D_{j+1} \times O_{j+1}} \colon D_{j+1} \times O_{j+1} \to D_j \times O_j$$

and there exists a normal form N_{j+1} such that $H_{j+1} = H_j \circ \tilde{\mathcal{F}}_j = N_{j+1} + P_{j+1}$ with

$$\|P_{j+1}\|_{r_{j+1}, s_{j+1}, h_{j+1}} \leqslant \frac{\sqrt{C_{10}}}{3 \cdot 2^{\bar{\nu}}} \left(\epsilon_j E_j + \left(\eta_j^2 + \kappa_j^d \, e^{-\kappa_j \sigma_j} \right) \epsilon_j \right).$$

Then we apply (iv) of lemma 2.3.6 to obtain

$$\|P_{j+1}\|_{r_{j+1}, s_{j+1}, h_{j+1}} \leqslant \epsilon_{j+1} = \alpha E_{j+1} r_{j+1} \sigma_{j+1}^{\bar{\nu}}.$$

Therefore

$$\mathcal{F}^{j+1} = \mathcal{F}_0 \circ \cdots \circ \mathcal{F}_j = \left(\Phi^j \circ (\Phi_j, \varphi_j); \varphi^j \circ \varphi_j \right) \colon D_{j+1} \times \mathbb{R}^d \to D_0 \times \mathbb{R}^d$$

is a transformation such that $H \circ \tilde{\mathcal{F}}^{j+1} = H_j \circ \tilde{\mathcal{F}}_j = N_{j+1} + P_{j+1}$ with all the required properties in (i) and (iii).
It remains the estimates on $\mathcal{F}^j$. By (2.3.65) and (2.3.66) we have[61]

$$\begin{aligned}
& \|\bar{W}_j(\mathcal{F}_j - \mathrm{id})\|_{r_{j+1}, s_{j+1}, 0} \ , \quad \|\bar{W}_j(D\mathcal{F}_j - \mathrm{Id})\bar{W}_j^{-1}\|_{r_{j+1}, s_{j+1}, 0} \leqslant \\
\leqslant \quad & \left(\frac{\sigma_0}{\sigma_j} \right)^{2\bar{\nu} - 1} \cdot \max \left(dC_5 \frac{\epsilon_j}{\alpha r_j \sigma_j^{\bar{\nu}}}, \frac{C_5}{C_0} \frac{\epsilon_j}{r_j h_j} \right) \\
\leqslant \quad & 2^{-j(2\bar{\nu} - 1)} \frac{dC_5^2}{4 C_0 C_4} u_j \qquad (2.3.98) \\
\overset{(2.3.95)}{\leqslant} \quad & \frac{dC_5^2}{4 C_0 C_4} \frac{\epsilon_0}{r_0 h_0} \cdot 2^{-\bar{\nu}(2\mu^j + j - 2) + j} \ . \qquad (2.3.99)
\end{aligned}$$

[61]Notice that $\frac{h_j}{\alpha \sigma_j^{\bar{\nu}}} = \frac{h_0}{\alpha \sigma_0^{\bar{\nu}}} \cdot \frac{1}{2^{j\bar{\nu}}} \leqslant \frac{h_0}{\alpha \sigma_0^{\bar{\nu}}} \leqslant \frac{C_5}{4 C_0 C_4}$.

Thus

$$\begin{aligned}
\|\bar W_0(\mathcal F^{j+1}-\mathcal F^j)\|_{r_{j+1},s_{j+1},0} &= \|\bar W_0(\mathcal F^j\circ\mathcal F_j-\mathcal F^j)\|_{r_{j+1},s_{j+1},0}\\
&\leqslant \|\bar W_0 D\mathcal F^j\bar W_j^{-1}\|_{r_j,s_j,0}\cdot\|\bar W_j(\mathcal F_j-\mathrm{id})\|_{r_{j+1},s_{j+1},0}\\
&\overset{(2.3.99)}{\leqslant} \frac{dC_5^2}{4C_0C_4}\frac{\epsilon_0}{r_0h_0}\cdot 2^{-\bar\nu(2\mu^j+j-2)+j}\cdot\|\bar W_0 D\mathcal F^j\bar W_j^{-1}\|_{r_j,s_j,0}.
\end{aligned}$$

Next, we need to bound $\|\bar W_0 D\mathcal F^j\bar W_j^{-1}\|$ uniformly on $D_j\times\mathbb R^d$. But for any $j\geqslant 0$, we have[62]

$$\begin{aligned}
\|\bar W_j\bar W_{j+1}^{-1}\| &= \left\|diag\left(\frac{r_{j+1}}{r_j}\mathbb 1_d,\frac{1}{2^{2\bar\nu-1}}\frac{\sigma_{j+1}}{\sigma_j}\mathbb 1_d,\frac{h_{j+1}}{h_j}\mathbb 1_d\right)\right\|\\
&=\max\left(\frac{r_{j+1}}{r_j},\frac{1}{2^{2\bar\nu-1}}\frac{\sigma_{j+1}}{\sigma_j},\frac{h_{j+1}}{h_j}\right)\\
&=\max\left(\eta_j,\frac{1}{4^{\bar\nu}}\right)\\
&=\frac{1}{4^{\bar\nu}}. \qquad (2.3.100)
\end{aligned}$$

and

$$\begin{aligned}
\bar W_0 D\mathcal F^j\bar W_j^{-1} &= \bar W_0 D\mathcal F_0\circ\cdots\circ D\mathcal F_{j-1}\bar W_j^{-1}\\
&=\left(\bar W_0 D\mathcal F_0\bar W_0^{-1}\right)\left(\bar W_0\bar W_1^{-1}\right)\cdots\left(\bar W_{j-1}D\mathcal F_{j-1}\bar W_{j-1}^{-1}\right)\left(\bar W_{j-1}\bar W_j^{-1}\right)\ ,
\end{aligned}$$

so that,

$$\begin{aligned}
\|\bar W_0 D\mathcal F^j\bar W_j^{-1}\| &\leqslant \|\bar W_0 D\mathcal F_0\bar W_0^{-1}\|\|\bar W_0\bar W_1^{-1}\|\cdots\|\bar W_{j-1}D\mathcal F_{j-1}\bar W_{j-1}^{-1}\|\|\bar W_{j-1}\bar W_j^{-1}\|\\
&\overset{(2.3.100)}{\leqslant} \|\bar W_0 D\mathcal F_0\bar W_0^{-1}\|\cdots\|\bar W_{j-1}D\mathcal F_{j-1}\bar W_{j-1}^{-1}\|\\
&\overset{(2.3.99)}{\leqslant} 4^{-j\bar\nu}\prod_{k=0}^{j-1}\left(1+\frac{dC_5^2}{4C_0C_4}\frac{\epsilon_0}{r_0h_0}\cdot 2^{-\bar\nu(2\mu^k+k-2)+k}\right)\\
&\leqslant 4^{-j\bar\nu}\prod_{k=0}^{\infty}\left(1+\frac{dC_5^2}{4C_0C_4C_6}\cdot 2^{-\bar\nu(2\mu^k+k-2)+k}\right)\\
&\leqslant 4^{-j\bar\nu}\exp\left(\frac{dC_5^2}{4C_0C_4C_6}\sum_{k=0}^{\infty}2^{-\bar\nu(2\mu^k+k-2)+k}\right)\\
&= 4^{-j\bar\nu}\frac{4C_0C_4C_7}{dC_5^2}\ .
\end{aligned}$$

[62] Recall that any $j\geqslant 0$, $\eta_j\leqslant\sqrt{c}\leqslant 1/\sqrt{4^{\bar\nu}C_{10}}<1/(3\cdot 4^{\bar\nu})$.

$$C_6 := \frac{C_5}{C_0}$$

$$C_7 := \frac{dC_5^2}{4C_0C_4}\exp\left(\frac{dC_5^2}{4C_0C_4C_6}\sum_{j=0}^{\infty}2^{-\bar\nu(2\mu^j+j-2)+j}\right)$$

Therefore,

$$\|\bar W_0(\mathcal{F}^{j+1}-\mathcal{F}^j)\|_{r_{j+1},s_{j+1},0} \leqslant C_7\frac{\epsilon_0}{r_0h_0}\cdot 2^{-\bar\nu(2\mu^j+3j-2)+j}\ .$$

Finally, using again (2.3.98) and (2.3.100), we get

$$\begin{aligned}\|\bar W_0\left(D\mathcal{F}^{j+1}-\mathrm{Id}\right)\bar W_{j+1}^{-1}\| &= \|\bar W_0\left(D\mathcal{F}^{j}\circ D\mathcal{F}_j-\mathrm{Id}\right)\bar W_{j+1}^{-1}\|\\ &= \|\bar W_0\left(D\mathcal{F}^{j}-\mathrm{Id}\right)\bar W_{j}^{-1}\circ \bar W_j D\mathcal{F}_j\bar W_j^{-1}\circ\bar W_j\bar W_{j+1}^{-1}+\\ &\qquad \bar W_0\left(D\mathcal{F}_j-\mathrm{Id}\right)\bar W_{j+1}^{-1}\|\\ &\leqslant \|\bar W_0\left(D\mathcal{F}^{j}-\mathrm{Id}\right)\bar W_{j}^{-1}\|\|\bar W_j D\mathcal{F}_j\bar W_j^{-1}\|\|\bar W_j\bar W_{j+1}^{-1}\|+\\ &\quad \|\left(\bar W_0\bar W_1^{-1}\right)\cdots\left(\bar W_{j-1}\bar W_j^{-1}\right)\bar W_j\left(D\mathcal{F}_j-\mathrm{Id}\right)\bar W_j^{-1}\circ\bar W_j\bar W_{j+1}^{-1}\|\\ &\leqslant \frac{1}{4^{\bar\nu}}\|\bar W_0\left(D\mathcal{F}^{j}-\mathrm{Id}\right)\bar W_{j}^{-1}\|\left(1+\frac{1}{2^{j(2\bar\nu-1)}}\frac{dC_5^2}{4C_0C_4}u_j\right)+\\ &\qquad \frac{1}{4^{(j+1)\bar\nu}}\frac{1}{2^{j(2\bar\nu-1)}}\frac{dC_5^2}{4C_0C_4}u_j\ .\end{aligned}$$

Therefore, letting

$$w_j := \log\left(\|\bar W_0\left(D\mathcal{F}^{j}-\mathrm{Id}\right)\bar W_{j}^{-1}\|+\frac{1}{4^{j\bar\nu}}\right),\quad z_j := \frac{1}{2^{j(2\bar\nu-1)}}\frac{dC_5^2}{4C_0C_4}u_j,\quad j\geqslant 0,$$

we get, for any $j \geqslant 1$,

$$w_{j+1}\leqslant w_j+\log\left(\frac{1}{4^{\bar\nu}}(1+z_j)\right)\quad\text{and}\quad w_1=\log\left(\frac{1}{4^{\bar\nu}}\right)\ ,$$

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

$$\begin{aligned} w_j &\leqslant \sum_{k=0}^{j-1} \log\left(\frac{1}{4^{\bar\nu}}(1+z_k)\right) \\ &\leqslant \log\left(\frac{1}{4^{j\bar\nu}}\prod_{k=0}^{j-1}(1+z_k)\right) \\ &\leqslant \log\left(\frac{1}{4^{j\bar\nu}}\right) + \log\left(\prod_{k=0}^{\infty}(1+z_k)\right) \\ &\leqslant \log\left(\frac{1}{4^{j\bar\nu}}\right) + \sum_{k=0}^{\infty} z_k \\ &\leqslant \log\left(\frac{1}{4^{j\bar\nu}}\right) + C_8 u_0 \ , \end{aligned}$$

i.e.

$$\|\bar W_0 \left(D\mathcal{F}^j - \mathrm{Id}\right) \bar W_j^{-1}\| \leqslant \frac{1}{4^{j\bar\nu}}\left(e^{C_8 u_0} - 1\right).$$

In particular, for any $j \geqslant 1$,

$$\begin{aligned} h_j \|\bar W_0 \left(\partial_\omega \mathcal{F}^j - \mathrm{Id}\right)\| &\leqslant \|\bar W_0 \left(D\mathcal{F}^j - \mathrm{Id}\right) \bar W_j^{-1}\| \\ &\leqslant \frac{1}{4^{j\bar\nu}}\left(e^{C_8 u_0} - 1\right) \ , \end{aligned}$$

i.e. [63]

$$h_0 \|\bar W_0 \left(\partial_\omega \mathcal{F}^j - \mathrm{Id}\right)\| \leqslant e^{C_8 u_0} - 1 \leqslant C_8 u_0 \, e^{C_8 u_0} \leqslant C_8 \, e^{C_8/C_6} u_0 = C_9 u_0 \ ,$$

i.e.

$$\|\bar W_0 \left(\mathcal{F}^j - \mathrm{id}\right)\|_{L,\mathbb{R}^d} \leqslant C_9 \frac{\epsilon_0}{r_0 h_0^2} \quad \text{uniformly on} \quad D_j \times \mathbb{R}^d \ .$$

∎

2.3.3.3 Deduction of Theorem 2.1.4

We set $P_0 = P$, $s_0 = s$, $r_0 = r$, $h_0 = h$, and $\epsilon_0 = \epsilon = \|P\|_{r,s,h}$; thus Lemma 2.3.7 applies. Hence, by (2.3.97), $(\mathcal{F}_j)_j$ is a Cauchy sequence and therefore converges uniformly to some $\mathcal{F} = (\Phi, \varphi)$ on

$$\bigcap_{j\geqslant 0} D_j \times \mathbb{R}^d = T_* \times \mathbb{R}^d, \text{ where } T_* := \{0\} \times \mathbb{T}^d_{\frac{s}{2}},$$

[63] Recall that $u_0 \leqslant 1/C_6$ and $e^a - 1 \leqslant a\, e^a$, $\forall a \geqslant 0$.

with the map $x \mapsto \Phi(0,x,\omega)$ real analytic on $\mathbb{T}^d_{\frac{s}{2}}$ for each given $\omega \in \mathbb{R}^d$ (by Weierstrass's theorem) and for any $j \geqslant 1$,

$$\begin{aligned}\|\bar{W}_0(\mathcal{F}^j - \mathrm{id})\|_{r_j,s_j,h_j} &\leqslant \|\bar{W}_0(\mathcal{F}^j - \mathcal{F}^{j-1})\|_{r_j,s_j,h_j} + \cdots + \|\bar{W}_0(\mathcal{F}^2 - \mathcal{F}^0)\|_{r_0,s_0,h_0} \\ &\leqslant C_{10}u_0,\end{aligned}$$

$$C_{10} := C_7 \sum_{j=0}^{\infty} 2^{-\bar{\nu}(2\mu^j+3j-2)+j}$$

and letting $j \to \infty$, we get, uniformly on $T_* \times \mathbb{R}^d$,

$$\|\bar{W}_0(\mathcal{F} - \mathrm{id})\| \leqslant C_{10}\frac{\epsilon}{rh}. \tag{2.3.101}$$

Moreover, by letting $j \to \infty$ in (2.3.96), we get, uniformly on $T_* \times \mathbb{R}^d$,

$$\|\bar{W}_0(\mathcal{F} - \mathrm{id})\|_{L,\mathbb{R}^d} \leqslant C_9\frac{\epsilon}{rh}. \tag{2.3.102}$$

Let's prove that φ is a lipeomorphism from Ω onto itself. Indeed, for any $j \geqslant 0$,

$$\begin{aligned}\|D\varphi^{j+1} - \mathrm{Id}\|_0 + 1 &= \|(D\varphi^j - \mathrm{Id})D\varphi_j + (D\varphi_j - \mathrm{Id})|_0 + 1 \\ &\leqslant \|D\varphi^j - \mathrm{Id}\|_0\left(\|D\varphi_j - \mathrm{Id}\|_0 + 1\right) + \|D\varphi_j - \mathrm{Id})\|_0 + 1 \\ &= \left(\|D\varphi^j - \mathrm{Id}\|_0 + 1\right)\left(\|D\varphi_j - \mathrm{Id}\|_0 + 1\right)\ ,\end{aligned}$$

so that[64]

$$\begin{aligned}\|D\varphi^{j+1} - \mathrm{Id}\|_0 &\leqslant -1 + \left(\|D\varphi^1 - \mathrm{Id}\|_0 + 1\right)\prod_{k=0}^{j}\left(\|D\varphi_k - \mathrm{Id}\|_0 + 1\right) \\ &\overset{(2.3.67)}{\leqslant} -1 + \prod_{k=0}^{j}\left(1 + \frac{C_5}{2C_0}u_k\right) \\ &\leqslant \exp\left(\frac{C_5}{2C_0}\sum_{k=0}^{\infty}u_k\right) - 1 \\ &\overset{(2.3.95)}{\leqslant} \exp\left(\frac{C_5}{2C_0}u_0\sum_{k=0}^{\infty}2^{-\bar{\nu}(2\mu^k-k-2)}\right) - 1 \\ &\leqslant \exp\left(\frac{C_5}{2C_0C_6}\sum_{k=0}^{\infty}2^{-\bar{\nu}(2\mu^k-k-2)}\right) - 1 \\ &\leqslant e^{\log\left(\frac{3}{2}\right)} - 1 = \frac{1}{2} < 1\end{aligned}$$

[64] Recall that $\varphi^1 = \varphi_0 = \mathrm{id}$, so that $\|D\varphi^1 - \mathrm{Id}\|_0 = 0$.

Hence, φ is a lipeomorphism (Lipschitz continuous bijection with inverse Lipschitz continuous as well) from $\mathbb{R}^d$ onto itself closed to the identity. Furthermore, $\varphi \equiv \mathrm{id}$ outside of Ω since each φ_j is so, so that φ restricted to Ω is a lipeomorphism from Ω onto itself.
Next, we prove that for each $\omega \in \Omega_\alpha$, $\Phi(0,x,\omega)$ is an invariant Kronecker torus for $H_{|\varphi(\omega)}(y,x) := H(y,x,\varphi(\omega))$. Indeed, by letting $j \to \infty$ in *Iterative Lemma*, *(iii)*, we obtain

$$H_{|\varphi(\omega)} \circ \Phi(y,x,\omega) = H \circ \mathcal{F}(y,x,\omega) =: e_\infty(\omega) + \omega \cdot y,$$
$$\text{on} \quad \bigcap_{j \geqslant 0} D_j \times O_j = T_* \times \Omega_\alpha.$$

Thus,

$$\Phi^{-1} \circ \phi_{H_{|\varphi(\omega)}} \circ \Phi(y,x;\omega) = (y, \omega t + x) \quad i.e. \quad \phi_{H_{|\varphi(\omega)}} \circ \Phi(y,x;\omega) = \Phi(y, \omega t + x; \omega),$$

on $T_* \times \Omega_\alpha$.
It remains just to prove that the tori are Lagragian. Indeed, since $T_{(0,x)}T_* \cong \{0\} \times \mathbb{C}^d$ for any $x \in \mathbb{T}^d_{\frac{s}{2}}$, each Φ^j is symplectic and ϖ is smooth, then we have, for any $\omega \in \mathbb{R}^d$,

$$\Phi^* \varpi \|_{T^*} = \lim_{j \to \infty} (\Phi^j)^* \varpi \|_{T^*} = \varpi \|_{T^*} = 0.$$

∎

Remark 2.3.8 Notice that one could apply Lemma 2.2.2 as well to prove that φ is a lipeomorphism, provided that C_6 is chosen a little bigger. In fact, by (2.3.67), we have[65]

$$\|\varphi_j - \mathrm{id}\|_{L,\mathbb{R}^d} \leqslant \frac{C_5}{2C_0} u_j.$$

Thus, for

$$C_6 \geqslant \frac{C_5}{C_0} \sum_{j=0}^{\infty} 2^{-\bar{\nu}(2\mu^j - j - 2)} \ ,$$

by taking With $\mathcal{L}_j \equiv \mathrm{id}$, $g_j = \varphi_j$, $\delta = 1$, $l_j = \frac{C_5}{2C_0} u_j$, since $\sum_{j=0}^{\infty} l_j \leqslant \frac{C_5}{2C_0C_6} \sum_{j=0}^{\infty} 2^{-\bar{\nu}(2\mu^j - j - 2)} \leqslant \frac{1}{2}$, for any $j \geqslant 0$ so that we apply again Lemma 2.2.2 to get

$$\|\varphi^j - \mathrm{id}\|_{L,\mathbb{R}^d} \leqslant \frac{C_5}{2C_0} \sum_{k=0}^{\infty} u_k \leqslant \frac{1}{2}.$$

Therefore

$$\|\varphi - \mathrm{id}\|_{L,\mathbb{R}^d} \leqslant \frac{1}{2} < 1.$$

[65] Recall the notations in the proof of *Iterative Lemma.*

3 | Comparison of the KAM theorems on a mechanical Hamiltonian

We consider the simple mechanical Hamiltonian[66]

$$H_0(y,x;\varepsilon) := \frac{y^2}{2} + \varepsilon P_0(x) := \frac{y^2}{2} + \varepsilon\left(\cos x_1 + \sum_{j=1}^{d-1}\cos(x_{j+1}-x_j)\right),$$

and we choose

$$\boxed{s = \bar{s} = \frac{10s_*}{9} = \frac{10\hat{s}}{9} = 20\sigma = 1, \quad \varepsilon_0 = \infty, \quad \tau \geqslant d-1, \quad \alpha > 0, \quad y_0 = \omega \in \Delta_\alpha^\tau}.$$

Moreover, we have

$$H_0(y+y_0,x;\varepsilon) = \frac{\omega^2}{2} + \omega\cdot y + \frac{y^2}{2} + \varepsilon P_0(x) =: H(y,x;\omega;\varepsilon). \tag{3.0.1}$$

3.1 Application of Theorem 2.1.1

By (3.0.1), we have

$$K_0 = \frac{\omega^2}{2}, \quad Q(y,x) = \frac{y^2}{2}, \quad P(y,x;\varepsilon) = P_0(x), \quad T = \mathbb{1}_d.$$

Hence[67]

$$M = \|P\|_{r,s,\varepsilon_0} = \cosh s + (d-1)\cosh(2s), \quad \widehat{\mathrm{E}} = \max\left\{\frac{\omega^2}{2}, r|\omega|, \frac{dr^2}{2}, |\omega|^2\right\},$$

[66] As usual, $y^2 = y\cdot y = y_1^2 + \cdots + y_d^2$.
[67] See §3.3 for an idea to compute $\|P\|_{r,s,\varepsilon_0}$.

$$L = \frac{\mathrm{C}_\sharp}{3} \widehat{\mathrm{E}}^{10} r^{-10} \alpha^{-4} |\omega|^{-6} M,$$

$$\varepsilon_* := \mathrm{c}\, \widehat{\mathrm{E}}^{-9} \sigma^{4\tau+13} r^{10} \alpha^4 |\omega|^6 M^{-1} = \frac{\mathrm{c}\, \widehat{\mathrm{E}}^{-9} \sigma^{4\tau+13} r^{10} \alpha^4 |\omega|^6}{\cosh s + (d-1)\cosh(2s)}\, .$$

Therefore, Theorem 2.1.1 holds for

$$\boxed{|\varepsilon| < \varepsilon_* := \frac{\mathrm{c}\, \widehat{\mathrm{E}}^{-9} \sigma^{4\tau+13} r^{10} \alpha^4 |\omega|^6}{\cosh s + (d-1)\cosh(2s)}\, .} \tag{3.1.1}$$

3.2 Application of Theorem 2.1.2

By the very definition of H_0, we have

$$P = P_0, \quad K(y) = \frac{y^2}{2}, \quad K_y(y) = y, \quad T = K_{yy}(y) = \mathbb{1}_d,$$

so that[68]

$$M = \|P_0\|_{r,s,y_0} = \cosh s + (d-1)\cosh(2s), \qquad \|K_y\|_{r,y_0} \overset{def}{=} \sup_{|y-y_0|<r} |y| = r + |y_0| = r + |\omega|,$$

$$\|K_{yy}\|_{r,y_0} \overset{def}{=} \sup_{|a|=1} \|K_{yy}(\cdot)a\|_{r,y_0} = \sup_{|a|=1} \sup_{|y-y_0|<r} |K_{yy}(y)a| = \sup_{|a|=1} \sup_{|y-y_0|<r} |a| = 1,$$

$$\|T\| = 1, \qquad \widehat{\mathsf{E}} = \max\left\{r(r+|\omega|),\, |\omega|^2\right\}.$$

Therefore, Theorem 2.1.2 holds for

$$\boxed{|\varepsilon| < \varepsilon_* := \frac{\alpha^2}{M_0}\mu_*\, ,} \tag{3.2.1}$$

where

$$\mu_* := \max\left\{0 < \mu \leqslant \varepsilon_\sharp \ :\ \mathfrak{p}_1 \cdot \max\left\{1\, ,\, \mathfrak{p}_2\, \mu\, \left(\log \mu^{-1}\right)^{2\nu}\right\} \cdot \mu\, \left(\log \mu^{-1}\right)^{\nu} < 1\right\}\ ,$$

$$\varepsilon_\sharp := \min\left\{e^{-1}\, ,\, \exp\left(-\frac{\sigma}{5}\left(\frac{12\sqrt{2}}{5}\frac{\alpha}{r}\right)^{\frac{1}{\nu}}\right)\right\}\ ,$$

$$\mathfrak{p}_1 := \mathsf{C}_8\, \sigma_0^{-(3\nu+2d+1)}\ \max\left\{1, \frac{\alpha}{r}\right\}\ ,$$

$$\mathfrak{p}_2 := \mathsf{C}_{11}\, \sigma_0^{-(4\nu+2d)}\ ,$$

[68] See §3.3 for an idea to compute $\|P_0\|_{r,s,y_0}$ and $\|K_y\|_{r,y_0}$; for the later, writing $y_0 = (y_{01}, \cdots, y_{0d})$, one can just choose the family $y_a := (y_{01} + a\,\mathrm{sign}\,(y_{01}), \cdots, y_{0d} + a\,\mathrm{sign}\,(y_{0d})),\ 0 < a < r$.

3.3 Application of Theorem 2.1.4

By (3.0.1), we have

$$K_0(\omega)=\frac{\omega^2}{2},\quad P(y,x;\omega)=\frac{y^2}{2}+\varepsilon P_0(x).$$

We choose

$$0<r<\frac{2c\alpha s^\nu}{d},\quad h:=\frac{\epsilon C}{r},$$

with c, C as in (2.1.9) and $\epsilon=\|P\|_{r,s,h}$. Next, we compute ϵ.

$$\begin{aligned}\epsilon=\sup_{D^d_{r,s}\times\Omega^d_{\alpha,h}}|P|&\leqslant\frac{dr^2}{2}+\sup_{x\in\mathbb{T}^d_1}|\varepsilon|\left(|\cos x_1|+\sum_{j=1}^{d-1}|\cos(x_{j+1}-x_j)|\right)\\&\leqslant\frac{dr^2}{2}+|\varepsilon|(\cosh s+(d-1)\cosh(2s)),\end{aligned}$$

Now, choosing

$$y_a:=\left(a\,i^{\frac{1-\operatorname{sign}\varepsilon}{2}},a\,i^{\frac{1-\operatorname{sign}\varepsilon}{2}},\cdots,a\,i^{\frac{1-\operatorname{sign}\varepsilon}{2}}\right),\quad 0<a<r$$

and

$$x_b:=\left(b\,i,-b\,i,\cdots,\underbrace{(-1)^{j+1}b\,i}_{j^{th}\text{ term}},\cdots,(-1)^{d+1}b\,i\right),\quad 0<b<s,$$

we get

$$\begin{aligned}|P(y_a,x_b;\omega)|&=\left|\frac{da^2}{2}\operatorname{sign}(\varepsilon)+\varepsilon\left(\cosh b+\sum_{j=1}^{d-1}\cosh(2b)\right)\right|\\&=\frac{da^2}{2}+|\varepsilon|\left(\cosh b+(d-1)\cosh(2b)\right).\end{aligned}$$

Therefore,

$$\begin{aligned}\epsilon&\geqslant\sup_{\substack{0<a<r\\0<b<s}}|P(y_a,x_b;\omega)|\\&=\sup_{\substack{0<a<r\\0<b<s}}\frac{da^2}{2}+|\varepsilon|\left(\cosh b+(d-1)\cosh(2b)\right)\\&=\frac{dr^2}{2}+|\varepsilon|\left(\cosh s+(d-1)\cosh(2s)\right).\end{aligned}$$

Thus

$$\begin{aligned}\epsilon &= \frac{dr^2}{2} + |\varepsilon|(\cosh s + (d-1)\cosh(2s)) \\ &= \frac{1}{2}c\alpha r + |\varepsilon|(\cosh s + (d-1)\cosh(2s)).\end{aligned}$$

Consequently, if

$$\boxed{|\varepsilon| \leqslant \varepsilon_* := \frac{2c\alpha s^\nu - dr^2}{2(\cosh s + (d-1)\cosh(2s))}}, \tag{3.3.1}$$

then Theorem 2.1.4 holds.

3.4 Application of Theorem 2.1.6

We choose

$$u \equiv 0, \quad v \equiv y_0 = \omega.$$

Thus,

$$H_{yy} \equiv \mathbb{1}_d, \quad \mathcal{M} = \mathbb{1}_d, \quad \mathcal{T} = \mathbb{1}_d, \quad f \equiv 0, \quad g = \varepsilon \nabla_x P_0.$$

Therefore, we can take

$$\mathtt{E} = 1, \quad E_{j,k} = 0 \quad \text{if} \quad jk > 0 \quad \text{or} \quad k \geqslant 3, \quad E_{0,2} = 1, \quad \rho = U = \widetilde{V} = 0, ,$$

$$\widetilde{T} = 1, \quad V = |\omega|_1, \quad \mathtt{M} = \overline{\mathtt{M}} = 1, \quad F = 0.$$

Next, we compute $G = |\varepsilon| \, \|\nabla_x P_0\|_{r,s,y_0}$. We have,

$$\begin{aligned}G &= |\varepsilon| \sup_{|a|_1=1} \left\| -a_1 \frac{e^{ix_1} - e^{ix_2}}{2i} + \sum_{j=1}^{d-1} (a_j - a_{j+1}) \frac{e^{i(x_{d+1}-x_d)} - e^{-i(x_{d+1}-x_d)}}{2i} \right\|_{r,s,y_0} \\ &= |\varepsilon| \sup_{|a|_1=1} \left(|a_1| \frac{e^s + e^s}{2} + \sum_{j=1}^{d-1} |a_j - a_{j+1}| \frac{e^{2s} + e^{2s}}{2} \right) \\ &= |\varepsilon| \sup_{|a|_1=1} \left(|a_1| \, e^s + \sum_{j=1}^{d-1} (|a_j| + |a_{j+1}|) \, e^{2s} \right) \\ &\leqslant |\varepsilon| \min(e^s + (d-1) \, e^{2s}, \; 2 \, e^{2s}) \, .\end{aligned}$$

But then, taking $a_0 = (1, 0)$ if $d = 2$ and

$$a_0 = (0, 1, 0, \cdots, 0) \quad \text{if} \quad d \geqslant 3,$$

we obtain

$$G \geqslant |\varepsilon| \, \| \nabla_x P_0(\cdot) a_0 \|_{r,s,y_0} = |\varepsilon| \min(e^s + (d-1) \, e^{2s}, \; 2 \, e^{2s}).$$

Hence

$$G = |\varepsilon| \min(e^s + (d-1) \, e^{2s}, \; 2 \, e^{2s}) =: |\varepsilon| \widehat{G} \, .$$

It remains the choice of $E_{3,0}$. Writting

$$P_0 = \frac{1}{2}(e^{ix_1} + \, e^{-ix_1}) + \frac{1}{2} \sum_{j=1}^{d-1} e^{i(x_{j+1}-x_j)} + e^{-i(x_{j+1}-x_j)} =: \sum_{m \in \mathbb{Z}^d} P_{0m} \, e^{im\cdot x}, \tag{3.4.1}$$

we have, for any $j, k, l = 1, \cdots, d$,

$$\frac{\partial^3 P_0}{\partial x_j \partial x_k \partial x_l} = -i \sum_{m \in \mathbb{Z}^d} m_j m_k m_l P_{0m} \, e^{im\cdot x}, \tag{3.4.2}$$

so that

$$\begin{aligned}
\left\|\partial_x^3 P_0\right\|_{r,s,y_0} &\overset{def}{=} \sup_{|b|_1=|c|_1=1} \sum_{j=1}^d \sup_{|a|_1=1} \left\| \sum_{k,l=1}^d \frac{\partial^3 P_0}{\partial x_j \partial x_k \partial x_l}(\cdot) a_j b_k c_l \right\|_{r,s,y_0} \\
&= \sup_{\substack{|b|_1=1\\|c|_1=1}} \sum_{j=1}^d \sup_{|a|_1=1} \left(\left\| \sum_{l=1}^d \frac{\partial^3 P_0}{\partial x_j^2 \partial x_l}(\cdot) a_j b_j c_l \right\|_{r,s,y_0} + \right. \\
&\quad \left. \left\| \sum_{\substack{1\leqslant k\leqslant d\\k\neq j}} \frac{\partial^3 P_0}{\partial x_j^2 \partial x_k}(\cdot) a_j b_k c_j \right\|_{r,s,y_0} + \left\| \sum_{\substack{1\leqslant k\leqslant d\\k\neq j}} \frac{\partial^3 P_0}{\partial x_j \partial x_k^2}(\cdot) a_j b_k c_k \right\|_{r,s,y_0} \right) \\
&\overset{(3.4.2)}{=} \sup_{|b|_1=|c|_1=1} \sum_{j=1}^d \sup_{|a|_1=1} \left(|a_j||b_j| \sum_{m\in\mathbb{Z}^d} |m_j|^2 \left| \sum_{l=1}^d c_l m_l \right| |P_{0m}|\, e^{s|m|_1} + \right. \\
&\qquad |a_j||c_j| \sum_{m\in\mathbb{Z}^d} |m_j|^2 \left| \sum_{\substack{1\leqslant k\leqslant d\\k\neq j}} b_k m_k \right| |P_{0m}|\, e^{s|m|_1} + \\
&\qquad \left. |a_j| \sum_{m\in\mathbb{Z}^d} |m_j| \left| \sum_{\substack{1\leqslant k\leqslant d\\k\neq j}} b_k c_k m_k^2 \right| |P_{0m}|\, e^{s|m|_1} \right) \\
&\leqslant \sum_{m\in\mathbb{Z}^d} |P_{0m}|\, e^{s|m|_1} \sup_{\substack{|b|_1=1\\|c|_1=1}} \left(\sum_{j=1}^d |b_j||m_j|^2 \left| \sum_{l=1}^d c_l m_l \right| + \right. \\
&\qquad \left. \sum_{j=1}^d |c_j||m_j|^2 \left| \sum_{\substack{1\leqslant k\leqslant d\\k\neq j}} b_k m_k \right| + \sum_{j=1}^d |m_j| \left| \sum_{\substack{1\leqslant k\leqslant d\\k\neq j}} b_k c_k m_k^2 \right| \right) \\
&\leqslant \sum_{m\in\mathbb{Z}^d} |P_{0m}|\, e^{s|m|_1} (|m|^3 + |m|^3 + |m|_1 |m|^2) \\
&\overset{(3.4.1)}{=} 2\frac{1}{2} \left(3\, e^s + 4(d-1)\, e^{2s} \right) \\
&= 3\, e^s + 4(d-1)\, e^{2s} =: E_{3,0}.
\end{aligned}$$

Therefore

$$\widehat{V} = r, \qquad \mathtt{A}_1 = \mathtt{A}_2 = |\omega|, \qquad \mathtt{A}_3 = 0, \qquad \mathtt{A}_4 = 3\,e^s + 4(d-1)\,e^{2s},$$
$$\mathtt{A}_5 = \mathtt{A}_6 = \max\left\{3\,e^s + 4(d-1)\,e^{2s}\,, |\omega|^2\right\}, \qquad \mathtt{A}_7 = \alpha^{-2}\max\left\{3\,e^s + 4(d-1)\,e^{2s}\,, |\omega|^2\right\},$$
$$\mathtt{A}_8 = (s-\hat{s})^{2\tau}\max\left\{1\,,\,\frac{|\omega|}{r}\right\}, \qquad \mathtt{A}_9 = \max\left\{\mathtt{A}_7\,,\,\mathtt{A}_8\right\}, \qquad \mathtt{A}_* = \alpha^{-2}|\varepsilon|\widehat{G}.$$

Therefore, Theorem 2.1.6 holds for

$$\boxed{|\varepsilon| \leqslant \varepsilon_* := \frac{\alpha^2(s-\hat{s})^{2(2\tau+1)}}{109\cdot 2^{8\tau+13}\tau!^4\mathtt{A}_9\widehat{G}}\,.} \tag{3.4.3}$$

In particular, for $d = 2$, we have the following.

Corollary 3.4.1 *Consider the hamiltonian* $H(y_1, y_2, x_1, x_2; \varepsilon) := \frac{y_1^2+y_2^2}{2} + \varepsilon(\cos x_1 + \cos(x_2 - x_1))$ *and* $\omega := \left(\frac{\sqrt{5}-1}{2}, 1\right)$*. Then, for any* $|\varepsilon| < \varepsilon_*$*, there exists a Kronecker's invariant torus* $\mathscr{T}_{y_0,\omega}$ *for* H *i.e.*

KAM theorem	Parameters	ε_*
Kolmogorov	$r = 1,\ \sigma = 1/20$	9.18337×10^{-30}
Arnold	$r = 1,\ \sigma = 1/20$	2.02258×10^{-49}
Moser	$r = 1.73502 \times 10^{-15}, \quad \sigma = 1/20$ $h = 2.53148 \times 10^{-10} + 4.46141 \times 10^{20}\|\varepsilon\|$	6.12208×10^{-37}
Salamon–Zehnder	$r = 1,\ s = 1,\ \hat{s} = 1/10$	7.38385×10^{-27}

Table 3.1: Values of ε_* according to the KAM theorem

4 | Global symplectic extension of Arnold's theorem

4.1 Assumptions

Let $\alpha, r_0 > 0$, $\tau \geqslant d-1$, $0 < s_0 \leqslant 1$, $\mathscr{D} \subset \mathbb{R}^d$ be a non–empty, bounded domain[69] and consider the Hamiltonian parametrized by $\varepsilon \in \mathbb{R}$

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

where K, P are real–analytic functions with bounded holomorphic extensions to[70]

$$D_{r_0,s_0}(\mathscr{D}) := \bigcup_{y_0 \in \mathscr{D}} D_{r_0,s_0}(y_0)\,,$$

the norm being

$$\|\cdot\|_{r_0,s_0,\mathscr{D}} := \sup_{D_{r_0,s_0}(\mathscr{D})} |\cdot|\,.$$

Assume that

$$|\det K_{yy}(y)| \neq 0\,, \quad \forall\, y \in \mathscr{D}\,. \tag{4.1.1}$$

Define

$$\begin{aligned}
\Delta_\alpha^\tau &:= \left\{\omega \in \mathbb{R}^d : \quad |\omega \cdot k| \geqslant \frac{\alpha}{|k|_1^\tau}\,, \quad \forall\, k \in \mathbb{Z}^d \backslash \{0\}\right\}\,,\\
\mathscr{D}_{r_0,\alpha} &:= \left\{y_0 \in \mathscr{D} : \ \operatorname{dist}(y_0, \partial\mathscr{D}) \geqslant \frac{r_0}{32d} \ \text{ and } \ K_y(y_0) \in \Delta_\alpha^\tau\right\}\,,\\
T &: \mathscr{D}_{r_0,\alpha} \ni y_0 \mapsto K_{yy}(y_0)^{-1} \in \operatorname{Iso}(\mathbb{R}^d)\,.
\end{aligned}$$

[69] *i.e.* open and connected.
[70] Recall the notations in §1.2

Finally, for $\varepsilon \neq 0$ given, let[71]

$$
\begin{aligned}
M_0 &:= \|P\|_{r_0,s_0,\mathscr{D}_{r_0,\alpha}} \ ,\\
\mathsf{K}_0 &:= \|K_{yy}\|_{r_0,\mathscr{D}_{r_0,\alpha}} \ ,\\
\epsilon &:= \frac{\mathsf{K}_0|\varepsilon|M_0}{\alpha^2} \ ,\\
\mathsf{T}_0 &:= \|T\|_{\mathscr{D}_{r_0,\alpha}} = \sup_{y_0\in\mathscr{D}_{r_0,\alpha}} \|T(y_0)\| \ ,\\
\mathsf{K}_\infty &:= \mathsf{K}_0\, e^{\frac{1}{3}} \ ,\\
\mathsf{T}_\infty &:= \mathsf{T}_0\, e^{\frac{1}{3}} \ ,\\
\mathsf{C}_0 &:= 4\left(\frac{3}{2}\right)^{2\tau+d+2} \int_{\mathbb{R}^d} \left(|y|_1^{\tau+1} + d|y|_1^{2\tau+2}\right) e^{-|y|_1} dy \ ,\\
\mathsf{C}_1 &:= 2\left(\frac{3}{2}\right)^{\tau+d+1} \int_{\mathbb{R}^d} |y|_1^{\tau+1}\, e^{-|y|_1} dy \ ,\\
\mathsf{C}_2 &:= 2^{3d} d \ ,\\
\mathsf{C}_3 &:= d^2\mathsf{C}_1^2 + 6d\mathsf{C}_1 + \mathsf{C}_2 \ ,\\
\mathsf{C}_4 &:= \mathsf{C}_0 + \frac{24\,e}{3}\mathsf{C}_1 \ ,\\
\mathsf{C}_5 &:= \max\left\{4d\mathsf{T}_\infty\mathsf{K}_\infty \, , \, \frac{3}{2^{3\tau+7}}\mathsf{C}_1\right\} \ ,\\
0 < \sigma_0 &< \min\left\{\frac{s_0}{2}, \, 2^{-2(\tau-1)}\mathsf{C}_5\sqrt{2}\right\} \ ,\\
s_* &:= s_0 - 2\sigma_0 \ ,\\
\mathsf{C}_6 &:= \frac{16\mathsf{C}_5\sqrt{2}}{\sigma_0} \ ,\\
\mathsf{C}_7 &:= \max\left\{\mathsf{C}_3, \, \mathsf{C}_4\right\} \ ,\\
\mathsf{C}_8 &:= \left(1 + \frac{1}{3}\, e^{\frac{1}{3}}\right)^d - 1 \ ,\\
\mathsf{R} &:= 4\alpha\mathsf{T}_\infty \ ,\\
\lambda_0 &:= \log \epsilon^{-1} \ ,\\
r_1 &:= \frac{\mathsf{R}}{16\mathsf{C}_5\left(4\sigma_0^{-1}\lambda_0\right)^{\tau+1}} \ ,\\
\mathsf{a}_* &:= 3\cdot 2^{2\tau+\frac{5}{2}}\sigma_0^{-(3\tau+2d+4)}\mathsf{C}_5 \max\left\{2^{2(\tau+2)}\sigma_0^{d+1}\, , \, \mathsf{C}_7\lambda_0^{-(\tau+1)}\sqrt{2}\right\} \ ,\\
\mathsf{b}_* &:= 3\cdot 2^{2(\tau+1)}\mathsf{C}_5\ \sigma_0^{-(3\tau+2d+4)} \max\left\{\frac{16\alpha\mathsf{T}_\infty}{r_0}\sigma_0^{\tau+d+1}\, , \, \mathsf{C}_7\max\left\{1, \, \frac{r_0\mathsf{K}_\infty}{\alpha}\right\}\right\} \ ,\\
\mathsf{c}_* &:= \exp\left(-\frac{\sigma_0}{4}\left(\frac{\mathsf{R}}{r_0\sigma_0}\right)^{\frac{1}{\tau+1}}\right) \ ,\\
\mathsf{d}_* &:= 2^{2\tau+2d+3}\mathsf{C}_6^2 \ ,
\end{aligned}
$$

[71] Recall from footnote[13] that $\mathsf{C}_0, \mathsf{C}_1 > 1$.

$$
\begin{aligned}
\mathsf{e}_* &:= \frac{\lambda_0^{2(\tau+1)}}{\alpha^2 \mathsf{T}_\infty} \cdot \mathsf{a}_* \ , \\
\mathsf{f}_* &:= \frac{\lambda_0^{\tau+1}}{\alpha^2 \mathsf{T}_\infty} \cdot \mathsf{b}_* \ , \\
\mathsf{g}_* &:= \frac{\mathsf{b}_*}{\mathsf{T}_\infty \mathsf{K}_0} \ , \\
\mathsf{h}_* &:= \frac{4\mathsf{a}_* \ \mathsf{b}_* \ \mathsf{d}_* \sigma_0}{3\mathsf{T}_\infty^2 \mathsf{K}_0^2} \ , \\
\varepsilon_1 &:= \frac{1}{3}\mathsf{f}_* \sigma_0^{d+1} |\varepsilon| M_0 + \frac{4}{9}\mathsf{f}_*\mathsf{e}_*\mathsf{d}_*\sigma_0 |\varepsilon|^2 M_0^2 = \frac{1}{3}\mathsf{g}_* \ \sigma_0^{d+1} \ \epsilon \left(\log \epsilon^{-1}\right)^{\tau+1} + \frac{1}{3}\mathsf{h}_* \ \epsilon^2 \left(\log \epsilon^{-1}\right)^{3(\tau+1)} \ , \\
\varepsilon_2 &:= \frac{1}{3}\mathsf{f}_* \sigma_0 |\varepsilon| M_0 + \frac{4}{9}\mathsf{f}_*\mathsf{e}_*\mathsf{d}_*\sigma_0 |\varepsilon|^2 M_0^2 = \frac{1}{3}\mathsf{g}_* \ \sigma_0 \ \epsilon \left(\log \epsilon^{-1}\right)^{\tau+1} + \frac{1}{3}\mathsf{h}_* \ \epsilon^2 \left(\log \epsilon^{-1}\right)^{3(\tau+1)} \ , \\
\varepsilon_\sharp &:= \mathsf{c}_* \frac{\alpha^2}{\mathsf{K}_0 M_0} \ .
\end{aligned}
$$

4.2 Statement of the extension Theorem

Theorem 4.2.1 *Under the assumptions in §4.1, we have the following. For any given ε such that*[72]

$$
\boxed{\epsilon \leqslant \mathsf{c}_* \ , \qquad \mathsf{g}_* \ \epsilon \left(\log \epsilon^{-1}\right)^{\tau+1} \leqslant 1 \ , \qquad \mathsf{h}_* \ \epsilon^2 \left(\log \epsilon^{-1}\right)^{3(\tau+1)} \leqslant 1 \ ,} \tag{4.2.1}
$$

there exist $\mathscr{D}_* \subset \mathscr{D}$ having the same cardinality as $\mathscr{D}_{r_0,\alpha}$, a lipeomorphism $G_* \colon \mathscr{D}_{r_0,\alpha} \xrightarrow{onto} \mathscr{D}_*$, a C^∞ map $K_* \colon \mathscr{D} \to \mathbb{R}$ and a C^∞–symplectomorphism $\phi_* \colon \mathscr{D} \times \mathbb{T}^d \circlearrowleft$, real–analytic in $x \in \mathbb{T}^d_{s_*}$ and such that the following hold.

$$
\begin{aligned}
\partial_{y_*} K_* \circ G_* &= \partial_y K \qquad \text{on} \quad \mathscr{D}_{r_0,\alpha} \ , && (4.2.2) \\
\partial_{y_*}^\beta H \circ \phi_*(y_*, x) &= \partial_{y_*}^\beta K_*(y_*), \qquad \forall \ (y_*, x) \in \mathscr{D}_* \times \mathbb{T}^d, \quad \forall \ \beta \in \mathbb{N}_0^d && (4.2.3)
\end{aligned}
$$

and

$$
\begin{aligned}
|\operatorname{meas}(\mathscr{D}_*) - \operatorname{meas}(\mathscr{D}_{r_0,\alpha})| &\leqslant \mathsf{C}_8 \ \varepsilon_2 \ e^{\varepsilon_2} \operatorname{meas}(\mathscr{D}_{r_0,\alpha}) \ , && (4.2.4) \\
|\mathsf{W}_0(\phi_* - \mathrm{id})| &\leqslant \varepsilon_1 \qquad \text{on} \quad \mathscr{D}_* \times \mathbb{T}^d_{s_*} \ , && (4.2.5)
\end{aligned}
$$

where

$$
\mathsf{W}_0 := \operatorname{diag}\left(\frac{1}{4r_1}\mathbb{1}_d, \mathbb{1}_d\right) \ .
$$

[72] Notice that (4.2.1) equivals to: $|\varepsilon| \leqslant \varepsilon_\sharp$, $|\varepsilon| \, \mathsf{f}_* \, \|P\|_{r_0,s_0,\mathscr{D}_{r_0,\alpha}} \leqslant 1$ and $4|\varepsilon|^2 \, \mathsf{f}_* \, \mathsf{e}_* \, \mathsf{d}_*^2 \, \sigma_0 \, \|P\|^2_{r_0,s_0,\mathscr{D}_{r_0,\alpha}} \leqslant 3$.

Remark 4.2.2 From (4.2.3), on deduces that the d–tori

$$\mathcal{T}_{\omega_*,\varepsilon} := \phi_*\left(y_*, \mathbb{T}^d\right), \qquad y_* \in \mathscr{D}_*, \quad \omega_* := \partial_{y_*} K_*(y_*), \tag{4.2.6}$$

are non-degenrate invariant Kronecker tori for H *i.e.*

$$\phi_H^t \circ \phi_*(y_*, x) = \phi_*(y_*, x + \omega_* t), \qquad \forall\, x \in \mathbb{T}^d. \tag{4.2.7}$$

4.3 Proof of Theorem 4.2.1

KAM step Given $r, s, K, P, \mathscr{D}, \mathscr{D}_\sharp$ satisfying the assumptions 4.1, we seek for $r_1 < r$, $s_1 < s$, a set $\mathscr{D}'_\sharp \subset D_{r_1}(\mathscr{D}_\sharp)$ having the same cardinality as $\mathscr{D}_\sharp$ and a near–to–the–identity real–analytic symplectic transformation $\phi_1 : \mathscr{D} \times \mathbb{T}^d \circlearrowleft$ satisfying

$$\phi_1 \colon D_{r_1,s_1}(\mathscr{D}'_\sharp) \to D_{r,s}(\mathscr{D}_\sharp),$$

with $D_{r_1,s_1}(\mathscr{D}'_\sharp) \subset D_{r,s}(\mathscr{D}_\sharp)$ and ϕ_1 generated by an extension $y' \cdot x + \varepsilon \hat{g}(y', x)$ of a function of the form $y' \cdot x + \varepsilon g(y', x)$ *i.e.*

$$\phi_1 \colon \begin{cases} y = y' + \varepsilon \hat{g}_x(y', x) \\ x' = x + \varepsilon \hat{g}_{y'}(y', x), \end{cases} \tag{4.3.1}$$

such that

$$\begin{cases} H_1 := H \circ \phi_1 = K_1 + \varepsilon^2 P_1, \quad K_1 = K_1(y'), \quad \text{on } D_{r_1,s_1}(\mathscr{D}'_\sharp), \\ \det \partial^2_{y'} K_1(y_1) \neq 0, \quad \forall\, y_1 \in \mathscr{D}'_\sharp, \\ \partial_{y'} K_1(\mathscr{D}'_\sharp) = \partial_y K(\mathscr{D}_\sharp). \end{cases} \tag{4.3.2}$$

By Taylor's formula, we get[73]

$$\begin{aligned} H(y' + \varepsilon g_x(y', x), x) =& K(y') + \varepsilon P_0(y') + \varepsilon \left[K_y(y') \cdot g_x + T_\kappa P(y', \cdot) - P_0(y')\right] + \\ &+ \varepsilon^2 \left(P^{(1)} + P^{(2)} + P^{(3)}\right)(y', x) \\ =& K_1(y') + \varepsilon \left[K_y(y') \cdot g_x + T_\kappa P(y', \cdot) - P_0(y')\right] + \varepsilon^2 P'(y', x), \end{aligned} \tag{4.3.3}$$

[73] Recall that $\langle \cdot \rangle$ stands for the average over $\mathbb{T}^d$.

with $\kappa \in \mathbb{N}$, which will be chosen large enough so that $P^{(3)} = O(\varepsilon)$, $P_0(y') := \langle P(y', \cdot)\rangle$ and

$$\begin{cases} K_1 := K(y') + \varepsilon P_0(y') =: K(y') + \varepsilon \widetilde{K}(y') \\ P' := P^{(1)} + P^{(2)} + P^{(3)} \\ P^{(1)} := \frac{1}{\varepsilon^2}\left[K(y' + \varepsilon g_x) - K(y') - \varepsilon K_y(y') \cdot g_x\right] = \int_0^1 (1-t) K_{yy}(\varepsilon t g_x) \cdot g_x \cdot g_x dt \\ P^{(2)} := \frac{1}{\varepsilon}\left[P(y' + \varepsilon g_x, x) - P(y', x)\right] = \int_0^1 P_y(y' + \varepsilon t g_x, x) \cdot g_x dt \\ P^{(3)} := \frac{1}{\varepsilon}\left[P(y', x) - T_\kappa P(y', \cdot)\right] = \frac{1}{\varepsilon} \sum_{|n|_1 > \kappa} P_n(y')\, e^{in\cdot x}\,. \end{cases} \tag{4.3.4}$$

By the non–degeneracy condition in (4.1.1) and Lemma 2.2.7, for ε small enough (to be made precised below), there exists $\bar{r} \leqslant r$ such that for each $y_0 \in \mathscr{D}_\sharp$, there exists a unique $y_1 \in D_{\bar{r}}(y_0)$ satisfying $\partial_{y'} K_1(y_1) = \partial_y K(y_0)$ and $\det \partial^2_{y'} K_1(y_1) \neq 0$; $\mathscr{D}'_\sharp$ is precisely the set of those y_1 when y_0 runs in $\mathscr{D}_\sharp$. More precisely, $\mathscr{D}'_\sharp$ and $\mathscr{D}_\sharp$ are "diffeomorphic"[74], say via G, and, for each $y_1 \in \mathscr{D}'_\sharp$, the matrix $\partial^2_{y'} K_1(y_1)$ is invertible with inverse of the form

$$T_1(y_1) := \partial^2_{y'} K_1(y_1)^{-1} =: T(y_0) + \varepsilon \widetilde{T}(y_1), \quad y_1 = G(y_0).$$

Write

$$y_1 =: y_0 + \varepsilon \widetilde{y}\,, \quad y_1 = G(y_0)\,, \quad \forall\, y_1 \in \mathscr{D}'_\sharp. \tag{4.3.5}$$

In view of (4.3.3), in order to get the first part of (4.3.2), we need to find g such that $K_y(y') \cdot g_x + T_\kappa P(y', \cdot) - P_0(y')$ vanishes; such a g is indeed given by

$$g := \sum_{0 < |n|_1 \leqslant \kappa} \frac{-P_n(y')}{i K_y(y') \cdot n}\, e^{in\cdot x}, \tag{4.3.6}$$

provided that

$$K_y(y') \cdot n \neq 0, \quad \forall\, 0 < |n|_1 \leqslant \kappa, \quad \forall\, y' \in D_{r_1}(\mathscr{D}'_\sharp) \quad (\subset D_r(\mathscr{D}_\sharp))\,. \tag{4.3.7}$$

But, in fact, since $K_y(y_0)$ is rationally independent, for each $y_0 \in \mathscr{D}_\sharp$, then, given any $\kappa \in \mathbb{N}$, there exists $r' \leqslant r$ such that

$$K_y(y') \cdot n \neq 0, \quad \forall\, 0 < |n|_1 \leqslant \kappa, \quad \forall\, y' \in D_{r'}(\mathscr{D}_\sharp). \tag{4.3.8}$$

[74] *i.e.* there a exits a bijection from $\mathscr{D}_\sharp$ onto $\mathscr{D}'_\sharp$ which extends to a diffeomorphism on some neighborhood of $\mathscr{D}_\sharp$.

Then we invert the function $x \mapsto x + \varepsilon \hat{g}_{y'}(y', x)$ in order to define P_1. But, by Lemma 2.2.6, for ε small enough, the map $x \mapsto x + \varepsilon \hat{g}_{y'}(y', x)$ admits an real–analytic inverse of the form

$$\varphi(y', x'; \varepsilon) := x' + \varepsilon \widetilde{\varphi}(y', x'; \varepsilon), \tag{4.3.9}$$

so that the Arnod's symplectic transformation is given by

$$\phi_1 \colon (y', x') \mapsto \begin{cases} y = y' + \varepsilon \hat{g}_x(y', \varphi(y', x')) \\ x = \varphi(y', x'; \varepsilon) = x' + \varepsilon \widetilde{\varphi}(y', x'; \varepsilon). \end{cases} \tag{4.3.10}$$

Hence, (4.3.2) holds with

$$P_1(y', x') := P'(y', \varphi(y', x')). \tag{4.3.11}$$

Finally, we extend K_1.

Next, we make a quantitative evaluation of the above construction. Assume that[75] $H(y, x; \varepsilon) := K(y) + \varepsilon P(y, x)$, where K, P are real–analytic functions with bounded holomorphic extensions to $D_{r,s}(\mathscr{D})$ and

$$\begin{array}{lll} \mathscr{D}_\sharp \subset \left\{ y_0 \in \mathscr{D} : \ \operatorname{dist}(y_0, \partial \mathscr{D}) \geqslant \dfrac{r}{32d} \ \text{ and } \ K_y(y_0) \in \Delta_\alpha^\tau \right\} \ , & & \\ \det K_{yy}(y) \neq 0 \ , & T(y) := K_{yy}(y)^{-1} \ , \forall\, y \in \mathscr{D}_\sharp & \\ \|K_{yy}\|_{r, \mathscr{D}_\sharp} \leqslant \mathsf{K} < \mathsf{K}_\infty \ , & \|T\|_{\mathscr{D}_\sharp} \leqslant \mathsf{T} < \mathsf{T}_\infty \ , & \\ \|P\|_{r,s,\mathscr{D}_\sharp} \leqslant M \ , & \omega \in \Delta_\alpha^\tau \ , & r \leqslant r_0 \ , \end{array} \tag{4.3.12}$$

Fix $0 < 2\sigma < s \leqslant 1$ and fix $\varepsilon \neq 0$ in such away that,

$$\begin{array}{l} \lambda \geqslant \log \left(\dfrac{\alpha^2}{\mathsf{K} |\varepsilon| M} \right) > 1, \quad \kappa := 4\sigma^{-1} \lambda, \quad \bar{r} := \dfrac{1}{4\mathsf{C}_5} \min \left(\dfrac{\mathsf{R}}{\kappa^{\tau+1}} \, , \, r\sigma \right) , \\ \tilde{r} := \dfrac{\bar{r}}{4\mathsf{C}_5}, \quad \bar{s} := s - \dfrac{2}{3}\sigma, \quad s' := s - \sigma \, , \end{array} \tag{4.3.13}$$

so that [76]

$$\bar{r} \leqslant \frac{r\sigma}{4\mathsf{C}_5} = \frac{\sigma}{16 d \mathsf{T}_\infty \mathsf{K}_\infty} r \leqslant \frac{r}{32d} < \frac{r}{2} \quad \text{and} \quad \kappa > 8 \, . \tag{4.3.14}$$

[75] 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.

[76] Recall footnote 42.

Lemma 4.3.1 *Let*[77]

$$\begin{aligned}\bar{\mathsf{L}} &:= \mathsf{C}_4 \max\{\alpha, r\mathsf{K}\}\frac{M}{\alpha^2\bar{r}}\sigma^{-(2\tau+d+2)}\ ,\\ \mathsf{L} &:= M\max\left\{\frac{16\mathsf{T}_\infty}{r\bar{r}}\sigma^{-(\tau+d+1)}\ ,\ \mathsf{C}_7\max\{\alpha, r\mathsf{K}\}\frac{1}{\alpha^2\bar{r}}\sigma^{-2(\tau+d+1)}\right\}\\ &= M\max\left\{\frac{16\mathsf{T}_\infty}{r\bar{r}}\sigma^{-(\tau+d+1)}\ ,\ \frac{4}{\mathsf{K}r^2}\ ,\ \mathsf{C}_7\max(\alpha, r\mathsf{K})\frac{1}{\alpha^2\bar{r}}\sigma^{-2(\tau+d+1)}\right\}\ .\end{aligned}$$

Then

$$\begin{cases}\|g_x\|_{\bar{r},\bar{s},\mathscr{D}_\sharp} \leqslant \mathsf{C}_1\dfrac{M}{\alpha}\sigma^{-(\tau+d+1)}\ ,\\ \|g_{y'}\|_{\bar{r},\bar{s},\mathscr{D}_\sharp},\ \|\partial^2_{y'x}g\|_{\bar{r},\bar{s},\mathscr{D}_\sharp} \leqslant \bar{\mathsf{L}}\,,\\ \|\partial^2_{y'}\widetilde{K}\|_{\bar{r},\mathscr{D}_\sharp} \leqslant \mathsf{K}\mathsf{L}\,.\end{cases} \tag{4.3.15}$$

If $\varepsilon_* > 0$ *satisfies*

$$\varepsilon_* \leqslant \varepsilon_\sharp \quad \textit{and} \quad \varepsilon_*\mathsf{L} \leqslant \frac{\sigma}{3}\ , \tag{4.3.16}$$

then, for any $|\varepsilon| \leqslant \varepsilon_*$, *there exists a diffeomorphism* $G\colon D_{\tilde{r}}(\mathscr{D}_\sharp) \to G(D_{\tilde{r}}(\mathscr{D}_\sharp))$, $\partial_{y'}K_1 \circ G = \partial_y K$ *and such that* $\mathscr{D}'_\sharp := G(\mathscr{D}_\sharp) \subset B_{\bar{r}}(\mathscr{D}_\sharp)$,

$$\begin{cases}|\varepsilon|\|g_x\|_{\bar{r},\bar{s},\mathscr{D}_\sharp} \leqslant \dfrac{r}{3}\,, & \|G-\mathrm{id}\|_{\tilde{r},\mathscr{D}_\sharp} \leqslant \dfrac{\bar{r}}{2}\ ,\\ |\varepsilon|\|\widetilde{T}\|_{\mathscr{D}'_\sharp} \leqslant \mathsf{T}|\varepsilon|\mathsf{L}\,, & \|\partial_z G - \mathbb{1}_d\|_{\tilde{r},\mathscr{D}_\sharp} \leqslant |\varepsilon|\mathsf{L}\,,\\ \|P'\|_{\bar{r},\bar{s},\mathscr{D}_\sharp} \leqslant \mathsf{L}M\ , & B_{\bar{r}/4}(\mathscr{D}'_\sharp) \subset B_{\bar{r}/2}(\mathscr{D}_\sharp) \subset \mathscr{D}\end{cases} \tag{4.3.17}$$

and the following hold. g *has and extension* $\hat{g}\colon \mathbb{R}^d \times \mathbb{T}^d \to \mathbb{R}$ *and, for any* $|\varepsilon| \leqslant \varepsilon_*$ *and* $y' \in D_{\bar{r}/2}(\mathscr{D}_\sharp)$, *the map* $\psi_\varepsilon(x) := x + \varepsilon\hat{g}_{y'}(y', x)$ *has an analytic inverse* $\varphi(x') = x' + \varepsilon\widetilde{\varphi}(y', x'; \varepsilon)$ *such that, for all* $|\varepsilon| \leqslant \varepsilon_*$,

$$\|\widetilde{\varphi}\|_{\bar{r}/4,s',y_0} \leqslant \bar{\mathsf{L}} \qquad \text{and} \qquad \varphi = \mathrm{id} + \varepsilon\widetilde{\varphi} : D_{\bar{r}/4,s'}(\mathscr{D}'_\sharp) \to \mathbb{T}^d_{\bar{s}}\ ; \tag{4.3.18}$$

for any $y_0 \in \mathscr{D}_\sharp$ *and* $(y', x, \varepsilon) \in D_{\bar{r},\bar{s}}(y_0) \times D^1_{\varepsilon_*}(0)$, $|y' + \varepsilon g_x(y', x) - y_0| < \frac{5}{6}r$; *the map* ϕ_1 *is a symplectic diffeomorphism and*

$$\phi_1 = \big(y' + \varepsilon g_x(y', \varphi(y', x')), \varphi(y', x')\big) : D_{\bar{r}/4,s'}(\mathscr{D}'_\sharp) \to D_{2r/3,\bar{s}}(\mathscr{D}_\sharp), \tag{4.3.19}$$

with

$$\|\mathsf{W}\,\tilde{\phi}\|_{\bar{r}/4,s',\mathscr{D}'_\sharp} \leqslant \bar{\mathsf{L}}\,, \tag{4.3.20}$$

[77] Notice that $\mathsf{L} \geqslant \sigma^{-d}\bar{\mathsf{L}} \geqslant \bar{\mathsf{L}}$ since $\sigma \leqslant 1$. Notice also that $\mathsf{T}\mathsf{K} \geqslant 1$, so that $\frac{16\mathsf{T}_\infty}{r\bar{r}}\sigma^{-(\tau+d+1)} \geqslant \frac{16\mathsf{T}}{r^2} \geqslant \frac{4}{\mathsf{K}r^2}$.

where $\tilde\phi$ *is defined by the relation* $\phi_1 =: \mathrm{id} + \varepsilon\tilde\phi$,

$$\mathsf{W} := \begin{pmatrix} \frac{1}{\bar r}\mathbb{1}_d & 0 \\ 0 & \mathbb{1}_d \end{pmatrix}$$

and

$$\|P_1\|_{\bar r/4, s', \mathscr{D}'_\sharp} \leqslant \mathsf{L}M \;. \tag{4.3.21}$$

Moreover, K_1 *possesses a* C^∞*–extensions* $\widehat K_1\colon \mathbb{R}^d \to \mathbb{R}$ *such that for any* $n \in \mathbb{N}_0$*, there exists* $C_n \in \mathbb{N}$ *and for any* $\beta_1, \beta_2 \in \mathbb{N}_0^d$ *with* $|\beta_1|_1 + |\beta_2|_1 \leqslant n$*,*

$$\bar r^{|\beta_1|_1}\sigma^{|\beta_2|_1}\|\partial_{y'}^{\beta_1}\partial_{x'}^{\beta_2}\mathsf{W}(\phi_1 - \mathrm{id})\|_{0,0} \leqslant C_n|\varepsilon|\bar{\mathsf{L}}\,, \tag{4.3.22}$$

$$\bar r^{|\beta_1|_1}\|\partial_{y'}^{\beta_1}(\widehat K_1 - K)\|_0 \leqslant C_n|\varepsilon|M\;. \tag{4.3.23}$$

Proof We begin by extending the "diophantine condition w.r.t. K_y" uniformly to $D_{\bar r}(\mathscr{D}_\sharp)$ up to the order κ. Indeed, for any $y_0 \in \mathscr{D}_\sharp$, $0 < |n|_1 \leqslant \kappa$ and $y' \in D_{\bar r}(y_0)$,

$$\begin{aligned}
|K_y(y')\cdot n| &= |\omega\cdot n + (K_y(y') - K_y(y_0))\cdot n| \geqslant |\omega\cdot n|\left(1 - d\frac{\|K_{yy}\|_{\bar r,\mathscr{D}_\sharp}}{|\omega\cdot n|}|n|_1\bar r\right)\\
&\geqslant \frac{\alpha}{|n|_1^\tau}\left(1 - \frac{d\mathsf{K}}{\alpha}|n|_1^{\tau+1}\bar r\right) \geqslant \frac{\alpha}{|n|_1^\tau}\left(1 - \frac{d\mathsf{K}}{\alpha}\kappa^{\tau+1}\bar r\right) \geqslant \frac{\alpha}{|n|_1^\tau}\left(1 - \frac{d\mathsf{K}_\infty}{\alpha}\kappa^{\tau+1}\bar r\right)\\
&\geqslant \frac{\alpha}{2|n|_1^\tau},
\end{aligned} \tag{4.3.24}$$

so that, by Lemma 2.2.4–(i), we have

$$\begin{aligned}
\|g\|_{\bar r,\bar s,\mathscr{D}_\sharp} &\overset{def}{=} \sup_{D_{\bar r,\bar s}(\mathscr{D}_\sharp)}\left|\sum_{0<|n|_1\leqslant\kappa}\frac{P_n(y')}{K_y(y')\cdot n}\,e^{in\cdot x}\right| \leqslant \sum_{0<|n|_1\leqslant\kappa}\frac{\|P_n\|_{\bar r,\bar s,\mathscr{D}_\sharp}}{|K_y(y')\cdot n|}\,e^{\left(s-\frac{2}{3}\sigma\right)|n|_1}\\
&\leqslant \sum_{0<|n|_1\leqslant\kappa} M\,e^{-s|n|_1}\frac{2|n|_1^\tau}{\alpha}\,e^{\left(s-\frac{2}{3}\sigma\right)|n|_1} \leqslant \frac{2M}{\alpha}\sum_{n\in\mathbb{Z}^d}|n|_1^\tau\,e^{-\frac{2}{3}\sigma|n|_1}\\
&\leqslant \frac{2M}{\alpha}\int_{\mathbb{R}^d}|y|_1^\tau\,e^{-\frac{2}{3}\sigma|y|_1}dy\\
&= \left(\frac{3}{2\sigma}\right)^{\tau+d}\frac{2M}{\alpha}\int_{\mathbb{R}^d}|y|_1^\tau\,e^{-|y|_1}dy\\
&\leqslant \mathsf{C}_1\frac{M}{\alpha}\sigma^{-(\tau+d)}
\end{aligned}$$

and analogously,

$$\begin{aligned}
\|g_x\|_{\bar r,\bar s,\mathscr{D}_\sharp} &\overset{def}{=} \sup_{D_{\bar r,\bar s}(\mathscr{D}_\sharp)} \left| \sum_{0<|n|_1\leqslant\kappa} \frac{nP_n(y')}{K_y(y')\cdot n}\, e^{in\cdot x}\right| \leqslant \sum_{0<|n|_1\leqslant\kappa} \frac{\|P_n\|_{\bar r,\bar s,\mathscr{D}_\sharp}}{|K_y(y')\cdot n|}|n|_1\, e^{\left(s-\frac{2}{3}\sigma\right)|n|_1}\\
&\leqslant \sum_{0<|n|_1\leqslant\kappa} M\, e^{-s|n|_1}\frac{2|n|_1^{\tau+1}}{\alpha}\, e^{\left(s-\frac{2}{3}\sigma\right)|n|_1} \leqslant \frac{2M}{\alpha}\sum_{n\in\mathbb{Z}^d}|n|_1^{\tau+1}\, e^{-\frac{2}{3}\sigma|n|_1}\\
&\leqslant \frac{2M}{\alpha}\int_{\mathbb{R}^d}|y|_1^{\tau+1}\, e^{-\frac{2}{3}\sigma|y|_1}dy\\
&= \left(\frac{3}{2\sigma}\right)^{\tau+d+1}\frac{2M}{\alpha}\int_{\mathbb{R}^d}|y|_1^{\tau+1}\, e^{-|y|_1}dy\\
&\leqslant \mathsf{C}_1\frac{M}{\alpha}\sigma^{-(\tau+d+1)}\,,
\end{aligned}$$

$$\begin{aligned}
\|\partial_{y'}g\|_{\bar r,\bar s,\mathscr{D}_\sharp} &\overset{def}{=} \sup_{D_{\bar r,\bar s}(\mathscr{D}_\sharp)}\left|\sum_{0<|n|_1\leqslant\kappa}\left(\frac{\partial_yP_n(y')}{K_y(y')\cdot n}-P_n(y')\frac{K_{yy}(y')n}{(K_y(y')\cdot n)^2}\right)\, e^{in\cdot x}\right|\\
&\leqslant \sum_{0<|n|_1\leqslant\kappa}\sup_{D_{\bar r}(\mathscr{D}_\sharp)}\left(\frac{\|(P_y)_n\|_{\bar r,s,\mathscr{D}_\sharp}}{|K_y(y')\cdot n|}+d\|P_n\|_{r,s,\mathscr{D}_\sharp}\frac{\|K_{yy}\|_{r,\mathscr{D}_\sharp}|n|_1}{|K_y(y')\cdot n|^2}\right)\, e^{\left(s-\frac{2}{3}\sigma\right)|n|_1}\\
&\overset{(4.3.12)+(4.3.24)}{\leqslant} \sum_{0<|n|_1\leqslant\kappa}\left(\frac{M}{r-\bar r}\, e^{-s|n|_1}\frac{2|n|_1^\tau}{\alpha}+dM\, e^{-s|n|_1}\mathsf{K}|n|_1\left(\frac{2|n|_1^\tau}{\alpha}\right)^2\right)\, e^{\left(s-\frac{2}{3}\sigma\right)|n|_1}\\
&\overset{(4.3.14)}{\leqslant} \frac{4M}{\alpha^2r}\sum_{0<|n|_1\leqslant\kappa}\left(|n|_1^\tau\alpha+dr\mathsf{K}|n|_1^{2\tau+1}\right)\, e^{-\frac{2}{3}\sigma|n|_1}\\
&\leqslant \max(\alpha,r\mathsf{K})\frac{4M}{\alpha^2r}\sum_{0<|n|_1\leqslant\kappa}\left(|n|_1^\tau+d|n|_1^{2\tau+1}\right)\, e^{-\frac{2}{3}\sigma|n|_1}\\
&\leqslant \max(\alpha,r\mathsf{K})\frac{4M}{\alpha^2r}\int_{\mathbb{R}^d}\left(|y|_1^\tau+d|y|_1^{2\tau+1}\right)\, e^{-\frac{2}{3}\sigma|y|_1}dy\\
&= \left(\frac{3}{2\sigma}\right)^{2\tau+d+1}\max(\alpha,r\mathsf{K})\frac{4M}{\alpha^2r}\int_{\mathbb{R}^d}\left(|y|_1^\tau+d|y|_1^{2\tau+1}\right)\, e^{-|y|_1}dy\\
&\leqslant \mathsf{C}_0\max(\alpha,r\mathsf{K})\frac{M}{\alpha^2r}\sigma^{-(2\tau+d+1)}\\
&\leqslant \bar{\mathsf{L}}\,,
\end{aligned}$$

$$
\begin{aligned}
\|\partial^2_{y'x} g\|_{\bar r,\bar s,\mathscr{D}_\sharp} &\overset{def}{=} \sup_{D_{\bar r,\bar s}(\mathscr{D}_\sharp)} \left| \sum_{0<|n|_1\leqslant\kappa} \left(\frac{\partial_y P_n(y')}{K_y(y')\cdot n} - P_n(y')\frac{K_{yy}(y')n}{(K_y(y')\cdot n)^2}\right)\cdot n\, e^{in\cdot x}\right| \\
&\leqslant \sum_{0<|n|_1\leqslant\kappa} \sup_{D_{\bar r}(\mathscr{D}_\sharp)} \left(\frac{\|(P_y)_n\|_{\bar r,s,\mathscr{D}_\sharp}}{|K_y(y')\cdot n|} + d\|P_n\|_{r,s,\mathscr{D}_\sharp}\frac{\|K_{yy}\|_{r,\mathscr{D}_\sharp}|n|_1}{|K_y(y')\cdot n|^2}\right)|n|_1\, e^{\left(s-\frac{2}{3}\sigma\right)|n|_1} \\
&\leqslant \max(\alpha, r\mathsf{K})\frac{4M}{\alpha^2 r}\sum_{0<|n|_1\leqslant\kappa}\left(|n|_1^{\tau}+d|n|_1^{2\tau+1}\right)|n|_1\, e^{-\frac{2}{3}\sigma|n|_1} \\
&\leqslant \max(\alpha, r\mathsf{K})\frac{4M}{\alpha^2 r}\int_{\mathbb{R}^d}\left(|y|_1^{\tau}+d|y|_1^{2\tau+1}\right)|y|_1\, e^{-\frac{2}{3}\sigma|y|_1}dy \\
&= \left(\frac{3}{2\sigma}\right)^{2\tau+d+2}\max(\alpha, r\mathsf{K})\frac{4M}{\alpha^2 r}\int_{\mathbb{R}^d}\left(|y|_1^{\tau+1}+d|y|_1^{2\tau+2}\right)\, e^{-|y|_1}dy \\
&= \mathsf{C}_0\max(\alpha, r\mathsf{K})\frac{M}{\alpha^2 r}\sigma^{-(2\tau+d+2)} \\
&= \bar{\mathsf{L}}\,,
\end{aligned}
$$

and, for $|\varepsilon|<\varepsilon_*$,

$$
\|\partial_{y'}\widetilde K\|_{\bar r,\mathscr{D}_\sharp} = \|\,[P_y]\,\|_{\bar r,\mathscr{D}_\sharp} \leqslant \|P_y\|_{\bar r,\bar s,\mathscr{D}_\sharp} \leqslant \frac{M}{r-\bar r}\leqslant \frac{2M}{r}\,,
$$

$$
\|\partial^2_{y'}\widetilde K\|_{\bar r,\mathscr{D}_\sharp} = \|\,[P_{yy}]\,\|_{\bar r,\mathscr{D}_\sharp} \leqslant \|P_{yy}\|_{\bar r,\bar s,\mathscr{D}_\sharp} \leqslant \frac{M}{(r-\bar r)^2}\leqslant \frac{4M}{r^2}\leqslant \mathsf{K}\mathsf{L}\,.
$$

Now, we extend the generating function and K_1 to $\mathbb{R}^d\times\mathbb{T}^d$ and $\mathbb{R}^d$ respectively, by making use of a cut–off function. Let then $\chi_1\in C(\mathbb{C}^d)\cap C^\infty(\mathbb{R}^d)$, with $0\leqslant\chi_1\leqslant 1$, $\operatorname{supp}\chi_1\subset D_{\bar r}(\mathscr{D}_\sharp)$, $\chi_1\equiv 1$ on $D_{\bar r/2}(\mathscr{D}_\sharp)$ and satisfying (2.2.1). Thus, given $x\in\mathbb{T}^d_{\bar s}$, set $\hat g(y',x')=\chi_1(y')g(y',x')$, $\hat K_1 = K+\chi_1\cdot(K_1-K)$ if $y'\in D_{\bar r}(\mathscr{D}_\sharp)$, $\hat g(y',x')=0$, $\hat K_1=K$ if $y'\in\left(\mathbb{C}^d\backslash D_{\bar r}(\mathscr{D}_\sharp)\right)\cap\mathscr{D}$ and $\hat K_1=K\equiv 0$ on $\mathbb{C}^d\backslash\left(D_{\bar r}(\mathscr{D}_\sharp)\cup\mathscr{D}\right)$.

$$
\|\hat g\|_{0,0}\leqslant\|g\|_{\bar r,\bar s,\mathscr{D}_\sharp}\leqslant \mathsf{C}_1\frac{M}{\alpha}\sigma^{-(\tau+d)}\,, \tag{4.3.25}
$$

$$
\|\hat g_x\|_{0,0}\leqslant\|g_x\|_{\bar r,\bar s,\mathscr{D}_\sharp}\leqslant \mathsf{C}_1\frac{M}{\alpha}\sigma^{-(\tau+d+1)}\,, \tag{4.3.26}
$$

$$
\|\hat g_{y'}\|_{0,0}\leqslant\|\partial_{y'}\chi_1\|_0\|g\|_{\bar r,\bar s,\mathscr{D}_\sharp}+\|g_{y'}\|_{\bar r,\bar s,\mathscr{D}_\sharp}\overset{(2.2.1)}{\leqslant}\mathsf{C}_4\max(\alpha,r\mathsf{K})\frac{M}{\alpha^2\bar r}\sigma^{-(2\tau+d+1)}\leqslant\bar{\mathsf{L}}\,, \tag{4.3.27}
$$

$$
\|\hat g_{y'x}\|_{0,0}\leqslant\|\partial_{y'}\chi_1\|_0\|g_x\|_{\bar r,\bar s,\mathscr{D}_\sharp}+\|g_{y'x}\|_{\bar r,\bar s,\mathscr{D}_\sharp}\overset{(2.2.1)}{\leqslant}\mathsf{C}_4\max(\alpha,r\mathsf{K})\frac{M}{\alpha^2\bar r}\sigma^{-(2\tau+d+2)}\leqslant\bar{\mathsf{L}}\,. \tag{4.3.28}
$$

And generally, for any $n \in \mathbb{N}_0$, there exists $C_n \in \mathbb{N}$ and for any $\beta_1, \beta_2 \in \mathbb{N}_0^d$ with $|\beta_1|_1 + |\beta_2|_1 \leqslant n$,

$$\bar{r}^{|\beta_1|_1}\sigma^{|\beta_2|_1-1}\|\partial_{y'}^{\beta_1}\partial_{x'}^{\beta_2}\hat{g}_{y'}\|_{0,0} \leqslant C_n\overline{\mathsf{L}}\,. \tag{4.3.29}$$

(4.3.23) is a straightforward consequence of Leibniz's rule.
Next, we construct $\mathscr{D}_\sharp'$ in (4.3.2) for $|\varepsilon| \leqslant \varepsilon_*$. For, fix $|\varepsilon| \leqslant \varepsilon_*$, $y_0 \in \mathscr{D}_\sharp$ and consider

$$\begin{aligned} F\colon D_{\bar{r}}(y_0) \times D_{\bar{r}}(y_0) &\longrightarrow \qquad \mathbb{C}^d \\ (y,z) \quad &\longmapsto K_y(y) + \varepsilon \widetilde{K}_{y'}(y) - K_y(z). \end{aligned}$$

Then

- $F_y(y_0,y_0) = \partial_y^2 K(y_0) + \varepsilon\partial_{y'}^2\widetilde{K}(y_0) = T(y_0)^{-1}\left(\mathbb{1}_d + \varepsilon T(y_0)\partial_{y'}^2\widetilde{K}(y_0)\right) =: T(y_0)^{-1}(\mathbb{1}_d + \varepsilon A_0)$ and

$$\|\varepsilon A_0\| \leqslant \|T(y_0)\|\|\varepsilon\partial_{y'}^2\widetilde{K}(y_0)\| \leqslant \mathsf{T}\frac{4|\varepsilon|M}{r^2} \overset{(4.3.14)}{\leqslant} |\varepsilon|\frac{2\mathsf{T}_\infty M}{r\bar{r}} \leqslant \frac{1}{2}\varepsilon_*\mathsf{L} \overset{(4.3.16)}{\leqslant} \frac{\sigma}{6} < \frac{1}{2}.$$

  Hence, $F_y(y_0,y_0)$ is invertible, with inverse

$$T_0 := (\mathbb{1}_d + \varepsilon A_0)^{-1}T(y_0) = \left(\mathbb{1}_d + \sum_{k\geqslant 1}(-\varepsilon)^k A_0^k\right)T(y_0)$$

  satisfying

$$\|T_0\| \leqslant \frac{\|T(y_0)\|}{1-\|\varepsilon A_0\|} \leqslant 2\mathsf{T}. \tag{4.3.30}$$

- For any $(y,z) \in D_{\bar{r}}(y_0) \times D_{\bar{r}}(y_0)$,

$$\begin{aligned} \|\mathbb{1}_d - T_0F_y(y,z)\| &\leqslant \|T_0\|\|\partial_y^2K(y_0) - K_{yy}\|_{\bar{r},\mathscr{D}_\sharp} + |\varepsilon|\;\|T_0\|\;|\partial_{y'}^2\widetilde{K}(y_0)| + |\varepsilon|\;\|T_0\|\;\|\partial_{y'}^2\widetilde{K}\|_{\bar{r},\mathscr{D}_\sharp} \\ &\leqslant d\cdot 2\mathsf{T}\|K_{yyy}\|_{\bar{r},\mathscr{D}_\sharp}\cdot\bar{r} + 4|\varepsilon|\mathsf{T}\frac{4M}{r^2} \\ &\leqslant 2d\mathsf{T}_\infty\mathsf{K}_\infty\frac{\bar{r}}{r-\bar{r}} + 16\mathsf{T}_\infty\frac{|\varepsilon|M}{r^2} \\ &\overset{(4.3.14)}{\leqslant} \frac{\mathsf{C}_5}{2}\frac{2\bar{r}}{r} + |\varepsilon|\frac{16\mathsf{T}_\infty M}{r\bar{r}} \\ &\leqslant \mathsf{C}_5\frac{\bar{r}}{r} + \varepsilon_*\mathsf{L} \\ &\overset{(4.3.13)+(4.3.16)}{\leqslant} \frac{1}{4} + \frac{\sigma}{3} \\ &\leqslant \frac{1}{4} + \frac{1}{4} = \frac{1}{2}\,; \end{aligned}$$

- For any $z\in D_{\tilde r}(y_0)$,

$$
\begin{aligned}
2\|T_0\||F(y_0,z)| &\leqslant 4\mathsf{T}|K_y(z)-K_y(y_0)|+4\mathsf{T}|\varepsilon|\|\widetilde{K}_{y'}\|_{\bar r,\mathscr{D}_\sharp}\\
&\leqslant 4\mathsf{T}\|K_{yy}\|_{\bar r,\mathscr{D}_\sharp}\cdot\tilde r+4\mathsf{T}\frac{2|\varepsilon|M}{r}\\
&\leqslant 4\mathsf{T}_\infty\mathsf{K}_\infty\tilde r+\frac{8|\varepsilon|\mathsf{T}_\infty M}{r}\\
&\leqslant \frac{\mathsf{C}_5}{d}\frac{\bar r}{4\mathsf{C}_5}+\varepsilon_*\frac{\bar r}{2}\mathsf{L}\\
&\overset{(4.3.16)}{\leqslant} \frac{\bar r}{8}+\frac{\bar r}{12}\\
&<\frac{\bar r}{4}\,,
\end{aligned}
$$

  i.e.

$$2\|T_0\|\|F(y_0,\cdot)\|_{\tilde r,y_0}\leqslant\frac{\bar r}{4}\,.$$

Therefore, Lemma 2.2.6 applies. Hence, there exists a real–analytic map $G^{y_0}\colon D_{\tilde r}(y_0)\to D_{\bar r}(y_0)$ such that its graph coincides with $F^{-1}(\{0\})$ *i.e.* $y_1=y_1(z,y_0,\varepsilon):=G^{y_0}(z)$ is the unique $y\in D_{\bar r}(y_0)$ satisfying $0=F(y,z)=\partial_y K_1(y)-K_y(z)$, for any $z\in D_{\tilde r}(y_0)$. Moreover, $\forall\ z\in D_{\tilde r}(y_0)$,

$$|G^{y_0}(z)-y_0|\leqslant 2\|T_0\|\|F(y_0,\cdot)\|_{\tilde r,y_0}\leqslant\frac{\bar r}{4}\,, \tag{4.3.31}$$

$$|G^{y_0}(z)-z|\leqslant|G^{y_0}(z)-y_0|+|y_0-z|\leqslant\frac{\bar r}{4}+\tilde r<\frac{\bar r}{2}, \tag{4.3.32}$$

so that

$$D_{\bar r/4}(G^{y_0}(z))\subset D_{\bar r/2}(y_0). \tag{4.3.33}$$

Next, we prove that $\partial^2_{y'}K_1(y_1)$ is invertible, where $y_1=G^{y_0}(z)$ for some given $z\in D_{\tilde r}(y_0)$. Indeed, by Taylor's formula, we have,

$$
\begin{aligned}
\partial^2_{y'}K_1(y_1)&=K_{yy}(y_0)+\int_0^1 K_{yyy}(y_0+t(y_1-y_0))(y_1-y_0)dt+\varepsilon\partial^2_{y'}\widetilde{K}(y_1)\\
&=T(y_0)^{-1}\left(\mathbb{1}_d+T(y_0)\left(\int_0^1 K_{yyy}(y_0+t(y_1-y_0))(y_1-y_0)dt+\partial^2_{y'}\widetilde{K}(y_1)\right)\right)\\
&=:T(y_0)^{-1}(\mathbb{1}_d+\varepsilon A),
\end{aligned}
$$

and, by Cauchy's estimate, for any[78] $|\varepsilon| \leqslant \varepsilon_*$,

$$\begin{aligned}
|\varepsilon|\|A\| &\leqslant \|T(y_0)\| \left(d\|K_{yyy}\|_{\bar r,\mathscr{D}_\sharp}|y_1-y_0| + |\varepsilon|\|\partial_{y'}^2\widetilde K\|_{\bar r,\mathscr{D}_\sharp}\right)\\
&\leqslant \|T\|_{\mathscr{D}_\sharp}\left(\frac{d\|K_{yy}\|_{r,\mathscr{D}_\sharp}}{r-\bar r}|y_1-y_0| + |\varepsilon|\|\partial_{y'}^2\widetilde K\|_{\bar r,\mathscr{D}_\sharp}\right)\\
&\leqslant \mathsf{T}\left(\frac{2d\mathsf{K}}{r}\frac{\bar r}{2}+\frac{4M}{r^2}\right)\\
&\overset{(4.3.14)}{\leqslant} \mathsf{T}\left(\frac{\sigma}{16\mathsf{T}_\infty}+\frac{1}{4\mathsf{T}_\infty}\varepsilon_*\mathsf{L}\right)\\
&\overset{(4.3.16)}{\leqslant} \mathsf{T}\left(\frac{\sigma}{16\mathsf{T}_\infty}+\frac{1}{4\mathsf{T}_\infty}\frac{\sigma}{3}\right)\\
&< \frac{\sigma}{6}\\
&< \frac12 .
\end{aligned}$$

Hence $\partial_{y'}^2 K_1(y_1)$ is invertible with

$$\partial_{y'}^2 K_1(y_1)^{-1} = (\mathbb{1}_d+\varepsilon A)^{-1}T(y_0) = T(y_0)+\sum_{k\geqslant 1}(-\varepsilon)^k A^k T(y_0) =: T(y_0)+\varepsilon\widetilde T(y_1),$$

and

$$|\varepsilon|\|\widetilde T(y_1)\| \leqslant |\varepsilon|\frac{\|A\|}{1-|\varepsilon|\|A\|}\|T\|_{\mathscr{D}_\sharp} \leqslant 2|\varepsilon|\|A\|\|T\|_{\mathscr{D}_\sharp} \leqslant 2\frac{\sigma}{6}\mathsf{T} = \mathsf{T}\frac{\sigma}{3}\,.$$

Similarly, from

$$K_{yy}(z) = K_{yy}(y_0)\left(\mathbb{1}_d + T(y_0)\int_0^1 K_{yyy}(y_0+t(z-y_0))(z-y_0)dt\right)$$

and

$$\left\|T(y_0)\int_0^1 K_{yyy}(y_0+t(z-y_0))(z-y_0)dt\right\|_{r/\mathsf{C}_5,y_0} \leqslant \mathsf{T}\|K_{yyy}\|_{r/2,y_0}\frac{r}{\mathsf{C}_5} \leqslant \mathsf{T}\frac{d\mathsf{K}}{r-r/2}\frac{r}{\mathsf{C}_5} < \frac12$$

one has that, for any $z\in D_{r/\mathsf{C}_5}(y_0)$,

$$K_{yy}(z)^{-1} \text{ exists and } \|K_{yy}(z)^{-1}\| \leqslant \|K_{yy}(z)^{-1}-T(y_0)\|+\|T(y_0)\| \leqslant 2\frac12\mathsf{T}+\mathsf{T} = 2\mathsf{T} \quad (4.3.34)$$

[78] Recall footnote 42 .

Now, differentiating $F(G^{y_0}(z), z) = 0$, we get, for any $z \in D_{\tilde{r}}(y_0)$,

$$\partial_{y'}^2 K_1(G^{y_0}(z)) \cdot \partial_z G^{y_0}(z) = K_{yy}(z) \ .$$

Therefore G^{y_0} is a local diffeomorphism, with

$$\begin{aligned}\partial_z G^{y_0}(z) &= \partial_{y'}^2 K_1(G^{y_0}(z))^{-1} K_{yy}(z)\\ &= \Big(K_{yy}(z)^{-1}\Big(K_{yy}(z) + \varepsilon \partial_{y'}^2 \widetilde{K}(g^{y_0}(z))\Big)\Big)^{-1}\\ &= \Big(\mathbb{1}_d + \varepsilon K_{yy}(z)^{-1} \partial_{y'}^2 \widetilde{K}(g^{y_0}(z))\Big)^{-1}\end{aligned}$$

and

$$\|\varepsilon K_{yy}^{-1} \partial_{y'}^2 \widetilde{K}\|_{\tilde{r},y_0} \leqslant \|K_{yy}^{-1}\|_{\tilde{r},y_0} \|\varepsilon \partial_{y'}^2 \widetilde{K}\|_{\tilde{r},\mathscr{D}_\sharp} \leqslant 2\mathsf{T}\frac{|\varepsilon|\mathsf{L}}{4\mathsf{T}_\infty} \leqslant \frac{1}{2}|\varepsilon|\mathsf{L} \leqslant \frac{\sigma}{6} < \frac{1}{2}$$

so that

$$\|\partial_z G^{y_0} - \mathbb{1}_d\|_{\tilde{r},y_0} \leqslant 2\|\varepsilon K_{yy}^{-1} \partial_{y'}^2 \widetilde{K}\|_{\tilde{r},y_0} \leqslant |\varepsilon|\mathsf{L}. \tag{4.3.35}$$

Now, we show that the family $\{G^{y_0}\}_{y_0 \in \mathscr{D}_\sharp}$ is compatible so that, together, they define a global map on $D_{\tilde{r}}(\mathscr{D}_\sharp)$, say G and that, in fact, G is a real–analytic diffeomorphism. For, assume that $z \in D_{\tilde{r}}(y_0) \bigcap D_{\tilde{r}}(\hat{y}_0)$, for some $y_0, \hat{y}_0 \in \mathscr{D}_\sharp$. Then, we need to show that $G^{y_0}(z) = G^{\hat{y}_0}(z)$. But, we have

$$|G^{\hat{y}_0}(z) - y_0| \leqslant |G^{\hat{y}_0}(z) - \hat{y}_0| + |\hat{y}_0 - z| + |z - y_0| \overset{(4.3.31)}{\leqslant} \frac{\bar{r}}{2} + \tilde{r} + \tilde{r} < \bar{r}.$$

Hence, $z \in D_{\tilde{r}}(y_0)$, $G^{\hat{y}_0}(z) \in D_{\bar{r}}(y_0)$ and, by definitions, $F(G^{\hat{y}_0}(z), z) = 0 = F(G^{y_0}(z), z)$. Then, by unicity, we get $G^{y_0}(z) = G^{\hat{y}_0}(z)$. Thus, the map

$$G\colon D_{\tilde{r}}(\mathscr{D}_\sharp) \to \mathbb{C}^d \quad \text{such that} \quad G_{|D_{\tilde{r}}(y_0)} := G^{y_0}, \quad \forall\, y_0 \in \mathscr{D}_\sharp\ ,$$

is well–defined and, therefore, is a real–analytic local diffeomorphism. It remains only to check that G is injective to conclude that it is a global diffeomorphism. Let then $z \in D_{\tilde{r}}(y_0), \hat{z} \in D_{\tilde{r}}(\hat{y}_0)$ such that $G(z) = G(\hat{z})$, for some $y_0, \hat{y}_0 \in \mathscr{D}_\sharp$. Then, we have

$$|z - \hat{z}| < \frac{r}{\mathsf{C}_5} - \tilde{r}.$$

Indeed, if not then

$$\begin{aligned}0 = |G(z) - G(\hat{z})| &\geqslant -|G(z) - z| + |z - \hat{z}| - |\hat{z} - G(\hat{z})| \\ &\overset{(4.3.32)}{\geqslant} -\bar{r} + \frac{r}{\mathsf{C}_5} - \tilde{r} - \bar{r} \\ &\geqslant \frac{r}{\mathsf{C}_5} - 3\bar{r} \\ &\overset{(4.3.14)}{\geqslant} \frac{r}{\mathsf{C}_5} - 3\frac{r}{4\mathsf{C}_5} \\ &> 0\,,\end{aligned}$$

contradiction. Therefore,

$$|z - \hat{z}| < \frac{r}{\mathsf{C}_5} - \tilde{r}\ . \tag{4.3.36}$$

Thus,

$$|\hat{z} - y_0| \leqslant |\hat{z} - z| + |z - y_0| < \frac{r}{\mathsf{C}_5} - \tilde{r} + \tilde{r} = \frac{r}{\mathsf{C}_5}\ .$$

Hence, $z, \hat{z} \in D_{r/\mathsf{C}_5}(y_0)$. But $G(z) = G(\hat{z})$ is equivalent to $K_y(z) = K_y(\hat{z})$ and then,

$$0 = K_y(z) - K_y(\hat{z}) = \int_0^1 K_{yy}(\hat{z} + t(z - \hat{z}))dt(z - \hat{z})\ .$$

Thus, it is enough to show that $\int_0^1 K_{yy}(\hat{z} + t(z - \hat{z}))dt$ is invertible. But

$$\begin{aligned}\int_0^1 K_{yy}(\hat{z} + t(z - \hat{z}))dt &= K_{yy}(\hat{z}) + \int_0^1\int_0^1 K_{yyy}(\hat{z} + tt'(z - \hat{z}))tdt'dt \cdot (z - \hat{z}) \\ &\overset{(4.3.34)}{=} K_{yy}(\hat{z})\left(\mathbb{1}_d + K_{yy}(\hat{z})^{-1}\int_0^1\int_0^1 K_{yyy}(\hat{z} + tt'(z - \hat{z}))tdt'dt \cdot (z - \hat{z})\right)\end{aligned}$$

and

$$\begin{aligned}\left\|K_{yy}(\hat{z})^{-1}\int_0^1\int_0^1 K_{yyy}(\hat{z} + tt'(z - \hat{z}))tdt'dt \cdot (z - \hat{z})\right\| &\overset{(4.3.34)}{\leqslant} 2\mathsf{T} \cdot \frac{1}{2}\|K_{yyy}\|_{r/2,y_0}|z - \hat{z}| \\ &\overset{(4.3.36)}{\leqslant} \mathsf{T}\frac{2d\mathsf{K}}{r}\left(\frac{r}{\mathsf{C}_5} - \tilde{r}\right) \\ &< \frac{\mathsf{C}_5}{2r}\frac{r}{\mathsf{C}_5} \\ &= \frac{1}{2}.\end{aligned}$$

Therefore, $\int_0^1 K_{yy}(\hat z+t(z-\hat z))dt$ is invertible and then we get $z-\hat z=0$ *i.e.* G is injective. Next, we estimate P'. We have, for any $|\varepsilon|\leqslant\varepsilon_*$,

$$|\varepsilon|\|g_x\|_{\bar r,\bar s,\mathscr{D}_\sharp}\leqslant|\varepsilon|\mathsf{C}_1\frac{M}{\alpha}\sigma^{-(\tau+d+1)}\leqslant|\varepsilon|\bar r\mathsf{L}\overset{(4.3.16)}{\leqslant}\frac{r}{3}\frac{\sigma}{3}\leqslant\frac{r}{3}$$

so that, for any $y_0\in\mathscr{D}_\sharp$ and $(y',x)\in D_{\bar r,\bar s}(y_0)$,

$$|y'+\varepsilon g_x(y',x)-y_0|\leqslant\bar r+\frac{r}{3}\leqslant\frac{r}{2}+\frac{r}{3}=\frac{5r}{6}<r\,,$$

and thus

$$\begin{aligned}\|P^{(1)}\|_{\bar r,\bar s,\mathscr{D}_\sharp}&\leqslant d^2\|K_{yy}\|_{r,\mathscr{D}_\sharp}\|g_x\|^2_{\bar r,\bar s,\mathscr{D}_\sharp}\leqslant d^2\mathsf{K}\left(\mathsf{C}_1\frac{M}{\alpha}\sigma^{-(\tau+d+1)}\right)^2\\&=d^2\mathsf{C}_1^2\frac{\mathsf{K}M^2}{\alpha^2}\sigma^{-2(\tau+d+1)},\end{aligned}$$

$$\begin{aligned}\|P^{(2)}\|_{\bar r,\bar s,\mathscr{D}_\sharp}&\leqslant d\|P_y\|_{\frac{5r}{6},\bar s,\mathscr{D}_\sharp}\|g_x\|_{\bar r,\bar s,\mathscr{D}_\sharp}\leqslant d\frac{6M}{r}\mathsf{C}_1\frac{M}{\alpha}\sigma^{-(\tau+d+1)}\\&=6d\mathsf{C}_1\frac{M^2}{\alpha r}\sigma^{-(\tau+d+1)}\end{aligned}$$

and, by Lemma 2.2.4–(i), we have

$$\begin{aligned}|\varepsilon|\|P^{(3)}\|_{\bar r,s-\frac{\sigma}{2},\mathscr{D}_\sharp}&\leqslant\sum_{|n|_1>\kappa}\|P_n\|_{\bar r,\mathscr{D}_\sharp}e^{(s-\frac{\sigma}{2})|n|_1}\leqslant M\sum_{|n|_1>\kappa}e^{-\frac{\sigma|n|_1}{2}}\\&\leqslant M\,e^{-\frac{\kappa\sigma}{4}}\sum_{|n|_1>\kappa}e^{-\frac{\sigma|n|_1}{4}}\leqslant M\,e^{-\frac{\kappa\sigma}{4}}\sum_{|n|_1>0}e^{-\frac{\sigma|n|_1}{4}}\\&=M\,e^{-\frac{\kappa\sigma}{4}}\left(\left(\sum_{k\in\mathbb{Z}}e^{-\frac{\sigma|k|}{4}}\right)^d-1\right)=M\,e^{-\frac{\kappa\sigma}{4}}\left(\left(1+\frac{2\,e^{-\frac{\sigma}{4}}}{1-e^{-\frac{\sigma}{4}}}\right)^d-1\right)\\&=M\,e^{-\frac{\kappa\sigma}{4}}\left(\left(1+\frac{2}{e^{\frac{\sigma}{4}}-1}\right)^d-1\right)\leqslant M\,e^{-\frac{\kappa\sigma}{4}}\left(\left(1+\frac{2}{\frac{\sigma}{4}}\right)^d-1\right)\\&\leqslant\sigma^{-d}M\,e^{-\frac{\kappa\sigma}{4}}\left((\sigma+8)^d-\sigma^d\right)\leqslant d8^d\sigma^{-d}M\,e^{-\frac{\kappa\sigma}{4}}\\&=\mathsf{C}_2\sigma^{-d}M\,e^{-\frac{\kappa\sigma}{4}}\\&\overset{(4.3.13)}{\leqslant}\mathsf{C}_2\sigma^{-d}M\frac{\mathsf{K}|\varepsilon|M}{\alpha^2}\\&=\mathsf{C}_2\frac{\mathsf{K}|\varepsilon|M^2}{\alpha^2}\sigma^{-d}\,.\end{aligned}$$

Hence,[79]

$$\begin{aligned}
\|P'\|_{\bar r,\bar s,\mathscr{D}_\sharp} &\leqslant \|P^{(1)}\|_{\bar r,\bar s,\mathscr{D}_\sharp} + \|P^{(2)}\|_{\bar r,\bar s,\mathscr{D}_\sharp} + \|P^{(3)}\|_{\bar r,\bar s,\mathscr{D}_\sharp}\\
&\leqslant d^2\mathsf{C}_1^2\frac{\mathsf{K}M^2}{\alpha^2}\sigma^{-2(\tau+d+1)} + 6d\mathsf{C}_1\frac{M^2}{\alpha r}\sigma^{-(\tau+d+1)} + \mathsf{C}_2\frac{\mathsf{K}M^2}{\alpha^2}\sigma^{-d}\\
&= \left(d^2\mathsf{C}_1^2 r\mathsf{K} + 6d\mathsf{C}_1\alpha\sigma^{\tau+d+1} + \mathsf{C}_2 r\mathsf{K}\sigma^{2\tau+d+2}\right)\frac{M^2}{\alpha^2 r}\sigma^{-2(\tau+d+1)}\\
&\leqslant \left(d^2\mathsf{C}_1^2 + 6d\mathsf{C}_1 + \mathsf{C}_2\right)\max(\alpha, r\mathsf{K})\frac{M^2}{\alpha^2 r}\sigma^{-2(\tau+d+1)}\\
&= \mathsf{C}_3\max(\alpha, r\mathsf{K})\frac{M^2}{\alpha^2 r}\sigma^{-2(\tau+d+1)}
\end{aligned}$$

Now, we need to invert $x \mapsto x + \varepsilon\hat g_{y'}(y', x)$. But, thanks to (4.3.27), (4.3.28), (4.3.29) and (4.3.16), we can apply Lemma B.1 (see Appendix B), to conclude that for any given $y' \in \mathbb{C}^d$, the map $\psi_\varepsilon(y', \cdot)\colon \mathbb{T}^d_{\bar s} \ni x \mapsto \psi_\varepsilon(y', x)$ has an inverse $\varphi(y', x') = x' + \varepsilon\widetilde\varphi(y', x'; \varepsilon)$, C^∞ on $\mathbb{R}^d \times \mathbb{T}^d$ and real analytic on $D_{\bar r/4, s'}(\mathscr{D}_\sharp)$ such that (4.3.18) holds and (4.3.22) as well, using the multivariate Fàa Di Bruno formula (see [CS96], Theorem 2.1) and the real–analyticity of g_x. ∎

[79] Recall that $r \leqslant r_0$ and $\sigma < 1$.

Finally, we prove the convergence of the scheme by mimicking Lemma 2.3.2.

Lemma 4.3.2 *Let* $H_0 := H$*,* $K_0 := K$*,* $P_0 := P$*,* $\phi^0 = \phi_0 := \mathrm{id}$*, and* r_0*,* s_0*,* s_**,* σ_0*,* λ_0*,* W_0*,* M_0*,* K_0*,* K_∞*,* T_0*,* T_∞*,* d_**,* e_**,* f_**,* ε_1*,* ε_2 *and* $\varepsilon_\sharp$ *be as in* §4.1 *and for a given* $\varepsilon \neq 0$*, sequence of non–negative numbers* $(M_j)_j$ *and* $j \geqslant 0$*, define*[80]

$$
\begin{aligned}
\sigma_j &:= \frac{\sigma_0}{2^j}\,,\\
s_{j+1} &:= s_j - \sigma_j = s_* + \frac{\sigma_0}{2^j}\,,\\
\bar{s}_j &:= s_j - \frac{2\sigma_i}{3}\,,\\
\mathsf{K}_{j+1} &:= \mathsf{K}_0 \prod_{k=0}^{j}\Big(1+\frac{\sigma_k}{3}\Big) \leqslant \mathsf{K}_0\, e^{\frac{2\sigma_0}{3}} \leqslant \mathsf{K}_\infty\,,\\
\mathsf{T}_{j+1} &:= \mathsf{T}_0 \prod_{k=0}^{j}\Big(1+\frac{\sigma_k}{3}\Big) \leqslant \mathsf{T}_0\, e^{\frac{2\sigma_0}{3}} \leqslant \mathsf{T}_\infty\,,\\
\lambda_j &:= 2^j\lambda_0 = 2^j \log\left(\frac{\alpha^2}{\mathsf{K}_0|\varepsilon| M_0}\right)\,,\\
\kappa_j &:= 4\sigma_j^{-1}\lambda_j = 4^j\kappa_0 = 4^j\sigma_0^{-1}\log\left(\frac{\alpha^2}{\mathsf{K}_0|\varepsilon| M_0}\right)\,,\\
r_{j+1} &:= \frac{1}{16\mathsf{C}_5}\min\left(\frac{\mathsf{R}}{\kappa_j^{\tau+1}}\,,\, r_j\sigma_j\right)\,,\\
\tilde{r}_{j+1} &:= \frac{r_{j+1}}{\mathsf{C}_5}\,.
\end{aligned}
$$

Assume that ε_* *satisfies*

$$
\varepsilon_* \leqslant \varepsilon_\sharp\,,\quad \varepsilon_*\,\mathsf{f}_*\,\|P\|_{r_0,s_0,\mathscr{D}_{r_0,\alpha}} \leqslant 1 \quad \textit{and} \quad 4\varepsilon_*^2\,\mathsf{f}_*\,\mathsf{e}_*\,\mathsf{d}_*^2\,\sigma_0\,\|P\|^2_{r_0,s_0,\mathscr{D}_{r_0,\alpha}} \leqslant 3\,. \tag{4.3.37}
$$

where

$$
\begin{aligned}
\mathsf{d}_* &:= 2^{2\tau+2d+3}\mathsf{C}_6^2\,,\\
\mathsf{e}_* &:= 6\max\left\{\frac{4\mathsf{T}_\infty}{\mathsf{C}_6 r_1^2}\sigma_0^{-(\tau+d+2)}\,,\ \frac{\mathsf{C}_7}{4}\frac{1}{\alpha r_1}\sigma_0^{-(2\tau+2d+3)}\right\}\,,\\
\mathsf{f}_* &:= 3\max\left\{\frac{4\mathsf{T}_\infty}{r_0 r_1}\sigma_0^{-(\tau+d+2)}\,,\ \frac{\mathsf{C}_7}{4}\max(\alpha, r_0\mathsf{K}_\infty)\frac{1}{\alpha^2 r_1}\sigma_0^{-(2\tau+2d+3)}\right\}\,.
\end{aligned}
$$

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

Then, for any $|\varepsilon|\leqslant\varepsilon_*$, one can construct a sequence of diffeomorphisms

$$G_{j+1}\colon D_{\tilde r_{j+1}}(\mathscr{D}_j)\to G_{j+1}(D_{\tilde r_{j+1}}(\mathscr{D}_j))$$

and of C^∞–symplectomorphisms

$$\phi_j\colon \mathscr{D}\times\mathbb{T}^d \overset{into}{\longrightarrow} \mathscr{D}\times\mathbb{T}^d$$

such that

$$\partial_y K_{j+1}\circ G_{j+1}=\partial_y K_j\ , \tag{4.3.38}$$

$$\phi_{j+1}: D_{r_{j+1},s_{j+1}}(\mathscr{D}_{j+1})\to D_{r_j,s_j}(\mathscr{D}_j) \qquad \textit{is real–analytic}, \tag{4.3.39}$$

$$H_{j+1}:=H_j\circ\phi_{j+1}=:K_{j+1}+\varepsilon^{2^{j+1}}P_{j+1} \qquad \textit{on } D_{r_{j+1},s_{j+1}}(\mathscr{D}_{j+1}) \tag{4.3.40}$$

and converge uniformly. More precisely, given any $|\varepsilon|\leqslant\varepsilon_*$, we have the following:

(*i*) *the sequence $G^{j+1}:=G_{j+1}\circ G_j\circ\cdots\circ G_1$ converges uniformely on $\mathscr{D}_{r_0,\alpha}$ to a lipeomorphism $G_*\colon\mathscr{D}_{r_0,\alpha}\to\mathscr{D}_*:=G_*(\mathscr{D}_{r_0,\alpha})\subset\mathscr{D}$;*

(*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*) *$\phi^j:=\phi_0\circ\phi_1\circ\phi_2\circ\cdots\circ\phi_j$ converges uniformly on $\mathscr{D}\times\mathbb{T}^d$ to a C^∞–symplectomorphism $\phi_*\colon\mathscr{D}\times\mathbb{T}^d\overset{into}{\longrightarrow}\mathscr{D}\times\mathbb{T}^d$, with $\phi_*(y,\cdot)\colon\mathbb{T}^d_{s_*}\ni x\mapsto\phi_*(y,x)$ holomorphic, for any $y\in\mathscr{D}$;*

(*iv*) *K_j converges uniformly on $\mathscr{D}$ to a C^∞–map K_*, with*

$$\partial_{y_*}K_*\circ G_*=\partial_y K \quad \textit{on} \quad \mathscr{D}_{r_0,\alpha}\ ,$$
$$\partial^\beta_{y_*}H\circ\phi_*(y_*,x)=\partial^\beta_{y_*}K_*(y_*),\ \ \forall (y_*,x)\in\mathscr{D}_*\times\mathbb{T}^d\ ,\forall\,\beta\in\mathbb{N}_0^d\ .$$

Finally, the following estimates hold for any $|\varepsilon|\leqslant\varepsilon_*$ and for any $i\geqslant 1$:

$$|\varepsilon|^{2^i}M_i:=|\varepsilon|^{2^i}\|P_i\|_{r_i,s_i,\mathscr{D}_i}\leqslant\frac{(\,|2\varepsilon|^2\mathsf{e}_*\,d_*^2M_1)^{2^{i-1}}}{\mathsf{e}_*\,{d_*}^{i+1}}\ , \tag{4.3.41}$$

$$|\operatorname{meas}(\mathscr{D}_*)-\operatorname{meas}(\mathscr{D}_{r_0,\alpha})|\leqslant\mathsf{C}_8\varepsilon_2\,e^{\varepsilon_2}\operatorname{meas}(\mathscr{D}_{r_0,\alpha})\ , \tag{4.3.42}$$

$$|\mathsf{W}(\phi_*-\mathrm{id})|\leqslant\varepsilon_1 \qquad \textit{on} \quad \mathscr{D}_*\times\mathbb{T}^d_{s_*}\ . \tag{4.3.43}$$

Proof For $i \geqslant 0$, define

$$\begin{aligned}
\mathsf{W}_i &:= \operatorname{diag}\left(\frac{1}{4r_{i+1}}\mathbb{1}_d, \mathbb{1}_d\right),\\
\bar{\mathsf{L}}_i &:= \frac{\mathsf{C}_4}{4}\max\{\alpha, r_i\mathsf{K}_\infty\}\frac{M_i}{\alpha^2 r_{i+1}}\sigma_i^{-(2\tau+d+2)},\\
\mathsf{L}_i &:= M_i \max\left\{\frac{4\mathsf{T}_\infty}{r_i r_{i+1}}\sigma_i^{-(\tau+d+1)}, \frac{\mathsf{C}_7}{4}\max\{\alpha, r_i\mathsf{K}_\infty\}\frac{1}{\alpha^2 r_{i+1}}\sigma_i^{-2(\tau+d+1)}\right\}\\
&\geqslant M_i \max\left\{\frac{4\mathsf{T}_\infty}{r_i r_{i+1}}\sigma_i^{-(\tau+d+1)}, \frac{4}{\mathsf{K}_i r_i^2}, \frac{\mathsf{C}_7}{4}\max\{\alpha, r_i\mathsf{K}_i\}\frac{1}{\alpha^2 r_{i+1}}\sigma_i^{-2(\tau+d+1)}\right\}.
\end{aligned}$$

Let us assume (*inductive hypothesis*) that we can iterate $j \geqslant 1$ times the KAM step, obtaining j diffeomorphisms

$$G_{i+1}\colon D_{\tilde{r}_{i+1}}(\mathscr{D}_i) \to G_{i+1}(D_{\tilde{r}_{i+1}}(\mathscr{D}_j))$$

and j C^∞–symplectomorphisms

$$\phi_{i+1}\colon \mathscr{D}\times\mathbb{T}^d \overset{into}{\longrightarrow} \mathscr{D}\times\mathbb{T}^d, \tag{4.3.44}$$

satisfying $(4.3.38)_{j=i} \div (4.3.40)_{j=i}$, for $0 \leqslant i \leqslant j-1$, with

$$\begin{cases}
\|\partial_y^2 K_i\|_{r_i,\mathscr{D}_i} \leqslant \mathsf{K}_i\,,\\
\|T_i\|_{\mathscr{D}_i} \leqslant \mathsf{T}_i\,,\\
\|P_i\|_{r_i,s_i,\mathscr{D}_i} \leqslant M_i\,,\\
|\varepsilon|^{2^i}\mathsf{L}_i \leqslant \frac{\sigma_i}{3}\\
\lambda_i \geqslant \log\left(\frac{\alpha^2}{\mathsf{K}_0|\varepsilon|^{2^i}M_i}\right).
\end{cases} \tag{4.3.45}$$

Observe that for $j=1$, it is $i=0$ and (4.3.45) is implied by the definitions of K_0, T_0, M_0 and by condition (4.3.37).
Because of (4.3.37) and (4.3.45), (4.3.16) holds for H_i and Lemma 4.3.1 can be applied to

H_i and one has, for $0 \leqslant i \leqslant j-1$ and for any $|\varepsilon| \leqslant \varepsilon_*$ (see (4.3.15), (4.3.17) and (4.3.21)):

$$\|G_{i+1} - \mathrm{id}\|_{\tilde r_{i+1},\mathscr{D}_i} \leqslant 2r_{i+1} \;, \tag{4.3.46}$$

$$\|\partial_z G_{i+1} - \mathbb{1}_d\|_{\tilde r_{i+1},\mathscr{D}_i} \leqslant |\varepsilon|^{2^i}\mathsf{L}_i \;, \tag{4.3.47}$$

$$\|\mathsf{K}_{i+1}\|_{r_{i+1},\mathscr{D}_{i+1}} \leqslant \|\mathsf{K}_i\|_{r_i,\mathscr{D}_i} + |\varepsilon|^{2^i} M_i \;, \tag{4.3.48}$$

$$\|\partial_y^2\mathsf{K}_{i+1}\|_{r_{i+1},\mathscr{D}_{i+1}} \leqslant \|\partial_y^2\mathsf{K}_i\|_{r_i,\mathscr{D}_i} + \mathsf{K}_i|\varepsilon|^{2^i}\mathsf{L}_i \;, \tag{4.3.49}$$

$$\|T_{i+1}\|_{\mathscr{D}_{i+1}} \leqslant \|T_i\|_{\mathscr{D}_i} + \mathsf{T}_i|\varepsilon|^{2^i}\mathsf{L}_i \;, \tag{4.3.50}$$

$$\|\mathsf{W}_i(\phi_{i+1} - \mathrm{id})\|_{r_{i+1},s_{i+1},\mathscr{D}_{i+1}} \leqslant |\varepsilon|^{2^i}\overline{\mathsf{L}}_i \;, \tag{4.3.51}$$

$$\|P_{i+1}\|_{r_{i+1},s_{i+1},\mathscr{D}_{i+1}} \leqslant M_{i+1} := M_i\mathsf{L}_i \;. \tag{4.3.52}$$

Let $0 \leqslant i \leqslant j-1$. Then, by definition,

$$r_1 = \frac{\mathsf{R}}{16\mathsf{C}_5\kappa_0^{\tau+1}} \leqslant \frac{\mathsf{R}}{16\mathsf{C}_5} < \frac{\alpha}{\mathsf{K}_\infty} \,. \tag{4.3.53}$$

and, since $\sigma_0 < 2^{-2(\tau-1)}\mathsf{C}_5\sqrt{2}$ *i.e.*

$$\frac{16\mathsf{C}_5\sqrt{2}}{2^{2(\tau+1)}\sigma_0} > 1 \;, \tag{4.3.54}$$

we have

$$\begin{aligned}
r_{i+1} &= \min\left(\frac{\mathsf{R}}{16\mathsf{C}_5\kappa_i^{\tau+1}} \,,\, \frac{r_i\sigma_i}{16\mathsf{C}_5}\right)\\
&= \min\left(\frac{\mathsf{R}}{16\mathsf{C}_5\kappa_i^{\tau+1}} \,,\, \frac{\mathsf{R}\sigma_i}{(16\mathsf{C}_5)^2\kappa_{i-1}^{\tau+1}} \,,\, \frac{r_{i-1}\sigma_{i-1}\sigma_i}{(16\mathsf{C}_5)^2}\right)\\
&\;\vdots\\
&= \min\left(\frac{\mathsf{R}}{16\mathsf{C}_5\kappa_i^{\tau+1}} \,,\, \frac{\mathsf{R}\sigma_i}{(16\mathsf{C}_5)^2\kappa_{i-1}^{\tau+1}} \,,\, \cdots \,,\, \frac{r_1\sigma_1\cdots\sigma_i}{(16\mathsf{C}_5)^i}\right)\\
&= \min\left(\frac{\mathsf{R}}{16\mathsf{C}_5\kappa_i^{\tau+1}} \,,\, \frac{\mathsf{R}\sigma_i}{(16\mathsf{C}_5)^2\kappa_{i-1}^{\tau+1}} \,,\, \cdots \,,\, \frac{\mathsf{R}\sigma_1\cdots\sigma_i}{(16\mathsf{C}_5)^{i+1}\kappa_0^{\tau+1}}\right)\\
&= \frac{\mathsf{R}\sigma_1\cdots\sigma_i}{(16\mathsf{C}_5)^{i+1}\,\kappa_0^{\tau+1}}\min\left(\left(\frac{16\mathsf{C}_5\cdot 2^{\frac12(i+1)}}{2^{2(\tau+1)}\sigma_0}\right)^i \,,\, \left(\frac{16\mathsf{C}_5\cdot 2^{\frac12 i}}{2^{2(\tau+1)}\sigma_0}\right)^{i-1} \,,\, \cdots \,,\, \left(\frac{16\mathsf{C}_5\cdot 2^{\frac12}}{2^{2(\tau+1)}\sigma_0}\right)^0\right)\\
&\overset{(4.3.54)}{=} \frac{\mathsf{R}\sigma_1\cdots\sigma_i}{(16\mathsf{C}_5)^{i+1}\,\kappa_0^{\tau+1}}\\
&= 2^{-\frac{i^2}{2}}\mathsf{C}_6^{-i}r_1 \,.
\end{aligned}$$

Thus,

$$\begin{aligned}|\varepsilon|\mathsf{L}_0(3\sigma_0^{-1}) &= 3|\varepsilon|M_0\max\left(\frac{4\mathsf{T}_\infty}{r_0r_1}\sigma_0^{-(\tau+d+2)}\,,\ \frac{\mathsf{C}_7}{4}\max(\alpha,r_0\mathsf{K}_\infty)\frac{1}{\alpha^2r_1}\sigma_0^{-(2\tau+2d+3)}\right)\\ &= \mathsf{f}_*|\varepsilon|M_0 \overset{(4.3.37)}{\leqslant} 1\end{aligned}$$

and for $i\geqslant 1$[81],

$$\begin{aligned}|\varepsilon|^{2^i}\mathsf{L}_i(3\sigma_i^{-1}) &\overset{(4.3.53)}{=} 3|\varepsilon|^{2^i}M_i\max\left(\frac{4\mathsf{T}_\infty}{r_ir_{i+1}}\sigma_i^{-(\tau+d+2)}\,,\ \frac{\mathsf{C}_7}{4}\frac{1}{\alpha r_{i+1}}\sigma_i^{-(2\tau+2d+3)}\right)\\ &= 3|\varepsilon|^{2^i}M_i\max\left(\frac{4\mathsf{T}_\infty\sqrt{2}}{\mathsf{C}_6r_1^2}\sigma_0^{-(\tau+d+2)}\ 2^{i^2}\left(2^{\tau+d+1}\mathsf{C}_6^2\right)^i\,,\right.\\ &\qquad\qquad\left.\frac{\mathsf{C}_7}{4}\frac{1}{\alpha r_1}\sigma_0^{-(2\tau+2d+3)}\ 2^{\frac{i^2}{2}}\left(2^{2\tau+2d+3}\mathsf{C}_6\right)^i\right)\\ &\leqslant 3\left(2^{2\tau+2d+3}\mathsf{C}_6^2\right)^i\ 2^{i^2}|\varepsilon|^{2^i}M_i\max\left(\frac{4\mathsf{T}_\infty}{\mathsf{C}_6r_1^2}\sigma_0^{-(\tau+d+2)}\,,\ \frac{\mathsf{C}_7}{4}\frac{1}{\alpha r_1}\sigma_0^{-(2\tau+2d+3)}\right)\\ &= \mathsf{e}_*\mathsf{d}_*^i\ 2^{i^2-1}|\varepsilon|^{2^i}M_i\\ &\leqslant \mathsf{e}_*\mathsf{d}_*^i|2\varepsilon|^{2^i}M_i =: \frac{\theta_i}{\mathsf{d}_*}\,,\end{aligned}$$

so that

$$\mathsf{L}_i < \mathsf{e}_*\,\mathsf{d}_*^i M_i\,,$$

thus by (4.3.52), for any $1\leqslant i\leqslant j-1$,

$$|\varepsilon|^{2^{i+1}}M_{i+1} < \mathsf{e}_*\mathsf{d}_*^i|\varepsilon|^{2^{i+1}}M_i^2$$

i.e. $\theta_{i+1} < \theta_i^2$, which iterated, yields $\theta_i\leqslant\theta_1^{2^{i-1}}$ for $1\leqslant i\leqslant j$. Next, we show that, thanks to (4.3.37), (4.3.45) holds also for $i=j$. In fact, by (4.3.45) and (4.3.50), we have

$$\|T_{i+1}\|_{\mathscr{D}_{i+1}} \leqslant \|T_i\|_{\mathscr{D}_{i+1}} + \mathsf{T}_i|\varepsilon|^{2^i}\mathsf{L}_i \leqslant \mathsf{T}_i + \mathsf{T}_i\frac{\sigma_i}{3} = \mathsf{T}_{i+1}\,.$$

and similarly for $\|\partial_y^2K_{i+1}\|_{r_{i+1},\mathscr{D}_{i+1}}$. Now, we check the last relation in (4.3.45) for $i=j$. But, by definitions, for any $i\geqslant 0$,

$$M_{i+1} = M_i\mathsf{L}_i \geqslant M_i\max\{\alpha, r_i\mathsf{K}_\infty\}\frac{M_i}{\alpha^2r_{i+1}} \geqslant \frac{M_i^2\mathsf{K}_0}{\alpha^2}\,,$$

[81] Notice that $2^i\geqslant i^2-1$, $\forall\ i\in\mathbb{N}_0$.

i.e.

$$\frac{|\varepsilon|^{2^{i+1}} M_{i+1} \mathsf{K}_0}{\alpha^2} \geqslant \left(\frac{|\varepsilon|^{2^i} M_i \mathsf{K}_0}{\alpha^2} \right)^2 ,$$

which iterated yields, for any $i \geqslant 0$,

$$\frac{|\varepsilon|^{2^i} M_i \mathsf{K}_0}{\alpha^2} \geqslant \left(\frac{|\varepsilon| M_0 \mathsf{K}_0}{\alpha^2} \right)^{2^i} .$$

i.e.

$$\lambda_i = \log \left(\left(\frac{\alpha^2}{|\varepsilon| M_0 \mathsf{K}_0} \right)^{2^i} \right) \geqslant \log \left(\frac{\alpha^2}{|\varepsilon|^{2^i} M_i \mathsf{K}_0} \right) .$$

Now, by $(4.3.41)_{i=j}$,

$$|\varepsilon|^{2^j} \mathsf{L}_j(3\sigma_j^{-1}) \leqslant \frac{\theta_j}{\mathsf{d}_*} \leqslant \frac{1}{\mathsf{d}_*} \theta_1^{2^{j-1}} \leqslant \frac{1}{\mathsf{d}_*} (4\mathsf{e}_*\, d_*^2 \varepsilon_*^2 M_1)^{2^{j-1}} \overset{(4.3.37)}{\leqslant} \frac{1}{\mathsf{d}_*} < 1 ,$$

which implies the fourth inequality in (4.3.45) with $i = j$; the proof of the induction is finished and one can construct an *infinite sequence* of diffeomorphisms $G_{i+1} \colon D_{\tilde r_{i+1}}(\mathscr{D}_i) \to G_{i+1}(D_{\tilde r_{i+1}}(\mathscr{D}_i))$ and symplectomorphisms $\phi_i \colon \mathscr{D} \times \mathbb{T}^d \circlearrowleft$ satisfying (4.3.45), (4.3.46) ÷ (4.3.52), (4.3.41) and $(4.3.38)_{j=i}$ ÷ $(4.3.40)_{j=i}$ *for all* $i \geqslant 0$.

Next, we show that G^j converges. For any $j \geqslant 1$,

$$\|G^{j+1} - G^j\|_{\mathscr{D}_{r_0,\alpha}} = \|G_{j+1} \circ G^j - G^j\|_{\mathscr{D}_{r_0,\alpha}} \leqslant \|G_{j+1} - \mathrm{id}\|_{\mathscr{D}_j} \leqslant \|G_{j+1} - \mathrm{id}\|_{\tilde r_{j+1}, \mathscr{D}_j} \overset{(4.3.46)}{\leqslant} 2r_{j+1}.$$

Thus, G^j is Cauchy and therefore converges uniformly on $\mathscr{D}_{r_0,\alpha}$ to a map G_*.

Next, we prove that ϕ_j is convergent by showing that it is Cauchy as well. For any $j \geqslant 3$, we have, using again Cauchy's estimate,

$$\begin{aligned}
\|\mathsf{W}_{j-1}(\phi^j-\phi^{j-1})\|_{r_j,s_j,\mathscr{D}_j} &= \|\mathsf{W}_{j-1}\phi^{j-1}\circ\phi_j-\mathsf{W}_{j-1}\phi^{j-1}\|_{r_i,s_i,\mathscr{D}_i}\\
&\overset{(4.3.44)}{\leqslant} \|\mathsf{W}_{j-1}D\phi^{j-1}\mathsf{W}_{j-1}^{-1}\|_{2r_{j-1}/3,s_{j-1},\mathscr{D}_{j-1}}\,\|\mathsf{W}_{j-1}(\phi_j-\mathrm{id})\|_{r_j,s_j,\mathscr{D}_j}\\
&\overset{(4.3.51)}{\leqslant} \max\left(r_{j-1}\frac{3}{r_{j-1}},\frac{3}{2\sigma_{j-1}}\right)\|\mathsf{W}_{j-1}\phi^{j-1}\|_{r_{j-1},s_{j-1},\mathscr{D}_{j-1}}\cdot|\varepsilon|^{2^j}\overline{\mathsf{L}}_j\\
&= \frac{3}{2\sigma_{i-1}}\|\mathsf{W}_{j-1}\phi^{j-1}\|_{r_{j-1},s_{j-1},\mathscr{D}_{j-1}}\cdot|\varepsilon|^{2^j}\overline{\mathsf{L}}_j\\
&\leqslant \frac{3}{2\sigma_{j-1}}\|\mathsf{W}_{j-1}\phi^{0}\|_{r_0,s_0,\mathscr{D}_{r_0,\alpha}}\cdot|\varepsilon|^{2^j}\overline{\mathsf{L}}_j\\
&\leqslant \frac{3}{2\sigma_{j-1}}\left(\prod_{i=0}^{j-2}\|\mathsf{W}_{i+1}\mathsf{W}_i^{-1}\|\right)\|\mathsf{W}_0\phi_0\|_{r_0,s_0,\mathscr{D}_{r_0,\alpha}}\cdot|\varepsilon|^{2^j}\overline{\mathsf{L}}_j\\
&= \frac{3}{2\sigma_{j-1}}\left(\prod_{i=0}^{j-2}\frac{r_i}{r_{i+1}}\right)\|\mathsf{W}_0\phi_0\|_{r_0,s_0,y_0}\cdot|\varepsilon|^{2^j}\overline{\mathsf{L}}_j\\
&= \frac{3r_0}{2r_{j-1}\sigma_{j-1}}\|\mathsf{W}_0\phi_0\|_{r_0,s_0,\mathscr{D}_{r_0,\alpha}}\cdot|\varepsilon|^{2^j}\overline{\mathsf{L}}_j\ .
\end{aligned}$$

Therefore, for any $n \geqslant 0$, $j \geqslant 1$,

$$\begin{aligned}
\|\mathsf{W}_0(\phi^{n+j}-\phi^n)\|_{r_{n+j},s_{n+j},\mathscr{D}_{n+j}} &\leqslant \sum_{i=n}^{n+j} \|\mathsf{W}_0(\phi^{i+1}-\phi^i)\|_{r_{i+1},s_{i+1},\mathscr{D}_{i+1}}\\
&\leqslant \sum_{i=n}^{n+j}\left(\prod_{k=0}^{i}\|\mathsf{W}_k\mathsf{W}_{k+1}^{-1}\|\right)\|\mathsf{W}_i(\phi^{i+1}-\phi^i)\|_{r_{i+1},s_{i+1},\mathscr{D}_{i+1}}\\
&= \sum_{i=n}^{n+j}\left(\prod_{k=0}^{i}\frac{r_{k+1}}{r_k}\right)\|\mathsf{W}_i(\phi_{i+1}-\phi_i)\|_{r_{i+1},s_{i+1},\mathscr{D}_{i+1}}\\
&= \sum_{i=n}^{n+j}\frac{r_{i+1}}{r_0}\|\mathsf{W}_i(\phi^{i+1}-\phi^i)\|_{r_{i+1},s_{i+1},\mathscr{D}_{i+1}}\\
&\leqslant \frac{1}{2}\|\mathsf{W}_0\phi_0\|_{r_0,s_0,\mathscr{D}_{r_0,\alpha}}\sum_{i=n}^{n+j}\frac{r_{i+1}}{r_i}|\varepsilon|^{2^{i+1}}\overline{\mathsf{L}}_{i+1}3\sigma_i^{-1}\\
&\leqslant \frac{1}{2}\|\mathsf{W}_0\phi_0\|_{r_0,s_0,\mathscr{D}_{r_0,\alpha}}\sum_{i=n}^{n+j}\frac{r_{i+1}}{r_i}|\varepsilon|^{2^{i+1}}\mathsf{L}_{i+1}3\sigma_i^{-1}\\
&\leqslant \frac{1}{2}\|\mathsf{W}_0\phi_0\|_{r_0,s_0,\mathscr{D}_{r_0,\alpha}}\sum_{i=n}^{n+j}\theta_{i+1}\\
&= \frac{1}{2}\|\mathsf{W}_0\phi_0\|_{r_0,s_0,\mathscr{D}_{r_0,\alpha}}\sum_{i=n}^{n+j}\theta_1^{2^i}.
\end{aligned}$$

Hence ϕ_j converges uniformly on $\mathscr{D}_* \times \mathbb{T}^d_{s_*} \times (-\varepsilon_*, \varepsilon_*)$ to some ϕ_*, which is then real–analytic function on $\mathscr{D}_* \times \mathbb{T}^d_{s_*} \times (-\varepsilon_*, \varepsilon_*)$.
To estimate $\|\mathsf{W}_0(\phi_* - \mathrm{id})\|_{\mathscr{D}_*\times\mathbb{T}^d_{s_*}}$, observe that , for $i \geqslant 1$,[82]

$$|\varepsilon|^{2^i}\mathsf{L}_i = \frac{\sigma_0}{3\cdot 2^i}\ \mathsf{e}_*\, {d_*}^i|\varepsilon|^{2^i}M_i < \frac{1}{3\cdot 2^i d_*}(|2\varepsilon|^2\mathsf{e}_*\, d_*^2M_1)^{2^{i-1}} \leqslant \frac{1}{3d_*}\Big(\frac{4|\varepsilon|^2\mathsf{e}_*\, d_*^2M_1}{2}\Big)^i$$

and therefore

$$\sum_{i\geqslant 1}|\varepsilon|^{2^i}\mathsf{L}_i \leqslant \frac{1}{3d_*}\sum_{i\geqslant 1}\Big(\frac{4|\varepsilon|^2\mathsf{e}_*\, d_*^2M_1}{2}\Big)^i \leqslant \frac{1}{3}4|\varepsilon|^2\mathsf{e}_*\mathsf{d}_*M_1 \leqslant \frac{1}{3\mathsf{d}_*}\ .$$

[82]Notce that $2^{i-1} \geqslant i$, $\forall\, i \geqslant 0$.

Moreover,

$$\begin{aligned}
\|\mathsf{W}_0(\phi^i-\mathrm{id})\|_{r_i,s_i,\mathscr{D}_i} &\leqslant \|\mathsf{W}_0(\phi^{i-1}\circ\phi_i-\phi_i)\|_{r_i,s_i,\mathscr{D}_i}+\|\mathsf{W}_0(\phi_i-\mathrm{id})\|_{r_i,s_i,\mathscr{D}_i}\\
&\leqslant \|\mathsf{W}_0(\phi^{i-1}-\mathrm{id})\|_{r_{i-1},s_{i-1},\mathscr{D}_{i-1}}+\left(\prod_{j=0}^{i-2}\|\mathsf{W}_j\mathsf{W}_{j+1}^{-1}\|\right)\|\mathsf{W}_{i-1}(\phi_i-\mathrm{id})\|_{r_i,s_i,\mathscr{D}_i}\\
&= \|\mathsf{W}_0(\phi^{i-1}-\mathrm{id})\|_{r_{i-1},s_{i-1},\mathscr{D}_{i-1}}+\left(\prod_{j=0}^{i-2}\frac{r_{j+1}}{r_j}\right)\|\mathsf{W}_{i-1}(\phi_i-\mathrm{id})\|_{r_i,s_i,\mathscr{D}_i}\\
&= \|\mathsf{W}_0(\phi^{i-1}-\mathrm{id})\|_{r_{i-1},s_{i-1},\mathscr{D}_{i-1}}+\frac{r_{i-1}}{r_0}\|\mathsf{W}_{i-1}(\phi_i-\mathrm{id})\|_{r_i,s_i,\mathscr{D}_i}\\
&\leqslant \|\mathsf{W}_0(\phi^{i-1}-\mathrm{id})\|_{r_{i-1},s_{i-1},\mathscr{D}_{i-1}}+|\varepsilon|^{2^{i-1}}\overline{\mathsf{L}}_{i-1}\ ,
\end{aligned}$$

which iterated yields

$$\begin{aligned}
\|\mathsf{W}_0(\phi^i-\mathrm{id})\|_{r_i,s_i,\mathscr{D}_i} &\leqslant \sum_{k=0}^{i-1}|\varepsilon|^{2^k}\overline{\mathsf{L}}_k\\
&\leqslant |\varepsilon|\overline{\mathsf{L}}_0+\sum_{k\geqslant 1}|\varepsilon|^{2^k}\mathsf{L}_k\\
&\leqslant |\varepsilon|\overline{\mathsf{L}}_0+\frac{4}{3}|\varepsilon|^2\mathsf{e}_*\mathsf{d}_*M_1\\
&= |\varepsilon|\mathsf{C}_0\max(\alpha,r_0\mathsf{K}_\infty)\frac{M_0}{\alpha^2 r_0}\sigma_0^{-(2\tau+d+2)}+\frac{4}{9}|\varepsilon|^2\mathsf{f}_*\mathsf{e}_*\mathsf{d}_*\sigma_0M_0^2\\
&\leqslant \frac{1}{3}|\varepsilon|M_0\mathsf{f}_*\sigma_0^{d+1}+\frac{4}{9}|\varepsilon|^2\mathsf{f}_*\mathsf{e}_*\mathsf{d}_*\sigma_0M_0^2\\
&= \varepsilon_1\ .
\end{aligned}$$

Therefore, taking the limit over i completes the proof of (4.3.43).
Next, we show that $\|G_*-\mathrm{id}\|_{L,\mathscr{D}_{r_0,\alpha}}<1$, which will imply that[83] $G_*\colon \mathscr{D}_{r_0,\alpha}\xrightarrow{onto}\mathscr{D}_*$ is a lipeomorphism. Indeed, for any $j\geqslant 1$, there exists $\hat{r}_j>0$ such that the restricted maps

[83]See Proposition II.2. in [Zeh10].

$G_i\colon G^{i-1}(D_{\hat r_j}(\mathscr{D}_{r_0,\alpha}))\to\mathbb{C}$, $1\leqslant i\leqslant j$ with $G^0:=\mathrm{id}$, are well–defined[84] and, therefore,

$$\begin{aligned}
\|\partial_z G^j-\mathbb{1}_d\|_{\hat r_j,\mathscr{D}_{r_0,\alpha}} &\leqslant \|\partial_z G^j-\partial_z G^{j-1}\|_{\hat r_j,\mathscr{D}_{r_0,\alpha}}+\|\partial_z G^{j-1}-\mathbb{1}_d\|_{\hat r_j,\mathscr{D}_{r_0,\alpha}}\\
&=\|\partial_z G_j\circ\partial_z G^{j-1}-\partial_z G^{j-1}\|_{\hat r_j,\mathscr{D}_{r_0,\alpha}}+\|\partial_z G^{j-1}-\mathbb{1}_d\|_{\hat r_j,\mathscr{D}_{r_0,\alpha}}\\
&\leqslant \|\partial_z G_j-\mathbb{1}_d\|_{\tilde r_j,\mathscr{D}_{j-1}}\|\partial_z G^{j-1}\|_{\hat r_j,\mathscr{D}_{r_0,\alpha}}+\|\partial_z G^{j-1}-\mathbb{1}_d\|_{\hat r_j,\mathscr{D}_{r_0,\alpha}}\\
&\leqslant \|\partial_z G_j-\mathbb{1}_d\|_{\tilde r_j,\mathscr{D}_{j-1}}(\|\partial_z G^{j-1}-\mathbb{1}_d\|_{\hat r_j,\mathscr{D}_{r_0,\alpha}}+1)+\|\partial_z G^{j-1}-\mathbb{1}_d\|_{\hat r_j,\mathscr{D}_{r_0,\alpha}}\\
&=(\|\partial_z G_j-\mathbb{1}_d\|_{\tilde r_j,\mathscr{D}_{j-1}}+1)(\|\partial_z G^{j-1}-\mathbb{1}_d\|_{\hat r_j,\mathscr{D}_{r_0,\alpha}}+1)-1\\
&\overset{(4.3.47)}{\leqslant}(|\varepsilon|^{2^{j-1}}\mathsf{L}_{j-1}+1)(\|\partial_z G^{j-1}-\mathbb{1}_d\|_{\hat r_j,\mathscr{D}_{r_0,\alpha}}+1)-1
\end{aligned}$$

which iterated leads to[85]

$$\begin{aligned}
\|\partial_z G^j-\mathbb{1}_d\|_{\hat r_j,\mathscr{D}_{r_0,\alpha}} &\leqslant -1+\prod_{i=1}^{\infty}(|\varepsilon|^{2^{i-1}}\mathsf{L}_{i-1}+1)\\
&\leqslant -1+\exp\left(\sum_{i=0}^{\infty}|\varepsilon|^{2^i}\mathsf{L}_i\right)\\
&=-1+\exp\left(|\varepsilon|\mathsf{L}_0+\sum_{i=1}^{\infty}|\varepsilon|^{2^i}\mathsf{L}_i\right)\\
&\leqslant -1+\exp\left(\frac{\sigma_0}{3}\mathsf{f}_*|\varepsilon|M_0+\frac{1}{3}4|\varepsilon|^2\mathsf{e}_*\mathsf{d}_*M_1\right)\\
&=-1+e^{\varepsilon_2}\\
&\leqslant \varepsilon_2\,e^{\varepsilon_2}\\
&\overset{(4.3.37)}{\leqslant}\left(\frac{\sigma_0}{3}+\frac{1}{3\mathsf{d}_*}\right)\exp\left(\frac{\sigma_0}{3}+\frac{1}{3\mathsf{d}_*}\right)\\
&\leqslant\left(\frac16+\frac16\right)\exp\left(\frac16+\frac16\right)\\
&=\frac13\exp\left(\frac13\right)<1.
\end{aligned}$$

Thus, G_* is Lipschitz continuous, with

$$\|G_*-\mathrm{id}\|_{L,\mathscr{D}_{r_0,\alpha}}\leqslant\varepsilon_2\,e^{\varepsilon_2}\leqslant\frac13\exp\left(\frac13\right)<1,$$

[84] *i.e.* $G^{i-1}(D_{\hat r_j}(\mathscr{D}_{r_0,\alpha}))\subset D_{\tilde r_i}(\mathscr{D}_{i-1})=\mathrm{dom}\,(G_i)$, $\forall\,1\leqslant i\leqslant j$.

[85] Recall that $e^x-1\leqslant x\,e^x$, $\forall\,x\geqslant0$.

so that, by[86] Lemma D.1 (see Appendix D), we get

$$\begin{aligned} |\,\mathrm{meas}\,(\mathscr{D}_*) - \,\mathrm{meas}\,(\mathscr{D}_{r_0,\alpha})| &\leqslant \left(\left(1+\frac{1}{3}\exp\left(\frac{1}{3}\right)\right)^d - 1\right)\varepsilon_2\, e^{\varepsilon_2}\,\mathrm{meas}\,(\mathscr{D}_{r_0,\alpha}) \\ &= \mathsf{C}_8\varepsilon_2\, e^{\varepsilon_2}\,\mathrm{meas}\,(\mathscr{D}_{r_0,\alpha}), \end{aligned}$$

which proves (4.3.42), Lemma 4.3.2 and, whence, the extension Theorem. ∎

[86] With $\delta \coloneqq \varepsilon_2\, e^{\varepsilon_2} \leqslant \frac{1}{3}\exp\left(\frac{1}{3}\right)$.

5 A “sharp” version of Arnold’s theorem

5.1 Assumptions

Let $r_0 > 0$, $\tau \geqslant d-1$, $0 < s_* < s_0 \leqslant 1$, $y_0 \in \mathbb{R}^d$ and consider the hamiltonian parametrized by $\varepsilon \in \mathbb{R}$

$$H_0(y, x; \varepsilon) := K_0(y) + \varepsilon P_0(y, x),$$

with

$$K_0, P_0 \in \mathcal{B}_{r_0,s_0}(y_0)\,.$$

such that

$$\det(\partial_y^2 K_0(y_0)) \neq 0. \tag{5.1.1}$$

Set

$$T := \partial_y^2 K_0(y_0)^{-1}, \quad M_0 := \|P\|_{r_0,s_0,y_0}, \quad \mathsf{K}_0 := \|\partial_y^2 K_0\|_{r_0,y_0}, \quad \mathsf{T}_0 := \|T\|\,.$$

Finally, define[87]

$$
\begin{aligned}
\nu &:= \tau + 1 \ , \\
\hat{s} &:= s_0 - s_* \ , \\
\sigma_0 &:= \frac{\hat{s}}{2} \ , \\
\eta_0 &:= \mathsf{T}_0 \mathsf{K}_0 \ , \\
\mathsf{C}_0 &:= 4\sqrt{2}\left(\frac{3}{2}\right)^{2\nu+d} \int_{\mathbb{R}^d} \left(|y|_1^{\nu} + d|y|_1^{2\nu}\right) e^{-|y|_1} dy \ , \\
\mathsf{C}_1 &:= 2\left(\frac{3}{2}\right)^{\nu+d} \int_{\mathbb{R}^d} |y|_1^{\nu} e^{-|y|_1} dy \ , \\
\mathsf{C}_2 &:= 2^{3d} d \ , \\
\mathsf{C}_3 &:= \left(d^2\mathsf{C}_1^2 + 6d\mathsf{C}_1 + \mathsf{C}_2\right)\sqrt{2} \ , \\
\mathsf{C}_4 &:= \max\{\mathsf{C}_0, \mathsf{C}_3\} \ , \\
\mathsf{C}_6 &:= \max\left\{2^{2\nu}, \frac{3\cdot 2^5 d}{5}\right\} \ , \\
\mathsf{C}_7 &:= 3d\cdot 2^{6\nu+2d+3}\sqrt{2}\max\left\{640d^2, \mathsf{C}_4\right\} \ , \\
\mathsf{C}_8 &:= \left(2^{-d}\mathsf{C}_6\right)^{\frac{1}{8}} \ , \\
\mathsf{C}_9 &:= \frac{\mathsf{C}_6\mathsf{C}_7\mathsf{C}_8}{2^{2\nu+7}d} \ , \\
\mathsf{C}_{10} &:= 3\cdot 2^d\,\mathsf{C}_8 \ , \\
\mu_* &:= \max\left\{0 < \mu \leqslant e^{-1} \ : \ \mathsf{C}_7\,\mathsf{C}_8\,\eta_0^{\frac{17}{8}}\,\sigma_0^{-(4\nu+2d+1)}\,\mu\left(\log\mu^{-1}\right)^{2\nu} < 1\right\} \ .
\end{aligned}
$$

5.2 Statement of the KAM Theorem

Theorem 5.2.1 *Under the assumptions in §5.1, the following holds. Let*

$$\alpha \leqslant \frac{r_0}{\mathsf{T}_0} \tag{5.2.1}$$

and assume that

$$\omega := \partial_y K_0(y_0) \in \Delta_\alpha^\tau, \quad i.e. \quad |\omega\cdot k| \geqslant \frac{\alpha}{|k|_1^\tau}, \ \forall\ k \in \mathbb{Z}^d\backslash\{0\} \ . \tag{5.2.2}$$

[87] Recall from footnote 13 that $\mathsf{C}_0, \mathsf{C}_1 > 1$.

Assume

$$|\varepsilon| \leqslant \mu_* \, \frac{\alpha^2}{\mathsf{K}_0 \, M_0} \, . \tag{5.2.3}$$

Then, there exist $y_* \in \mathbb{R}^d$ *and an embedding* $\phi_* \colon \mathbb{T}^d \to D_{r_0,s_0}(y_0)$*, real–analytic on* $\mathbb{T}^d_{s_*}$ *and close to the trivial embedding*

$$\phi_0 \colon x \in \mathbb{T}^d \to (y_*, x) \in D_{r_0,s_0}(y_0),$$

and such that the d–torus

$$\mathcal{T}_{\omega,\varepsilon} := \phi_* \left(\mathbb{T}^d \right) \tag{5.2.4}$$

is a non-degenerate invariant Kronecker torus for H *i.e.*

$$\phi^t_H \circ \phi_*(x) = \phi_*(x + \omega t). \tag{5.2.5}$$

Moreover,

$$|y_* - y_0| \leqslant \frac{1}{\mathsf{C}_9} \sigma_0^{3\nu+2d+1} \frac{\alpha}{\mathsf{K}_0 \eta_0^{\frac{17}{8}}} \, , \tag{5.2.6}$$

$$|\mathsf{W}(\phi_* - \mathrm{id})| \leqslant \frac{1}{\mathsf{C}_{10} \, \eta_0^{\frac{1}{8}}} \, , \tag{5.2.7}$$

uniformly on $\{y_*\} \times \mathbb{T}^d_{s_*}$ *, where*

$$\mathsf{W} := \ \mathrm{diag} \left(\frac{\mathsf{K}_0}{\alpha} \mathbb{1}_d, \mathbb{1}_d \right) \, .$$

5.3 Proof of Theorem 5.2.1

Lemma 5.3.1 (KAM step) *Let* $r > 0$, $0 < 2\sigma < s \leqslant 1$ *and consider the hamiltonian parametrized by* $\varepsilon \in \mathbb{R}$

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

with

$$K, P \in \mathcal{B}_{r,s}(\mathsf{y}) \, .$$

Assume that[88,89]

$$\begin{aligned} &\det K_{yy}(\mathsf{y}) \neq 0 \, , && T := K_{yy}(\mathsf{y})^{-1} \, , \\ &\|K_{yy}\|_{r,\mathsf{y}} \leqslant \mathsf{K} \, , && \|T\| \leqslant \mathsf{T} \, , \\ &\|P\|_{r,s,\mathsf{y}} \leqslant M \, , && \omega := K_{yy}(\mathsf{y}) \in \Delta^\tau_\alpha \, . \end{aligned} \tag{5.3.1}$$

[88] 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, r, the roles of y_j, r_j in the iterative step.

[89] Notice that $\mathsf{T}\mathsf{K} \geqslant \mathsf{T}\|K_{yy}(\mathsf{y})\| \geqslant \|T\|\|K_{yy}(\mathsf{y})\| = \|T\|\|T^{-1}\| \geqslant 1$.

Fix $\varepsilon \neq 0$ and assume that

$$\lambda \geqslant \log\left(\sigma^{2\nu+d}\frac{\alpha^2}{|\varepsilon| M\mathsf{K}}\right) \geqslant 1\ . \tag{5.3.2}$$

Let

$$\begin{aligned} &\kappa := 4\sigma^{-1}\lambda, \quad \check{r} \leqslant \frac{5}{24d}\frac{r}{\mathsf{TK}}\ , \quad \bar{r} \leqslant \min\left\{\frac{\alpha}{2d\mathsf{K}\kappa^{\tau+1}}\ , \check{r}\right\}.\\ &\bar{s} := s - \frac{2}{3}\sigma, \quad s' := s - \sigma\ , \end{aligned} \tag{5.3.3}$$

and[90]

$$\begin{aligned} \overline{\mathsf{L}} &:= \frac{\mathsf{C}_0}{\sqrt{2}}\max\left\{1, \frac{\alpha}{r\mathsf{K}}\right\}\frac{M\mathsf{K}}{\alpha^2}\sigma^{-(2\nu+d)}\ ,\\ \mathsf{L} &:= M\max\left\{\frac{40d\mathsf{T}^2\mathsf{K}}{r^2}\sigma^{-(\nu+d)}\ , \frac{4}{\mathsf{K}r^2}\ , \frac{\mathsf{C}_4}{\sqrt{2}}\max\left\{1, \frac{\alpha}{r\mathsf{K}}\right\}\frac{\mathsf{K}}{\alpha^2}\sigma^{-2(\nu+d)}\right\}\\ &= M\max\left\{\frac{40d\mathsf{T}^2\mathsf{K}}{r^2}\sigma^{-(\nu+d)}\ , \frac{\mathsf{C}_4}{\sqrt{2}}\max\left\{1, \frac{\alpha}{r\mathsf{K}}\right\}\frac{\mathsf{K}}{\alpha^2}\sigma^{-2(\nu+d)}\right\}\ . \end{aligned}$$

Then, there exists a generating function $g \in \mathcal{B}_{\bar{r},\bar{s}}(\mathsf{y})$ with the following properties:

$$\begin{cases} \|g_x\|_{\bar{r},\bar{s},\mathsf{y}} \leqslant \mathsf{C}_1\dfrac{M}{\alpha}\sigma^{-(\nu+d)}\ ,\\ \|g_{y'}\|_{\bar{r},\bar{s},\mathsf{y}},\ \|\partial^2_{y'x}g\|_{\bar{r},\bar{s},\mathsf{y}} \leqslant \overline{\mathsf{L}}\ ,\\ \|\partial^2_{y'}\widetilde{K}\|_{\check{r},\mathsf{y}} \leqslant \mathsf{KL}\ , \end{cases} \tag{5.3.4}$$

where

$$\widetilde{K}(y') := \langle P(y', \cdot)\rangle\ .$$

If, in addition,

$$|\varepsilon|\mathsf{L} \leqslant \frac{\sigma}{3}\ , \tag{5.3.5}$$

then, there exists $\mathsf{y}' \in \mathbb{R}^d$ such that

$$\begin{cases} \partial_{y'}K'(\mathsf{y}') = \omega\ , & \det \partial^2_{y'}K'(\mathsf{y}') \neq 0\ ,\\ |\varepsilon|\|g_x\|_{\bar{r},\bar{s},\mathsf{y}} \leqslant \dfrac{r}{3}\ , & |\mathsf{y}' - \mathsf{y}| \leqslant \dfrac{8|\varepsilon|\mathsf{T}M}{r}\ ,\\ |\varepsilon|\|\widetilde{T}\| \leqslant \mathsf{T}|\varepsilon|\mathsf{L}\ , & \|P_+\|_{\bar{r},\bar{s},\mathsf{y}} \leqslant \mathsf{L}M\ , \end{cases} \tag{5.3.6}$$

[90]Notice that $\mathsf{L} \geqslant \sigma^{-d}\overline{\mathsf{L}} \geqslant \overline{\mathsf{L}}$ and $\frac{40d\mathsf{T}^2\mathsf{K}}{r^2}\sigma^{-(\nu+d)} > \frac{4\mathsf{T}}{r^2} \geqslant \frac{4}{\mathsf{K}r^2}$.

where

$$K' := K + \varepsilon \widetilde{K}\ , \qquad \big(\partial^2_{y'} K'(\mathsf{y}')\big)^{-1} =: T + \varepsilon\, \widetilde{T}\ , \qquad P_+(y', x) := P(y' + \varepsilon g_x(y', x), x)\ .$$

and the following hold. For $y' \in D_{\bar r}(\mathsf{y})$*, the map* $\psi_\varepsilon(x) := x + \varepsilon g_{y'}(y', x)$ *has an analytic inverse* $\varphi(x') = x' + \varepsilon\widetilde{\varphi}(y', x'; \varepsilon)$ *such that*

$$\|\widetilde{\varphi}\|_{\bar r, s', \mathsf{y}} \leqslant \bar{\mathsf{L}} \qquad \text{and} \quad \varphi = \mathrm{id} + \varepsilon\widetilde{\varphi} : D_{\bar r/2, s'}(\mathsf{y}') \to \mathbb{T}^d_{\bar s}\ ; \tag{5.3.7}$$

for any $(y', x) \in D_{\bar r, \bar s}(\mathsf{y})$, $|y' + \varepsilon g_x(y', x) - \mathsf{y}| < \frac{2}{3} r$*; the map* ϕ' *is a symplectic diffeomorphism and*

$$\phi' = \big(y' + \varepsilon g_x(y', \varphi(y', x')), \varphi(y', x')\big) : D_{\bar r/2, s'}(\mathsf{y}') \to D_{2r/3, \bar s}(\mathsf{y}), \tag{5.3.8}$$

with

$$\|\mathsf{W}\, \tilde{\phi}\|_{\bar r/2, s', \mathsf{y}'} \leqslant \sigma^d \mathsf{L}\,, \tag{5.3.9}$$

where $\tilde{\phi}$ *is defined by the relation* $\phi' =: \mathrm{id} + \varepsilon\tilde{\phi}$*,*

$$\mathsf{W} := \begin{pmatrix} \max\{\frac{\mathsf{K}}{\alpha}\ , \frac{1}{r}\}\ \mathbb{1}_d & 0 \\ \\ 0 & \mathbb{1}_d \end{pmatrix}$$

and

$$\|P'\|_{\bar r/2, s', \mathsf{y}'} \leqslant \mathsf{L} M\ , \tag{5.3.10}$$

with

$$P'(y', x') := P_+(y', \varphi(x')) = P \circ \phi'(y', x')\ .$$

Proof

Step 1: Construction of the Arnold's transformation We seek for a near–to–the–identity symplectic transformation

$$\phi' \colon D_{r_1, s_1}(\mathsf{y}') \to D_{r,s}(\mathsf{y}),$$

with $D_{r_1, s_1}(\mathsf{y}') \subset D_{r,s}(\mathsf{y})$, generated by a function of the form $y' \cdot x + \varepsilon g(y', x)$, so that

$$\phi' \colon \begin{cases} y = y' + \varepsilon g_x(y', x) \\ x' = x + \varepsilon g_{y'}(y', x)\,, \end{cases} \tag{5.3.11}$$

such that

$$\begin{cases} H' := H \circ \phi' = K' + \varepsilon^2 P'\ , \\ \partial_{y'} K'(\mathsf{y}') = \omega, \quad \det \partial^2_{y'} K'(\mathsf{y}') \neq 0\,. \end{cases} \tag{5.3.12}$$

By Taylor's formula, we get[91]

$$
\begin{aligned}
H(y'+\varepsilon g_x(y',x),x) =&K(y')+\varepsilon\widetilde{K}(y')+\varepsilon\Big[K'(y')\cdot g_x+T_\kappa P(y',\cdot)-\widetilde{K}(y')\Big]+\\
&+\varepsilon^2\left(P^{(1)}+P^{(2)}+P^{(3)}\right)(y',x)\\
=&K'(y')+\varepsilon\Big[K'(y')\cdot g_x+T_\kappa P(y',\cdot)-\widetilde{K}(y')\Big]+\varepsilon^2 P_+(y',x),
\end{aligned}
\tag{5.3.13}
$$

with $\kappa\in\mathbb{N}$, which will be chosen large enough so that $P^{(3)}=O(\varepsilon)$ and

$$
\begin{cases}
P_+ := P^{(1)}+P^{(2)}+P^{(3)}\\
P^{(1)} := \dfrac{1}{\varepsilon^2}\left[K(y'+\varepsilon g_x)-K(y')-\varepsilon K'(y')\cdot g_x\right]=\displaystyle\int_0^1(1-t)K_{yy}(\varepsilon t g_x)\cdot g_x\cdot g_x dt\\
P^{(2)} := \dfrac{1}{\varepsilon}\left[P(y'+\varepsilon g_x,x)-P(y',x)\right]=\displaystyle\int_0^1 P_y(y'+\varepsilon t g_x,x)\cdot g_x dt\\
P^{(3)} := \dfrac{1}{\varepsilon}\left[P(y',x)-T_\kappa P(y',\cdot)\right]=\dfrac{1}{\varepsilon}\displaystyle\sum_{|n|_1>\kappa}P_n(y')\,e^{in\cdot x}\,.
\end{cases}
\tag{5.3.14}
$$

By the non–degeneracy condition in (5.3.1), for ε small enough (to be made precised below), $\det\partial^2_{y'}K'(\mathsf{y})\neq 0$ and, therefore, by Lemma 2.2.6, there exists a unique $\mathsf{y}'\in D_r(\mathsf{y})$ such that the second part of (5.3.12) holds. In view of (5.3.13), in order to get the first part of (5.3.12), we need to find g such that $K_y(y')\cdot g_x+T_\kappa P(y',\cdot)-\widetilde{K}(y')$ vanishes; such a g is indeed given by

$$
g:=\sum_{0<|n|_1\leqslant\kappa}\frac{-P_n(y')}{iK_y(y')\cdot n}\,e^{in\cdot x},
\tag{5.3.15}
$$

provided that

$$
K_y(y')\cdot n\neq 0,\quad \forall\, 0<|n|_1\leqslant\kappa,\quad \forall\, y'\in D_{r_1}(\mathsf{y}')\quad (\subset D_r(\mathsf{y}))\,.
\tag{5.3.16}
$$

But, in fact, since $K_y(\mathsf{y})$ is rationally independent, then, given any $\kappa\in\mathbb{N}$, there exists $\bar r\leqslant r$ such that

$$
K_y(y')\cdot n\neq 0,\quad \forall\, 0<|n|_1\leqslant\kappa,\quad \forall\, y'\in D_{\bar r}(\mathsf{y}).
\tag{5.3.17}
$$

The last step is to invert the function $x\mapsto x+\varepsilon g_{y'}(y',x)$ in order to define P'. But, by Lemma 2.2.6, for ε small enough, the map $x\mapsto x+\varepsilon g_{y'}(y',x)$ admits an real–analytic inverse of the form

$$
\varphi(y',x';\varepsilon):=x'+\varepsilon\widetilde{\varphi}(y',x';\varepsilon),
\tag{5.3.18}
$$

[91]Recall that $\langle\cdot\rangle$ stands for the average over $\mathbb{T}^d$.

so that the Arnod's symplectic transformation is given by

$$\phi' \colon (y',x') \mapsto \begin{cases} y = y' + \varepsilon g_x(y',\varphi(y',x')) \\ x = \varphi(y',x';\varepsilon) = x' + \varepsilon\widetilde{\varphi}(y',x';\varepsilon). \end{cases} \tag{5.3.19}$$

Hence, (5.3.12) holds with

$$P'(y',x') := P_+(y',\varphi(y',x')). \tag{5.3.20}$$

Step 2: Quantitative estimates
First of all, notice that [92]

$$\bar r \leqslant \check r \leqslant \frac{5r}{24d} < \frac{r}{2}\ . \tag{5.3.21}$$

We begin by extending the "diophantine condition w.r.t. K_y" uniformly to $D_{\bar r}(\mathsf{y})$ up to the order κ. Indeed, by the Mean Value Inequality and $K_y(\mathsf{y}) = \omega \in \Delta^\tau_\alpha$, we get, for any $0 < |n|_1 \leqslant \kappa$ and any $y' \in D_{\bar r}(\mathsf{y})$,

$$\begin{aligned} |K_y(y')\cdot n| &= |\omega\cdot n + (K_y(y') - K_y(\mathsf{y}))\cdot n| \geqslant |\omega\cdot n|\left(1 - d\frac{\|K_{yy}\|_{\bar r,\mathsf{y}}}{|\omega\cdot n|}|n|_1\bar r\right) \\ &\geqslant \frac{\alpha}{|n|_1^\tau}\left(1 - \frac{d\mathsf{K}}{\alpha}|n|_1^{\tau+1}\bar r\right) \geqslant \frac{\alpha}{|n|_1^\tau}\left(1 - \frac{d\mathsf{K}}{\alpha}\kappa^{\tau+1}\bar r\right) \geqslant \frac{\alpha}{2|n|_1^\tau}, \end{aligned} \tag{5.3.22}$$

so that, by Lemma 2.2.4–(i), we have

$$\begin{aligned} \|g_x\|_{\bar r,\bar s,\mathsf{y}} &\overset{def}{=} \sup_{D_{\bar r,\bar s}(\mathsf{y})}\left|\sum_{0<|n|_1\leqslant\kappa}\frac{nP_n(y')}{K_y(y')\cdot n}\,e^{in\cdot x}\right| \leqslant \sum_{0<|n|_1\leqslant\kappa}\frac{\|P_n\|_{\bar r,\bar s,\mathsf{y}}}{|K_y(y')\cdot n|}|n|_1\,e^{\left(s-\frac{2}{3}\sigma\right)|n|_1} \\ &\leqslant \sum_{0<|n|_1\leqslant\kappa} M\,e^{-s|n|_1}\frac{2|n|_1^\nu}{\alpha}\,e^{\left(s-\frac{2}{3}\sigma\right)|n|_1} \leqslant \frac{2M}{\alpha}\sum_{n\in\mathbb{Z}^d}|n|_1^\nu\,e^{-\frac{2}{3}\sigma|n|_1} \\ &\leqslant \frac{2M}{\alpha}\int_{\mathbb{R}^d}|y|_1^\nu\,e^{-\frac{2}{3}\sigma|y|_1}dy \\ &= \left(\frac{3}{2\sigma}\right)^{\nu+d}\frac{2M}{\alpha}\int_{\mathbb{R}^d}|y|_1^\nu\,e^{-|y|_1}dy \\ &= \mathsf{C}_1\frac{M}{\alpha}\sigma^{-(\nu+d)}\,, \end{aligned}$$

[92] Recall footnote 42.

$$\begin{aligned}
\|\partial_{y'} g\|_{\bar r,\bar s,\mathsf{y}} &\overset{def}{=} \sup_{D_{\bar r,\bar s}(\mathsf{y})}\left|\sum_{0<|n|_1\leqslant\kappa}\left(\frac{\partial_y P_n(y')}{K_y(y')\cdot n}-P_n(y')\frac{K_{yy}(y')n}{(K_y(y')\cdot n)^2}\right)e^{in\cdot x}\right|\\
&\leqslant \sum_{0<|n|_1\leqslant\kappa}\sup_{D_{\bar r}(\mathsf{y})}\left(\frac{\|(P_y)_n\|_{\bar r,s,\mathsf{y}}}{|K_y(y')\cdot n|}+d\|P_n\|_{r,s,\mathsf{y}}\frac{\|K_{yy}\|_{r,\mathsf{y}}|n|_1}{|K_y(y')\cdot n|^2}\right)e^{\left(s-\frac{2}{3}\sigma\right)|n|_1}\\
&\overset{(5.3.1)+(5.3.22)}{\leqslant} \sum_{0<|n|_1\leqslant\kappa}\left(\frac{M}{r-\bar r}e^{-s|n|_1}\frac{2|n|_1^\tau}{\alpha}+dM\,e^{-s|n|_1}\mathsf{K}|n|_1\left(\frac{2|n|_1^\tau}{\alpha}\right)^2\right)e^{\left(s-\frac{2}{3}\sigma\right)|n|_1}\\
&\overset{(5.3.21)}{\leqslant}\frac{4M}{\alpha^2 r}\sum_{0<|n|_1\leqslant\kappa}\left(|n|_1^\tau\alpha+dr\mathsf{K}|n|_1^{2\tau+1}\right)e^{-\frac{2}{3}\sigma|n|_1}\\
&\leqslant \max\{\alpha,r\mathsf{K}\}\frac{4M}{\alpha^2 r}\sum_{0<|n|_1\leqslant\kappa}\left(|n|_1^\tau+d|n|_1^{2\tau+1}\right)e^{-\frac{2}{3}\sigma|n|_1}\\
&\leqslant \max\left\{1,\frac{\alpha}{r\mathsf{K}}\right\}\frac{4M\mathsf{K}}{\alpha^2}\int_{\mathbb{R}^d}\left(|y|_1^\tau+d|y|_1^{2\tau+1}\right)e^{-\frac{2}{3}\sigma|y|_1}dy\\
&=\left(\frac{3}{2\sigma}\right)^{2\tau+d+1}\max\left\{1,\frac{\alpha}{r\mathsf{K}}\right\}\frac{4M\mathsf{K}}{\alpha^2}\int_{\mathbb{R}^d}\left(|y|_1^\tau+d|y|_1^{2\tau+1}\right)e^{-|y|_1}dy\\
&\leqslant\frac{\mathsf{C}_0}{\sqrt{2}}\max\left\{1,\frac{\alpha}{r\mathsf{K}}\right\}\frac{M\mathsf{K}}{\alpha^2}\sigma^{-(2\tau+d+1)}\\
&\leqslant \bar{\mathsf{L}}\,,
\end{aligned}$$

and, analogously,

$$\begin{aligned}
\|\partial^2_{y'x} g\|_{\bar r,\bar s,\mathsf{y}} &\overset{def}{=} \sup_{D_{\bar r,\bar s}(\mathsf{y})}\left|\sum_{0<|n|_1\leqslant\kappa}\left(\frac{\partial_y P_n(y')}{K_y(y')\cdot n}-P_n(y')\frac{K_{yy}(y')n}{(K_y(y')\cdot n)^2}\right)\cdot n\,e^{in\cdot x}\right|\\
&\leqslant \sum_{0<|n|_1\leqslant\kappa}\sup_{D_{\bar r}(\mathsf{y})}\left(\frac{\|(P_y)_n\|_{\bar r,s,\mathsf{y}}}{|K_y(y')\cdot n|}+d\|P_n\|_{r,s,\mathsf{y}}\frac{\|K_{yy}\|_{r,\mathsf{y}}|n|_1}{|K_y(y')\cdot n|^2}\right)|n|_1\,e^{\left(s-\frac{2}{3}\sigma\right)|n|_1}\\
&\leqslant \max\{\alpha,r\mathsf{K}\}\frac{4M}{\alpha^2 r}\sum_{0<|n|_1\leqslant\kappa}\left(|n|_1^\tau+d|n|_1^{2\tau+1}\right)|n|_1\,e^{-\frac{2}{3}\sigma|n|_1}\\
&\leqslant \max\left\{1,\frac{\alpha}{r\mathsf{K}}\right\}\frac{4M\mathsf{K}}{\alpha^2}\int_{\mathbb{R}^d}\left(|y|_1^\tau+d|y|_1^{2\tau+1}\right)|y|_1\,e^{-\frac{2}{3}\sigma|y|_1}dy\\
&=\left(\frac{3}{2\sigma}\right)^{2\tau+d+2}\max\left\{1,\frac{\alpha}{r\mathsf{K}}\right\}\frac{4M\mathsf{K}}{\alpha^2}\int_{\mathbb{R}^d}\left(|y|_1^{\tau+1}+d|y|_1^{2\tau+2}\right)e^{-|y|_1}dy\\
&=\frac{\mathsf{C}_0}{\sqrt{2}}\max\left\{1,\frac{\alpha}{r\mathsf{K}}\right\}\frac{M\mathsf{K}}{\alpha^2}\sigma^{-(2\nu+d)}\\
&=\bar{\mathsf{L}}\,,
\end{aligned}$$

and, for $|\varepsilon| < \varepsilon_*$,

$$\|\widetilde{K}_y\|_{r/2,\mathsf{y}} = \|\,[P_y]\,\|_{r/2,\mathsf{y}} \leqslant \|P_y\|_{r/2,\bar{s},\mathsf{y}} \leqslant \frac{M}{r-\frac{r}{2}} \leqslant \frac{2M}{r}\ ,$$

$$\|\partial^2_{y'}\widetilde{K}\|_{r/2,\mathsf{y}} = \|\,[P_{yy}]\,\|_{r/2,\mathsf{y}} \leqslant \|P_{yy}\|_{r/2,\bar{s},\mathsf{y}} \leqslant \frac{M}{(r-\frac{r}{2})^2} \leqslant \frac{4M}{r^2} \leqslant \mathsf{KL}\ .$$

Next, we prove the existence and uniqueness of y' in (5.3.12). Consider then

$$\begin{aligned} F\colon D_{\check{r}}(\mathsf{y}) \times D^1_{2|\varepsilon|}(0) &\longrightarrow \qquad \mathbb{C}^d \\ (y,\eta) \quad &\longmapsto K_y(y) + \eta \widetilde{K}_{y'}(y) - K_y(\mathsf{y}). \end{aligned}$$

Then

- $F(\mathsf{y},0) = 0, \quad F_y(\mathsf{y},0)^{-1} = K_{yy}(\mathsf{y})^{-1} = T;$
- For any $(y,\eta) \in D_{\check{r}}(\mathsf{y}) \times D^1_{2|\varepsilon|}(0)$,

$$\begin{aligned} \|\mathbb{1}_d - TF_y(y,\eta)\| &\leqslant \|\mathbb{1}_d - TK_{yy}\| + |\eta|\ \|T\|\ \|\partial^2_{y'}\widetilde{K}\|_{r/2,\mathsf{y}} \\ &\leqslant d\|T\|\|K_{yyy}\|_{\check{r},\mathsf{y}}\check{r} + 2|\varepsilon|\mathsf{T}\frac{4M}{r^2} \\ &\leqslant d\mathsf{TK}\frac{\check{r}}{r-\check{r}} + 8\mathsf{T}\frac{|\varepsilon|M}{r^2} \\ &\overset{(5.3.21)}{\leqslant} d\mathsf{TK}\frac{2\check{r}}{r} + |\varepsilon|\frac{8\mathsf{T}M}{r^2} \\ &\leqslant 2d\mathsf{TK}\frac{\bar{r}}{r} + \frac{1}{2}|\varepsilon|\mathsf{L} \\ &\overset{(5.3.21)+(5.3.5)}{\leqslant} \frac{5}{12} + \frac{\sigma}{6} \\ &\leqslant \frac{5}{12} + \frac{1}{12} = \frac{1}{2}\ ; \end{aligned}$$

- Recalling $\sigma \leqslant \frac{1}{2}$, we have

$$\begin{aligned}2\|T\|\|F(\mathsf{y},\cdot)\|_{2|\varepsilon|,0} &= 2\|T\| \sup_{B^1_{2|\varepsilon|}(0)} |\eta \widetilde{K}_{y'}(\mathsf{y})| \\ &\leqslant 2\mathsf{T}\frac{4|\varepsilon|M}{r} \\ &\leqslant \frac{5\cdot 2^{\nu+d}}{8d}\frac{r}{\mathsf{TK}}\sigma^{\nu+d}|\varepsilon|\mathsf{L} \\ &= 3\cdot 2^d\,(2\sigma)^\nu\,\check{r}\,\sigma^d|\varepsilon|\mathsf{L} \\ &\leqslant 3\cdot 2^d\,\check{r}\,\sigma^d|\varepsilon|\mathsf{L} \qquad (5.3.23)\\ &\overset{(5.3.5)}{\leqslant} 3\,\check{r}\,(2\sigma)^d\,\frac{\sigma}{3} \\ &\leqslant \frac{\check{r}}{2}\,.\end{aligned}$$

Therefore, Lemma 2.2.6 applies. Hence, there exists a function $g\colon D^1_{2|\varepsilon|}(0) \to D_{\bar{r}}(\mathsf{y})$ such that its graph coincides with $F^{-1}(\{0\})$. In particular, $\mathsf{y}' := g(\varepsilon)$ is the unique $y \in D_{\bar{r}}(\mathsf{y})$ satisfying $0 = F(y,\varepsilon) = \partial_y K'(y) - \omega$ *i.e.* the second part of (5.3.12). Moreover,

$$|\mathsf{y}'-\mathsf{y}| \leqslant 2\|T\|\|F(\mathsf{y},\cdot)\|_{2|\varepsilon|,0} \leqslant \frac{8|\varepsilon|\mathsf{T}M}{r} \overset{(5.3.23)}{\leqslant} 3\cdot 2^d\,\check{r}\,\sigma^d|\varepsilon|\mathsf{L} \leqslant \frac{\check{r}}{2}\,, \qquad (5.3.24)$$

so that

$$D_{\frac{\check{r}}{2}}(\mathsf{y}') \subset D_{\check{r}}(\mathsf{y}). \qquad (5.3.25)$$

Next, we prove that $\partial_y^2 K'(\mathsf{y}')$ is invertible. Indeed, by Taylor' formula, we have

$$\begin{aligned}\partial_y^2 K'(\mathsf{y}') &= K_{yy}(\mathsf{y}) + \int_0^1 K_{yyy}(\mathsf{y}+t\varepsilon\widetilde{y})\cdot\varepsilon\widetilde{y}dt + \varepsilon\widetilde{K}_{yy}(\mathsf{y}') \\ &= T^{-1}\left(\mathbb{1}_d + \varepsilon T\left(\int_0^1 K_{yyy}(\mathsf{y}+t\varepsilon\widetilde{y})\cdot\widetilde{y}dt + \widetilde{K}_{yy}(\mathsf{y}')\right)\right) \\ &=: T^{-1}(\mathbb{1}_d + \varepsilon A),\end{aligned}$$

and, by Cauchy's estimate,

$$\begin{aligned}
|\varepsilon|\|A\| &\leqslant \|T\| \left(d\|K_{yyy}\|_{r/2,\mathsf{y}} |\varepsilon| |\mathsf{y}' - \mathsf{y}| + |\varepsilon| \|\partial_{y'}^2 \widetilde{K}\|_{r/2,\mathsf{y}} \right) \\
&\leqslant \|T\| \left(\frac{d\|K_{yy}\|_{r,\mathsf{y}}}{r - \frac{r}{2}} |\varepsilon| |\mathsf{y}' - \mathsf{y}| + |\varepsilon| \|\widetilde{K}_{yy}\|_{r/2,\mathsf{y}} \right) \\
&\overset{(5.3.24)}{\leqslant} \mathsf{T} \left(\frac{2d\mathsf{K}}{r} \frac{8|\varepsilon|\mathsf{T}M}{r} + \frac{4|\varepsilon|M}{r^2} \right) \\
&\leqslant \frac{4|\varepsilon|\mathsf{T}M}{r^2} (4d\mathsf{T}\mathsf{K} + 1) \\
&\leqslant \frac{20d|\varepsilon|\mathsf{T}^2\mathsf{K}M}{r^2} \\
&\leqslant \frac{1}{2} |\varepsilon| \mathsf{L} \\
&\overset{(5.3.5)}{\leqslant} \frac{\sigma}{6} \\
&\leqslant \frac{1}{2}.
\end{aligned}$$

Hence $\partial_{y'}^2 K'(\mathsf{y}')$ is invertible with

$$\partial_{y'}^2 K'(\mathsf{y}')^{-1} = (\mathbb{1}_d + \varepsilon A)^{-1} T = T + \sum_{k \geqslant 1} (-\varepsilon)^k A^k T =: T + \varepsilon \widetilde{T},$$

and

$$|\varepsilon| \|\widetilde{T}\| \leqslant |\varepsilon| \frac{\|A\|}{1 - |\varepsilon| \|A\|} \|T\| \leqslant 2|\varepsilon| \|A\| \|T\| \leqslant |\varepsilon| \mathsf{L}\mathsf{T} \leqslant 2\frac{\sigma}{6} \mathsf{T} = \mathsf{T} \frac{\sigma}{3} .$$

Next, we prove estimate on P_+. We have,

$$|\varepsilon| \|g_x\|_{\bar{r},\bar{s},\mathsf{y}} \leqslant |\varepsilon| \mathsf{C}_1 \frac{M}{\alpha} \sigma^{-(\tau+d+1)} \leqslant |\varepsilon| \frac{r}{3} \mathsf{L} \overset{(5.3.5)}{\leqslant} \frac{r}{3} \frac{\sigma}{3} \leqslant \frac{r}{3}$$

so that, for any $(y', x) \in D_{\bar{r},\bar{s}}(\mathsf{y})$,

$$|y' + \varepsilon g_x(y', x) - \mathsf{y}| \leqslant \bar{r} + \frac{r}{3} < \frac{r}{8d} + \frac{r}{3} < \frac{2r}{3} < r\,,$$

and thus

$$\begin{aligned}
\|P^{(1)}\|_{\bar{r},\bar{s},\mathsf{y}} \leqslant d^2 \|K_{yy}\|_{r,\mathsf{y}} \|g_x\|_{\bar{r},\bar{s},\mathsf{y}}^2 &\leqslant d^2 \mathsf{K} \left(\mathsf{C}_1 \frac{M}{\alpha} \sigma^{-(\nu+d)} \right)^2 \\
&= d^2 \mathsf{C}_1^2 \frac{\mathsf{K}M^2}{\alpha^2} \sigma^{-2(\nu+d)},
\end{aligned}$$

$$\begin{aligned}\|P^{(2)}\|_{\bar r,\bar s,\mathsf{y}} &\leqslant d\|P_y\|_{\frac{5r}{6},\bar s,\mathsf{y}}\|g_x\|_{\bar r,\bar s,\mathsf{y}} \leqslant d\frac{6M}{r}\mathsf{C}_1\frac{M}{\alpha}\sigma^{-(\nu+d)}\\ &= 6d\mathsf{C}_1\frac{M^2}{\alpha r}\sigma^{-(\nu+d)}\end{aligned}$$

and by Lemma 2.2.4–(i), we have,

$$\begin{aligned}|\varepsilon|\|P^{(3)}\|_{\bar r,s-\frac{\sigma}{2},\mathsf{y}} &\leqslant \sum_{|n|_1>\kappa}\|P_n\|_{\bar r,\mathsf{y}}\, e^{(s-\frac{\sigma}{2})|n|_1} \leqslant M\sum_{|n|_1>\kappa} e^{-\frac{\sigma|n|_1}{2}}\\ &\leqslant M\, e^{-\frac{\kappa\sigma}{4}}\sum_{|n|_1>\kappa} e^{-\frac{\sigma|n|_1}{4}} \leqslant M\, e^{-\frac{\kappa\sigma}{4}}\sum_{|n|_1>0} e^{-\frac{\sigma|n|_1}{4}}\\ &= M\, e^{-\frac{\kappa\sigma}{4}}\left(\left(\sum_{k\in\mathbb{Z}} e^{-\frac{\sigma|k|}{4}}\right)^d - 1\right) = M\, e^{-\frac{\kappa\sigma}{4}}\left(\left(1+\frac{2\, e^{-\frac{\sigma}{4}}}{1-\, e^{-\frac{\sigma}{4}}}\right)^d - 1\right)\\ &= M\, e^{-\frac{\kappa\sigma}{4}}\left(\left(1+\frac{2}{e^{\frac{\sigma}{4}}-1}\right)^d - 1\right) \leqslant M\, e^{-\frac{\kappa\sigma}{4}}\left(\left(1+\frac{2}{\frac{\sigma}{4}}\right)^d - 1\right)\\ &\leqslant \sigma^{-d}M\, e^{-\frac{\kappa\sigma}{4}}\left((\sigma+8)^d-\sigma^d\right) \leqslant d8^d\sigma^{-d}M\, e^{-\frac{\kappa\sigma}{4}}\\ &= \mathsf{C}_2\sigma^{-d}M\, e^{-\lambda}\\ &\overset{(5.3.2)}{\leqslant} \mathsf{C}_2\sigma^{-d}M\sigma^{-(2\nu+d)}\frac{|\varepsilon|M\mathsf{K}}{\alpha^2}\\ &= \mathsf{C}_2M\frac{|\varepsilon|M\mathsf{K}}{\alpha^2}\sigma^{-2(\nu+d)}\,.\end{aligned}$$

Hence[93],

$$\begin{aligned}\|P_+\|_{\bar r,\bar s,\mathsf{y}} &\leqslant \|P^{(1)}\|_{\bar r,\bar s,\mathsf{y}} + \|P^{(2)}\|_{\bar r,\bar s,\mathsf{y}} + \|P^{(3)}\|_{\bar r,\bar s,\mathsf{y}}\\ &\leqslant d^2\mathsf{C}_1^2\frac{\mathsf{K}M^2}{\alpha^2}\sigma^{-2(\nu+d)} + 6d\mathsf{C}_1\frac{M^2}{\alpha r}\sigma^{-(\nu+d)} + \mathsf{C}_2M\frac{|\varepsilon|M\mathsf{K}}{\alpha^2}\sigma^{-2(\nu+d)}\\ &= \left(d^2\mathsf{C}_1^2r\mathsf{K} + 6d\mathsf{C}_1\alpha\sigma^{\nu+d} + \mathsf{C}_2r\mathsf{K}\right)\frac{M^2}{\alpha^2 r}\sigma^{-2(\tau+d+1)}\\ &\leqslant \left(d^2\mathsf{C}_1^2 + 6d\mathsf{C}_1 + \mathsf{C}_2\right)\max\{\alpha, r\mathsf{K}\}\frac{M^2}{\alpha^2 r}\sigma^{-2(\tau+d+1)}\\ &\overset{(5.3.1)}{\leqslant} \frac{\mathsf{C}_3}{\sqrt{2}}\max\left\{1,\frac{\alpha}{r\mathsf{K}}\right\}\frac{M^2\mathsf{K}}{\alpha^2}\sigma^{-2(\nu+d)}\\ &\leqslant \mathsf{L}M\,.\end{aligned}$$

[93] Recall that $\sigma < 1$.

The proof of the claims on ϕ' and P' are proven in a similar way as in Lemma 2.3.1. ∎

Finally, we prove the convergence of the scheme by mimicking Lemma 2.3.2.
Let $H_0 := H$, $K_0 := K$, $P_0 := P$ and r_0, s_0, s_*, σ_0, W_0, M_0, K_0, T_0, μ_0 be as in §5.1 and for a given $\varepsilon \neq 0$ and $j \geqslant 0$, define[94]

$$\begin{aligned}
\sigma_j &:= \frac{\sigma_0}{2^j}\,,\\
s_{j+1} &:= s_j - \sigma_j = s_* + \frac{\sigma_0}{2^j}\,,\\
\bar{s}_j &:= s_j - \frac{2\sigma_i}{3}\,,\\
\mathsf{K}_{j+1} &:= \mathsf{K}_0 \prod_{k=0}^{j}(1+\frac{\sigma_k}{3}) \leqslant \mathsf{K}_0\, e^{\frac{2\sigma_0}{3}} \leqslant \mathsf{K}_0\sqrt{2}\,,\\
\mathsf{T}_{j+1} &:= \mathsf{T}_0 \prod_{k=0}^{j}(1+\frac{\sigma_k}{3}) \leqslant \mathsf{T}_0\, e^{\frac{2\sigma_0}{3}} \leqslant \mathsf{T}_0\sqrt{2}\,,\\
\lambda_0 &:= \log \mu_0^{-1}\,,\\
\mathsf{e}_* &:= \mathsf{C}_7\, \sigma_0^{-(4\nu+2d+1)} \lambda_0^{2\nu}\, \eta_0^2\,,\\
\mathsf{d}_* &:= 2^{2\nu+2d+1}\, \mathsf{C}_6^2\, \eta_0^2\,,\\
\kappa_0 &:= 4\sigma_0^{-1}\lambda_0\,,\\
\kappa_j &:= 4^j \kappa_0\,,\\
\widehat{\alpha} &:= \frac{\alpha}{\sqrt{|\varepsilon|}}\,,\\
\hat{r}_0 &:= \frac{r_0}{\sqrt{|\varepsilon|}}\,,\\
\hat{r}_{j+1} &:= \frac{1}{2}\min\left\{\frac{\widehat{\alpha}}{2d\sqrt{2}\mathsf{K}_0\kappa_j^{\nu}}\,,\ \frac{5}{48d}\frac{\widehat{r}_j}{\eta_0}\right\}\,,\\
r_{j+1} &:= \hat{r}_{j+1}\sqrt{|\varepsilon|}\,,\\
\check{r}_j &:= \frac{5}{48d}\frac{r_j}{\eta_0}\,,
\end{aligned}$$

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

$$\begin{aligned}
\widehat{M}_0 &:= M_0\,,\\
\widehat{M}_{j+1} &:= \mathsf{e}_*\mathsf{d}_*^{j-1}\frac{\mathsf{K}_0\widehat{M}_j^2}{\widehat{\alpha}^2}\,,\\
\mu_j &:= \frac{\mathsf{K}_0\widehat{M}_j}{\widehat{\alpha}^2}\,,\\
\theta_j &:= \mathsf{e}_*\ \mathsf{d}_*^j\ \mu_j\,,\\
\mathsf{W}_j &:= \ \mathrm{diag}\ \left(\max\left\{\frac{\mathsf{K_j}}{\widehat{\alpha}}\,,\frac{\sqrt{|\varepsilon|}}{r_j}\right\}\ \mathbb{1}_d\,,\sqrt{|\varepsilon|}\ \mathbb{1}_d\right)\,,\\
\mathsf{L}_j &:= M_i\max\left\{\frac{80d\sqrt{2}\ \mathsf{T}_0\ \eta_0}{r_j^2}\sigma_j^{-(\nu+d)}\,,\ \mathsf{C}_4\max\left\{1,\frac{\alpha}{r_j\mathsf{K}_j}\right\}\frac{\mathsf{K}_0}{\alpha^2}\sigma_j^{-2(\nu+d)}\right\}\\
&= M_j\max\left\{\frac{80d\sqrt{2}\ \mathsf{T}_0\ \eta_0}{r_j^2}\sigma_j^{-(\nu+d)}\,,\ \frac{4}{\mathsf{K}_jr_j^2}\,,\ \mathsf{C}_4\max\left\{1,\frac{\alpha}{r_j\mathsf{K}_j}\right\}\frac{\mathsf{K}_0}{\alpha^2}\sigma_j^{-2(\nu+d)}\right\}\,.
\end{aligned}$$

Thus, for any $j \geqslant 0$,

$$\theta_{j+1} = \mathsf{e}_*\ \mathsf{d}_*^{j+1}\ \mu_{j+1} = \mathsf{e}_*\ \mathsf{d}_*^{j+1}\frac{\mathsf{K}_0\widehat{M}_{j+1}}{\widehat{\alpha}^2} = \mathsf{e}_*\ \mathsf{d}_*^{j+1}\frac{\mathsf{K}_0}{\widehat{\alpha}^2}\ \mathsf{e}_*\mathsf{d}_*^{j-1}\frac{\mathsf{K}_0\widehat{M}_j^2}{\widehat{\alpha}^2} = \left(\mathsf{e}_*\ \mathsf{d}_*^j\ \mu_j\right)^2 = \theta_j^2$$

i.e.

$$\theta_j = \theta_0^{2^j}\ .$$

The very first step being quite different from all the others, it has to be done separately. Hence,

Lemma 5.3.2 *Under the above assumptions and notations, if*

$$|\varepsilon| \leqslant \left(\frac{r_0}{\widehat{\alpha}\mathsf{T}_0}\right)^2 \qquad \textit{and} \qquad \max\{\,e\ \mu_0\ ,\,\theta_0\} \leqslant 1\ , \tag{5.3.26}$$

then, there exist $y_1 \in \mathscr{D}$ and a real–analytic symplectomorphism

$$\phi_0 : D_{r_1,s_1}(y_1) \to D_{r_0,s_0}(y_0)\ , \tag{5.3.27}$$

such that, for $H_1 := H_0 \circ \phi_0$, we have

$$\begin{cases} H_1 =: K_1 + \varepsilon^2 P_1\ , \\ \partial_{y_1}K_1(y_1) = \omega\ , \quad \partial_{y_1}^2 K_1(y_1) \neq 0 \end{cases} \tag{5.3.28}$$

and

$$|y_1 - y_0| \leqslant \frac{8|\varepsilon|\mathsf{T}_0 M_0}{r_0}\,, \tag{5.3.29}$$

$$\|K_1\|_{\check{r}_1, y_1} \leqslant \mathsf{K}_1\,, \qquad \|T_1\| \leqslant \mathsf{T}_1\,, \qquad T_1 := \partial^2_{y_1} K_1(y_1)^{-1}\,, \tag{5.3.30}$$

$$\varepsilon^2 M_1 := \varepsilon^2 \|P_1\|_{r_1, s_1, y_1} \leqslant |\varepsilon| \widehat{M_1}\,, \tag{5.3.31}$$

$$\|\mathsf{W}_0(\phi_0 - \mathrm{id})\|_{r_1, s_1, y_1} \leqslant \sigma_0^d\, |\varepsilon| \mathsf{L}_0 \cdot \sqrt{|\varepsilon|}\,. \tag{5.3.32}$$

Proof By

$$\kappa_0 \overset{(5.3.26)}{\geqslant} 4\sigma_0^{-1} \geqslant 8 \tag{5.3.33}$$

and

$$\frac{\widehat{\alpha}}{2d\sqrt{2}\mathsf{K}_0 k_0^\nu} \overset{(5.3.33)}{\leqslant} \frac{1}{2d \cdot 8^\nu \sqrt{2}\mathsf{K}_0}\,\frac{r_0}{\mathsf{T}_0\sqrt{|\varepsilon|}} < \frac{\hat{r}_0}{4\mathsf{C}_5}\,,$$

we get

$$\hat{r}_1 = \frac{1}{2}\min\left\{\frac{\widehat{\alpha}}{2d\sqrt{2}\mathsf{K}_0\kappa_0^\nu}\,,\, \frac{\hat{r}_0}{4\mathsf{C}_5}\right\} = \frac{\widehat{\alpha}}{4d\sqrt{2}\mathsf{K}_0\kappa_0^\nu} \tag{5.3.34}$$

and, thus

$$\begin{aligned}
|\varepsilon|\mathsf{L}_0(3\sigma_0^{-1}) &\leqslant 3|\varepsilon|M_0 \max\left\{\frac{80d\sqrt{2}\,\mathsf{T}_0\,\eta_0}{r_0^2}\sigma_0^{-(\nu+d)}\,,\, \mathsf{C}_4\max\left\{1, \frac{\alpha}{r_0\mathsf{K}_0}\right\}\frac{\mathsf{K}_0}{\alpha^2}\sigma_0^{-2(\nu+d)}\right\}\sigma_0^{-1}\\
&\leqslant 3\max\left\{80d\sqrt{2}\,\eta_0\frac{\alpha\,\mathsf{T}_0}{r_0}\frac{\alpha}{r_0\mathsf{K}_0}\,,\, \mathsf{C}_4\max\left\{1, \frac{\alpha}{r_0\mathsf{K}_0}\right\}\right\}\sigma_0^{-2(\nu+d)-1}\frac{\mathsf{K}_0 M_0}{\widehat{\alpha}^2}\\
&\overset{(5.3.26)}{\leqslant} 3\max\left\{80d\sqrt{2}\,,\, \mathsf{C}_4\right\}\sigma_0^{-2(\nu+d)-1}\mu_0\,\eta_0\\
&\leqslant \mathsf{e}_*\,\mu_0\\
&= \theta_0 \overset{(5.3.26)}{\leqslant} 1.
\end{aligned} \tag{5.3.35}$$

Therefore, Lemma 5.3.2 is a straightforward consequence of Lemma 5.3.1. ∎

Once the first step is completed, all the following steps do not need any other condition. Actually, they are "completely" independent upon ε, and, therefore, the first condition in (5.3.26) is useless. To be precise, the following holds.

Lemma 5.3.3 *Assume* (5.3.28) ÷ (5.3.31) *with some* $\varepsilon \neq 0$ *and*

$$\max\left\{e\,\mu_0\,,\, \mathsf{C}_8\,\eta_0^{\frac{1}{8}}\,\theta_0\right\} < 1\,. \tag{5.3.36}$$

Then, one can construct a sequence of symplectic transformations

$$\phi_{j-1} : D_{r_j,s_j}(y_j) \to D_{r_{j-1},s_{j-1}}(y_{j-1}) \, , \qquad j \geqslant 2 \tag{5.3.37}$$

so that

$$H_j := H_{j-1} \circ \phi_{j-1} =: K_j + \varepsilon^{2^j} P_j \tag{5.3.38}$$

converges uniformly. More precisely, $\varepsilon^{2^{j-1}} P_{j-1}$*,* $\phi^{j-1} := \phi_1 \circ \phi_2 \circ \cdots \circ \phi_{j-1}$*,* K_{j-1}*,* y_{j-1} *converge uniformly on* $\{y_*\} \times \mathbb{T}^d_{s_*}$ *to, respectively,* 0*,* ϕ^**,* K_**,* y_* *which are real–analytic on* $\mathbb{T}^d_{s_*}$ *and* $H_1 \circ \phi^* = K_*$ *with* $\det \partial_y^2 K_*(y_*) \neq 0$*. Finally, the following estimates hold for any* $i \geqslant 1$*:*

$$|\varepsilon|^{2^i} M_i := |\varepsilon|^{2^i} \|P_i\|_{r_i,s_i,y_i} \leqslant |\varepsilon| \widehat{M_i} \ , \tag{5.3.39}$$

$$|y_{i+1} - y_i| \leqslant \frac{8\sqrt{2}\mathsf{T}_0 |\varepsilon|^{2^i} M_i}{r_i} \ , \tag{5.3.40}$$

$$|\mathsf{W}(\phi^* - \mathrm{id})| \leqslant \frac{\theta_0^2}{3 \cdot 2^{2d+1}\, \mathsf{d}_*} \sqrt{|\varepsilon|} \quad \text{on} \quad \{y_*\} \times \mathbb{T}^d_{s_*} \ . \tag{5.3.41}$$

Proof First of all, notice that, for any $i \geqslant 1$,

$$\begin{aligned}
\hat{r}_{i+1} &= \min\left\{ \frac{\widehat{\alpha}}{4d\sqrt{2}\mathsf{K}_0 \kappa_i^\nu} \, , \, \frac{5}{96d}\frac{\hat{r}_i}{\eta_0} \right\} \\
&= \min\left\{ \frac{\hat{r}_1}{4^{i\nu}} \, , \, \frac{5}{96d\eta_0}\hat{r}_i \right\} \\
&= \min\left\{ \frac{\hat{r}_1}{4^{\nu i}} \, , \, \frac{5}{96d\eta_0}\frac{\hat{r}_1}{4^{\nu(i-1)}} \, , \, \left(\frac{5}{96d\eta_0}\right)^2 \hat{r}_{i-1} \right\} \\
&\;\vdots \\
&= \min\left\{ \frac{\hat{r}_1}{4^{\nu i}} \, , \, \frac{5}{96d\eta_0}\frac{\hat{r}_1}{4^{\nu(i-1)}} \, , \cdots , \, \left(\frac{5}{96d\eta_0}\right)^i \hat{r}_1 \right\} \\
&= \frac{\hat{r}_1}{4^{\nu i}} \min\left\{ \left(\frac{5 \cdot 4^\nu}{96d\eta_0}\right)^0 , \cdots , \left(\frac{5 \cdot 4^\nu}{96d\eta_0}\right)^i \right\} \\
&= \frac{\hat{r}_1}{4^{\nu i}} \min{}^i \left\{ \frac{5 \cdot 4^\nu}{96d\eta_0} \, , \, 1 \right\} \\
&= \hat{r}_1 \min{}^i \left\{ \frac{1}{2^{2\nu}} \, , \, \frac{5}{96d\eta_0} \right\} \\
&= \frac{\hat{r}_1}{\mathbf{a}_1^i} \, ,
\end{aligned}$$

where

$$\mathbf{a}_1 := \max\left\{2^{2\nu}\,,\ \frac{96d\eta_0}{5}\right\} \leqslant \max\left\{2^{2\nu}\,,\ \frac{96d}{5}\right\}\cdot\eta_0 = \mathsf{C}_6\,\eta_0\ . \tag{5.3.42}$$

For a given $j \geqslant 2$, let $(\mathscr{P}^j)$ be the following assertion: there exist $j-1$ symplectic transformations[95]

$$\phi_i : D_{r_{i+1},s_{i+1}}(y_{i+1}) \to D_{2r_i/3,s_i}(y_i), \quad \text{for} \quad 1 \leqslant i \leqslant j-1, \tag{5.3.43}$$

and $j-1$ Hamiltonians $H_{i+1} = H_i \circ \phi_i = K_{i+1} + \varepsilon^{2^{i+1}} P_{i+1}$ real–analytic on $D_{r_{i+1},s_{i+1}}(y_{i+1})$ such that, for any $1 \leqslant i \leqslant j-1$,

$$\begin{cases} \|\partial_y^2 K_i\|_{r_i,y_i} \leqslant \mathsf{K}_i\,, \\ \|T_i\| \leqslant \mathsf{T}_i\,, \\ \partial_y K_i(y_i) = \omega\ , \quad \partial_y^2 K_i(y_i) \neq 0\,, \\ |\varepsilon|^{2^i}\|P_i\|_{r_i,s_i,y_i} \leqslant |\varepsilon|\widehat{M}_i\,, \\ \kappa_i \geqslant 4\sigma_i^{-1}\log\left(\sigma_i^{2\nu+d}\mu_i^{-1}\right)\,, \\ |\varepsilon|^{2^i}\mathsf{L}_i \leqslant \frac{\sigma_i}{3} \end{cases} \tag{5.3.44}$$

[95]Compare (5.3.8).

and

$$\begin{cases} \partial_y K_{i+1}(y_{i+1}) = \omega \ , \quad \partial_y^2 K_{i+1}(y_{i+1}) \neq 0 \ , \\ \\ |y_{i+1} - y_i| \leqslant \dfrac{8\sqrt{2}\mathsf{T}_0 |\varepsilon|^{2^i} M_i}{r_i} \ , \\ \\ \|T_{i+1}\| \leqslant \|T_i\| + \mathsf{T}_i |\varepsilon|^{2^i} \mathsf{L}_i \ , \\ \\ \|K_{i+1}\|_{r_{i+1},y_{i+1}} \leqslant \|K_i\|_{r_i,y_i} + |\varepsilon|^{2^i} M_i \ , \\ \\ \|\partial_y^2 K_{i+1}\|_{r_{i+1},y_{i+1}} \leqslant \|\partial_y^2 K_i\|_{r_i,y_i} + \mathsf{K}_i |\varepsilon|^{2^i} \mathsf{L}_i \ , \\ \\ \|\mathsf{W}_i(\phi_i - \mathrm{id})\|_{r_{i+1},s_{i+1},y_{i+1}} \leqslant \sigma_i^d \ |\varepsilon|^{2^i} \mathsf{L}_i \cdot \sqrt{|\varepsilon|} \ , \\ \\ M_{i+1} := \|P_{i+1}\|_{r_{i+1},s_{i+1},y_{i+1}} \leqslant M_i \mathsf{L}_i \ . \end{cases} \tag{5.3.45}$$

Assume ($\mathscr{P}^j$), for some $j \geqslant 2$ and let's check ($\mathscr{P}^{j+1}$). Fix then $1 \leqslant i \leqslant j-1$. Thus

$$\|\partial_y^2 K_{i+1}\|_{r_{i+1},y_{i+1}} \overset{(5.3.45)}{\leqslant} \|\partial_y^2 K_i\|_{r_i,y_i} + \mathsf{K}_i |\varepsilon|^{2^i} \mathsf{L}_i \overset{(5.3.44)}{\leqslant} \mathsf{K}_i + \mathsf{K}_i \frac{\sigma_i}{3} = \mathsf{K}_{i+1} < \mathsf{K}_0 \sqrt{2}$$

and, similarly,

$$\|T_{i+1}\| \leqslant \mathsf{T}_{i+1},$$

which prove the two first relations in (5.3.44) for $i = j$. Also

$$\frac{\alpha}{r_i \mathsf{K}_i} > \frac{\alpha}{r_1 \mathsf{K}_0 \sqrt{2}} = \frac{\widehat{\alpha}}{\hat{r}_1 \mathsf{K}_0 \sqrt{2}} = 4d\kappa_0^\nu \overset{(5.3.33)}{>} 1 \ , \tag{5.3.46}$$

so that

$$
\begin{aligned}
|\varepsilon|^{2^i}\mathsf{L}_i(3\sigma_i^{-1}) &= 3|\varepsilon|^{2^i}M_i\max\left\{\frac{80d\sqrt{2}\mathsf{T}_0\eta_0}{r_i^2}\sigma_i^{-(\nu+d)}\,,\ \mathsf{C}_4\max\left\{1,\frac{\alpha}{r_i\mathsf{K}_i}\right\}\frac{\mathsf{K}_0}{\alpha^2}\sigma_i^{-2(\nu+d)}\right\}\sigma_i^{-1}\\
&\overset{(5.3.46)}{\leqslant} 3|\varepsilon|^{2^i}M_i\max\left\{\frac{80d\sqrt{2}\mathsf{T}_0\eta_0}{r_i^2}\,,\ \mathsf{C}_4\frac{1}{\alpha r_i}\right\}\sigma_i^{-2(\nu+d)-1}\\
&= 3\max\left\{80d\sqrt{2}\mathsf{T}_0\eta_0\frac{\widehat{\alpha}}{\widehat{r}_i}\,,\ \mathsf{C}_4\right\}\sigma_i^{-2(\nu+d)-1}\frac{|\varepsilon|^{2^i}M_i}{|\varepsilon|\widehat{\alpha}\widehat{r}_i}\\
&= 3\max\left\{640d^2\eta_0^2\mathbf{a}^{i-1}\kappa_0^{\nu}\,,\ \mathsf{C}_4\right\}\sigma_i^{-2(\nu+d)-1}\frac{|\varepsilon|^{2^i}M_i}{|\varepsilon|\widehat{\alpha}^2}4d\sqrt{2}\mathsf{K}_0\kappa_0^{\nu}\mathbf{a}^{i-1}\\
&\overset{(5.3.33)}{\leqslant} 12d\sqrt{2}\max\left\{640d^2\,,\ \mathsf{C}_4\right\}\mathsf{K}_0\sigma_i^{-2(\nu+d)-1}\frac{|\varepsilon|^{2^i}M_i}{|\varepsilon|\widehat{\alpha}^2}\eta_0^2\mathbf{a}^{2(i-1)}\kappa_0^{2\nu}\\
&\overset{(5.3.42)}{\leqslant} 12d\sqrt{2}\max\left\{640d^2\,,\ \mathsf{C}_4\right\}\mathsf{K}_0\sigma_i^{-2(\nu+d)-1}\frac{|\varepsilon|^{2^i}M_i}{|\varepsilon|\widehat{\alpha}^2}\eta_0^{2i}\mathsf{C}_6^{2(i-1)}\kappa_0^{\nu}\\
&= 3d\cdot 2^{6\nu+2d+3}\sqrt{2}\max\left\{640d^2\,,\ \mathsf{C}_4\right\}\mathsf{K}_0\sigma_0^{-(4\nu+2d+1)}\left(2^{2\nu+2d+1}\mathsf{C}_6^2\eta_0^2\right)^{i-1}\frac{|\varepsilon|^{2^i}M_i}{|\varepsilon|\widehat{\alpha}^2}\left(\log\mu_0^{-1}\right)^{2\nu}\eta_0^2\\
&\overset{(5.3.44)}{\leqslant} \mathsf{C}_7\sigma_0^{-(4\nu+2d+1)}\left(\log\mu_0^{-1}\right)^{2\nu}\eta_0^2\,\mathsf{d}_*^{i-1}\frac{\mathsf{K}_0\widehat{M}_i}{\widehat{\alpha}^2}\\
&= \mathsf{e}_*\,\mathsf{d}_*^{i-1}\mu_i\\
&= \frac{\theta_i}{\mathsf{d}_*}\\
&= \frac{\theta_0^{2^i}}{\mathsf{d}_*}\\
&\overset{(5.3.36)}{\leqslant} \frac{1}{\mathsf{d}_*} < 1\ .
\end{aligned}
$$

Moreover,

$$
|\varepsilon|^{2^i}\mathsf{L}_i < \mathsf{e}_*\,\mathsf{d}_*^{i-1}\mu_i\ ,
$$

thus by last relation in (5.3.45), for any $1 \leqslant i \leqslant j-1$,

$$
|\varepsilon|^{2^{i+1}}M_{i+1} \leqslant |\varepsilon|^{2^i}\mathsf{L}_i\ |\varepsilon|^{2^i}M_i < \mathsf{e}_*\mathsf{d}_*^{i-1}\ \mu_i\ |\varepsilon|^{2^i}M_i \overset{(5.3.44)}{\leqslant} \mathsf{e}_*\mathsf{d}_*^{i-1}\ \mu_i\ |\varepsilon|\ \widehat{M}_i = |\varepsilon|\ \widehat{M}_{i+1}\ ,
$$

which proves the fourth relation in (5.3.44) for $i = j$. Hence, by exactly the same computation as above, one gets

$$
|\varepsilon|^{2^{i+1}}\mathsf{L}_{i+1}(3\sigma_{i+1}^{-1}) \leqslant \frac{\theta_{i+1}}{\mathsf{d}_*} = \frac{\theta_0^{2^{i+1}}}{\mathsf{d}_*} < 1\ ,
$$

which proves the last relation in (5.3.44) for $i = j$. It remains only to check that the fifth relation in (5.3.44) holds as well for $i = j$ in order to apply Lemma 5.3.1 to H_i, $1 \leqslant i \leqslant j$ and get (5.3.45) and, consequently, $(\mathscr{P}^{j+1})$. But in fact, we have[96]

$$\begin{aligned} 4\sigma_j^{-1}\log\left(\sigma_j^{2\nu+d}\mu_j^{-1}\right) &\leqslant 4\sigma_j^{-1}\log\left(\mu_j^{-1}\right) \\ &= 4\sigma_j^{-1}\log\left(\frac{1}{\mathsf{e}_*\,\mathsf{d}_*^j}\theta_0^{-2^j}\right) \\ &\leqslant 4\sigma_j^{-1}\log\left(\left(\frac{\theta_0}{\mathsf{e}_*}\right)^{-2^j}\right) \\ &= 4\sigma_j^{-1}\log\left(\mu_0^{-2^j}\right) \\ &= 4^j\cdot 4\sigma_0^{-1}\log\left(\mu_0^{-1}\right) \\ &= \kappa_j\ . \end{aligned}$$

To finish the proof of the induction *i.e.* one can construct an *infinite sequence* of Arnold's transformations satisfying (5.3.44) and (5.3.45) *for all* $i \geqslant 1$, one needs only to check $(\mathscr{P}^2)$. Thanks to[97] (5.3.28) $\div$ (5.3.31), we just need to check the two last inequalities in $(5.3.44)_{i=1}$. But, in fact, one proves those two relations by exactly the same computation as above. Then, we apply Lemma 5.3.1 to H_1 to get $(5.3.43)_{i=1}$ and $(5.3.45)_{i=1}$, which achieves the proof of $(\mathscr{P}^2)$.
Next, we prove that ϕ^j is convergent by proving that it is Cauchy. For any $j \geqslant 3$, we have,

[96] Notice that, since $\sigma_0 < 1$ then $\mathsf{e}_* \overset{(5.3.36)}{\geqslant} \mathsf{C}_7 > 1$.
[97] Observe that for $j = 2$, $i = 1$.

using again Cauchy's estimate,[98]

$$\begin{aligned}
\|\mathsf{W}_{j-1}(\phi^{j-1}-\phi^{j-2})\|_{r_j,s_j,y_j} &= \|\mathsf{W}_{j-1}\phi^{j-2}\circ\phi_{j-1}-\mathsf{W}_{j-1}\phi^{j-2}\|_{r_j,s_j,y_j}\\
&\overset{(5.3.43)}{\leqslant} \|\mathsf{W}_{j-1}D\phi^{j-2}\mathsf{W}_{j-1}^{-1}\|_{2r_{j-1}/3,s_{j-1},y_{j-1}}\,\|\mathsf{W}_{j-1}(\phi_{j-1}-\mathrm{id})\|_{r_j,s_j,y_j}\\
&\overset{(5.3.45)}{\leqslant} \max\left(r_{j-1}\frac{3}{r_{j-1}},\frac{3}{2\sigma_{j-1}}\right)\|\mathsf{W}_{j-1}\phi^{j-2}\|_{r_{j-1},s_{j-1},y_{j-1}}\times\\
&\qquad\times\|\mathsf{W}_{j-1}(\phi_{j-1}-\mathrm{id})\|_{r_j,s_j,y_j}\\
&= \frac{3}{2\sigma_{j-1}}\|\mathsf{W}_{j-1}\phi^{j-2}\|_{r_{j-1},s_{j-1},y_{j-1}}\,\|\mathsf{W}_{j-1}(\phi_{j-1}-\mathrm{id})\|_{r_j,s_j,y_j}\\
&\leqslant \frac{1}{2}\|\mathsf{W}_{j-1}\phi^{j-2}\|_{r_{j-1},s_{j-1},y_{j-1}}\cdot\sigma_{j-1}^d\left(|\varepsilon|^{2^{j-1}}\mathsf{L}_{j-1}3\sigma_{i-1}^{-1}\right)\sqrt{|\varepsilon|}\\
&\leqslant \frac{1}{2}\|\mathsf{W}_{j-1}\phi_1\|_{r_2,s_2,y_2}\cdot\sigma_{j-1}^d\,\theta_{j-1}\cdot\sqrt{|\varepsilon|}\\
&\leqslant \frac{1}{2}\left(\prod_{i=1}^{j-2}\|\mathsf{W}_{i+1}\mathsf{W}_i^{-1}\|\right)\|\mathsf{W}_1\phi_1\|_{r_2,s_2,y_2}\cdot\sigma_{j-1}^d\,\theta_{j-1}\cdot\sqrt{|\varepsilon|}\\
&\overset{(5.3.46)}{=} \frac{1}{2}\left(\prod_{i=1}^{j-2}\frac{r_i}{r_{i+1}}\right)\|\mathsf{W}_1\phi_1\|_{r_2,s_2,y_2}\cdot\sigma_{j-1}^d\,\theta_{j-1}\cdot\sqrt{|\varepsilon|}\\
&= \frac{r_1}{2r_{j-1}}\|\mathsf{W}_1\phi_1\|_{r_2,s_2,y_2}\cdot\sigma_{j-1}^d\,\theta_{j-1}\cdot\sqrt{|\varepsilon|}\\
&\overset{(5.3.42)}{\leqslant} \frac{1}{2}\sigma_2^d\,\mathsf{C}_6\,\eta_0\,\|\mathsf{W}_1\phi_1\|_{r_2,s_2,y_2}\cdot\left(2^{-d}\mathsf{C}_6\,\eta_0\right)^{j-3}\cdot\,\theta_0^{2^{j-1}}\cdot\sqrt{|\varepsilon|}\\
&\leqslant \frac{1}{2}\sigma_2^d\,\mathsf{C}_6\,\eta_0\,\|\mathsf{W}_1\phi_1\|_{r_2,s_2,y_2}\cdot\left(2^{-d}\mathsf{C}_6\,\eta_0\right)^{2^{j-4}}\cdot\,\theta_0^{2^{j-1}}\cdot\sqrt{|\varepsilon|}\\
&= \frac{1}{2}\sigma_2^d\,\mathsf{C}_6\,\eta_0\,\|\mathsf{W}_1\phi_1\|_{r_2,s_2,y_2}\cdot\left(\left(2^{-d}\mathsf{C}_6\,\eta_0\right)^{\frac{1}{8}}\theta_0\right)^{2^{j-1}}\cdot\sqrt{|\varepsilon|}\\
&= \frac{1}{2}\sigma_2^d\,\mathsf{C}_6\,\eta_0\,\|\mathsf{W}_1\phi_1\|_{r_2,s_2,y_2}\cdot\left(\mathsf{C}_8\,\eta_0^{\frac{1}{8}}\,\theta_0\right)^{2^{j-1}}\cdot\sqrt{|\varepsilon|}\,.
\end{aligned}$$

[98] Recall that $2^{i-1}\geqslant i$, $\forall\, i\geqslant 0$.

Therefore, for any $n \geqslant 1$, $j \geqslant 0$,

$$\begin{aligned}
\|\mathsf{W}_1(\phi^{n+j+1}-\phi^n)\|_{r_{n+j+2},s_{n+j+2},y_{n+j+2}} &\leqslant \sum_{i=n}^{n+j} \|\mathsf{W}_1(\phi^{i+1}-\phi^i)\|_{r_{i+2},s_{i+2},y_{i+2}} \\
&\leqslant \sum_{i=n}^{n+j}\left(\prod_{k=1}^{i}\|\mathsf{W}_k\mathsf{W}_{k+1}^{-1}\|\right)\|\mathsf{W}_{i+1}(\phi^{i+1}-\phi^i)\|_{r_{i+2},s_{i+2},y_{i+2}} \\
&\overset{(5.3.46)}{=} \sum_{i=n}^{n+j}\prod_{k=1}^{i}\max\left\{1\ ,\ \frac{r_{k+1}}{r_k}\right\}\|\mathsf{W}_{i+1}(\phi^{i+1}-\phi^i)\|_{r_{i+2},s_{i+2},y_{i+2}} \\
&= \sum_{i=n}^{n+j}\|\mathsf{W}_{i+1}(\phi^{i+1}-\phi^i)\|_{r_{i+2},s_{i+2},y_{i+2}} \\
&\leqslant \frac{1}{2}\sigma_2^d\,\mathsf{C}_6\,\eta_0\,\|\mathsf{W}_1\phi_1\|_{r_2,s_2,y_2}\cdot\sqrt{|\varepsilon|}\sum_{i=n}^{n+j}\left(\mathsf{C}_8\,\eta_0^{\frac{1}{8}}\,\theta_0\right)^{2^{i+1}} .
\end{aligned}$$

Hence, by (5.3.36), ϕ^j converges uniformly on $\{y_*\}\times\mathbb{T}^d_{s_*}$ to some ϕ^*, which is then real–analytic map in $x\in\mathbb{T}^d_{s_*}$.
To estimate $|\mathsf{W}_0(\phi^*-\mathrm{id})|$ on $\{y_*\}\times\mathbb{T}^d_{s_*}$, observe that , for $i\geqslant 1$,[99]

$$\sigma_i^d\,|\varepsilon|^{2^i}\mathsf{L}_i \leqslant \frac{\sigma_0^{d+1}}{3\cdot 2^{i(d+1)}}\,\frac{\theta_0^{2^i}}{\mathsf{d}_*} \leqslant \frac{1}{3\cdot 2^{(d+1)(i+1)}\mathsf{d}_*}\theta_0^{i+1} = \frac{1}{3\mathsf{d}_*}\left(\frac{\theta_0}{2^{d+1}}\right)^{i+1}$$

and therefore

$$\sum_{i\geqslant 1}\sigma_i^d\,|\varepsilon|^{2^i}\mathsf{L}_i \leqslant \frac{1}{3\mathsf{d}_*}\sum_{i\geqslant 1}\left(\frac{\theta_0}{2^{d+1}}\right)^{i+1} \leqslant \frac{\theta_0^2}{3\cdot 2^{2d+1}\,\mathsf{d}_*} .$$

Moreover, for any $i\geqslant 1$,

$$\begin{aligned}
\|\mathsf{W}_1(\phi^i-\mathrm{id})\|_{r_{i+1},s_{i+1},y_{i+1}} &\leqslant \|\mathsf{W}_1(\phi^{i-1}\circ\phi_i-\phi_i)\|_{r_{i+1},s_{i+1},y_{i+1}} + \|\mathsf{W}_1(\phi_i-\mathrm{id})\|_{r_{i+1},s_{i+1},y_{i+1}} \\
&\leqslant \|\mathsf{W}_1(\phi^{i-1}-\mathrm{id})\|_{r_i,s_i,y_i} + \left(\prod_{j=0}^{i-1}\|\mathsf{W}_j\mathsf{W}_{j+1}^{-1}\|\right)\|\mathsf{W}_i(\phi_i-\mathrm{id})\|_{r_{i+1},s_{i+1},y_{i+1}} \\
&= \|\mathsf{W}_1(\phi^{i-1}-\mathrm{id})\|_{r_i,s_i,y_i} + \|\mathsf{W}_i(\phi_i-\mathrm{id})\|_{r_{i+1},s_{i+1},y_{i+1}} \\
&= \|\mathsf{W}_1(\phi^{i-1}-\mathrm{id})\|_{r_i,s_i,y_i} + \|\mathsf{W}_i(\phi_i-\mathrm{id})\|_{r_{i+1},s_{i+1},y_{i+1}} \\
&\leqslant \|\mathsf{W}_1(\phi^{i-1}-\mathrm{id})\|_{r_i,s_i,y_i} + \sigma_i^d\,|\varepsilon|^{2^i}\mathsf{L}_i\sqrt{|\varepsilon|}\ ,
\end{aligned}$$

[99] Recall that $2^i \geqslant i+1$, $\forall\, i\geqslant 0$ and $\sigma_0\leqslant\frac{1}{2}$.

which iterated yields

$$\begin{aligned}\|\mathsf{W}_1(\phi^i - \mathrm{id})\|_{r_i,s_i,y_i} &\leqslant \sqrt{|\varepsilon|}\sum_{k=1}^{i-1}\sigma_k^d\,|\varepsilon|^{2^k}\mathsf{L}_k\\ &\leqslant \sqrt{|\varepsilon|}\sum_{k\geqslant 1}\sigma_k^d\,|\varepsilon|^{2^k}\mathsf{L}_k\\ &\leqslant \frac{\theta_0^2}{3\cdot 2^{2d+1}\,\mathsf{d}_*}\sqrt{|\varepsilon|}\,.\end{aligned}$$

Therefore, taking the limit over i completes the proof of (5.3.41), Lemma 5.3.3.
Now, to complete the proof of the Theorem, just set $\phi_* := \phi_0\circ\phi^*$ and observe that, uniformely on $\{y_*\}\times\mathbb{T}^d_{s_*}$,

$$\begin{aligned}|\mathsf{W}_0(\phi_* - \mathrm{id})| &\leqslant |\mathsf{W}_0(\phi_0\circ\phi^* - \phi^*)| + |\mathsf{W}_0(\phi^* - \mathrm{id})|\\ &\leqslant \|\mathsf{W}_0(\phi_0 - \mathrm{id})\|_{r_1,s_1,y_1} + \|\mathsf{W}_0\mathsf{W}_1^{-1}\|\;|\mathsf{W}_1(\phi^* - \mathrm{id})|\\ &\leqslant \sigma_0^d\,|\varepsilon|\mathsf{L}_0\sqrt{|\varepsilon|} + \frac{\theta_0^2}{3\cdot 2^{2d+1}\,\mathsf{d}_*}\sqrt{|\varepsilon|}\\ &\leqslant \left(\frac{\sigma_0^{d+1}}{3}\,\theta_0 + \frac{\theta_0^2}{3\cdot 2^{2d+1}\,\mathsf{d}_*}\right)\sqrt{|\varepsilon|}\\ &\leqslant \left(\frac{1}{3\cdot 2^{d+1}}\,\theta_0 + \frac{\theta_0^2}{3\cdot 2^{2d+1}\,\mathsf{d}_*}\right)\sqrt{|\varepsilon|}\\ &\leqslant \frac{\theta_0}{3\cdot 2^d}\sqrt{|\varepsilon|}\,.\end{aligned}$$

Moreover, for any $i \geqslant 1$,

$$
\begin{aligned}
|y_i - y_0| &\leqslant \sum_{j=0}^{i-1} |y_{j-1} - y_j| \\
&\overset{(5.3.40)}{\leqslant} 8\sqrt{2}\mathsf{T}_0 \sum_{j=0}^{i-1} \frac{|\varepsilon|^{2^i} M_i}{r_i} \\
&\overset{(5.3.39)}{\leqslant} \frac{8\sqrt{2}\mathsf{T}_0}{r_1} \sum_{j=0}^{\infty} \mathbf{a}_1^{i-1} |\varepsilon| \widehat{M}_i \\
&\overset{(5.3.42)}{\leqslant} \frac{64 d \eta_0 \kappa_0^{\nu} |\varepsilon|}{\alpha} \sum_{j=0}^{\infty} (\mathsf{C}_6\ \eta_0)^{i-1} \widehat{M}_i \\
&= \frac{64 d \kappa_0^{\nu}}{\mathsf{C}_6 \mathsf{e}_* \mathsf{K}_0} \frac{\widehat{\alpha}^2 |\varepsilon|}{\alpha} \sum_{j=0}^{\infty} \left(\frac{\mathsf{C}_6\ \eta_0}{\mathsf{d}_*} \right)^i \theta_0^{2^i} \\
&\leqslant \frac{64 d \kappa_0^{\nu}}{\mathsf{C}_6 \mathsf{e}_* \mathsf{K}_0} \alpha \sum_{j=0}^{\infty} \theta_0^{i+1} \\
&\leqslant \frac{64 d \kappa_0^{\nu}}{\mathsf{C}_6 \mathsf{e}_* \mathsf{K}_0} \alpha \cdot 2\theta_0 \\
&= \frac{2^{2\nu+7} d}{\mathsf{C}_6} \sigma_0^{-\nu} \mu_0 \lambda_0^{\nu} \frac{\alpha}{\mathsf{K}_0} \\
&\leqslant \frac{2^{2\nu+7} d}{\mathsf{C}_6} \sigma_0^{-\nu} \frac{1}{\mathsf{C}_7 \eta_0^2} \sigma_0^{4\nu+2d+1} \mathsf{e}_* \mu_0 \frac{\alpha}{\mathsf{K}_0} \\
&= \frac{2^{2\nu+7} d}{\mathsf{C}_6 \mathsf{C}_7} \sigma_0^{3\nu+2d+1} \theta_0 \frac{\alpha}{\mathsf{K}_0 \eta_0^2} \\
&\overset{(5.3.36)}{\leqslant} \frac{2^{2\nu+7} d}{\mathsf{C}_6 \mathsf{C}_7 \mathsf{C}_8} \sigma_0^{3\nu+2d+1} \frac{\alpha}{\mathsf{K}_0 \eta_0^{\frac{17}{8}}} \\
&= \frac{1}{\mathsf{C}_9} \sigma_0^{3\nu+2d+1} \frac{\alpha}{\mathsf{K}_0 \eta_0^{\frac{17}{8}}} ,
\end{aligned}
$$

and then passing to the limit, we get

$$
|y_* - y_0| \leqslant \frac{1}{\mathsf{C}_9} \sigma_0^{3\nu+2d+1} \frac{\alpha}{\mathsf{K}_0 \eta_0^{\frac{17}{8}}} .
$$

∎

Part II

“Sharp” measure estimates of Kolmogorov’s sets

6 | “Explicit” integrability on a Cantor–like set and a “sharp” measure estimate

6.1 Assumptions

Let $\tau \geqslant d-1 \geqslant 1$ and set[100]

$$\begin{aligned}
\nu &:= \tau + 1 \, , \\
\mathsf{C}_0 &:= 4\sqrt{2}\left(\frac{3}{2}\right)^{2\nu+d} \int_{\mathbb{R}^d} \left(|y|_1^{\nu} + d|y|_1^{2\nu}\right)\, e^{-|y|_1} dy \, , \\
\mathsf{C}_1 &:= 2\left(\frac{3}{2}\right)^{\nu+d} \int_{\mathbb{R}^d} |y|_1^{\nu}\, e^{-|y|_1} dy \, , \\
\mathsf{C}_2 &:= 2^{3d} d \, , \\
\mathsf{C}_3 &:= \left(d^2 \mathsf{C}_1^2 + 6d\mathsf{C}_1 + \mathsf{C}_2\right)\sqrt{2} \, , \\
\mathsf{C}_4 &:= \max\left\{\mathsf{C}_0, \, \mathsf{C}_3\right\} \, , \\
\mathsf{C}_5 &:= 3d^2 \cdot 2^{6\nu+2d+11} \max\left\{2^7 d\sqrt{2}\, , \, 8^{-\nu}\mathsf{C}_4\right\} \, , \\
\mathsf{C}_6 &:= 2^{\nu + \frac{3}{4}d + \frac{53}{8}} d^{\frac{5}{4}} \, , \\
\mathsf{C}_7 &:= 2\, e\, d\left(\frac{3}{2}\right)^{d-1} \, , \\
\mathsf{C}_8 &:= \frac{\mathsf{C}_5}{3 \cdot 2^d} \, , \\
\mathsf{C}_9 &:= 2^{3(\nu+1)}\, d\, \sqrt{2}\, \mathsf{C}_6 \, .
\end{aligned}$$

[100] Notice that each C_i is greater than 1 and depends only upon d and τ.

6.2 Statement of the extension Theorem

Theorem 6.2.1 *Under the assumptions and notations in §6.1, we have the following. Let $\mathscr{D} \subset \mathbb{R}^d$ be a non–empty, bounded domain.*[101] *Consider the Hamiltonian parametrized by $\varepsilon \in \mathbb{R}$*

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

where K, P are real–analytic functions defined on $\mathscr{D} \times \mathbb{T}^d$ with bounded holomorphic extensions to[102]

$$D_{\mathsf{r}_0,s_0}(\mathscr{D}) := \bigcup_{y_0 \in \mathscr{D}} D_{\mathsf{r}_0,s_0}(y_0)\,,$$

for some $\mathsf{r}_0 > 0$ and $0 < s_0 \leqslant 1$, the norm being

$$\|\cdot\|_{\mathsf{r}_0,s_0,\mathscr{D}} := \sup_{D_{\mathsf{r}_0,s_0}(\mathscr{D})} |\cdot|\,.$$

Let $\alpha > 0$, $\delta > 0$ and[103]

$$\begin{aligned}
\Delta_\alpha^\tau &:= \left\{\omega \in \mathbb{R}^d : \quad |\omega \cdot k| \geqslant \frac{\alpha}{|k|_1^\tau}\,, \quad \forall\, k \in \mathbb{Z}^d \backslash \{0\}\right\}\,,\\
\mathscr{D}_\delta &:= \{y \in \mathscr{D} : B_\delta(y) \subseteq \mathscr{D}\}\,,\\
\mathscr{D}_{\delta,\alpha} &:= \{y_0 \in \mathscr{D}_\delta : \; K_y(y_0) \in \Delta_\alpha^\tau\}\,.
\end{aligned}$$

Assume that

$$|\det K_{yy}(y)| \neq 0\,, \quad \forall\, y \in \mathscr{D}_{\delta,\alpha}\,. \tag{6.2.1}$$

Fix

$$0 \leqslant \sigma_0 < \min\left\{\frac{1}{2}\,, \frac{d}{2^{2\nu-7}}\right\} s_0$$

[101] *i.e.* open and connected.
[102] Recall the notations in §1.2
[103] Notice that $\mathscr{D}_\delta$ is closed, connected, with non–empty interior of $\mathscr{D}_\delta$ provided that δ is small enough.

and define[104]

$$
\begin{aligned}
&s_* := s_0 - \max\left\{2\,,\,\frac{2^{2\nu-7}}{d}\right\}\sigma_0\,,\\
&r_0 := \min\left\{\mathsf{r}_0,\, 32d\delta\right\}\,,\\
&M_0 := \|P\|_{\mathsf{r}_0,s_0,\mathscr{D}}\,,\\
&\mathsf{K}_0 := \|K_{yy}\|_{\mathsf{r}_0,\mathscr{D}}\,,\\
&\mathsf{T}_0 := \|T\|_{\mathscr{D}} := \sup_{y_0\in\mathscr{D}}\|T(y_0)\|\,,\\
&\eta_0 := \mathsf{T}_0\mathsf{K}_0\,,\\
&r_* := \frac{\sigma_0^\nu}{\mathsf{C}_9}\left(\frac{\sigma_0}{\eta_0}\right)^{\frac{5}{4}}\frac{\alpha}{\mathsf{K}_0}\,,\\
&\mu_* := \sup\left\{\mu \leqslant e^{-1} : 2\,\mathsf{C}_5\,\mathsf{C}_6\,\sigma_0^{4\nu+2d+\frac{13}{4}}\,\eta_0^{\frac{13}{4}}\,\mu\,\left(\log\mu^{-1}\right)^{2\nu}\leqslant 1\right\}\,,
\end{aligned}
$$

where $T(y) := K_{yy}(y)^{-1}$. *Finally, assume*

$$
\boxed{\alpha \leqslant \frac{r_0\sigma_0}{\mathsf{T}_0} \qquad \text{and} \qquad |\varepsilon| \leqslant \mu_*\frac{\alpha^2}{\mathsf{K}_0\,M_0}\,.} \tag{6.2.2}
$$

Then, there exist $\mathscr{D}_* \subset \mathscr{D}_{\delta-r_*}$ *having the same cardinality as* $\mathscr{D}_{\delta,\alpha}$*, a lipeomorphism* $G^*\colon \mathscr{D}_{\delta,\alpha} \xrightarrow{onto} \mathscr{D}_*$*, a function* $K_* \in C^\infty_W(\mathscr{D}_*,\mathbb{R})$ *and a* C^∞_W*–symplectic transformation*[105] $\phi_*\colon \mathscr{D}_*\times\mathbb{T}^d \to \mathscr{K} := \phi_*(\mathscr{D}_*\times\mathbb{T}^d) \subset \mathscr{D}\times\mathbb{T}^d$ *and real–analytic in* $x\in\mathbb{T}^d_{s_*}$*, such that*[106]

$$
\partial_{y_*}K_*\circ G^* = \partial_y K \qquad\qquad \text{on} \quad \mathscr{D}_{\delta,\alpha}\,, \tag{6.2.3}
$$

$$
\partial^\beta_{y_*}(H\circ\phi_*)(y_*,x) = \partial^\beta_{y_*}K_*(y_*), \qquad \forall\,(y_*,x)\in\mathscr{D}_*\times\mathbb{T}^d, \quad \forall\,\beta\in\mathbb{N}_0^d \tag{6.2.4}
$$

and

$$
\|G^* - \mathrm{id}\|_{\mathscr{D}_{\delta,\alpha}} \leqslant r_*\,, \tag{6.2.5}
$$

$$
\|G^* - \mathrm{id}\|_{L,\mathscr{D}_{\delta,\alpha}} \leqslant \frac{e\,\sigma_0^{\nu+d}}{\mathsf{C}_6}\,, \tag{6.2.6}
$$

$$
\begin{aligned}
\operatorname{meas}(\mathscr{D}\times\mathbb{T}^d\backslash\mathscr{K}) \leqslant \left(1+\frac{d\ e\,\sigma_0^{\nu+d}}{\mathsf{C}_6}\right)^d\Big(&\operatorname{meas}\left((B_{\delta\sigma_0}(\mathscr{D})\backslash\mathscr{D})\times\mathbb{T}^d\right)+\\
&+\operatorname{meas}\left((\mathscr{D}\backslash\mathscr{D}_\delta)\times\mathbb{T}^d\right)+\operatorname{meas}\left((\mathscr{D}_\delta\backslash\mathscr{D}_{\delta,\alpha})\times\mathbb{T}^d\right)\Big)\,.
\end{aligned} \tag{6.2.7}
$$

---

[104] Notice that $\eta_0 \geqslant 1$.

[105] Which means that the Whitney–gradient $\nabla\phi_* = \partial\phi_*/\partial(y_*,x)$ satisfies $(\nabla\phi_*)\mathbb{J}(\nabla\phi_*)^T = \mathbb{J}$ uniformly on $\mathscr{D}_*\times\mathbb{T}^d$, where $\mathbb{J} = \begin{pmatrix} 0 & -\mathbb{1}_d\\ \mathbb{1}_d & 0\end{pmatrix}$.

[106] Notice that the derivatives are taken in the sense of Whitney.

Now, by applying Theorem F.1 (see Appendix F) to (6.2.7), we get the following measure estimate of the unstable set $\mathscr{D}\times\mathbb{T}^d\backslash\mathscr{K}$.

Theorem 6.2.2 *Let the notations and assumptions in Theorem 6.2.1 hold, with*

$$0<\alpha\leqslant\frac{32d\sigma_0}{\mathsf{T}_0}\min\left\{\frac{\mathsf{r}_0}{32d},\ \frac{R(\mathscr{D})}{3},\ \mathrm{minfoc}\,(\partial\mathscr{D})\right\},\qquad |\varepsilon|\leqslant\mu_*\frac{\alpha^2}{\mathsf{K}_0\,M_0}\,. \tag{6.2.8}$$

in place of (6.2.2) *and*

$$\delta:=\frac{\alpha\mathsf{T}_0}{32d\sigma_0}\,,$$

where[107]

$$R(\mathscr{D}):=\sup\{R>0:B_R(y)\subseteq\mathscr{D}\,,\ \textit{for some } y\in\mathscr{D}\}\,.$$

Futhermore, assume that the boundary $\partial\mathscr{D}$ of $\mathscr{D}$ is a smooth hypersurface of $\mathbb{R}^d$. Then, the conclusions in Theorem 6.2.1 still hold. Moreover,

$$\mathrm{meas}\,(\mathscr{D}\times\mathbb{T}^d\backslash\mathscr{K})\leqslant(3\pi)^d\frac{\mathsf{T}_0}{32d\sigma_0}\left(2\mathcal{H}^{d-1}(\partial\mathscr{D})\,\alpha+C(d,\sigma_0,\mathsf{T}_0,\mathbf{R}^{\partial\mathscr{D}})\,\alpha^2+\mathrm{meas}\,(\mathscr{D}_\delta\backslash\mathscr{D}_{\delta,\alpha})\right), \tag{6.2.9}$$

where[108] $\mathbf{R}^{\partial\mathscr{D}}$ *denotes the curvature tensor of $\partial\mathscr{D}$, $\mathbf{k}_{2j}(\mathbf{R}^{\partial\mathscr{D}})$, the $(2j)$–th integrated mean curvature of $\partial\mathscr{D}$ in $\mathbb{R}^d$ and*

$$C(d,\sigma_0,\mathsf{T}_0,\mathbf{R}^{\partial\mathscr{D}}):=\frac{\mathsf{T}_0}{16d\sigma_0}\sum_{j=1}^{\left[\frac{d-1}{2}\right]}\frac{\mathbf{k}_{2j}(\mathbf{R}^{\partial\mathscr{D}})}{1\cdot3\cdots(2j+1)}\left(\frac{\alpha\mathsf{T}_0}{32d\sigma_0}\right)^{2j-1}\,.$$

Remark 6.2.3 **(i)** Notice that (6.2.7) is mainly a consequence of (6.2.5); the crucial part of the proof is that one can actually extend a Lipschitz continuous function to a global Lipschitz continuous function without increasing neither the Lipschitz constant nor the sup–norm (see Theorem C.1 in Appendix C).
(ii) The following estimates hold as well:

$$|\,\mathrm{meas}\,(\mathscr{D}_*)-\mathrm{meas}\,(\mathscr{D}_{\delta,\alpha})|\leqslant\frac{\mathsf{C}_7}{2\mathsf{C}_6}\,\sigma_0^{\nu+d+\frac54}\,\eta_0^{-\frac54}\,\mathrm{meas}\,(\mathscr{D}_{\delta,\alpha})\,, \tag{6.2.10}$$

$$|\mathsf{W}_0(\phi_*-\mathrm{id})|\leqslant\frac{1}{3\cdot2^{d+1}\mathsf{C}_6}\left(\frac{\sigma_0}{\eta_0}\right)^{\frac54}\qquad\text{on}\quad\mathscr{D}_*\times\mathbb{T}^d_{s_*}\,, \tag{6.2.11}$$

[107]Notice that the first condition in (6.2.8) then reads $0<\delta\leqslant\min\left\{\frac{\mathsf{r}_0}{32d},\frac{R(\mathscr{D})}{3},\mathrm{minfoc}\,(\partial\mathscr{D})\right\}$. The condition $\delta\leqslant\frac{R(\mathscr{D})}{3}$ ensures that the interior of $\mathscr{D}_\delta$ is non–empty.
[108]See Appendix F for the definitions.

where

$$\mathsf{W}_0 := \operatorname{diag}\left(\frac{\mathsf{K}_0}{\alpha}\mathbb{1}_d, \mathbb{1}_d\right) .$$

Notice that the constant in (6.2.10) is of order 1 and not α; that is why we need Minty's Theorem (see **(i)** above).

(iii) Notice that the Theorem is consistent for $\sigma_0 = 0$. In fact, in that case

$$\varepsilon = \alpha \overset{(6.2.2)}{=} 0.$$

Hence, the Hamiltonian H is integrable. Moreover,

$$\mathscr{D}_{\delta,\alpha} = \mathscr{D}_\delta \ , \qquad G^* \overset{(6.2.5)}{=} \mathrm{id} \ , \qquad \phi_* \overset{(6.2.11)}{=} \mathrm{id} \ .$$

Thus,

$$\mathscr{D}_* = G^*(\mathscr{D}_{\delta,\alpha}) = \mathscr{D}_{\delta,\alpha} = \mathscr{D}_\delta \ .$$

Therefore, we get $\mathscr{K} = \phi_*(\mathscr{D}_* \times \mathbb{T}^d) = \mathscr{D}_\delta \times \mathbb{T}^d$, for any $\delta > 0$, as expected.

6.3 Proof of Theorem 6.2.1

Lemma 6.3.1 (KAM step) *Let $r > 0$, $0 < 2\sigma < s \leqslant 1$ and consider the Hamiltonian parametrized by $\varepsilon \in \mathbb{R}$*

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

where K, P are real–analytic functions with bounded holomorphic extensions to $D_{r,s}(\mathscr{D}_\sharp)$. Assume that[109]

$$\begin{aligned} &\det K_{yy}(y) \neq 0 \ , && T(y) := K_{yy}(y)^{-1} \ , \quad \forall\, y \in \mathscr{D}_\sharp \ , \\ &\|K_{yy}\|_{r,\mathscr{D}_\sharp} \leqslant \mathsf{K} \ , && \|T\|_{\mathscr{D}_\sharp} \leqslant \mathsf{T} \ , \\ &\|P\|_{r,s,\mathscr{D}_\sharp} \leqslant M \ , && K_y(\mathscr{D}_\sharp) \subset \Delta_\alpha^\tau \ . \end{aligned} \tag{6.3.1}$$

Fix $\varepsilon \neq 0$ and assume that

$$\lambda \geqslant \log\left(\sigma^{2\nu+d}\frac{\alpha^2}{|\varepsilon| M \mathsf{K}}\right) \geqslant 1 \ . \tag{6.3.2}$$

Let

$$\begin{aligned} &\kappa := 4\sigma^{-1}\lambda, \quad \bar{r} \leqslant \min\left\{\frac{\alpha}{2d\mathsf{K}\kappa^\nu} \ , \ \frac{r\sigma}{16d\mathsf{T}\mathsf{K}}\right\} , \\ &\tilde{r} := \frac{\bar{r}}{16d\mathsf{T}\mathsf{K}}, \qquad \bar{s} := s - \frac{2}{3}\sigma, \quad s' := s - \sigma \ , \end{aligned} \tag{6.3.3}$$

[109] 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.

and[110]

$$\begin{aligned}\overline{\mathsf{L}} &:= \frac{\mathsf{C}_0}{\sqrt{2}}\max\left\{1,\frac{\alpha}{r\mathsf{K}}\right\}\frac{\mathsf{K}_0 M}{\alpha^2}\sigma^{-(2\nu+d)}\,,\\ \mathsf{L} &:= M\max\left\{\frac{16\mathsf{T}}{r\bar r}\sigma^{-(\nu+d)}\,,\,\frac{\mathsf{C}_4}{\sqrt{2}}\max\left\{1,\frac{\alpha}{r\mathsf{K}}\right\}\frac{\mathsf{K}}{\alpha^2}\sigma^{-2(\nu+d)}\right\}\\ &= M\max\left\{\frac{16\mathsf{T}}{r\bar r}\sigma^{-(\nu+d)}\,,\,\frac{4}{\mathsf{K}r^2}\,,\,\frac{\mathsf{C}_4}{\sqrt{2}}\max\left\{1,\frac{\alpha}{r\mathsf{K}}\right\}\frac{\mathsf{K}}{\alpha^2}\sigma^{-2(\nu+d)}\right\}\,.\end{aligned}$$

Then, there exists a generating function $(y',x)\mapsto y'\cdot x+g(y',x)$*, with* $g\in\mathcal{B}_{\bar r,\bar s}(\mathscr{D}_\sharp)$ *and satisfying the following inequalities:*

$$\begin{cases}\|g_x\|_{\bar r,\bar s,\mathscr{D}_\sharp}\leqslant \mathsf{C}_1\dfrac{M}{\alpha}\sigma^{-(\nu+d)}\,,\\ \|g_{y'}\|_{\bar r,\bar s,\mathscr{D}_\sharp},\ \|\partial^2_{y'x}g\|_{\bar r,\bar s,\mathscr{D}_\sharp}\leqslant\overline{\mathsf{L}}\,,\\ \|\partial^2_{y'}\widetilde K\|_{\bar r,\mathscr{D}_\sharp}\leqslant\mathsf{KL}\,,\end{cases}\tag{6.3.4}$$

where

$$\widetilde K(y'):=\langle P(y',\cdot)\rangle\,.$$

If, in addition,

$$|\varepsilon|\mathsf{L}\leqslant\frac{\sigma}{3}\,,\tag{6.3.5}$$

then, there exists a diffeomorphism $G\colon D_{\tilde r}(\mathscr{D}_\sharp)\to G(D_{\tilde r}(\mathscr{D}_\sharp))$*, such that ,*

$$\begin{cases}\partial_{y'}K'\circ G=\partial_y K\,, & (\partial^2_{y'}K')\circ G\neq 0\quad \textit{on } \mathscr{D}_\sharp\,,\\ |\varepsilon|\|g_x\|_{\bar r,\bar s,\mathscr{D}_\sharp}\leqslant\dfrac{r}{3}\,, & \|G-\mathrm{id}\|_{\tilde r,\mathscr{D}_\sharp}\leqslant\sigma^{\nu+d}\bar r|\varepsilon|\mathsf{L}\,,\\ |\varepsilon|\|\widetilde T\|_{\mathscr{D}'_\sharp}\leqslant\mathsf{T}|\varepsilon|\mathsf{L}\,, & \|\partial_z G-\mathbb{1}_d\|_{\tilde r,\mathscr{D}_\sharp}\leqslant\sigma^{\nu+d}|\varepsilon|\mathsf{L}\,,\\ \|P_+\|_{\bar r,\bar s,\mathscr{D}_\sharp}\leqslant\mathsf{L}M\,, & B_{\bar r/2}(\mathscr{D}'_\sharp)\subset B_{\bar r}(\mathscr{D}_\sharp)\subset\mathscr{D}\,,\end{cases}\tag{6.3.6}$$

where

$$\begin{aligned}&\mathscr{D}'_\sharp:=G(\mathscr{D}_\sharp)\,, && K':=K+\varepsilon\widetilde K\,,\\ &\left(\partial^2_{y'}K'(\mathsf{y}')\right)^{-1}=:T\circ G^{-1}(\mathsf{y}')+\varepsilon\,\widetilde T(\mathsf{y}')\,, && \forall\ \mathsf{y}'\in\mathscr{D}'_\sharp\\ &P_+(y',x):=P(y'+\varepsilon g_x(y',x),x)\,.\end{aligned}$$

and the following hold. For any $y'\in D_{\tilde r}(\mathscr{D}_\sharp)$*, the map* $\psi_\varepsilon(x):=x+\varepsilon g_{y'}(y',x)$ *has an analytic inverse* $\varphi(x')=x'+\varepsilon\widetilde\varphi(y',x';\varepsilon)$ *such that*

$$\|\widetilde\varphi\|_{\bar r,s',\mathscr{D}_\sharp}\leqslant\overline{\mathsf{L}}\qquad\text{and}\quad\varphi=\mathrm{id}+\varepsilon\widetilde\varphi: D_{\bar r/2,s'}(\mathscr{D}'_\sharp)\to\mathbb{T}^d_{\bar s}\,;\tag{6.3.7}$$

[110] Notice that $\mathsf{L}\geqslant\sigma^{-d}\overline{\mathsf{L}}\geqslant\overline{\mathsf{L}}$ since $\sigma\leqslant 1$. Notice also that $\mathsf{TK}\geqslant 1$, so that $\frac{16\mathsf{T}}{r\bar r}\sigma^{-(\nu+d)}\geqslant\frac{16\mathsf{T}}{r^2}\geqslant\frac{4}{\mathsf{K}r^2}$.

for any $y_0 \in \mathscr{D}_\sharp$ *and* $(y', x) \in D_{\bar r, \bar s}(y_0)$, $|y' + \varepsilon g_x(y', x) - y_0| < \frac{2}{3}r$*; the map* ϕ' *is a symplectic diffeomorphism and*

$$\phi' = \big(y' + \varepsilon g_x(y', \varphi(y', x')), \varphi(y', x')\big) : D_{\bar r/2, s'}(\mathscr{D}'_\sharp) \to D_{2r/3, \bar s}(\mathscr{D}_\sharp), \tag{6.3.8}$$

with

$$\|\mathsf{W}\,\tilde\phi\|_{\bar r/2, s', \mathscr{D}'_\sharp} \leqslant \sigma^d \mathsf{L}\,, \tag{6.3.9}$$

where $\tilde\phi$ *is defined by the relation* $\phi' =: \mathrm{id} + \varepsilon\tilde\phi$*,*

$$\mathsf{W} := \begin{pmatrix} \max\{\frac{\mathsf{K}}{\alpha}\,, \frac{1}{r}\}\ \mathbb{1}_d & 0 \\ \\ 0 & \mathbb{1}_d \end{pmatrix}$$

and

$$\|P'\|_{\bar r/2, s', \mathscr{D}'_\sharp} \leqslant \mathsf{L}M\ , \tag{6.3.10}$$

with

$$P'(y', x') := P_+(y', \varphi(x')) = P \circ \phi'(y', x')\ .$$

Proof

Step 1: Construction of the Arnold's transformation We seek for $r_1 < r/2$, $s_1 < s$, a set $\mathscr{D}'_\sharp \subset D_{2r_1}(\mathscr{D}_\sharp)$ having the same cardinality as $\mathscr{D}_\sharp$ and a near–to–the–identity real–analytic symplectic transformation $\phi_1 : \mathscr{D} \times \mathbb{T}^d \circlearrowleft$ satisfying

$$\phi' \colon D_{r_1, s_1}(\mathscr{D}'_\sharp) \to D_{r,s}(\mathscr{D}_\sharp),$$

with $D_{r_1, s_1}(\mathscr{D}'_\sharp) \subset D_{r,s}(\mathscr{D}_\sharp)$ and ϕ' generated by $y' \cdot x + \varepsilon g(y', x)$ *i.e.*

$$\phi' \colon \begin{cases} y = y' + \varepsilon\ g_x(y', x) \\ x' = x + \varepsilon g_{y'}(y', x)\,, \end{cases} \tag{6.3.11}$$

such that

$$\begin{cases} H' := H \circ \phi' = K' + \varepsilon^2 P' & \text{on } D_{r_1, s_1}(\mathscr{D}'_\sharp)\,, \\ \det \partial^2_{y'} K'(y') \neq 0\,, & \forall\ y' \in \mathscr{D}'_\sharp\,, \\ \partial_{y'} K'(\mathscr{D}'_\sharp) = \partial_y K(\mathscr{D}_\sharp)\,. \end{cases} \tag{6.3.12}$$

By Taylor's formula, we get[111]

$$\begin{aligned} H(y' + \varepsilon g_x(y', x), x) =& K(y') + \varepsilon\widetilde K(y') + \varepsilon\left[K_y(y') \cdot g_x + T_\kappa P(y', \cdot) - \widetilde K(y')\right] + \\ &+ \varepsilon^2\left(P^{(1)} + P^{(2)} + P^{(3)}\right)(y', x) \\ =& K'(y') + \varepsilon\left[K_y(y') \cdot g_x + T_\kappa P(y', \cdot) - \widetilde K(y')\right] + \varepsilon^2 P'(y', x), \end{aligned} \tag{6.3.13}$$

[111]Recall that $\langle\cdot\rangle$ stands for the average over $\mathbb{T}^d$.

with $\kappa \in \mathbb{N}$, which will be chosen large enough so that $P^{(3)} = O(\varepsilon)$ and

$$\begin{cases} P_+ := P^{(1)} + P^{(2)} + P^{(3)} \\ P^{(1)} := \frac{1}{\varepsilon^2}\left[K(y' + \varepsilon g_x) - K(y') - \varepsilon K_y(y') \cdot g_x\right] = \int_0^1 (1-t) K_{yy}(\varepsilon t g_x) \cdot g_x \cdot g_x dt \\ P^{(2)} := \frac{1}{\varepsilon}\left[P(y' + \varepsilon g_x, x) - P(y', x)\right] = \int_0^1 P_y(y' + \varepsilon t g_x, x) \cdot g_x dt \\ P^{(3)} := \frac{1}{\varepsilon}\left[P(y', x) - T_\kappa P(y', \cdot)\right] = \frac{1}{\varepsilon} \sum_{|n|_1 > \kappa} P_n(y')\, e^{in\cdot x}\,. \end{cases} \tag{6.3.14}$$

By the non–degeneracy condition in (4.1.1) and Lemma 2.2.7, for ε small enough (to be made precised below), there exists $\bar{r} \leqslant r$ such that for each $\mathsf{y} \in \mathscr{D}_\sharp$, there exists a unique $\mathsf{y}' \in D_{\bar{r}}(\mathsf{y})$ satisfying $\partial_{y'} K'(\mathsf{y}') = \partial_y K(\mathsf{y})$ and $\det \partial^2_{y'} K'(\mathsf{y}') \neq 0$; $\mathscr{D}'_\sharp$ is precisely the set of these y' when y runs in $\mathscr{D}_\sharp$. More precisely, $\mathscr{D}'_\sharp$ and $\mathscr{D}_\sharp$ are "diffeomorphic"[112], say via G, and, for each $\mathsf{y}' \in \mathscr{D}'_\sharp$, the matrix $\partial^2_{y'} K_1(y_1)$ is invertible with inverse of the form

$$T'(\mathsf{y}') := \partial^2_{y'} K'(\mathsf{y}')^{-1} =: T(y_0) + \varepsilon \widetilde{T}(\mathsf{y}'), \quad \mathsf{y}' = G(\mathsf{y}).$$

In view of (6.3.13), in order to get the first part of (6.3.12), we need to find g such that $K_y(y') \cdot g_x + T_\kappa P(y', \cdot) - \widetilde{K}(y')$ vanishes; such a g is indeed given by

$$g(y', x) := \sum_{0 < |n|_1 \leqslant \kappa} \frac{-P_n(y')}{i K_y(y') \cdot n}\, e^{in\cdot x}, \tag{6.3.15}$$

provided that

$$K_y(y') \cdot n \neq 0, \quad \forall\, 0 < |n|_1 \leqslant \kappa, \quad \forall\, y' \in D_{r_1}(\mathscr{D}'_\sharp) \quad (\subset D_r(\mathscr{D}_\sharp))\,. \tag{6.3.16}$$

But, in fact, since $K_y(\mathsf{y})$ is rationally independent, for each $\mathsf{y} \in \mathscr{D}_\sharp$, then, given any $\kappa \in \mathbb{N}$, there exists $r' \leqslant r$ such that

$$K_y(y') \cdot n \neq 0, \quad \forall\, 0 < |n|_1 \leqslant \kappa, \quad \forall\, y' \in D_{r'}(\mathscr{D}_\sharp). \tag{6.3.17}$$

Then we invert the function $x \mapsto x + \varepsilon g_{y'}(y', x)$ in order to define P'. But, by Lemma 2.2.6, for ε small enough, the map $x \mapsto x + \varepsilon g_{y'}(y', x)$ admits an real–analytic inverse of the form

$$\varphi(y', x'; \varepsilon) := x' + \varepsilon \widetilde{\varphi}(y', x'; \varepsilon), \tag{6.3.18}$$

[112] *i.e.* there a exits a bijection from $\mathscr{D}_\sharp$ onto $\mathscr{D}'_\sharp$ which extends to a diffeomorphism on some neighborhood of $\mathscr{D}_\sharp$.

so that the Arnod's symplectic transformation is given by

$$\phi_1\colon (y',x') \mapsto \begin{cases} y = y' + \varepsilon g_x(y', \varphi(y',x')) \\ x = \varphi(y',x';\varepsilon) = x' + \varepsilon\widetilde{\varphi}(y',x';\varepsilon). \end{cases} \tag{6.3.19}$$

Hence, (6.3.12) holds with

$$P'(y',x') \coloneqq P'(y', \varphi(y',x')). \tag{6.3.20}$$

Step 2 Above all, notice that[113]

$$\bar{r} \leqslant \frac{r\sigma}{16d\mathsf{TK}} \leqslant \frac{r}{32d} < \frac{r}{2}\,. \tag{6.3.21}$$

We begin by extending the "diophantine condition w.r.t. K_y" uniformly to $D_{\bar{r}}(\mathscr{D}_\sharp)$ up to the order κ. Indeed, for any $\mathsf{y} \in \mathscr{D}_\sharp$, $0 < |n|_1 \leqslant \kappa$ and $y' \in D_{\bar{r}}(\mathsf{y})$,

$$\begin{aligned} |K_y(y')\cdot n| &= |\omega\cdot n + (K_y(y') - K_y(\mathsf{y}))\cdot n| \geqslant |\omega\cdot n| \left(1 - d\frac{\|K_{yy}\|_{\bar{r},\mathscr{D}_\sharp}}{|\omega\cdot n|}|n|_1\bar{r}\right) \\ &\geqslant \frac{\alpha}{|n|_1^\tau}\left(1 - \frac{d\mathsf{K}}{\alpha}|n|_1^{\tau+1}\bar{r}\right) \geqslant \frac{\alpha}{|n|_1^\tau}\left(1 - \frac{d\mathsf{K}}{\alpha}\kappa^{\tau+1}\bar{r}\right) \\ &\geqslant \frac{\alpha}{2|n|_1^\tau}, \end{aligned} \tag{6.3.22}$$

so that, by Lemma 2.2.4–(i), we have

$$\begin{aligned} \|g\|_{\bar{r},\bar{s},\mathscr{D}_\sharp} &\overset{def}{=} \sup_{D_{\bar{r},\bar{s}}(\mathscr{D}_\sharp)} \left| \sum_{0<|n|_1\leqslant\kappa} \frac{P_n(y')}{K_y(y')\cdot n}\, e^{in\cdot x} \right| \leqslant \sum_{0<|n|_1\leqslant\kappa} \frac{\|P_n\|_{\bar{r},\bar{s},\mathscr{D}_\sharp}}{|K_y(y')\cdot n|}\, e^{\left(s-\frac{2}{3}\sigma\right)|n|_1} \\ &\leqslant \sum_{0<|n|_1\leqslant\kappa} M\, e^{-s|n|_1}\frac{2|n|_1^\tau}{\alpha}\, e^{\left(s-\frac{2}{3}\sigma\right)|n|_1} \leqslant \frac{2M}{\alpha}\sum_{n\in\mathbb{Z}^d}|n|_1^\tau\, e^{-\frac{2}{3}\sigma|n|_1} \\ &\leqslant \frac{2M}{\alpha}\int_{\mathbb{R}^d}|y|_1^\tau\, e^{-\frac{2}{3}\sigma|y|_1}dy \\ &= \left(\frac{3}{2\sigma}\right)^{\tau+d}\frac{2M}{\alpha}\int_{\mathbb{R}^d}|y|_1^\tau\, e^{-|y|_1}dy \\ &\leqslant \mathsf{C}_1\frac{M}{\alpha}\sigma^{-(\tau+d)} \end{aligned}$$

[113] Recall footnote [42].

and analogously,

$$\begin{aligned}
\|g_x\|_{\bar r,\bar s,\mathscr{D}_\sharp} &\overset{def}{=} \sup_{D_{\bar r,\bar s}(\mathscr{D}_\sharp)} \left| \sum_{0<|n|_1\leqslant\kappa} \frac{nP_n(y')}{K_y(y')\cdot n}\, e^{in\cdot x}\right| \leqslant \sum_{0<|n|_1\leqslant\kappa} \frac{\|P_n\|_{\bar r,\bar s,\mathscr{D}_\sharp}}{|K_y(y')\cdot n|} |n|_1\, e^{\left(s-\frac{2}{3}\sigma\right)|n|_1}\\
&\leqslant \sum_{0<|n|_1\leqslant\kappa} M\, e^{-s|n|_1} \frac{2|n|_1^{\tau+1}}{\alpha}\, e^{\left(s-\frac{2}{3}\sigma\right)|n|_1} \leqslant \frac{2M}{\alpha}\sum_{n\in\mathbb{Z}^d} |n|_1^{\tau+1}\, e^{-\frac{2}{3}\sigma|n|_1}\\
&\leqslant \frac{2M}{\alpha}\int_{\mathbb{R}^d} |y|_1^{\tau+1}\, e^{-\frac{2}{3}\sigma|y|_1} dy\\
&= \left(\frac{3}{2\sigma}\right)^{\tau+d+1} \frac{2M}{\alpha}\int_{\mathbb{R}^d} |y|_1^{\tau+1}\, e^{-|y|_1} dy\\
&\leqslant \mathsf{C}_1 \frac{M}{\alpha}\sigma^{-(\tau+d+1)}\,,
\end{aligned}$$

$$\begin{aligned}
\|\partial_{y'}g\|_{\bar r,\bar s,\mathscr{D}_\sharp} &\overset{def}{=} \sup_{D_{\bar r,\bar s}(\mathscr{D}_\sharp)} \left| \sum_{0<|n|_1\leqslant\kappa} \left(\frac{\partial_y P_n(y')}{K_y(y')\cdot n} - P_n(y')\frac{K_{yy}(y')n}{(K_y(y')\cdot n)^2}\right) e^{in\cdot x}\right|\\
&\leqslant \sum_{0<|n|_1\leqslant\kappa} \sup_{D_{\bar r}(\mathscr{D}_\sharp)} \left(\frac{\|(P_y)_n\|_{\bar r,s,\mathscr{D}_\sharp}}{|K_y(y')\cdot n|} + d\|P_n\|_{r,s,\mathscr{D}_\sharp}\frac{\|K_{yy}\|_{r,\mathscr{D}_\sharp}|n|_1}{|K_y(y')\cdot n|^2}\right) e^{\left(s-\frac{2}{3}\sigma\right)|n|_1}\\
&\overset{(6.3.1)+(6.3.22)}{\leqslant} \sum_{0<|n|_1\leqslant\kappa} \left(\frac{M}{r-\bar r}\, e^{-s|n|_1}\frac{2|n|_1^{\tau}}{\alpha} + dM\, e^{-s|n|_1}\mathsf{K}|n|_1 \left(\frac{2|n|_1^{\tau}}{\alpha}\right)^2\right) e^{\left(s-\frac{2}{3}\sigma\right)|n|_1}\\
&\overset{(6.3.21)}{\leqslant} \frac{4M}{\alpha^2 r}\sum_{0<|n|_1\leqslant\kappa} \left(|n|_1^{\tau}\alpha + dr\mathsf{K}|n|_1^{2\tau+1}\right) e^{-\frac{2}{3}\sigma|n|_1}\\
&\leqslant \max(\alpha, r\mathsf{K})\frac{4M}{\alpha^2 r}\sum_{0<|n|_1\leqslant\kappa} \left(|n|_1^{\tau} + d|n|_1^{2\tau+1}\right) e^{-\frac{2}{3}\sigma|n|_1}\\
&\leqslant \max(\alpha, r\mathsf{K})\frac{4M}{\alpha^2 r}\int_{\mathbb{R}^d} \left(|y|_1^{\tau} + d|y|_1^{2\tau+1}\right) e^{-\frac{2}{3}\sigma|y|_1} dy\\
&= \left(\frac{3}{2\sigma}\right)^{2\tau+d+1} \max(\alpha, r\mathsf{K})\frac{4M}{\alpha^2 r}\int_{\mathbb{R}^d} \left(|y|_1^{\tau} + d|y|_1^{2\tau+1}\right) e^{-|y|_1} dy\\
&\leqslant \frac{\mathsf{C}_0}{\sqrt{2}}\max\left\{1, \frac{\alpha}{r\mathsf{K}}\right\}\frac{\mathsf{K}M}{\alpha^2}\sigma^{-(2\tau+d+1)}\\
&\leqslant \bar{\mathsf{L}}\,,
\end{aligned}$$

$$\begin{aligned}
\|\partial^2_{y'x} g\|_{\bar r,\bar s,\mathscr{D}_\sharp} &\overset{def}{=} \sup_{D_{\bar r,\bar s}(\mathscr{D}_\sharp)} \left| \sum_{0<|n|_1\leqslant \kappa} \left(\frac{\partial_y P_n(y')}{K_y(y')\cdot n} - P_n(y') \frac{K_{yy}(y')n}{(K_y(y')\cdot n)^2} \right)\cdot n\, e^{in\cdot x} \right| \\
&\leqslant \sum_{0<|n|_1\leqslant \kappa} \sup_{D_{\bar r}(\mathscr{D}_\sharp)} \left(\frac{\|(P_y)_n\|_{\bar r,s,\mathscr{D}_\sharp}}{|K_y(y')\cdot n|} + d\|P_n\|_{r,s,\mathscr{D}_\sharp} \frac{\|K_{yy}\|_{r,\mathscr{D}_\sharp}|n|_1}{|K_y(y')\cdot n|^2} \right) |n|_1\, e^{\left(s-\frac{2}{3}\sigma\right)|n|_1} \\
&\leqslant \max(\alpha, r\mathsf{K}) \frac{4M}{\alpha^2 r} \sum_{0<|n|_1\leqslant\kappa} \left(|n|_1^\tau + d|n|_1^{2\tau+1}\right) |n|_1\, e^{-\frac{2}{3}\sigma|n|_1} \\
&\leqslant \max(\alpha, r\mathsf{K}) \frac{4M}{\alpha^2 r} \int_{\mathbb{R}^d} \left(|y|_1^\tau + d|y|_1^{2\tau+1}\right) |y|_1\, e^{-\frac{2}{3}\sigma|y|_1} dy \\
&= \left(\frac{3}{2\sigma}\right)^{2\tau+d+2} \max(\alpha, r\mathsf{K}) \frac{4M}{\alpha^2 r} \int_{\mathbb{R}^d} \left(|y|_1^{\tau+1} + d|y|_1^{2\tau+2}\right) e^{-|y|_1} dy \\
&= \frac{\mathsf{C}_0}{\sqrt{2}} \max\left\{1, \frac{\alpha}{r\mathsf{K}}\right\} \frac{\mathsf{K}M}{\alpha^2} \sigma^{-(2\nu+d)} \\
&= \bar{\mathsf{L}}\,,
\end{aligned}$$

and

$$\begin{aligned}
\|\partial_{y'}\widetilde{K}\|_{\bar r,\mathscr{D}_\sharp} &= \|\langle P_y\rangle\|_{\bar r,\mathscr{D}_\sharp} \leqslant \|P_y\|_{\bar r,\bar s,\mathscr{D}_\sharp} \leqslant \frac{M}{r-\bar r} \leqslant \frac{2M}{r}\;, \\
\|\partial^2_{y'}\widetilde{K}\|_{\bar r,\mathscr{D}_\sharp} &= \|\langle P_{yy}\rangle\|_{\bar r,\mathscr{D}_\sharp} \leqslant \|P_{yy}\|_{\bar r,\bar s,\mathscr{D}_\sharp} \leqslant \frac{M}{(r-\bar r)^2} \leqslant \frac{4M}{r^2} \leqslant \mathsf{K}\mathsf{L}
\end{aligned}$$

Next, we construct $\mathscr{D}'_\sharp$ in (6.3.12). For, fix $\mathsf{y}\in\mathscr{D}_\sharp$ and consider

$$\begin{aligned}
F\colon D_{\bar r}(\mathsf{y})\times D_{\bar r}(\mathsf{y}) &\longrightarrow \qquad \mathbb{C}^d \\
(y,z) \quad &\longmapsto K_y(y) + \varepsilon \widetilde{K}_{y'}(y) - K_y(z).
\end{aligned}$$

Then

- $F_y(\mathsf{y},\mathsf{y}) = \partial^2_y K(\mathsf{y}) + \varepsilon\partial^2_{y'}\widetilde{K}(\mathsf{y}) = T(\mathsf{y})^{-1}\left(\mathbb{1}_d + \varepsilon T(\mathsf{y})\partial^2_{y'}\widetilde{K}(\mathsf{y})\right) =: T(\mathsf{y})^{-1}(\mathbb{1}_d+\varepsilon A_0)$ and

  $$\|\varepsilon A_0\| \leqslant \|T(\mathsf{y})\|\|\varepsilon\partial^2_{y'}\widetilde{K}(\mathsf{y})\| \leqslant \mathsf{T}\frac{4|\varepsilon|M}{r^2} \overset{(6.3.21)}{\leqslant} |\varepsilon|\frac{2\mathsf{T}M}{r\bar r} \leqslant \frac{1}{2}|\varepsilon|\mathsf{L} \overset{(6.3.5)}{\leqslant} \frac{\sigma}{6} < \frac{1}{2}.$$

  Hence, $F_y(\mathsf{y},\mathsf{y})$ is invertible, with inverse

  $$T_0 := (\mathbb{1}_d+\varepsilon A_0)^{-1}T(\mathsf{y}) = \left(\mathbb{1}_d + \sum_{k\geqslant 1}(-\varepsilon)^k A_0^k\right) T(\mathsf{y})$$

  satisfying

  $$\|T_0\| \leqslant \frac{\|T(\mathsf{y})\|}{1-\|\varepsilon A_0\|} \leqslant 2\mathsf{T}. \tag{6.3.23}$$

- For any $(y,z)\in D_{\tilde r}(\mathsf{y})\times D_{\tilde r}(\mathsf{y})$,

$$\begin{aligned}
\|\mathbb{1}_d - T_0F_y(y,z)\| &\leqslant \|T_0\|\|\partial_y^2 K(\mathsf{y}) - K_{yy}\|_{\bar r,\mathscr{D}_\sharp} + |\varepsilon|\ \|T_0\|\ |\partial_{y'}^2\widetilde K(\mathsf{y})| + |\varepsilon|\ \|T_0\|\ \|\partial_{y'}^2\widetilde K\|_{\bar r,\mathscr{D}_\sharp}\\
&\leqslant d\cdot 2\mathsf{T}\|K_{yyy}\|_{\bar r,\mathscr{D}_\sharp}\cdot \bar r + 4|\varepsilon|\mathsf{T}\frac{4M}{r^2}\\
&\leqslant 2d\mathsf{T}\mathsf{K}\frac{\bar r}{r-\bar r} + 16\mathsf{T}\frac{|\varepsilon| M}{r^2}\\
&\overset{(6.3.21)}{\leqslant} 2d\mathsf{T}\mathsf{K}\frac{2\bar r}{r} + |\varepsilon|\frac{16\mathsf{T}M}{r\bar r}\\
&\overset{(6.3.21)+(6.3.5)}{\leqslant} \frac{1}{4}+\frac{\sigma}{3}\\
&\leqslant \frac{1}{4}+\frac{1}{4}=\frac{1}{2}\ ;
\end{aligned}$$

- For any $z\in D_{\tilde r}(\mathsf{y})$,

$$\begin{aligned}
2\|T_0\||F(\mathsf{y},z)| &\leqslant 4\mathsf{T}|K_y(z)-K_y(\mathsf{y})| + 4\mathsf{T}|\varepsilon|\|\widetilde K_{y'}\|_{\bar r,\mathscr{D}_\sharp}\\
&\leqslant 4\mathsf{T}\|K_{yy}\|_{\bar r,\mathscr{D}_\sharp}\cdot\tilde r + 4\mathsf{T}\frac{2|\varepsilon| M}{r}\\
&\leqslant 4\mathsf{T}\mathsf{K}\tilde r + \frac{8|\varepsilon|\mathsf{T}_\infty M}{r}\\
&\leqslant \frac{\bar r}{4d} + |\varepsilon|\frac{\bar r}{2}\mathsf{L}\\
&\overset{(6.3.5)}{\leqslant} \frac{\bar r}{8}+\frac{\bar r}{12}\\
&< \frac{\bar r}{4}\ ,
\end{aligned}$$

i.e.

$$2\|T_0\|\|F(\mathsf{y},\cdot)\|_{\tilde r,\mathsf{y}} \leqslant \frac{\bar r}{4}\ .$$

Therefore, Lemma 2.2.6 applies. Hence, there exists a real–analytic map $G^{\mathsf{y}}\colon D_{\tilde r}(\mathsf{y})\to D_{\bar r}(\mathsf{y})$ such that its graph coincides with $F^{-1}(\{0\})$ *i.e.* $y' = y'(z,\mathsf{y},\varepsilon) := G^{\mathsf{y}}(z)$ is the unique $y\in D_{\bar r}(\mathsf{y})$ satisfying $0 = F(y,z) = \partial_y K'(y) - K_y(z)$, for any $z\in D_{\tilde r}(\mathsf{y})$. Moreover, $\forall\ z\in D_{\tilde r}(\mathsf{y})$,

$$|G^{\mathsf{y}}(z)-\mathsf{y}| \leqslant 2\|T_0\|\|F(\mathsf{y},\cdot)\|_{\tilde r,\mathsf{y}} \leqslant \frac{\bar r}{4}\ , \tag{6.3.24}$$

$$|G^{\mathsf{y}}(z)-z| \leqslant |G^{\mathsf{y}}(z)-\mathsf{y}| + |\mathsf{y}-z| \leqslant \frac{\bar r}{4}+\tilde r < \frac{\bar r}{2}, \tag{6.3.25}$$

so that

$$D_{\bar r/4}(G^{\mathsf y}(z)) \subset D_{\bar r/2}(\mathsf y). \tag{6.3.26}$$

Next, we prove that $\partial^2_{y'}K'(y')$ is invertible, where $y' = G^{\mathsf y}(z)$ for some given $z \in D_{\tilde r}(\mathsf y)$. Indeed, by Taylor's formula, we have,

$$\begin{aligned}
\partial^2_{y'}K'(y') &= K_{yy}(\mathsf y) + \int_0^1 K_{yyy}(\mathsf y + t(y'-\mathsf y))(y'-\mathsf y)dt + \varepsilon\partial^2_{y'}\widetilde K(y')\\
&= T(\mathsf y)^{-1}\left(\mathbb{1}_d + T(\mathsf y)\left(\int_0^1 K_{yyy}(\mathsf y + t(y'-\mathsf y))(y'-\mathsf y)dt + \partial^2_{y'}\widetilde K(y')\right)\right)\\
&=: T(\mathsf y)^{-1}(\mathbb{1}_d + \varepsilon A),
\end{aligned}$$

and, by Cauchy's estimate,[114]

$$\begin{aligned}
|\varepsilon|\|A\| &\leqslant \|T(y_0)\|\left(d\|K_{yyy}\|_{\bar r,\mathscr D_\sharp}|y'-\mathsf y| + |\varepsilon|\|\partial^2_{y'}\widetilde K\|_{\bar r,\mathscr D_\sharp}\right)\\
&\leqslant \|T\|_{\mathscr D_\sharp}\left(\frac{d\|K_{yy}\|_{r,\mathscr D_\sharp}}{r-\bar r}|y'-\mathsf y| + |\varepsilon|\|\partial^2_{y'}\widetilde K\|_{\bar r,\mathscr D_\sharp}\right)\\
&\leqslant \mathsf T\left(\frac{2d\mathsf K}{r}\frac{\bar r}{2} + \frac{4M}{r^2}\right)\\
&\overset{(6.3.21)}{\leqslant} \mathsf T\left(\frac{\sigma}{16\mathsf T} + \frac{1}{4\mathsf T}|\varepsilon|\mathsf L\right)\\
&\overset{(6.3.5)}{\leqslant} \mathsf T\left(\frac{\sigma}{16\mathsf T} + \frac{1}{4\mathsf T}\frac{\sigma}{3}\right)\\
&= \frac{\sigma}{6}\\
&< \frac12.
\end{aligned}$$

Hence $\partial^2_{y'}K'(y')$ is invertible with

$$\partial^2_{y'}K'(y')^{-1} = (\mathbb{1}_d + \varepsilon A)^{-1}T(\mathsf y) = T(\mathsf y) + \sum_{k\geqslant 1}(-\varepsilon)^k A^k T(\mathsf y) =: T(\mathsf y) + \varepsilon\widetilde T(y'),$$

and

$$|\varepsilon|\|\widetilde T(y')\| \leqslant |\varepsilon|\frac{\|A\|}{1-|\varepsilon|\|A\|}\|T\|_{\mathscr D_\sharp} \leqslant 2|\varepsilon|\|A\|\|T\|_{\mathscr D_\sharp} \leqslant 2\frac{\sigma}{6}\mathsf T = \mathsf T\frac{\sigma}{3}\,.$$

Similarly, from

$$K_{yy}(z) = K_{yy}(\mathsf y)\left(\mathbb{1}_d + T(\mathsf y)\int_0^1 K_{yyy}(\mathsf y + t(z-\mathsf y))(z-\mathsf y)dt\right)$$

[114] Recall footnote 42 .

and

$$\left\| T(\mathsf{y}) \int_0^1 K_{yyy}(\mathsf{y} + t(z - \mathsf{y}))(z - \mathsf{y}) dt \right\|_{r/(4d\mathsf{TK}),\mathsf{y}} \leqslant \mathsf{T} \| K_{yyy} \|_{r/2,\mathsf{y}} \frac{r}{4d\mathsf{TK}} \leqslant \mathsf{T} \frac{d\mathsf{K}}{r - r/2} \frac{r}{4d\mathsf{TK}} = \frac{1}{2}$$

one has that, for any $z \in D_{r/(4d\mathsf{TK})}(\mathsf{y})$,

$$K_{yy}(z)^{-1} \text{ exists and } \|K_{yy}(z)^{-1}\| \leqslant \|K_{yy}(z)^{-1} - T(\mathsf{y})\| + \|T(\mathsf{y})\| \leqslant 2\frac{1}{2}\mathsf{T} + \mathsf{T} = 2\mathsf{T} \ . \quad (6.3.27)$$

Now, differentiating $F(G^{\mathsf{y}}(z), z) = 0$, we get, for any $z \in D_{\tilde{r}}(\mathsf{y})$,

$$\partial_{y'}^2 K'(G^{\mathsf{y}}(z)) \cdot \partial_z G^{\mathsf{y}}(z) = K_{yy}(z) \ .$$

Therefore G^{y} is a local diffeomorphism, with

$$\begin{aligned} \partial_z G^{\mathsf{y}}(z) &= \partial_{y'}^2 K'(G^{\mathsf{y}}(z))^{-1} K_{yy}(z) \\ &= \left(K_{yy}(z)^{-1} \left(K_{yy}(z) + \varepsilon \partial_{y'}^2 \widetilde{K}(G^{\mathsf{y}}(z)) \right) \right)^{-1} \\ &= \left(\mathbb{1}_d + \varepsilon K_{yy}(z)^{-1} \partial_{y'}^2 \widetilde{K}(G^{\mathsf{y}}(z)) \right)^{-1} \end{aligned}$$

and

$$\|\varepsilon K_{yy}^{-1} \partial_{y'}^2 \widetilde{K}\|_{\tilde{r},\mathsf{y}} \leqslant \|K_{yy}^{-1}\|_{\tilde{r},\mathsf{y}} \|\varepsilon \partial_{y'}^2 \widetilde{K}\|_{\tilde{r},\mathscr{D}_\sharp} \leqslant 2\mathsf{T} \frac{|\varepsilon|\mathsf{L}}{4\mathsf{T}} \sigma^{\nu+d} \leqslant \frac{1}{2} \sigma^{\nu+d} |\varepsilon| \mathsf{L} < \frac{\sigma}{6} < \frac{1}{2}$$

so that

$$\|\partial_z G^{\mathsf{y}} - \mathbb{1}_d\|_{\tilde{r},\mathsf{y}} \leqslant 2\|\varepsilon K_{yy}^{-1} \partial_{y'}^2 \widetilde{K}\|_{\tilde{r},\mathsf{y}} \leqslant \sigma^{\nu+d} |\varepsilon| \mathsf{L}. \quad (6.3.28)$$

Now, we show that the family $\{G^{\mathsf{y}}\}_{\mathsf{y} \in \mathscr{D}_\sharp}$ is compatible so that, together, they define a global map on $D_{\tilde{r}}(\mathscr{D}_\sharp)$, say G and that, in fact, G is a real–analytic diffeomorphism. For, assume that $z \in D_{\tilde{r}}(\mathsf{y}) \bigcap D_{\tilde{r}}(\hat{\mathsf{y}})$, for some $\mathsf{y}, \hat{\mathsf{y}} \in \mathscr{D}_\sharp$. Then, we need to show that $G^{\mathsf{y}}(z) = G^{\hat{\mathsf{y}}}(z)$. But, we have

$$|G^{\hat{\mathsf{y}}}(z) - \mathsf{y}| \leqslant |G^{\hat{\mathsf{y}}}(z) - \hat{\mathsf{y}}| + |\hat{\mathsf{y}} - z| + |z - \mathsf{y}| \overset{(6.3.24)}{\leqslant} \frac{\bar{r}}{2} + \tilde{r} + \tilde{r} < \bar{r}.$$

Hence, $z \in D_{\tilde{r}}(\mathsf{y})$, $G^{\hat{\mathsf{y}}}(z) \in D_{\bar{r}}(\mathsf{y})$ and, by definitions, $F(G^{\hat{\mathsf{y}}}(z), z) = 0 = F(G^{\mathsf{y}}(z), z)$. Then, by unicity, we get $G^{\mathsf{y}}(z) = G^{\hat{\mathsf{y}}}(z)$. Thus, the map

$$G \colon D_{\tilde{r}}(\mathscr{D}_\sharp) \to \mathbb{C}^d \quad \text{such that} \quad G_{|D_{\tilde{r}}(\mathsf{y})} := G^{\mathsf{y}}, \quad \forall\, \mathsf{y} \in \mathscr{D}_\sharp \ ,$$

is well–defined and, therefore, is a real–analytic local diffeomorphism. It remains only to check that G is injective to conclude that it is a global diffeomorphism. Let then $z \in D_{\tilde{r}}(\mathsf{y}), \hat{z} \in D_{\tilde{r}}(\hat{\mathsf{y}})$ such that $G(z) = G(\hat{z})$, for some $\mathsf{y}, \hat{\mathsf{y}} \in \mathscr{D}_\sharp$. Then, we have

$$|z - \hat{z}| < \frac{r}{4d\mathsf{TK}} - \tilde{r}.$$

Indeed, if not then

$$\begin{aligned} 0 = |G(z) - G(\hat{z})| &\geqslant -|G(z) - z| + |z - \hat{z}| - |\hat{z} - G(\hat{z})| \\ &\overset{(6.3.25)}{\geqslant} -\bar{r} + \frac{r}{4d\mathsf{TK}} - \tilde{r} - \bar{r} \\ &\geqslant \frac{r}{4d\mathsf{TK}} - 3\bar{r} \\ &\overset{(6.3.21)}{\geqslant} \frac{r}{4d\mathsf{TK}} - 3\frac{r}{16d\mathsf{TK}} \\ &> 0\,, \end{aligned}$$

contradiction. Therefore,

$$|z - \hat{z}| < \frac{r}{4d\mathsf{TK}} - \tilde{r}\ . \tag{6.3.29}$$

Thus,

$$|\hat{z} - \mathsf{y}| \leqslant |\hat{z} - z| + |z - \mathsf{y}| < \frac{r}{4d\mathsf{TK}} - \tilde{r} + \tilde{r} = \frac{r}{4d\mathsf{TK}}\ .$$

Hence, $z, \hat{z} \in D_{r/(4d\mathsf{TK})}(\mathsf{y})$. But $G(z) = G(\hat{z})$ is equivalent to $K_y(z) = K_y(\hat{z})$ and then,

$$0 = K_y(z) - K_y(\hat{z}) = \int_0^1 K_{yy}(\hat{z} + t(z - \hat{z}))dt(z - \hat{z})\ .$$

Thus, it is enough to show that $\int_0^1 K_{yy}(\hat{z} + t(z - \hat{z}))dt$ is invertible. But

$$\begin{aligned} \int_0^1 K_{yy}(\hat{z} + t(z - \hat{z}))dt &= K_{yy}(\hat{z}) + \int_0^1 \int_0^1 K_{yyy}(\hat{z} + tt'(z - \hat{z}))tdt'dt \cdot (z - \hat{z}) \\ &\overset{(6.3.27)}{=} K_{yy}(\hat{z}) \left(\mathbb{1}_d + K_{yy}(\hat{z})^{-1} \int_0^1 \int_0^1 K_{yyy}(\hat{z} + tt'(z - \hat{z}))tdt'dt \cdot (z - \hat{z}) \right) \end{aligned}$$

and

$$\left\| K_{yy}(\hat{z})^{-1}\int_0^1\int_0^1 K_{yyy}(\hat{z}+tt'(z-\hat{z}))tdt'dt\cdot(z-\hat{z})\right\| \overset{(6.3.27)}{\leqslant} 2\mathsf{T}\cdot\frac{1}{2}\|K_{yyy}\|_{r/2,\mathsf{y}}|z-\hat{z}|$$
$$\overset{(6.3.29)}{\leqslant} \mathsf{T}\frac{2d\mathsf{K}}{r}\left(\frac{r}{4d\mathsf{TK}}-\tilde{r}\right)$$
$$< \frac{2d\mathsf{TK}}{r}\frac{r}{4d\mathsf{TK}}$$
$$= \frac{1}{2}.$$

Therefore, $\int_0^1 K_{yy}(\hat{z}+t(z-\hat{z}))dt$ is invertible and then we get $z-\hat{z}=0$ *i.e.* G is injective. Next, we estimate $\|G-\mathrm{id}\|_{\tilde{r},\mathscr{D}_\sharp}$. The strategy is to show that the expression $(K_y+\varepsilon\widetilde{K}_{y'})^{-1}\circ K_y$ defines a map on $D_{\tilde{r}}(\mathsf{y})$ by means of the Inversion Function Lemma 2.2.7; hence, we will get an explicit formula for G:

$$G=(K_y+\varepsilon\widetilde{K}_{y'})^{-1}\circ K_y \qquad \text{on} \quad D_{\tilde{r}}(\mathsf{y})\ . \tag{6.3.30}$$

But, the proof is part of the above computation: for any $y\in D_{\bar{r}}(\mathsf{y})$,

$$\|\mathbb{1}_d-T_0F_y(y,\mathsf{y})\|\leqslant\frac{1}{2}$$

implies, using Lemma 2.2.7, that $K_y+\varepsilon\widetilde{K}_{y'}$ admits an inverse defined on $D_{r_\sharp}(K_y(\mathsf{y})+\varepsilon\widetilde{K}_{y'}(\mathsf{y}))$, where

$$r_\sharp:=\frac{\bar{r}}{4\mathsf{T}}<\frac{\bar{r}}{2\|T_0\|}\ .$$

Moreover, for any $y\in D_{\tilde{r}}(\mathsf{y})$,

$$|K_y(y)-(K_y(\mathsf{y})+\varepsilon\widetilde{K}_{y'}(\mathsf{y}))|\leqslant\|K_{yy}\|_{\tilde{r},\mathscr{D}_\sharp}\cdot\tilde{r}+\|\varepsilon\widetilde{K}_{y'}\|_{\tilde{r},\mathscr{D}_\sharp}$$
$$\leqslant \mathsf{K}\frac{\bar{r}}{16d\mathsf{TK}}+\frac{2|\varepsilon|M}{r}$$
$$\leqslant\frac{\bar{r}}{16d\mathsf{T}}+\frac{\bar{r}}{8\mathsf{T}}|\varepsilon|\mathsf{L}$$
$$<\frac{\bar{r}}{4\mathsf{T}}=r_\sharp\ ,$$

and thus, (6.3.30) is proven. Hence, for any $y \in D_{\tilde{r}}(\mathsf{y})$,

$$\begin{aligned}
|G(y) - y| &= |(\partial_{y'} K')^{-1}(K_y(y)) - (\partial_{y'} K')^{-1}(K_y(y) + \varepsilon \widetilde{K}_{y'}(y))| \\
&\leqslant \int_0^1 \|\partial_{y'}((\partial_{y'} K')^{-1})(K_y(y) + t\varepsilon \widetilde{K}_{y'}(y))\| dt \ \|\varepsilon \widetilde{K}_{y'}\|_{\bar{r},\mathscr{D}_\sharp} \\
&\leqslant \|(\partial^2_{y'} K')^{-1}\|_{\bar{r},\mathscr{D}_\sharp} \frac{2|\varepsilon| M}{r} \\
&< \frac{16\mathsf{T} M}{r} |\varepsilon| \\
&< \sigma^{\nu+d} \bar{r} |\varepsilon| \mathsf{L} \,.
\end{aligned}$$

Now, we estimate P_+. We have,

$$|\varepsilon| \|g_x\|_{\bar{r},\bar{s},\mathscr{D}_\sharp} \leqslant |\varepsilon| \mathsf{C}_1 \frac{M}{\alpha} \sigma^{-(\nu+d)} \leqslant |\varepsilon| \frac{r}{3} \mathsf{L} \overset{(6.3.5)}{\leqslant} \frac{r}{3}\frac{\sigma}{3} \leqslant \frac{r}{12}$$

so that, for any $\mathsf{y} \in \mathscr{D}_\sharp$ and $(y', x) \in D_{\bar{r},\bar{s}}(\mathsf{y})$,

$$|y' + \varepsilon g_x(y', x) - \mathsf{y}| \leqslant \bar{r} + \frac{r}{3} \leqslant \frac{r}{32d} + \frac{r}{12} < \frac{2r}{3} < r \,,$$

and thus

$$\begin{aligned}
\|P^{(1)}\|_{\bar{r},\bar{s},\mathscr{D}_\sharp} &\leqslant d^2 \|K_{yy}\|_{r,\mathscr{D}_\sharp} \|g_x\|^2_{\bar{r},\bar{s},\mathscr{D}_\sharp} \leqslant d^2 \mathsf{K} \left(\mathsf{C}_1 \frac{M}{\alpha} \sigma^{-(\nu+d)} \right)^2 \\
&= d^2 \mathsf{C}_1^2 \frac{\mathsf{K} M^2}{\alpha^2} \sigma^{-2(\nu+d)},
\end{aligned}$$

$$\begin{aligned}
\|P^{(2)}\|_{\bar{r},\bar{s},\mathscr{D}_\sharp} &\leqslant d \|P_y\|_{\frac{5r}{6},\bar{s},\mathscr{D}_\sharp} \|g_x\|_{\bar{r},\bar{s},\mathscr{D}_\sharp} \leqslant d \frac{6M}{r} \mathsf{C}_1 \frac{M}{\alpha} \sigma^{-(\nu+d)} \\
&= 6d \mathsf{C}_1 \frac{M^2}{\alpha r} \sigma^{-(\nu+d)}
\end{aligned}$$

and, by Lemma 2.2.4–(i), we have

$$\begin{aligned}
|\varepsilon|\|P^{(3)}\|_{\bar r,s-\frac{\sigma}{2},\mathscr{D}_\sharp} &\leqslant \sum_{|n|_1>\kappa}\|P_n\|_{\bar r,\mathscr{D}_\sharp}\,e^{(s-\frac{\sigma}{2})|n|_1}\leqslant M\sum_{|n|_1>\kappa}e^{-\frac{\sigma|n|_1}{2}}\\
&\leqslant M\,e^{-\frac{\kappa\sigma}{4}}\sum_{|n|_1>\kappa}e^{-\frac{\sigma|n|_1}{4}}\leqslant M\,e^{-\frac{\kappa\sigma}{4}}\sum_{|n|_1>0}e^{-\frac{\sigma|n|_1}{4}}\\
&= M\,e^{-\frac{\kappa\sigma}{4}}\left(\left(\sum_{k\in\mathbb{Z}}e^{-\frac{\sigma|k|}{4}}\right)^d-1\right)=M\,e^{-\frac{\kappa\sigma}{4}}\left(\left(1+\frac{2\,e^{-\frac{\sigma}{4}}}{1-\,e^{-\frac{\sigma}{4}}}\right)^d-1\right)\\
&= M\,e^{-\frac{\kappa\sigma}{4}}\left(\left(1+\frac{2}{e^{\frac{\sigma}{4}}-1}\right)^d-1\right)\leqslant M\,e^{-\frac{\kappa\sigma}{4}}\left(\left(1+\frac{2}{\frac{\sigma}{4}}\right)^d-1\right)\\
&\leqslant \sigma^{-d}M\,e^{-\frac{\kappa\sigma}{4}}\Big((\sigma+8)^d-\sigma^d\Big)\leqslant d8^d\sigma^{-d}M\,e^{-\frac{\kappa\sigma}{4}}\\
&= \mathsf{C}_2\sigma^{-d}M\,e^{-\lambda}\\
&\overset{(6.3.2)}{\leqslant}\mathsf{C}_2\sigma^{-d}M\sigma^{-(2\nu+d)}\frac{\mathsf{K}|\varepsilon|M}{\alpha^2}\\
&= \mathsf{C}_2M\frac{\mathsf{K}|\varepsilon|M}{\alpha^2}\sigma^{-2(\nu+d)}\,.
\end{aligned}$$

Hence,[115]

$$\begin{aligned}
\|P_+\|_{\bar r,\bar s,\mathscr{D}_\sharp} &\leqslant \|P^{(1)}\|_{\bar r,\bar s,\mathscr{D}_\sharp}+\|P^{(2)}\|_{\bar r,\bar s,\mathscr{D}_\sharp}+\|P^{(3)}\|_{\bar r,\bar s,\mathscr{D}_\sharp}\\
&\leqslant d^2\mathsf{C}_1^2\frac{\mathsf{K}M^2}{\alpha^2}\sigma^{-2(\nu+d)}+6d\mathsf{C}_1\frac{M^2}{\alpha r}\sigma^{-(\nu+d)}+\mathsf{C}_2M\frac{|\varepsilon|M\mathsf{K}}{\alpha^2}\sigma^{-2(\nu+d)}\\
&= \Big(d^2\mathsf{C}_1^2r\mathsf{K}+6d\mathsf{C}_1\alpha\sigma^{\nu+d}+\mathsf{C}_2r\mathsf{K}\Big)\frac{M^2}{\alpha^2r}\sigma^{-2(\nu+d)}\\
&\leqslant \Big(d^2\mathsf{C}_1^2+6d\mathsf{C}_1+\mathsf{C}_2\Big)\max\{\alpha,r\mathsf{K}\}\frac{M^2}{\alpha^2r}\sigma^{-2(\nu+d)}\\
&\overset{(6.3.1)}{\leqslant}\frac{\mathsf{C}_3}{\sqrt{2}}\max\left\{1,\frac{\alpha}{r\mathsf{K}}\right\}\frac{M^2\mathsf{K}}{\alpha^2}\sigma^{-2(\nu+d)}\\
&\leqslant \mathsf{L}M\,.
\end{aligned}$$

The proof of the claims on ϕ' and P' are proven in a similar way as in Lemma 2.3.1. ▮

[115] Recall that $r\leqslant r_0$ and $\sigma<1$.

Let $H_0 \coloneqq H$, $K_0 \coloneqq K$, $P_0 \coloneqq P$, $\phi^0 = \phi_0 \coloneqq \mathrm{id}$ and r_0, s_0, s_*, σ_0, M_0, K_0, T_0 and η_0 be as in §6.1. For a given $\varepsilon \neq 0$ and $j \geqslant 0$, define[116]

$$\begin{aligned}
\sigma_j &\coloneqq \frac{\sigma_0}{2^j}\,,\\
s_{j+1} &\coloneqq s_j - \sigma_j = s_* + \frac{\sigma_0}{2^j}\,,\\
\bar{s}_j &\coloneqq s_j - \frac{2\sigma_j}{3}\,,\\
\mathsf{K}_{j+1} &\coloneqq \mathsf{K}_0 \prod_{k=0}^{j}(1+\frac{\sigma_k}{3}) \leqslant \mathsf{K}_0\, e^{\frac{2\sigma_0}{3}} < \mathsf{K}_0\sqrt{2}\,,\\
\mathsf{T}_{j+1} &\coloneqq \mathsf{T}_0 \prod_{k=0}^{j}(1+\frac{\sigma_k}{3}) \leqslant \mathsf{T}_0\, e^{\frac{2\sigma_0}{3}} < \mathsf{T}_0\sqrt{2}\,,\\
\lambda_0 &\coloneqq \log \mu_0^{-1}\,,\\
\mathsf{e}_* &\coloneqq \mathsf{C}_5 \sigma_0^{-(4\nu+2d+2)} \eta_0^2 \lambda_0^{2\nu}\,,\\
\mathsf{d}_* &\coloneqq \frac{2^{2\nu+2d+13} d^2 \eta_0^2}{\sigma_0^2}\,,\\
\kappa_0 &\coloneqq 4\sigma_0^{-1}\lambda_0\,,\\
\kappa_j &\coloneqq 4^j \kappa_0\,,\\
\hat{r}_0 &\coloneqq \frac{r_0}{\sqrt{|\varepsilon|}}\,,\\
\hat{r}_{j+1} &\coloneqq \frac{1}{2}\min\left\{\frac{\widehat{\alpha}}{2d\sqrt{2}\mathsf{K}_0\kappa_j^{\nu}}\,,\,\frac{\hat{r}_j\sigma_j}{32d\eta_0}\right\}\,,\\
r_{j+1} &\coloneqq \hat{r}_{j+1}\sqrt{|\varepsilon|}\,,\\
\tilde{r}_{j+1} &\coloneqq \frac{r_{j+1}}{8d\mathsf{T}_j\mathsf{K}_j}\,,\\
\widehat{M}_0 &\coloneqq M_0\,,\\
\widehat{M}_1 &\coloneqq 2\mathsf{e}_* \frac{\mathsf{K}_0 \widehat{M}_0^2}{\widehat{\alpha}^2}\,,
\end{aligned}$$

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

$$
\begin{aligned}
\widehat{M}_{j+2} &:= 8\mathsf{e}_*(4\mathsf{d}_*)^j \frac{\mathsf{K}_0 \widehat{M}_{j+1}^2}{\widehat{\alpha}^2} ,\\
\mu_j &:= \frac{\mathsf{K}_0 \widehat{M}_j}{\widehat{\alpha}^2} ,\\
\theta_j &:= 8\mathsf{e}_* \, (4\mathsf{d}_*)^j \, \mu_j ,\\
\mathscr{D}_0 &:= \mathscr{D}_{\delta,\alpha} ,\\
\mathsf{W}_0 &:= \operatorname{diag}\left(\frac{\mathsf{K}_0}{\alpha} \mathbb{1}_d, \mathbb{1}_d\right) ,\\
\mathsf{W}_{j+1} &:= \operatorname{diag}\left(\max\left\{\frac{\mathsf{K_j}}{\alpha} , \frac{1}{r_j}\right\} \mathbb{1}_d , \mathbb{1}_d\right) ,\\
\mathsf{L}_j &:= M_j \max\left\{\frac{32\sqrt{2}\mathsf{T}_0}{r_j r_{j+1}} \sigma_j^{-(\nu+d)} , \, \mathsf{C}_4 \max\left\{1, \frac{\alpha}{r_j \mathsf{K}_j}\right\} \frac{\mathsf{K}_0}{\alpha^2} \sigma_j^{-2(\nu+d)}\right\}\\
&= M_j \max\left\{\frac{32\sqrt{2}\mathsf{T}_0}{r_j r_{j+1}} \sigma_j^{-(\nu+d)} , \, \frac{4}{\mathsf{K}_j r_j^2} , \, \mathsf{C}_4 \max\left\{1, \frac{\alpha}{r_j \mathsf{K}_j}\right\} \frac{\mathsf{K}_0}{\alpha^2} \sigma_j^{-2(\nu+d)}\right\} .
\end{aligned}
$$

Thus,

$$
\theta_1 = 8\mathsf{e}_* \, (4\mathsf{d}_*) \, \mu_1 = 32\mathsf{e}_* \, \mathsf{d}_* \frac{\mathsf{K}_0 \widehat{M}_1}{\widehat{\alpha}^2} = 32\mathsf{e}_* \, \mathsf{d}_* \frac{\mathsf{K}_0}{\widehat{\alpha}^2} \, 2\mathsf{e}_* \frac{\mathsf{K}_0 \widehat{M}_0^2}{\widehat{\alpha}^2} = \mathsf{d}_* \, (8\mathsf{e}_* \, \mu_0)^2 = \mathsf{d}_* \, \theta_0^2
$$

and, for any $j \geqslant 1$,

$$
\begin{aligned}
\theta_{j+1} &= 8\mathsf{e}_* \, (4\mathsf{d}_*)^{j+1} \, \mu_{j+1} = 8\mathsf{e}_*(4\mathsf{d}_*)^{j+1} \frac{\mathsf{K}_0 \widehat{M}_{j+1}}{\widehat{\alpha}^2}\\
&= 8\mathsf{e}_*(4\mathsf{d}_*)^{j+1} \frac{\mathsf{K}_0}{\widehat{\alpha}^2} \, 8\mathsf{e}_*(4\mathsf{d}_*)^{j-1} \frac{\mathsf{K}_0 \widehat{M}_j^2}{\widehat{\alpha}^2} = \left(8\mathsf{e}_*(4\mathsf{d}_*)^j \, \mu_j\right)^2 = \theta_j^2
\end{aligned}
$$

i.e.

$$
\theta_j = \theta_1^{2^{j-1}} = (\sqrt{\mathsf{d}_*} \, \theta_0)^{2^j} .
$$

The very first step being quite different from all the others, it has to be done separately. Hence,

Lemma 6.3.2 *Under the above assumptions and notations, if*

$$
|\varepsilon| \leqslant \left(\frac{r_0 \sigma_0}{\widehat{\alpha} \mathsf{T}_0}\right)^2 \qquad \textit{and} \qquad \max\{\, e \, \mu_0 \, , \, 16d \, \eta_0 \, \theta_0\} \leqslant 1 \, , \tag{6.3.31}
$$

then, there exist $\mathscr{D}_1 \subset \mathscr{D}$, a real–analytic diffeomorphism

$$G_1 \colon D_{\tilde{r}_1}(\mathscr{D}_{\delta,\alpha}) \to G_1(D_{\tilde{r}_1}(\mathscr{D}_{\delta,\alpha}))$$

and a real–analytic symplectomorphism

$$\phi_1 : D_{r_1,s_1}(\mathscr{D}_1) \to D_{r_0,s_0}(\mathscr{D}_0) \tag{6.3.32}$$

such that

$$G_1(\mathscr{D}_{\delta,\alpha}) = \mathscr{D}_1 \ , \tag{6.3.33}$$
$$\partial_{y_1} K_1 \circ G_1 = \partial_y K_0 \ , \tag{6.3.34}$$
$$H_1 := H_0 \circ \phi_1 =: K_1 + \varepsilon^2 P_1 \qquad \text{on } D_{r_1,s_1}(\mathscr{D}_1) \tag{6.3.35}$$

and

$$\mathscr{D}_1 \subset \mathscr{D}_{r_1} \ , \tag{6.3.36}$$
$$\|K_1\|_{r_1,\mathscr{D}_1} \leqslant \mathsf{K}_1 \ , \qquad \|T_1\|_{\mathscr{D}_1} \leqslant \mathsf{T}_1 \ , \qquad T_1 := (\partial^2_{y_1} K_1)^{-1} \ , \tag{6.3.37}$$
$$2\varepsilon^2 M_1 := 2\varepsilon^2 \|P_1\|_{r_1,s_1,\mathscr{D}_1} \leqslant |\varepsilon| \widehat{M}_1 \ , \tag{6.3.38}$$
$$\|G_1 - \mathrm{id}\|_{\tilde{r}_1,\mathscr{D}_{\delta,\alpha}} \leqslant 2\, r_1\, \sigma_0^{\nu+d}\, |\varepsilon| \mathsf{L}_0 \ , \tag{6.3.39}$$
$$\|\partial_z G_1 - \mathbb{1}_d\|_{\tilde{r}_1,\mathscr{D}_{\delta,\alpha}} \leqslant \sigma_0^{\nu+d} |\varepsilon| \mathsf{L}_0 \ , \tag{6.3.40}$$
$$\|\mathsf{W}_1(\phi_1 - \mathrm{id})\|_{r_1,s_1,\mathscr{D}_1} \leqslant \sigma_0^d\, |\varepsilon| \mathsf{L}_0 \ . \tag{6.3.41}$$

Proof By

$$\kappa_0 \overset{(6.3.31)}{\geqslant} 4\sigma_0^{-1} \geqslant 8 \tag{6.3.42}$$

and

$$\frac{\widehat{\alpha}}{2d\mathsf{K}_0\sqrt{2}\kappa_0^\nu} \overset{(6.3.42)+(6.3.31)}{\leqslant} \frac{1}{2d \cdot 8^\nu \mathsf{K}_0 \sqrt{2}} \frac{r_0\sigma_0}{\mathsf{T}_0\sqrt{|\varepsilon|}} < \frac{\hat{r}_0\sigma_0}{32 d\eta_0} \ ,$$

we get

$$\hat{r}_1 = \frac{1}{2} \min\left\{ \frac{\widehat{\alpha}}{2d\sqrt{2}\mathsf{K}_0\kappa_0^\nu} \ , \ \frac{\hat{r}_0\sigma_0}{32 d\eta_0} \right\} = \frac{\widehat{\alpha}}{4d\sqrt{2}\mathsf{K}_0\kappa_0^\nu} \tag{6.3.43}$$

and, thus

$$\begin{aligned}
|\varepsilon|\mathsf{L}_0(3\sigma_0^{-1}) &\leqslant 3|\varepsilon|M_0 \max\left\{\frac{32\sqrt{2}\mathsf{T}_0}{r_0 r_1}\sigma_0^{-(\nu+d)}\,,\ \mathsf{C}_4 \max\left\{1, \frac{\alpha}{r_0\mathsf{K}_0}\right\}\frac{\mathsf{K}_0}{\alpha^2}\sigma_0^{-2(\nu+d)}\right\}\sigma_0^{-1}\\
&\leqslant 3\max\left\{32\sqrt{2}\mathsf{T}_0\frac{\alpha}{r_1}\frac{\alpha}{r_0\mathsf{K}_0}\,,\ \mathsf{C}_4\max\left\{1, \frac{\alpha}{r_0\mathsf{K}_0}\right\}\right\}\sigma_0^{-2(\nu+d)-1}\frac{\mathsf{K}_0 M_0}{\widehat{\alpha}^2}\\
&\overset{(6.3.31)}{\leqslant} 3\max\left\{32\sqrt{2}\mathsf{T}_0\frac{\alpha}{r_1}\,,\ \mathsf{C}_4\right\}\sigma_0^{-2(\nu+d)-1}\mu_0\\
&= 3\max\left\{256 d\eta_0\kappa_0^{\nu}\,,\ \mathsf{C}_4\right\}\sigma_0^{-2(\nu+d)-1}\mu_0\\
&\overset{(6.3.42)}{\leqslant} 3\max\left\{256 d\,,\ 8^{-\nu}\mathsf{C}_4\right\}\eta_0\kappa_0^{\nu}\sigma_0^{-2(\nu+d)-1}\mu_0\\
&\leqslant \mathsf{e}_*\,\mu_0\\
&= \theta_0 \overset{(6.3.31)}{\leqslant} 1.
\end{aligned} \tag{6.3.44}$$

Therefore, Lemma 6.3.2 is a straightforward consequence of Lemma 6.3.1. ∎

Lemma 6.3.3 *Assume* (6.3.35) ÷ (6.3.38) *with some* $\varepsilon \neq 0$ *and*

$$\max\left\{e\,\mu_0\,,\ 2\mathsf{C}_6\,\eta_0^{\frac{5}{4}}\,\sigma_0^{-\frac{5}{4}}\,\theta_0\right\} \leqslant 1\,. \tag{6.3.45}$$

Then, one can construct a sequence of real–analytic diffeomorphisms

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

and of real–analytic symplectic transformations

$$\phi_j : D_{r_j,s_j}(\mathscr{D}_j) \to D_{r_{j-1},s_{j-1}}(\mathscr{D}_{j-1})\,, \tag{6.3.46}$$

such that

$$\begin{aligned}
&G_j(\mathscr{D}_{j-1}) = \mathscr{D}_j \subset \mathscr{D}_{r_j}\,,\\
&\partial_y K_{j+1}\circ G_{j+1} = \partial_y K_j\,,\\
&H_j := H_{j-1}\circ\phi_j =: K_j + \varepsilon^{2^j} P_j \qquad \text{on} \quad D_{r_j,s_j}(\mathscr{D}_j)\,,
\end{aligned}$$

converge uniformly. More precisely, we have the following:

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

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

$$\phi^*\colon \mathscr{D}_* \times \mathbb{T}^d \xrightarrow{into} B_{r_1}(\mathscr{D}_1) \times \mathbb{T}^d,$$

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*

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

Finally, the following estimates hold for any $i \geqslant 2$*:*

$$\|G_i - \mathrm{id}\|_{\tilde r_i,\mathscr{D}_{i-1}} \leqslant 2\ r_i\ \sigma_{i-1}^{\nu+d}\ |\varepsilon|^{2^{i-1}} \mathsf{L}_{i-1}\ , \tag{6.3.47}$$

$$\|\partial_z G_i - \mathbb{1}_d\|_{\tilde r_i,\mathscr{D}_{i-1}} \leqslant \sigma_{i-1}^{\nu+d}\ |\varepsilon|^{2^{i-1}} \mathsf{L}_{i-1}\ , \tag{6.3.48}$$

$$2^{i^2}|\varepsilon|^{2^i} M_i := 2^{i^2}|\varepsilon|^{2^i}\|P_i\|_{r_i,s_i,\mathscr{D}_i} \leqslant |\varepsilon|\widehat{M}_i\ , \tag{6.3.49}$$

$$|\operatorname{meas}(\mathscr{D}_*) - \operatorname{meas}(\mathscr{D}_{\delta,\alpha})| \leqslant \mathsf{C}_7\ \sigma_0^{\nu+d}\ \theta_0\ \operatorname{meas}(\mathscr{D}_{\delta,\alpha})\ , \tag{6.3.50}$$

$$|\mathsf{W}_2(\phi^* - \mathrm{id})| \leqslant \frac{\theta_0}{3\cdot 2^d} \qquad \text{on} \quad \mathscr{D}_* \times \mathbb{T}^d_{s_*}\ . \tag{6.3.51}$$

Proof First of all, notice that

$$\hat r_{i+1} = \hat r_1 \frac{\sigma_1\cdots\sigma_i}{(64d\eta_0)^i} = 2^{-\frac{i^2}{2}}\left(\frac{\sigma_0}{64d\sqrt{2}\eta_0}\right)^i \hat r_1\ , \qquad \forall\ i \geqslant 0\ . \tag{6.3.52}$$

Indeed, for any $j \geqslant 1$, we have

$$\frac{\sigma_1\cdots\sigma_j}{(64d\eta_0)^j} \leqslant \left(\frac{\sigma_1}{64d\eta_0}\right)^j \leqslant \left(\frac{\sigma_0}{2^7 d}\right)^j \leqslant \frac{1}{4^{j\nu}}\ , \tag{6.3.53}$$

so that

$$\hat r_2 = \min\left\{\frac{\widehat\alpha}{4d\sqrt{2}\mathsf{K}_0\kappa_1^\nu}\ ,\ \frac{\hat r_1\sigma_1}{64d\eta_0}\right\} = \hat r_1 \min\left\{\frac{1}{4^\nu}\ ,\ \frac{\sigma_1}{64d\eta_0}\right\} \overset{(6.3.53)}{=} \hat r_1\frac{\sigma_1}{64d\eta_0}\ ,$$

and if

$$\hat r_{i+1} = \hat r_1 \frac{\sigma_1\cdots\sigma_i}{(64d\eta_0)^i}\ , \qquad i \geqslant 1\ ,$$

then

$$\begin{aligned}\hat{r}_{i+2} &= \min\left\{\frac{\widehat{\alpha}}{4d\sqrt{2}\mathsf{K}_0\kappa_{i+1}^{\nu}}\,,\,\frac{\hat{r}_{i+1}\sigma_{i+1}}{64d\eta_0}\right\}\\ &= \min\left\{\frac{\hat{r}_1}{4^{\nu(i+1)}}\,,\,\hat{r}_1\frac{\sigma_1\cdots\sigma_{i+1}}{(64d\eta_0)^{i+1}}\right\}\\ &\overset{(6.3.53)}{=} \hat{r}_1\frac{\sigma_1\cdots\sigma_{i+1}}{(64d\eta_0)^{i+1}}\,,\end{aligned}$$

and (6.3.52) is proven.

For a given $j \geqslant 2$, let $(\mathscr{P}^j)$ be the following assertion: there exist $j-1$ real–analytic diffeomorphisms

$$G_{i+1}\colon D_{\tilde{r}_{i+1}}(\mathscr{D}_i)\to G_{i+1}(D_{\tilde{r}_{i+1}}(\mathscr{D}_i))\,,\quad \text{for}\quad 1\leqslant i\leqslant j-1\,,$$

$j-1$ real–analytic symplectic transformations

$$\phi_{i+1} : D_{r_{i+1},s_{i+1}}(\mathscr{D}_{i+1}) \to D_{2r_i/3,s_i}(\mathscr{D}_i), \tag{6.3.54}$$

and $j-1$ Hamiltonians $H_{i+1} = H_i\circ\phi_{i+1} = K_{i+1}+\varepsilon^{2^{i+1}}P_{i+1}$ real–analytic on $D_{r_{i+1},s_{i+1}}(\mathscr{D}_{i+1})$ such that, for any $1\leqslant i\leqslant j-1$,

$$\begin{cases} G_i(\mathscr{D}_{i-1}) = \mathscr{D}_i \subset \mathscr{D}_{r_i}\,,\\ \|\partial_y^2 K_i\|_{r_i,\mathscr{D}_i} \leqslant \mathsf{K}_i\,,\\ \|T_i\|_{\mathscr{D}_i} \leqslant \mathsf{T}_i\,,\\ 2^{i^2}|\varepsilon|^{2^i}\|P_i\|_{r_i,s_i,\mathscr{D}_i} \leqslant |\varepsilon|\widehat{M}_i\,,\\ \kappa_i \geqslant 4\sigma_i^{-1}\log\left(\sigma_i^{2\nu+d}\mu_i^{-1}\right)\,,\\ |\varepsilon|^{2^i}\mathsf{L}_i \leqslant \frac{\sigma_i}{3}\end{cases} \tag{6.3.55}$$

and

$$\begin{cases} \partial_y K_{i+1} \circ G_{i+1} = \partial_y K_i \ , \\ \|G_{i+1} - \mathrm{id}\|_{\tilde{r}_{i+1},\mathscr{D}_i} \leqslant 2 r_{i+1} \sigma_i^{\nu+d} |\varepsilon|^{2^i} \mathsf{L}_i \ , \\ \|\partial_z G_{i+1} - \mathbb{1}_d\|_{\tilde{r}_{i+1},\mathscr{D}_i} \leqslant \sigma_i^{\nu+d} |\varepsilon|^{2^i} \mathsf{L}_i \ , \\ \|T_{i+1}\|_{\mathscr{D}_{i+1}} \leqslant \|T_i\|_{\mathscr{D}_i} + \mathsf{T}_i |\varepsilon|^{2^i} \mathsf{L}_i \ , \\ \|K_{i+1}\|_{r_{i+1},\mathscr{D}_{i+1}} \leqslant \|K_i\|_{r_i,\mathscr{D}_i} + |\varepsilon|^{2^i} M_i \ , \\ \|\partial_y^2 K_{i+1}\|_{r_{i+1},\mathscr{D}_{i+1}} \leqslant \|\partial_y^2 K_i\|_{r_i,\mathscr{D}_i} + \mathsf{K}_i |\varepsilon|^{2^i} \mathsf{L}_i \ , \\ \|\mathsf{W}_{i+1}(\phi_{i+1} - \mathrm{id})\|_{r_{i+1},s_{i+1},\mathscr{D}_{i+1}} \leqslant \sigma_i^d \ |\varepsilon|^{2^i} \mathsf{L}_i \ , \\ M_{i+1} := \|P_{i+1}\|_{r_{i+1},s_{i+1},\mathscr{D}_{i+1}} \leqslant M_i \mathsf{L}_i \ . \end{cases} \tag{6.3.56}$$

Assume $(\mathscr{P}^j)$, for some $j \geqslant 2$ and let us check $(\mathscr{P}^{j+1})$. Fix then $1 \leqslant i \leqslant j-1$. Thus

$$\|\partial_y^2 K_{i+1}\|_{r_{i+1},\mathscr{D}_{i+1}} \overset{(6.3.56)}{\leqslant} \|\partial_y^2 K_i\|_{r_i,\mathscr{D}_i} + \mathsf{K}_i |\varepsilon|^{2^i} \mathsf{L}_i \overset{(6.3.55)}{\leqslant} \mathsf{K}_i + \mathsf{K}_i \frac{\sigma_i}{3} = \mathsf{K}_{i+1} < \sqrt{2} \mathsf{K}_0$$

and, similarly,

$$\|T_{i+1}\|_{\mathscr{D}_{i+1}} \leqslant \mathsf{T}_{i+1},$$

which prove the second and third relations in (6.3.55) for $i = j$. Therefore

$$\frac{\alpha}{r_{i+1} \mathsf{K}_{i+1}} > \frac{\alpha}{r_1 \mathsf{K}_0 \sqrt{2}} = \frac{\widehat{\alpha}}{\hat{r}_1 \mathsf{K}_0 \sqrt{2}} = 4 d \kappa_0^\nu > 1 \tag{6.3.57}$$

so that

$$
\begin{aligned}
|\varepsilon|^{2^i}\mathsf{L}_i(3\sigma_i^{-1}) &= 3|\varepsilon|^{2^i} M_i \max\left\{\frac{32\sqrt{2}\mathsf{T}_0}{r_i r_{i+1}}\sigma_i^{-(\nu+d)}\,,\ \mathsf{C}_4 \max\left\{1, \frac{\alpha}{r_i \mathsf{K}_i}\right\}\frac{\mathsf{K}_0}{\alpha^2}\sigma_i^{-2(\nu+d)}\right\}\sigma_i^{-1} \\
&\overset{(6.3.57)}{\leqslant} 3|\varepsilon|^{2^i} M_i \max\left\{\frac{32\sqrt{2}\mathsf{T}_0}{r_i r_{i+1}}\,,\ \mathsf{C}_4\frac{1}{\alpha r_i}\right\}\sigma_i^{-2(\nu+d)-1} \\
&= 3\max\left\{32\sqrt{2}\mathsf{T}_0\frac{\widehat{\alpha}}{\hat{r}_{i+1}}\,,\ \mathsf{C}_4\right\}\sigma_i^{-2(\nu+d)-1}\frac{|\varepsilon|^{2^i}M_i}{\alpha r_i} \\
&\overset{(6.3.52)}{=} 3\max\left\{128 d\sqrt{2}\eta_0\kappa_0^{\nu}\cdot 2^{\frac{i^2}{2}}\left(\frac{64d\sqrt{2}\eta_0}{\sigma_0}\right)^i\,,\ \mathsf{C}_4\right\}\sigma_i^{-2(\nu+d)-1}\frac{|\varepsilon|^{2^i}M_i}{\alpha r_i} \\
&\overset{(6.3.42)}{\leqslant} 3\max\left\{128 d\sqrt{2}\,,\ 8^{-\nu}\mathsf{C}_4\right\}\sigma_i^{-2(\nu+d)-1}\frac{|\varepsilon|^{2^i}M_i}{\alpha r_i}2^{\frac{i^2}{2}}\left(\frac{64d\sqrt{2}\eta_0}{\sigma_0}\right)^i\eta_0\kappa_0^{\nu} \\
&= 3d\cdot 2^{4\nu+2d+7}\sqrt{2}\max\left\{128 d\sqrt{2}\,,\ 8^{-\nu}\mathsf{C}_4\right\}\sigma_0^{-(3\nu+2d+2)}\left(\frac{2^{2\nu+2d+7}d\sqrt{2}\eta_0}{\sigma_0}\right)^{i-1}\times \\
&\qquad\times\frac{2^{\frac{i^2}{2}}|\varepsilon|^{2^i}M_i}{\alpha r_i}\eta_0^2\lambda_0^{\nu} \\
&\overset{(6.3.52)}{=} 3d^2\cdot 2^{6\nu+2d+10}\max\left\{128 d\sqrt{2}\,,\ 8^{-\nu}\mathsf{C}_4\right\}\mathsf{K}_0\sigma_0^{-(4\nu+2d+2)}\left(\frac{2^{2\nu+2d+14}d^2\eta_0^2}{\sigma_0^2}\right)^{i-1}\times \\
&\qquad\times\frac{2^{\frac{i^2}{2}+\frac{(i-1)^2}{2}}|\varepsilon|^{2^i}M_i}{|\varepsilon|\widehat{\alpha}^2}\eta_0^2\lambda_0^{2\nu} \\
&= \mathsf{C}_5\sigma_0^{-(4\nu+2d+2)}\eta_0^2\lambda_0^{2\nu}\mathsf{d}_*^{i-1}\mathsf{K}_0\frac{2^{i^2}|\varepsilon|^{2^i}M_i}{|\varepsilon|\widehat{\alpha}^2} \\
&\overset{(6.3.55)}{\leqslant} \mathsf{C}_5\sigma_0^{-(4\nu+2d+2)}\eta_0^2\lambda_0^{2\nu}\mathsf{d}_*^{i-1}\frac{\mathsf{K}_0\widehat{M}_i}{\widehat{\alpha}^2} \\
&= \mathsf{e}_*\ \mathsf{d}_*^{i-1}\mu_i \\
&= \frac{\theta_i}{\mathsf{d}_*} \\
&= \frac{(\sqrt{\mathsf{d}_*}\ \theta_0)^{2^i}}{\mathsf{d}_*} \\
&\overset{(6.3.45)}{\leqslant} \frac{1}{\mathsf{d}_*} < 1\,.
\end{aligned}
\tag{6.3.58}
$$

Moreover,

$$
|\varepsilon|^{2^i}\mathsf{L}_i < \mathsf{e}_*\ \mathsf{d}_*^{i-1}\mu_i\ ,
$$

thus by last relation in (6.3.56), for any $1 \leqslant i \leqslant j-1$,

$$2^{(i+1)^2}|\varepsilon|^{2^{i+1}} M_{i+1} \leqslant (2^{2i+1}|\varepsilon|^{2^i}\mathsf{L}_i)(2^{i^2}|\varepsilon|^{2^i} M_i) \overset{(6.3.55)}{\leqslant} (8\mathsf{e}_*(4\mathsf{d}_*)^{i-1}\ \mu_i)(|\varepsilon|\ \widehat{M}_i) = |\varepsilon|\ \widehat{M}_{i+1}\ ,$$

which proves the fourth relation in (6.3.55) for $i=j$. Hence, by exactly the same computation as above, one gets

$$|\varepsilon|^{2^{i+1}}\mathsf{L}_{i+1}(3\sigma_{i+1}^{-1}) \leqslant \frac{\theta_{i+1}}{\mathsf{d}_*} = \frac{(\sqrt{\mathsf{d}_*}\ \theta_0)^{2^{i+1}}}{\mathsf{d}_*} < 1\ ,$$

which proves the last relation in (6.3.55) for $i=j$. It remains only to check that the fifth relation in (6.3.55) holds as well for $i=j$ in order to apply Lemma 6.3.1 to H_i, $1 \leqslant i \leqslant j$ and get (6.3.56) and, consequently, $(\mathscr{P}^{j+1})$. But in fact, we have, for any $i \geqslant 1$,

$$\mathsf{d}_*^{\frac{1}{2}} \leqslant \mathsf{e}_* \implies 4^i\mathsf{d}_*^{\frac{1}{2}i} \leqslant (8\mathsf{e}_*)^i \leqslant (8\mathsf{e}_*)^{2^i-1} \implies 8\mathsf{e}_*(4\mathsf{d}_*)^i \leqslant (8\mathsf{e}_*)^{2^i}\mathsf{d}_*^{\frac{1}{2}i} \leqslant (8\mathsf{e}_*)^{2^i}\mathsf{d}_*^{2^{i-2}} \implies 8\mathsf{e}_*(4\mathsf{d}_*)^i\mathsf{d}_*^{-2^{i-1}} \leqslant (8\mathsf{e}_*)^{2^i}\mathsf{d}_*^{-2^{i-2}} < (8\mathsf{e}_*)^{2^i},$$

so that

$$\begin{aligned}
4\sigma_i^{-1}\log\left(\sigma_i^{2\nu+d}\mu_i^{-1}\right) &\leqslant 4\sigma_i^{-1}\log\left(\mu_i^{-1}\right)\\
&= 4\sigma_i^{-1}\log\left(8\mathsf{e}_*\ (4\mathsf{d}_*)^i(\sqrt{\mathsf{d}_*}\ \theta_0)^{-2^i}\right)\\
&\leqslant 4\sigma_i^{-1}\log\left(\left(\frac{\theta_0}{8\mathsf{e}_*}\right)^{-2^i}\right)\\
&= 4\sigma_i^{-1}\log\left(\mu_0^{-2^i}\right)\\
&= 4^i\cdot 4\sigma_0^{-1}\log\left(\mu_0^{-1}\right)\\
&= \kappa_i\ .
\end{aligned} \tag{6.3.59}$$

To finish the proof of the induction *i.e.* one can construct an *infinite sequence* of Arnold's transformations satisfying (6.3.55) and (6.3.56) *for all* $i \geqslant 1$, one needs only to check $(\mathscr{P}^2)$. But,[117] (6.3.35) ÷ (6.3.38), $(6.3.58)_{i=1}$ and $(6.3.59)_{i=1}$ imply $(6.3.55)_{i=1}$. Thus, we apply Lemma 6.3.1 to H_1 to achieve the proof of $(\mathscr{P}^2)$.
Next, we show that G^j converges. For any $j \geqslant 1$,

$$\begin{aligned}
\|G^{j+1}-G^j\|_{\mathscr{D}_0} &= \|G_{j+1}\circ G^j - G^j\|_{\mathscr{D}_0}\\
&= \|G_{j+1}-\mathrm{id}\|_{\mathscr{D}_j}\\
&\leqslant \|G_{j+1}-\mathrm{id}\|_{\tilde{r}_{j+1},\mathscr{D}_j}\\
&\overset{(6.3.56)}{\leqslant} 2r_{j+1}\sigma_j^{\nu+d}|\varepsilon|^{2^j}\mathsf{L}_j.
\end{aligned} \tag{6.3.60}$$

[117] Observe that for $j=2$, $i=1$.

Thus, G^j is Cauchy and therefore converges uniformly on $\mathscr{D}_{\delta,\alpha}$ to a map G^*.
Next, we prove that ϕ^j is convergent by showing that it is Cauchy as well. For any $j \geqslant 4$, we have, using again Cauchy's estimate,

$$\begin{aligned}
\|\mathsf{W}_{j-1}(\phi^{j-1}-\phi^{j-2})\|_{r_j,s_j,\mathscr{D}_j} &= \|\mathsf{W}_{j-1}\phi^{j-2}\circ\phi_{j-1}-\mathsf{W}_{j-1}\phi^{j-2}\|_{r_{j-1},s_{j-1},\mathscr{D}_{j-1}}\\
&\overset{(6.3.54)}{\leqslant} \|\mathsf{W}_{j-1}D\phi^{j-2}\mathsf{W}_{j-1}^{-1}\|_{2r_{j-2}/3,s_{j-2},\mathscr{D}_{j-2}}\|\mathsf{W}_{j-1}(\phi_{j-1}-\mathrm{id})\|_{r_{j-1},s_{j-1},\mathscr{D}_{j-1}}\\
&\overset{(6.3.56)}{\leqslant} \max\left(r_{j-1}\frac{3}{r_{j-1}},\frac{3}{2\sigma_{j-1}}\right)\|\mathsf{W}_{j-1}\phi^{j-2}\|_{r_{j-1},s_{j-1},\mathscr{D}_{j-1}}\times\\
&\qquad\times\|\mathsf{W}_{j-1}(\phi_{j-1}-\mathrm{id})\|_{r_{j-1},s_{j-1},\mathscr{D}_{j-1}}\\
&= \frac{3}{2\sigma_{j-1}}\|\mathsf{W}_{j-1}\phi^{j-2}\|_{r_{j-1},s_{j-1},\mathscr{D}_{j-1}}\|\mathsf{W}_{j-1}(\phi_{j-1}-\mathrm{id})\|_{r_{j-1},s_{j-1},\mathscr{D}_{j-1}}\\
&\leqslant \frac{1}{2}\|\mathsf{W}_{j-1}\phi^{j-2}\|_{r_{j-1},s_{j-1},\mathscr{D}_{j-1}}\cdot\sigma_{j-2}^d\left(|\varepsilon|^{2^{j-2}}\mathsf{L}_{j-2}3\sigma_{j-2}^{-1}\right)\\
&\leqslant \frac{1}{2}\|\mathsf{W}_{j-1}\phi_2\|_{r_2,s_2,\mathscr{D}_2}\cdot\sigma_{j-2}^d\,\theta_{j-2}\\
&\leqslant \frac{1}{2}\left(\prod_{i=2}^{j-2}\|\mathsf{W}_{i+1}\mathsf{W}_i^{-1}\|\right)\|\mathsf{W}_2\phi_2\|_{r_2,s_2,\mathscr{D}_2}\cdot\sigma_{j-2}^d\,\theta_{j-2}\\
&\overset{(6.3.57)}{=} \frac{1}{2}\left(\prod_{i=2}^{j-2}\frac{r_{i-1}}{r_i}\right)\|\mathsf{W}_2\phi_2\|_{r_2,s_2,\mathscr{D}_2}\cdot\sigma_{j-2}^d\,\theta_{j-2}\\
&= \frac{r_1}{2r_{j-2}}\|\mathsf{W}_2\phi_2\|_{r_2,s_2,\mathscr{D}_2}\cdot\sigma_{j-2}^d\,\theta_{j-2}\\
&= \frac{1}{2}\sigma_1^d\,\|\mathsf{W}_2\phi_2\|_{r_2,s_2,\mathscr{D}_2}\cdot 2^{\frac{1}{2}(j-3)^2}\left(\frac{2^{6-d}d\sqrt{2}\eta_0}{\sigma_0}\right)^{j-3}\cdot(\sqrt{\mathsf{d}_*}\,\theta_0)^{2^{j-2}}\\
&\leqslant \frac{1}{2}\sigma_1^d\,\|\mathsf{W}_2\phi_2\|_{r_2,s_2,\mathscr{D}_2}\cdot 2^{2^{j-3}}\left(\frac{2^{6-d}d\sqrt{2}\eta_0}{\sigma_0}\right)^{2^{j-4}}\cdot(\sqrt{\mathsf{d}_*}\,\theta_0)^{2^{j-2}}\\
&= \frac{1}{2}\sigma_1^d\,\|\mathsf{W}_2\phi_2\|_{r_2,s_2,\mathscr{D}_2}\cdot\left(\left(\frac{2^{8-d}d\sqrt{2}\eta_0}{\sigma_0}\right)^{\frac{1}{4}}\sqrt{\mathsf{d}_*}\,\theta_0\right)^{2^{j-2}}\\
&= \frac{1}{2}\sigma_1^d\,\|\mathsf{W}_2\phi_2\|_{r_2,s_2,\mathscr{D}_2}\cdot\left(\mathsf{C}_6\eta_0^{\frac{5}{4}}\sigma_0^{-\frac{5}{4}}\theta_0\right)^{2^{j-2}}.
\end{aligned}$$

Therefore, for any $n \geqslant 2$, $j \geqslant 0$,

$$\begin{aligned}
\|\mathsf{W}_2(\phi^{n+j+1}-\phi^n)\|_{r_{n+j+1},s_{n+j+1},\mathscr{D}_{n+j+1}} &\leqslant \sum_{i=n}^{n+j} \|\mathsf{W}_2(\phi^{i+1}-\phi^i)\|_{r_{i+1},s_{i+1},\mathscr{D}_{i+1}}\\
&\leqslant \sum_{i=n}^{n+j}\left(\prod_{k=2}^{i}\|\mathsf{W}_k\mathsf{W}_{k+1}^{-1}\|\right)\|\mathsf{W}_{i+1}(\phi^{i+1}-\phi^i)\|_{r_{i+1},s_{i+1},\mathscr{D}_{i+1}}\\
&\overset{(6.3.57)}{=} \sum_{i=n}^{n+j}\prod_{k=2}^{i}\max\left\{1\,,\frac{r_{k+1}}{r_k}\right\}\|\mathsf{W}_{i+1}(\phi^{i+1}-\phi^i)\|_{r_{i+1},s_{i+1},\mathscr{D}_{i+1}}\\
&= \sum_{i=n}^{n+j}\|\mathsf{W}_{i+1}(\phi^{i+1}-\phi^i)\|_{r_{i+1},s_{i+1},\mathscr{D}_{i+1}}\\
&\leqslant \frac{1}{2}\sigma_1^d\,\|\mathsf{W}_2\phi_2\|_{r_2,s_2,\mathscr{D}_2}\sum_{i=n}^{n+j}\left(\mathsf{C}_6\eta_0^{\frac{5}{4}}\sigma_0^{-\frac{5}{4}}\theta_0\right)^{2^{i+1}} .
\end{aligned} \tag{6.3.61}$$

Hence ϕ^j converges uniformly on $\mathscr{D}_* \times \mathbb{T}^d$ to some ϕ^*, which is then real–analytic function in $x \in \mathbb{T}^d_{s_*}$.
To estimate $|\mathsf{W}_2(\phi^* - \mathrm{id})|$ on $\mathscr{D}_* \times \mathbb{T}^d_{s_*}$, observe that , for $i \geqslant 1$,[118]

$$\sigma_i^d\,|\varepsilon|^{2^i}\mathsf{L}_i \leqslant \frac{\sigma_0^{d+1}}{3\cdot 2^{i(d+1)}}\,\frac{(\sqrt{\mathsf{d}_*}\,\theta_0)^{2^i}}{\mathsf{d}_*} \leqslant \frac{1}{3\cdot 2^{(d+1)(i+1)}\mathsf{d}_*}(\sqrt{\mathsf{d}_*}\,\theta_0)^{i+1} = \frac{1}{3\mathsf{d}_*}\left(\frac{\sqrt{\mathsf{d}_*}\,\theta_0}{2^{d+1}}\right)^{i+1}$$

and therefore

$$\sum_{i\geqslant 1}|\varepsilon|^{2^i}\mathsf{L}_i \leqslant \frac{1}{3\mathsf{d}_*}\sum_{i\geqslant 1}\left(\frac{\sqrt{\mathsf{d}_*}\,\theta_0}{2}\right)^{i+1} \leqslant \frac{(\sqrt{\mathsf{d}_*}\,\theta_0)^2}{6\,\mathsf{d}_*} = \frac{\theta_0^2}{6}\,,$$

$$\sum_{i\geqslant 1}\sigma_i^d\,|\varepsilon|^{2^i}\mathsf{L}_i \leqslant \frac{1}{3\mathsf{d}_*}\sum_{i\geqslant 1}\left(\frac{\sqrt{\mathsf{d}_*}\,\theta_0}{2^{d+1}}\right)^{i+1} \leqslant \frac{(\sqrt{\mathsf{d}_*}\,\theta_0)^2}{3\cdot 2^{2d+1}\,\mathsf{d}_*} = \frac{\theta_0^2}{3\cdot 2^{2d+1}}\ .$$

Moreover, for any $i \geqslant 2$,

$$\begin{aligned}
\|\mathsf{W}_2(\phi^i-\mathrm{id})\|_{r_i,s_i,\mathscr{D}_i} &\leqslant \|\mathsf{W}_2(\phi^{i-1}\circ\phi_i-\phi_i)\|_{r_i,s_i,\mathscr{D}_i}+\|\mathsf{W}_2(\phi_i-\mathrm{id})\|_{r_i,s_i,\mathscr{D}_i}\\
&\leqslant \|\mathsf{W}_2(\phi^{i-1}-\mathrm{id})\|_{r_{i-1},s_{i-1},\mathscr{D}_{i-1}}+\left(\prod_{j=0}^{i-1}\|\mathsf{W}_j\mathsf{W}_{j+1}^{-1}\|\right)\|\mathsf{W}_i(\phi_i-\mathrm{id})\|_{r_i,s_i,\mathscr{D}_i}\\
&= \|\mathsf{W}_2(\phi^{i-1}-\mathrm{id})\|_{r_{i-1},s_{i-1},\mathscr{D}_{i-1}}+\|\mathsf{W}_i(\phi_i-\mathrm{id})\|_{r_i,s_i,\mathscr{D}_i}\\
&= \|\mathsf{W}_2(\phi^{i-1}-\mathrm{id})\|_{r_{i-1},s_{i-1},\mathscr{D}_{i-1}}+\|\mathsf{W}_i(\phi_i-\mathrm{id})\|_{r_i,s_i,\mathscr{D}_i}\\
&\leqslant \|\mathsf{W}_2(\phi^{i-1}-\mathrm{id})\|_{r_{i-1},s_{i-1},\mathscr{D}_{i-1}}+\sigma_{i-1}^d\,|\varepsilon|^{2^{i-1}}\mathsf{L}_{i-1}\ ,
\end{aligned}$$

[118] Recall that $2^i \geqslant i+1$, $\forall\, i \geqslant 0$ and $\sigma_0 \leqslant \frac{1}{2}$.

which iterated yields

$$\begin{aligned}\|\mathsf{W}_2(\phi^i-\mathrm{id})\|_{r_i,s_i,\mathscr{D}_i} &\leqslant \sum_{k=1}^{i-1}\sigma_k^d\,|\varepsilon|^{2^k}\mathsf{L}_k\\ &\leqslant \sum_{k\geqslant 1}\sigma_k^d\,|\varepsilon|^{2^k}\mathsf{L}_k\\ &\leqslant \frac{\theta_0^2}{3\cdot 2^{2d+1}}\,.\end{aligned}$$

Therefore, taking the limit over i completes the proof of (6.3.51).
Next, we show that $\|G^*-\mathrm{id}\|_{L,\mathscr{D}_{\delta,\alpha}}<1$, which will imply that[119] $G^*\colon \mathscr{D}_{\delta,\alpha}\overset{onto}{\longrightarrow}\mathscr{D}_*$ is a lipeomorphism. Indeed, for any $j\geqslant 2$, we have

$$\begin{aligned}\|G^j-\mathrm{id}\|_{L,\mathscr{D}_{\delta,\alpha}}+1 &= \|(G_j-\mathrm{id})\circ G^{j-1}+(G^{j-1}-\mathrm{id})\|_{L,\mathscr{D}_{\delta,\alpha}}+1\\ &\leqslant \|G_j-\mathrm{id}\|_{L,G^{j-1}(\mathscr{D}_{\delta,\alpha})}\|G^{j-1}\|_{L,\mathscr{D}_{\delta,\alpha}}+\|G^{j-1}-\mathrm{id}\|_{L,\mathscr{D}_{\delta,\alpha}}+1\\ &\leqslant \|G_j-\mathrm{id}\|_{L,G^{j-1}(\mathscr{D}_{\delta,\alpha})}(\|G^{j-1}-\mathrm{id}\|_{L,\mathscr{D}_{\delta,\alpha}}+1)+\|G^{j-1}-\mathrm{id}\|_{L,\mathscr{D}_{\delta,\alpha}}+1\\ &= (\|G_j-\mathrm{id}\|_{L,\mathscr{D}_{j-1}}+1)(\|G^{j-1}-\mathrm{id}\|_{L,\mathscr{D}_{\delta,\alpha}}+1)\\ &\leqslant (\|\partial_z G_j-\mathbb{1}_d\|_{\tilde r_j,\mathscr{D}_{j-1}}+1)(\|G^{j-1}-\mathrm{id}\|_{L,\mathscr{D}_{\delta,\alpha}}+1)\\ &\overset{(6.3.48)+(6.3.40)}{\leqslant} (\sigma_{j-1}^{\nu+d}|\varepsilon|^{2^{j-1}}\mathsf{L}_{j-1}+1)(\|G^{j-1}-\mathrm{id}\|_{L,\mathscr{D}_{\delta,\alpha}}+1)\end{aligned}$$

which iterated leads to[120]

$$\begin{aligned}\|G^j-\mathbb{1}_d\|_{L,\mathscr{D}_{\delta,\alpha}} &\leqslant -1+\prod_{i=1}^{\infty}(\sigma_{j-1}^{\nu+d}|\varepsilon|^{2^{i-1}}\mathsf{L}_{i-1}+1)\\ &\leqslant -1+\exp\left(\sum_{i=0}^{\infty}\sigma_i^{\nu+d}|\varepsilon|^{2^i}\mathsf{L}_i\right)\\ &\leqslant -1+\exp\left(\sigma_0^{\nu+d}|\varepsilon|\mathsf{L}_0+\sigma_0^{\nu+d}\sum_{i=1}^{\infty}|\varepsilon|^{2^i}\mathsf{L}_i\right)\\ &\leqslant -1+\exp\left(\sigma_0^{\nu+d}\theta_0+\sigma_0^{\nu+d}\frac{\theta_0^2}{6}\right)\\ &\leqslant -1+\exp\left(2\sigma_0^{\nu+d}\theta_0\right)\\ &\leqslant 2\sigma_0^{\nu+d}\theta_0\exp\left(2\sigma_0^{\nu+d}\theta_0\right)\\ &\overset{(6.3.45)}{<} \frac{e\;\sigma_0^{\nu+d}}{\mathsf{C}_6}<1\,. \qquad (6.3.62)\end{aligned}$$

[119] See Proposition II.2. in [Zeh10].
[120] Recall that $e^x-1\leqslant x\,e^x$, $\forall\, x\geqslant 0$.

Thus, letting $n \to \infty$, we get that G^* is Lipschitz continuous, with

$$\|G^* - \mathrm{id}\|_{L,\mathscr{D}_{\delta,\alpha}} \leqslant 2\sigma_0^{\nu+d}\theta_0 \exp\left(2\sigma_0^{\nu+d}\theta_0\right) < 2\, e\, \sigma_0^{\nu+d}\theta_0 < \frac{e\, \sigma_0^{\nu+d}}{\mathsf{C}_6} < 1,$$

so that, by[121] Lemma D.1 (see Appendix D), we get

$$\begin{aligned}
|\operatorname{meas}(\mathscr{D}_*) - \operatorname{meas}(\mathscr{D}_{\delta,\alpha})| &\leqslant \left(\left(1 + 2\sigma_0^{\nu+d}\theta_0 \exp\left(2\sigma_0^{\nu+d}\theta_0\right)\right)^d - 1\right) \operatorname{meas}(\mathscr{D}_{\delta,\alpha}) \\
&\leqslant d \cdot 2\sigma_0^{\nu+d}\theta_0 \exp\left(2\sigma_0^{\nu+d}\theta_0\right) \left(1 + 2\sigma_0^{\nu+d}\theta_0 \exp\left(2\sigma_0^{\nu+d}\theta_0\right)\right)^{d-1} \operatorname{meas}(\mathscr{D}_{\delta,\alpha}) \\
&\leqslant 2\, e\, d \left(\frac{3}{2}\right)^{d-1} \sigma_0^{\nu+d}\theta_0 \operatorname{meas}(\mathscr{D}_{\delta,\alpha}) \\
&= \mathsf{C}_7\, \sigma_0^{\nu+d}\, \theta_0\, \operatorname{meas}(\mathscr{D}_{\delta,\alpha}),
\end{aligned}$$

which proves (6.3.50).
Next, we show that $\phi^* \in C^\infty_W(\mathscr{D}_* \times \mathbb{T}^d)$. For any $n, j \geqslant 1$, we have

$$\begin{aligned}
\|G^{n+j} - G^j\|_{\mathscr{D}_{\delta,\alpha}} &\leqslant \sum_{k=j}^{n+j-1} \|G^{k+1} - G^k\|_{\mathscr{D}_{\delta,\alpha}} \\
&\overset{(6.3.60)}{\leqslant} 2 \sum_{k=j}^{n+j-1} r_{k+1}\sigma_k^{\nu+d}|\varepsilon|^{2^k}\mathsf{L}_k \\
&\leqslant 2r_{j+1}\sigma_j^{\nu} \sum_{k\geqslant 1} \sigma_k^d |\varepsilon|^{2^k}\mathsf{L}_k \\
&\leqslant 2r_{j+1}\sigma_j^{\nu}\, \frac{\theta_0^2}{3 \cdot 2^{2d+1}} \\
&\overset{(6.3.45)}{<} \sigma_j^{\nu} \frac{r_{j+1}}{16 d\eta_0} \\
&< \sigma_j^{\nu}\, \tilde{r}_{j+1}\, .
\end{aligned}$$

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

$$\|G^* - G^j\|_{\mathscr{D}_{\delta,\alpha}} < \sigma_j^{\nu}\, \tilde{r}_{j+1} < \frac{\tilde{r}_{j+1}}{4}\, . \tag{6.3.63}$$

Hence,[122] for any $j \geqslant 1$,

$$G^j\left(D_{\frac{\tilde{r}_{j+1}}{8}}(\mathscr{D}_{\delta,\alpha})\right) \overset{(6.3.62)}{\subset} D_{\frac{\tilde{r}_{j+1}}{4}}\left(G^j(\mathscr{D}_{\delta,\alpha})\right) \overset{(6.3.63)}{\subset} D_{\frac{\tilde{r}_{j+1}}{2}}(\mathscr{D}_*) \overset{(6.3.63)}{\subset} D_{\tilde{r}_{j+1}}(\mathscr{D}_j) \subset D_{r_j}(\mathscr{D}_j)\, . \tag{6.3.64}$$

[121] With $\delta := 2\sigma_0^{\nu+d}\theta_0 \exp\left(2\sigma_0^{\nu+d}\theta_0\right)$.
[122] Recall that, by definition, $G^j(\mathscr{D}_{\delta,\alpha}) = \mathscr{D}_j$ and $G^*(\mathscr{D}_{\delta,\alpha}) = \mathscr{D}_*$.

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

$$\begin{aligned}\sum_{j\geqslant 3}\|\mathsf{W}_2(\phi^j-\phi^{j-1})\|_{\tilde r_{j+1}/2,s_j,\mathscr{D}_*}\left(\frac{\tilde r_{j+1}}{2}\right)^{-n} &\overset{(6.3.64)}{\leqslant} 2^{n+4}d\,\eta_0^n\sum_{j\geqslant 3}\|\mathsf{W}_2(\phi^j-\phi^{j-1})\|_{r_j,s_j,\mathscr{D}_j}\,r_{j+1}^{-n}\\ &\overset{(6.3.61)+(6.3.52)}{\leqslant} 2^{n+3}d\,\eta_0^n\,\sigma_1^d\,r_1^{-n}\,\|\mathsf{W}_2\phi_2\|_{r_2,s_2,\mathscr{D}_2}\times\\ &\qquad\times\sum_{j\geqslant 3}\left(\mathsf{C}_6\eta_0^{\frac54}\sigma_0^{-\frac54}\theta_0\right)^{2^j}2^{n\frac{j^2}{2}}\left(\frac{64d\sqrt{2}\eta_0}{\sigma_0}\right)^{nj}\\ &< \infty\,,\end{aligned}$$

since, for j sufficiently large,

$$\left(\mathsf{C}_6\eta_0^{\frac54}\sigma_0^{-\frac54}\theta_0\right)^{2^j}2^{n\frac{j^2}{2}}\left(\frac{64d\sqrt{2}\eta_0}{\sigma_0}\right)^{nj}<\left(\sqrt{2}\mathsf{C}_6\eta_0^{\frac54}\sigma_0^{-\frac54}\theta_0\right)^{2^j}\qquad\text{and}\qquad \sqrt{2}\mathsf{C}_6\,\eta_0^{\frac54}\,\sigma_0^{-\frac54}\,\theta_0\overset{(6.3.45)}{\leqslant}\frac{1}{\sqrt{2}}<1.$$

Thus, writing

$$\phi^j=(\phi^j-\phi^{j-1})+\cdots+(\phi^3-\phi^2)\,,\qquad j\geqslant 3\,,$$

and invoking Lemma E.2 (see Appendix E), we conclude that $\phi^*\in C^\infty_W(\mathscr{D}_*\times\mathbb{T}^d)$.

Finally, we prove $G^*\in C^\infty_W(\mathscr{D}_{\delta,\alpha})$ 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

$$\begin{aligned}\sum_{j\geqslant 1}\|G^{j+1}-G^j\|_{\tilde r_{j+1}/8,\mathscr{D}_{\delta,\alpha}}\left(\frac{\tilde r_{j+1}}{8}\right)^{-n} &= 8^n\sum_{j\geqslant 1}\|(G_{j+1}-\mathrm{id})\circ G^j\|_{\tilde r_{j+1}/8,\mathscr{D}_{\delta,\alpha}}\,\tilde r_{j+1}^{-n}\\ &\overset{(6.3.64)}{\leqslant} 8^n\sum_{j\geqslant 2}\|G_{j+1}-\mathrm{id}\|_{\tilde r_{j+1},\mathscr{D}_j}\,\tilde r_{j+1}^{-n}\\ &\leqslant 2^{3n+1}\sum_{j\geqslant 1}r_{j+1}\,\tilde r_{j+1}^{-n}\,\sigma_j^{\nu+d}|\varepsilon|^{2^j}\mathsf{L}_j\\ &<\infty\,,\end{aligned}$$

which proves that $G^*\in C^\infty_W(\mathscr{D}_{\delta,\alpha})$.

Now, to complete the proof of Theorem 6.2.1, observe that, uniformly on $\mathscr{D}_* \times \mathbb{T}^d_{s_*}$,

$$\begin{aligned}|\mathsf{W}_1(\phi_* - \mathrm{id})| &\leqslant |\mathsf{W}_1(\phi_1 \circ \phi^* - \phi^*)| + |\mathsf{W}_1(\phi^* - \mathrm{id})|\\ &\leqslant \|\mathsf{W}_1(\phi_1 - \mathrm{id})\|_{r_1,s_1,\mathscr{D}_1} + \|\mathsf{W}_1\mathsf{W}_2^{-1}\|\ |\mathsf{W}_2(\phi^* - \mathrm{id})|\\ &\leqslant \sigma_0^d\ |\varepsilon|\mathsf{L}_0 + \frac{\theta_0^2}{3\cdot 2^{2d+1}}\\ &\leqslant \frac{\sigma_0^{d+1}}{3}\ \theta_0 + \frac{\theta_0^2}{3\cdot 2^{2d+1}}\\ &\leqslant \frac{1}{3\cdot 2^{d+1}}\ \theta_0 + \frac{\theta_0^2}{3\cdot 2^{2d+1}}\\ &\leqslant \frac{\theta_0}{3\cdot 2^d}\,.\end{aligned}$$

Moreover, setting $G_0 \coloneqq \mathrm{id}$, we have for any $i \geqslant 3$,

$$\begin{aligned}|G^i - \mathrm{id}|_{\mathscr{D}_{\delta,\alpha}} &\leqslant \sum_{j=0}^{i-1} |G^{j+1} - G^j|_{\mathscr{D}_{\delta,\alpha}}\\ &= \sum_{j=0}^{i-1} |G_j - \mathrm{id}|_{\mathscr{D}_{j-1}}\\ &\overset{(6.3.47)+(6.3.39)}{\leqslant} 2\sum_{j=0}^{i-1} r_{j+1}\sigma_j^{\nu+d}|\varepsilon|^{2^j}\mathsf{L}_j\\ &\leqslant 2\sigma_0^\nu\ r_1 \sum_{j=0}^{\infty} \sigma_j^d|\varepsilon|^{2^j}\mathsf{L}_j\\ &\leqslant \frac{2\ \sigma_0^\nu\ \theta_0}{3\cdot 2^d}\ r_1\ ,\end{aligned}$$

and then passing to the limit, we get

$$|G^* - \mathrm{id}|_{\mathscr{D}_{\delta,\alpha}} = |G^* - \mathrm{id}|_{\mathscr{D}_{\delta,\alpha}} \leqslant \frac{2\ \sigma_0^\nu\ \theta_0}{3\cdot 2^d}\ r_1 \leqslant \frac{\sigma_0^\nu}{\mathsf{C}_9}\left(\frac{\sigma_0}{\eta_0}\right)^{\frac{5}{4}}\frac{\alpha}{\mathsf{K}_0} = r_*\ .$$

Finally, we prove (6.2.7). By Theorem C.1, $G^* - \mathrm{id}$ can be extended to a global Lipschitz continuous function $f\colon \mathbb{R}^d \circlearrowleft$, with[123]

$$\sup_{\mathbb{R}^d} |f|_2 = \sup_{\mathscr{D}_{\delta,\alpha}} |G^* - id|_2\ , \tag{6.3.65}$$

$$\sup_{\substack{y,y'\in\mathbb{R}^d\\ y\neq y'}} \frac{|f(y) - f(y')|_2}{|y - y'|_2} = \sup_{\substack{y,y'\in\mathscr{D}_{\delta,\alpha}\\ y\neq y'}} \frac{|(G^* - id)(y) - (G^* - id)(y')|_2}{|y - y'|_2}\ . \tag{6.3.66}$$

[123] Where $|y|_2 \coloneqq \sqrt{y_1^2 + \cdots + y_d^2}$, and recall that $|y| \leqslant |y|_2 \leqslant \sqrt{d}\ |y|$, for any $y \in \mathbb{R}^d$.

Hence,

$$\begin{aligned}\|f\|_{\mathbb{R}^d} &\overset{def}{=} \sup_{\mathbb{R}^d}|f| \\ &\leqslant \sup_{\mathbb{R}^d}|f|_2 \\ &\overset{(6.4.1)}{=} \sup_{\mathscr{D}_{\delta,\alpha}}|G^*-id|_2 \\ &\leqslant \sqrt{d}\sup_{\mathscr{D}_{\delta,\alpha}}|G^*-id| \\ &\overset{(6.2.5)}{\leqslant} \sqrt{d}\, r_* \\ &\overset{(6.2.2)}{<} \frac{1}{32d}\frac{r_0\sigma_0}{\eta_0} \\ &\leqslant \delta\sigma_0 \end{aligned} \tag{6.3.67}$$

and

$$\begin{aligned}\|f\|_{L,\mathbb{R}^d} &\overset{def}{=} \sup_{\substack{y,y'\in\mathbb{R}^d\\ y\neq y'}} \frac{|f(y)-f(y')|}{|y-y'|} \\ &\leqslant \sup_{\substack{y,y'\in\mathbb{R}^d\\ y\neq y'}} \frac{|f(y)-f(y')|_2}{|y-y'|_2/\sqrt{d}} \\ &\overset{(6.3.66)}{=} \sqrt{d}\sup_{\substack{y,y'\in\mathscr{D}_{\delta,\alpha}\\ y\neq y'}} \frac{|(G^*-id)(y)-(G^*-id)(y')|_2}{|y-y'|_2} \\ &\leqslant \sqrt{d}\sup_{\substack{y,y'\in\mathscr{D}_{\delta,\alpha}\\ y\neq y'}} \frac{\sqrt{d}|(G^*-id)(y)-(G^*-id)(y')|}{|y-y'|} \\ &= d\|G^*-\mathrm{id}\|_{L,\mathscr{D}_{\delta,\alpha}} \\ &\overset{(6.2.6)}{\leqslant} d\frac{e\;\sigma_0^{\nu+d}}{\mathsf{C}_6} < \frac{1}{2}\ . \end{aligned} \tag{6.3.68}$$

Set $g := f + \mathrm{id}$. Then, by Lemma G.1,

$$\mathscr{D} \subset g\big(\overline{B}_{\delta\sigma_0}(\mathscr{D})\big)\ . \tag{6.3.69}$$

Notice also that, by (6.3.68),[124] g is a homeomorphism of $\mathbb{R}^d$. Consequently,

$$\begin{aligned}
\text{meas}\,(\mathscr{D}\times\mathbb{T}^d\backslash\mathscr{K}) &= \text{meas}\,(\mathscr{D}\times\mathbb{T}^d)-\text{meas}\,\big(\phi_*(\mathscr{D}_*\times\mathbb{T}^d)\big)\\
&= \text{meas}\,(\mathscr{D}\times\mathbb{T}^d)-\text{meas}\,(\mathscr{D}_*\times\mathbb{T}^d)\\
&\overset{(6.3.69)}{\leqslant} \text{meas}\,\big(g(\overline{B}_{\delta\sigma_0}(\mathscr{D}))\times\mathbb{T}^d\big)-\text{meas}\,(\mathscr{D}_*\times\mathbb{T}^d)\\
&= (2\pi)^d\,\text{meas}\,\big(g(\overline{B}_{\delta\sigma_0}(\mathscr{D}))\backslash g(\mathscr{D}_{\delta,\alpha})\big)\\
&= (2\pi)^d\,\text{meas}\,\big(g(\overline{B}_{\delta\sigma_0}(\mathscr{D})\backslash\mathscr{D}_{\delta,\alpha})\big) \qquad \text{(because g is injective)}\\
&\leqslant (2\pi)^d\|g\|^d_{L,\mathbb{R}^d}\,\text{meas}\,\big(\overline{B}_{\delta\sigma_0}(\mathscr{D})\backslash\mathscr{D}_{\delta,\alpha}\big)\\
&\leqslant (2\pi)^d(1+\|f\|_{L,\mathbb{R}^d})^d\,\text{meas}\,\big(\overline{B}_{\delta\sigma_0}(\mathscr{D})\backslash\mathscr{D}_{\delta,\alpha}\big)\\
&\overset{(6.3.68)}{\leqslant} (2\pi)^d\left(1+\frac{d\ e\ \sigma_0^{\nu+d}}{\mathsf{C}_6}\right)^d\text{meas}\,\big(\overline{B}_{\delta\sigma_0}(\mathscr{D})\backslash\mathscr{D}_{\delta,\alpha}\big)\\
&= (2\pi)^d\left(1+\frac{d\ e\ \sigma_0^{\nu+d}}{\mathsf{C}_6}\right)^d\Big(\text{meas}\,\big(\overline{B}_{\delta\sigma_0}(\mathscr{D})\backslash\mathscr{D}\big)+\text{meas}\,(\mathscr{D}\backslash\mathscr{D}_\delta)+\\
&\qquad\qquad +\text{meas}\,(\mathscr{D}_\delta\backslash\mathscr{D}_{\delta,\alpha})\Big)
\end{aligned}$$

∎

6.4 Sharp measure estimate of the complement of $\mathscr{K}$ in an arbitrary set

The strategy here is to localize the *Kolmogorov set* and then sum them up. Thus, we start by examining the cube case.

6.4.1 Local analysis: the case where $\mathscr{D}$ is a cube

Theorem 6.4.1 *Let*

$$\mathscr{D}=B_R(y_0)\ , \qquad R>0$$

[124] See [Zeh10, Proposition II.2.].

and let the assumptions in Theorem 6.2.1 hold, with

$$\mathsf{C} := \frac{1}{32} + \frac{d\sqrt{d}}{\mathsf{C}_9} + \sum_{0\neq k\in\mathbb{Z}^d} \frac{1}{|k|_1^\nu} \ ,$$
$$\delta \leqslant \min\left\{ \frac{\mathsf{T}_0}{32d\sigma_0}\,\alpha \ , \ \frac{\mathsf{r}_0}{32d} \ , \ \frac{R}{4} \right\} \ .$$

Furthermore, assume that

$$K_y\colon \mathscr{D} \to \Omega := K_y(\mathscr{D})$$

is a diffeomorphism. Then,

$$\operatorname{meas}\left(\mathscr{D}\times\mathbb{T}^d\backslash\mathscr{K}\right) \leqslant \mathsf{C}\,(6\pi)^d\,\frac{\vartheta_0\mathsf{T}_0}{\sigma_0}\,R^{d-1}\alpha \ ,$$

with[125]

$$\vartheta_0 := \frac{\mathsf{K}_0^d}{\varrho_0} \geqslant 1 \ , \qquad \varrho_0 := \inf_{y\in\mathscr{D}_\delta} |\det K_{yy}(y)| > 0 \ .$$

Proof We shall denote the *Euclidean norm* by[126]

$$|y|_2 := \sqrt{y_1^2 + \cdots + y_d^2} \ .$$

Recall that

$$\mathscr{D}_\delta \overset{def}{=} B_{R-\delta}(y_0) \overset{def}{=} \{y\in\mathbb{R}^d \ : \ |y-y_0| = \max_{1\leqslant j\leqslant d} |y_j - y_{0j}| < R-\delta\}$$

and $\phi_*(\mathscr{D}_*\times\mathbb{T}^d) \subset \mathscr{D}\times\mathbb{T}^d$. Therefore,

$$\begin{aligned}
\operatorname{meas}\left(\mathscr{D}\times\mathbb{T}^d\backslash\phi_*(\mathscr{D}_*\times\mathbb{T}^d)\right) &= \operatorname{meas}\left(\mathscr{D}\times\mathbb{T}^d\right) - \operatorname{meas}\left(\phi_*(\mathscr{D}_*\times\mathbb{T}^d)\right)\\
&= \operatorname{meas}\left(\mathscr{D}\times\mathbb{T}^d\right) - \operatorname{meas}\left(\mathscr{D}_*\times\mathbb{T}^d\right)\\
&= (2\pi)^d\big(\operatorname{meas}(\mathscr{D}) - \operatorname{meas}(\mathscr{D}_*)\big)\\
&\leqslant (2\pi)^d\big(\operatorname{meas}(\mathscr{D}\backslash\mathscr{D}_\delta) + \operatorname{meas}(\mathscr{D}_\delta\backslash\mathscr{D}_*)\big)\\
&= (2\pi)^d\big((2R)^d - (2R-2\delta)^d + \operatorname{meas}(\mathscr{D}_\delta\backslash G_*(\mathscr{D}_0))\big)\\
&\leqslant (2\pi)^d\big(2^d d\,R^{d-1}\delta + \operatorname{meas}(\mathscr{D}_\delta\backslash G_*(\mathscr{D}_0))\big) \ .
\end{aligned}$$

---

[125] Indeed, pick any matrix $A = [a_{ij}]_{1\leqslant i,j\leqslant d}$. Then $\|A\| = \max_{1\leqslant i\leqslant d} |a_{i1}| + \cdots + |a_{id}|$ and $|\det A| = |\sum_{\xi\in\,\Xi_d} a_{1\xi(1)}\cdots a_{d\xi(d)}| \leqslant \sum_{\xi\in\,\Xi_d} |a_{1\xi(1)}|\cdots|a_{d\xi(d)}| \leqslant \prod_{1=1}^d (|a_{i1}| + \cdots + |a_{id}|) \leqslant \|A\|^d$, where Ξ_d is the set of permutations of $\{1,\cdots,d\}$.

[126] Recall that $|y| \leqslant |y|_2 \leqslant \sqrt{d}\,|y|$, for any $y\in\mathbb{R}^d$.

It remains to estimate meas $(\mathscr{D}_\delta \backslash G_*(\mathscr{D}_0))$. Firstly, thanks to Theorem C.1 (see Appendix C), $G_* - \mathrm{id}$ can be extended to a global Lipschitz continuous function $f \colon \mathbb{R}^d \circlearrowleft$, with

$$\sup_{\mathbb{R}^d} |f|_2 = \sup_{\mathscr{D}_0} |G_* - id|_2 \ , \tag{6.4.1}$$

$$\sup_{\substack{y,y' \in \mathbb{R}^d \\ y \neq y'}} \frac{|f(y) - f(y')|_2}{|y - y'|_2} = \sup_{\substack{y,y' \in \mathscr{D}_0 \\ y \neq y'}} \frac{|(G_* - id)(y) - (G_* - id)(y')|_2}{|y - y'|_2} \ . \tag{6.4.2}$$

Hence,

$$\begin{aligned} \|f\|_{\mathbb{R}^d} &\overset{def}{=} \sup_{\mathbb{R}^d} |f| \\ &\leqslant \sup_{\mathbb{R}^d} |f|_2 \\ &\overset{(6.4.1)}{=} \sup_{\mathscr{D}_0} |G_* - id|_2 \\ &\leqslant \sqrt{d} \sup_{\mathscr{D}_0} |G_* - id| \\ &\overset{(6.2.5)}{\leqslant} \sqrt{d}\, r_* =: \hat{r} \end{aligned} \tag{6.4.3}$$

and

$$\begin{aligned} \|f\|_{L,\mathbb{R}^d} &\overset{def}{=} \sup_{\substack{y,y' \in \mathbb{R}^d \\ y \neq y'}} \frac{|f(y) - f(y')|}{|y - y'|} \\ &\leqslant \sup_{\substack{y,y' \in \mathbb{R}^d \\ y \neq y'}} \frac{|f(y) - f(y')|_2}{|y - y'|_2/\sqrt{d}} \\ &\overset{(6.4.2)}{=} \sqrt{d} \sup_{\substack{y,y' \in \mathscr{D}_0 \\ y \neq y'}} \frac{|(G_* - id)(y) - (G_* - id)(y')|_2}{|y - y'|_2} \\ &\leqslant \sqrt{d} \sup_{\substack{y,y' \in \mathscr{D}_0 \\ y \neq y'}} \frac{\sqrt{d}|(G_* - id)(y) - (G_* - id)(y')|}{|y - y'|} \\ &= d \|G_* - \mathrm{id}\|_{L,\mathscr{D}_0} \\ &\overset{(6.2.6)}{\leqslant} d \frac{e\, \sigma_0^{\nu+d}}{\mathsf{C}_6} < \frac{1}{2} \ . \end{aligned} \tag{6.4.4}$$

Set $g := f + \mathrm{id}$. Then, by Lemma G.1,

$$\mathscr{D}_\delta \subset g\left(\overline{B}_{\hat{r}}(\mathscr{D}_\delta)\right) = g\left(\overline{B_{R-\delta+\hat{r}}(y_0)}\right) \ . \tag{6.4.5}$$

Notice also that, by (6.4.4),[127] g is a homeomorphism of $\mathbb{R}^d$. Consequently,

$$\begin{aligned}
\operatorname{meas}(\mathscr{D}_\delta \backslash G_*(\mathscr{D}_0)) &\overset{(6.4.5)}{\leqslant} \operatorname{meas}\left(g\left(\overline{B}_{R-\delta+\hat{r}}(y_0)\right)\backslash G_*(\mathscr{D}_0)\right)\\
&\overset{\text{def}}{=} \operatorname{meas}\left(g\left(\overline{B}_{R-\delta+\hat{r}}(y_0)\right)\backslash g(\mathscr{D}_0)\right)\\
&= \operatorname{meas}\left(g\left(\overline{B}_{R-\delta+\hat{r}}(y_0)\backslash \mathscr{D}_0\right)\right) \qquad \text{(because } g \text{ is injective)}\\
&\leqslant \|g\|^d_{L,\mathbb{R}^d} \operatorname{meas}\left(\overline{B}_{R-\delta+\hat{r}}(y_0)\backslash \mathscr{D}_0\right)\\
&\leqslant (1+\|f\|_{L,\mathbb{R}^d})^d \Big(\operatorname{meas}\left(\overline{B}_{R-\delta+\hat{r}}(y_0)\backslash \mathscr{D}_\delta\right) + \operatorname{meas}(\mathscr{D}_\delta \backslash \mathscr{D}_0) \Big)\\
&\overset{(6.2.6)}{\leqslant} \left(\frac{3}{2}\right)^d \Big(2^d (R-\delta+\hat{r})^d - 2^d (R-\delta)^d + \operatorname{meas}(\mathscr{D}_\delta \backslash \mathscr{D}_0) \Big)\\
&\leqslant \left(\frac{3}{2}\right)^d \left(2^d d R^{d-1} \sqrt{d}\, \frac{\sigma_0^\nu}{\mathsf{C}_9} \left(\frac{\sigma_0}{\eta_0}\right)^{\frac{5}{4}} \frac{\alpha}{\mathsf{K}_0} + \operatorname{meas}(\mathscr{D}_\delta \backslash \mathscr{D}_0) \right)\\
&\leqslant R^{d-1} \frac{3^d\, d\, \sqrt{d}}{\mathsf{C}_9} \frac{\alpha}{\mathsf{K}_0} + \left(\frac{3}{2}\right)^d \operatorname{meas}(\mathscr{D}_\delta \backslash \mathscr{D}_0)\ .
\end{aligned}$$

Finally,

$$\begin{aligned}
\operatorname{meas}(\mathscr{D}_\delta \backslash \mathscr{D}_0) &= \int_{\mathscr{D}_\delta \backslash \mathscr{D}_0} dy\\
&= \int_{\{y \in \mathscr{D}_\delta \,:\, K_y(y) \notin \Delta^\tau_\alpha\}} dy\\
&= \int_{\{z \in K_y(B_{R-\delta}(y_0)) \,:\, z \notin \Delta^\tau_\alpha\}} |\det K_{yy}|^{-1} dz\\
&\leqslant \frac{1}{\varrho_0} \int_{\left\{z \in B_{(R-\delta)\|K_{yy}\|_{r_0,\mathscr{D}_0}}(K_y(y_0)) \,:\, z \notin \Delta^\tau_\alpha\right\}} dz\\
&= \frac{1}{\varrho_0} \int_{\bigcup\limits_{0\neq k \in \mathbb{Z}^d} \left\{ z \in B_{(R-\delta)\mathsf{K}_0}(K_y(y_0)) \;:\; |k\cdot z| < \frac{\alpha}{|k|_1^\tau} \right\}} dz\\
&\leqslant \frac{1}{\varrho_0} \sum_{0\neq k\in\mathbb{Z}^d} \int_{\left\{ z \in B_{(R-\delta)\mathsf{K}_0}(K_y(y_0)) \,:\, |k\cdot z| < \frac{\alpha}{|k|_1^\tau} \right\}} dz\\
&\leqslant \frac{1}{\varrho_0} \sum_{0\neq k\in\mathbb{Z}^d} (2(R-\delta)\mathsf{K}_0)^{d-1} \frac{2\alpha}{|k|_1^\nu}\\
&= \mathbf{a}_1\, 2^d\, (R-\delta)^{d-1}\, \alpha\ ,
\end{aligned}$$

[127] See [Zeh10, Proposition II.2.].

where

$$\mathbf{a}_1 := \frac{\mathsf{K}_0^{d-1}}{\varrho_0} \sum_{0 \neq k \in \mathbb{Z}^d} \frac{1}{|k|_1^\nu} \ .$$

Putting all together, we get

$$\begin{aligned}
\text{meas}\left(\mathscr{D} \times \mathbb{T}^d \backslash \phi_*(\mathscr{D}_* \times \mathbb{T}^d)\right) &\leqslant (2\pi)^d \left(2^d d \, R^{d-1}\delta + R^{d-1} \frac{3^d \, d\sqrt{d}}{\mathsf{C}_9} \frac{\alpha}{\mathsf{K}_0} + \mathbf{a}_1 \, 3^d \, (R-\delta)^{d-1} \, \alpha \right) \\
&= (2\pi)^d \left(2^d d \, \frac{\mathsf{K}_0 \delta}{\alpha} + \frac{3^d \, d\sqrt{d}}{\mathsf{C}_9} + 3^d \mathsf{K}_0 \, \mathbf{a}_1 \right) \frac{R^{d-1}\alpha}{\mathsf{K}_0} \\
&\leqslant (6\pi)^d \left(\mathsf{K}_0 \, \frac{d\delta}{\alpha} + \frac{d\sqrt{d}}{\mathsf{C}_9} + \sum_{0 \neq k \in \mathbb{Z}^d} \frac{1}{|k|_1^\nu} \right) \frac{\mathsf{K}_0^{d-1}}{\varrho_0} R^{d-1}\alpha \\
&\leqslant (6\pi)^d \left(\frac{\eta_0}{32\sigma_0} + \frac{d\sqrt{d}}{\mathsf{C}_9} + \sum_{0 \neq k \in \mathbb{Z}^d} \frac{1}{|k|_1^\nu} \right) \frac{\mathsf{K}_0^{d-1}}{\varrho_0} R^{d-1}\alpha \\
&\leqslant (6\pi)^d \left(\frac{1}{32} + \frac{d\sqrt{d}}{\mathsf{C}_9} + \sum_{0 \neq k \in \mathbb{Z}^d} \frac{1}{|k|_1^\nu} \right) \frac{\mathsf{T}_0 \mathsf{K}_0^d}{\sigma_0 \varrho_0} R^{d-1}\alpha \\
&= \mathsf{C} \, (6\pi)^d \, \frac{\vartheta_0 \mathsf{T}_0}{\sigma_0} \, R^{d-1}\alpha \ .
\end{aligned}$$

∎

6.4.2 Global analysis

Let $\mathfrak{D}$ be any non–empty, bounded subset of $\mathbb{R}^d$. Consider the Hamiltonian parametrized by $\varepsilon \in \mathbb{R}$

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

where K, P are real–analytic functions defined on $\mathfrak{D} \times \mathbb{T}^d$ with bounded holomorphic extensions to[128]

$$D_{\mathsf{r}_0, s_0}(\mathfrak{D}) := \bigcup_{y_0 \in \mathfrak{D}} D_{\mathsf{r}_0, s_0}(y_0) \, ,$$

for some $\mathsf{r}_0 > 0$ and $0 < s_0 \leqslant 1$, the norm being

$$\| \cdot \|_{\mathsf{r}_0, s_0, \mathfrak{D}} := \sup_{D_{\mathsf{r}_0, s_0}(\mathfrak{D})} | \cdot | \, .$$

[128]Recall the notations in §1.2

Assume that

$$\varrho := \inf_{y\in\mathfrak{D}} |\det K_{yy}(y)| > 0\ . \tag{6.4.6}$$

Fix $0 < s_* < s_0$ and define[129]

$$\begin{aligned}
\sigma_0 &:= 2^{7-2\nu} d(s_0 - s_*)\ ,\\
M &:= \|P\|_{\mathsf{r}_0,s_0,\mathfrak{D}}\ ,\\
\mathsf{T} &:= \|T\|_{\mathfrak{D}} := \sup_{y\in\mathfrak{D}} \|T(y)\|\ ,\\
\mathsf{K} &:= \|K_{yy}\|_{\mathsf{r}_0,\mathfrak{D}}\ ,\\
\eta &:= \mathsf{T}\mathsf{K} \geqslant 1\ ,\\
\vartheta &:= \frac{\mathsf{K}^d}{\varrho} \geqslant 1\ ,\\
\delta &:= \frac{\mathsf{T}}{32 d\sigma_0}\alpha\ ,\\
R_0 &:= \frac{\mathsf{r}_0}{64 d\eta^2}\ ,\\
r_* &:= \frac{\sigma_0^\nu}{\mathsf{C}_9}\left(\frac{\sigma_0}{\eta}\right)^{\frac{5}{4}} \frac{\alpha}{\mathsf{K}}\ ,\\
\mu_* &:= \sup\left\{\mu \leqslant e^{-1} : 2\ \mathsf{C}_5\ \mathsf{C}_6\ \sigma_0^{4\nu+2d+\frac{13}{4}}\ \eta^{\frac{13}{4}}\ \mu\ \left(\log \mu^{-1}\right)^{2\nu} \leqslant 1\right\}\ ,
\end{aligned}$$

where $T(y) := K_{yy}(y)^{-1}$. Let

$$\begin{aligned}
\mathsf{C} &:= \frac{1}{32} + \frac{d\sqrt{d}}{\mathsf{C}_9} + \sum_{0\neq k\in\mathbb{Z}^d} \frac{1}{|k|_1^\nu}\ ,\\
\widehat{\mathsf{C}} &:= 64\ d\ \mathsf{C}\ .
\end{aligned}$$

Given $R > 0$, define the set $\mathscr{C}_R$ of coverings of $\mathfrak{D}$ by cubes as follows: $F \in \mathscr{C}_R$ *iff* there exists $n_F \in \mathbb{N}$ and n_F cubes, say F_i, $1 \leqslant i \leqslant n_F$, of equal side–length $2R$, centered at a point $y_i \in \mathfrak{D}$ and such that

$$F = \left\{F_i : 1 \leqslant i \leqslant n_F\right\} \qquad \text{and} \qquad \mathfrak{D} \subset \bigcup_{i=1}^{n_F} F_i\ . \tag{6.4.7}$$

Then define

$$\mathscr{R} := \left\{0 < R \leqslant R_0 : \mathscr{C}_R \neq \varnothing \text{ and } \inf_{F\in\mathscr{C}_R} n_F (2R)^d \leqslant 2^d \operatorname{meas}(\mathfrak{D})\right\} \tag{6.4.8}$$

[129] Recall footnote[125].

and

$$n_{\mathfrak{D}} := \min_{R\in\mathscr{R}} \min\left\{ n_F : F \in \mathscr{C}_R \quad \text{and} \quad n_F R^d \leqslant \text{meas}\,(\mathfrak{D}) \right\} . \tag{6.4.9}$$

Pick any $R_0' \in \mathscr{R}$ and $F^0 \in \mathscr{C}_{R_0'}$ such that $n_{F^0} = n_{\mathfrak{D}}$. Then

$$F^0 = \left\{ F_i^0 := B_{R_0'}(p_i') \ , \quad \text{for some} \quad p_i' \in \mathfrak{D} \ , \quad 1 \leqslant i \leqslant n_{\mathfrak{D}} \right\} .$$

Thus, the following holds.

Theorem 6.4.2 *Let the above assumptions and notations hold. Assume*

$$\boxed{\alpha \leqslant 8d\frac{R_*\sigma_0}{\mathsf{T}} \qquad \textit{and} \qquad |\varepsilon| \leqslant \mu_* \frac{\alpha^2}{\mathsf{K}\, M} \ ,} \tag{6.4.10}$$

where $R_* \in \{R_0, R_0'\}$.

Part I: Description of the local Kolmogorov's sets $\mathscr{K}^i$

There exists $n_* \in \mathbb{N}$ *and and* $\mathbf{p}_i \in \mathfrak{D}$, $1 \leqslant i \leqslant n_*$, *such that*

$$\mathfrak{D} \subset \bigcup_{i=1}^{n_*} B_{R_*}(\mathbf{p}_i) \ ,$$

and the following holds. Pick any $1 \leqslant i \leqslant n_*$. *Define*

$$\begin{aligned} H^i &:= K^i + \varepsilon P^i := H_{|B_{R_*}(\mathbf{p}_i)\times\mathbb{T}^d} \ , \\ \Delta_\alpha^\tau &:= \left\{ \omega \in \mathbb{R}^d : \quad |\omega \cdot k| \geqslant \frac{\alpha}{|k|_1^\tau} \ , \quad \forall\, k \in \mathbb{Z}^d \backslash \{0\} \right\} \ , \\ \mathscr{D}_0^i &:= \{ y_0 \in B_{R_*-\delta}(\mathbf{p}_i) : \ K_y(y_0) \in \Delta_\alpha^\tau \} \ . \end{aligned}$$

Then, there exist $\mathscr{D}_*^i \subset B_{R_*-\delta+r_*}(\mathbf{p}_i)$ *having the same cardinality as* $\mathscr{D}_0^i$, *a lipeomorphism* $G^{*i} \colon \mathscr{D}_0^i \overset{onto}{\longrightarrow} \mathscr{D}_*^i$, *with* $(G^{*i})^{-1} \in C_W^\infty(\mathscr{D}_*^i)$, *a function* $K_*^i \in C_W^\infty(\mathscr{D}_*^i, \mathbb{R})$ *and a* C_W^∞*–symplectic transformation*[130] $\phi_*^i \colon \mathscr{D}_*^i \times \mathbb{T}^d \to \mathscr{K}^i := \phi_*^i(\mathscr{D}_*^i \times \mathbb{T}^d) \subset B_{R_*}(\mathbf{p}_i) \times \mathbb{T}^d$ *and real–analytic in* $x \in \mathbb{T}_{s_*}^d$, *such that*[131]

$$\partial_{y_*} K_*^i \circ G^{*i} = \partial_y K^i \qquad\qquad \textit{on} \quad \mathscr{D}_0^i \ , \tag{6.4.11}$$

$$\partial_{y_*}^\beta (H^i \circ \phi_*^i)(y_*, x) = \partial_{y_*}^\beta K_*^i(y_*), \qquad \forall\, (y_*, x) \in \mathscr{D}_*^i \times \mathbb{T}^d, \quad \forall\, \beta \in \mathbb{N}_0^d \tag{6.4.12}$$

[130]Which means that the Whitney–gradient $\nabla\phi_*^i = \partial\phi_*^i/\partial(y_*, x)$ satisfies $(\nabla\phi_*^i)\mathbb{J}(\nabla\phi_*)^T = \mathbb{J}$ uniformly on $\mathscr{D}_*^i \times \mathbb{T}^d$, where $\mathbb{J} = \begin{pmatrix} 0 & -\mathbb{1}_d \\ \mathbb{1}_d & 0 \end{pmatrix}$.

[131]Notice that the derivatives are taken in the sense of Whitney.

and

$$\|G^{*i} - \mathrm{id}\|_{\mathscr{D}_0^i} \leqslant r_* \,, \tag{6.4.13}$$

$$\|G^{*i} - \mathrm{id}\|_{L,\mathscr{D}_0^i} \leqslant \frac{e\,\sigma_0^{\nu+d}}{\mathsf{C}_6} \,, \;. \tag{6.4.14}$$

Part II: Sharp measure estimate of the complement of $\mathscr{K}$
Set

$$\mathscr{K} := \bigcup_{i=1}^{n_*} \mathscr{K}^i \subset B_{R_*-\delta+r_*}(\mathfrak{D}) \times \mathbb{T}^d \,.$$

Case 1: $R_* = R_0$.
In that case,[132]

$$1 \leqslant n_* \leqslant \left(\left[\frac{\operatorname{diam} \mathfrak{D}}{R_0} \right] + 1 \right)^d \tag{6.4.15}$$

and

$$\operatorname{meas}\left(\left(\bigcup_{i=1}^{n_*} B_{R_0}(\mathbf{p}_i) \times \mathbb{T}^d \right) \backslash \mathscr{K} \right) \leqslant \widehat{\mathsf{C}}\;(6\pi)^d\, \frac{\vartheta\,\eta^2\,\mathsf{T}}{\sigma_0\;\mathsf{r}_0} \left(\operatorname{diam} \mathfrak{D} + \frac{\mathsf{r}_0}{2^6 d\eta^2} \right)^d \alpha \,. \tag{6.4.16}$$

Case 2: $R_* = R'_0$.
In that case, $n_* = n_{\mathfrak{D}}$, $\mathsf{p}_i = p'_i$ *and*

$$\operatorname{meas}\left(\left(\bigcup_{i=1}^{n_{\mathfrak{D}}} B_{R'_0}(p'_i) \times \mathbb{T}^d \right) \backslash \mathscr{K} \right) \leqslant \mathsf{C}\;(12\pi)^d\, n_{\mathfrak{D}}^{\frac{1}{d}}\, \frac{\vartheta\,\mathsf{T}}{\sigma_0}\; \operatorname{meas}(\mathfrak{D})^{\frac{d-1}{d}}\, \alpha \,. \tag{6.4.17}$$

We shall need the following elementary Lemmas in the proof.

Lemma 6.4.3 (Covering Lemma) *For any $R > 0$, there exist*[133]

$$1 \leqslant N \leqslant \left(\left[\frac{\operatorname{diam} \mathfrak{D}}{R} \right] + 1 \right)^d$$

and $p_i \in \mathfrak{D}$, $1 \leqslant i \leqslant N$ such that

$$\mathfrak{D} \subseteq \bigcup_{i=1}^{N} B_R(p_i) \,.$$

[132] $n_* = N$ and $\mathsf{p}_i = p_i$, where N and p_i are the ones appearing in Lemma 6.4.3, with $R = R_0$.
[133] $[x]$ denotes the integer part function $\max\{n \in \mathbb{Z} : n \leqslant x\}$, while $\lceil x \rceil$ denotes the "ceiling function" $\min\{n \in \mathbb{Z} : n \geqslant x\}$.

Proof Let $\rho := \operatorname{diam}\mathfrak{D}$ and $z_i := \inf\{y_i : y \in E\}$. Then $\mathfrak{D} \subseteq z + B_\rho(0)$. Now, let $0 < R' < R$ close enough to R so that

$$\left\lceil \frac{\rho}{R'} \right\rceil = \left\lfloor \frac{\rho}{R} \right\rfloor + 1 =: \widehat{N}\ .$$

Then, $\mathfrak{D}$ can be covered by $\widehat{N}$ closed, contiguous cubes Δ_i, $1 \leqslant j \leqslant \widehat{N}^d$, of equal side–length $2R'$. Let i_j those indices such that $\Delta_{i_j} \bigcap \mathfrak{D}$ and pick $p_j \in \Delta_{i_j} \bigcap \mathfrak{D}$; let N be the number of such cubes. But then, one has $\Delta_{i_j} \subseteq B_R(p_j)$, for each $1 \leqslant j \leqslant N \leqslant \widehat{N}$. The Lemma is therefore proved. ∎

Lemma 6.4.4 *Let*[134] $A\colon D_R(y_0) \to \mathcal{S}_d(\mathbb{C}^d)$ *be a matrix–valued function. Assume that*

$$a := \sup_{y \in D_R(y_0)} \|A(y)\| < 1.$$

Then, for any $y \in D_R(y_0)$, the eigenvalues $\mathbb{1}_d + A(y)$ are bounded in modulus from below by $1 - a$. hence, in particular,

$$|\det\left(\mathbb{1}_d + A(y)\right)| \geqslant (1-a)^d\ , \qquad \forall\, y \in D_R(y_0)\ . \tag{6.4.18}$$

Proof Let $y \in D_R(y_0)$ and $v \neq 0$ an eigenvector of $\mathbb{1}_d + A(y)$ with associated eigenvalue λ. Then

$$\begin{aligned} |\lambda|\|v\| &= \|v + A(y)v\| \\ &\geqslant \|v\| - \|A(y)v\| \\ &\geqslant \|v\| - \|A(y)\|\|v\| \\ &\geqslant (1-a)\|v\| > 0\ . \end{aligned}$$

Thus, the Lemma is proven since the determinant is equal to the product of the eigenvalues, counted with multiplicities. ∎

Proof of Theorem 6.4.2 Set

$$r_0 := \min\{\mathsf{r}_0\ ,\ 32d\delta\}\ .$$

Then,

$$\frac{\mathsf{T}}{32d\sigma_0}\,\alpha \overset{(6.4.10)}{\leqslant} \frac{R_*}{4} \leqslant \frac{R_0}{4} < \frac{\mathsf{r}_0}{32d}\ .$$

[134] $\mathcal{S}_d(\mathbb{C}^d)$ denotes the symmetric matrices of order d, with entries in $\mathbb{C}^d$.

Hence,

$$\delta = \min\left\{\frac{\mathsf{T}}{32d\sigma_0}\,\alpha\ ,\ \frac{\mathsf{r}_0}{32d}\ ,\ \frac{R_*}{4}\right\}\ , \tag{6.4.19}$$
$$r_0 = 32d\delta\ ,$$

so that

$$\alpha = \frac{r_0\sigma_0}{\mathsf{T}}\ . \tag{6.4.20}$$

Thus, thanks to (6.4.19) and (6.4.20), we need only to check that K_y is injective on F_i^0 in order to apply Theorem 6.4.1 to H_i. But, for any $y \in D_{\mathsf{r}_0/(4\eta)}(\mathsf{p}_i)$,

$$\begin{aligned}\|\mathbb{1}_d - T(\mathsf{p}_i)K_{yy}(y)\| &\leqslant \mathsf{T}\|K_{yy}(y) - K_{yy}(\mathsf{p}_i)\|\\ &\leqslant \mathsf{T}\|K_{yyy}\|_{\mathsf{p}_i,\mathsf{r}_0/2}\,\frac{\mathsf{r}_0}{4\eta}\\ &\leqslant \mathsf{T}\frac{\|K_{yy}\|_{\mathsf{p}_i,\mathsf{r}_0}}{\mathsf{r}_0/2}\frac{\mathsf{r}_0}{4\eta}\\ &\leqslant \mathsf{T}\mathsf{K}\frac{1}{2\eta}\\ &= \frac{1}{2}.\end{aligned}$$

Thus, by Lemma 2.2.7, $g := (K_y)^{-1}$ is a real analytic diffeomorphism on $D_{r'}(z_i)$, where

$$z_i := K_y(\mathsf{p}_i) \qquad \text{and} \qquad r' := \frac{\mathsf{r}_0}{8\eta\mathsf{T}} \leqslant \frac{1}{2\|T(\mathsf{p}_i)\|}\frac{\mathsf{r}_0}{4\eta};.$$

Furthermore,

$$\sup_{D_{r'}(z_i)} \|g_z\| \leqslant 2\|T(\mathsf{p}_i)\| \leqslant 2\mathsf{T}\ . \tag{6.4.21}$$

Set $T' := g_z(z_i)^{-1} = K_{yy}(\mathsf{p}_i)$. Then, for any $z \in D_{r'/(8\eta)}(z_i)$,

$$\begin{aligned}\|\mathbb{1}_d - T'g_z(z)\| &\leqslant \|T'\|\|g_z(z) - g_z(z_i)\|\\ &\leqslant \mathsf{K}\|g_{zz}\|_{z_i,r'/2}\,\frac{r'}{8\eta}\\ &\leqslant \mathsf{T}\frac{\|g_z\|_{z_i,r'}}{r'/2}\frac{r'}{8\eta}\\ &\overset{(6.4.21)}{\leqslant} 2\mathsf{T}\mathsf{K}\frac{1}{4\eta}\\ &= \frac{1}{2}.\end{aligned}$$

Thus, again by Lemma 2.2.7, the inverse of g, *i.e.* K_y, is a real analytic diffeomorphism on $D_{r''}(\mathsf{p}_i)$ (since $g(z_i) = \mathsf{p}_i$), where

$$R_* \leqslant R_0 < r'' := \frac{\mathsf{r}_0}{64\eta^2} = \frac{r'}{16\eta\mathsf{K}} \leqslant \frac{1}{2\|T'\|}\frac{r'}{8\eta} \ .$$

Moreover, in exactly the same way as above, one gets

$$\sup_{y\in D_{R_0}(\mathsf{p}_i)} \|\mathbb{1}_d - T(\mathsf{p}_i)K_{yy}(y)\| \leqslant 2\mathsf{T}\frac{\mathsf{K}}{\mathsf{r}_0/2}R_0 \overset{(6.4.8)}{\leqslant} \frac{1}{32d\eta} < \frac{1}{2} \ . \tag{6.4.22}$$

Hence,

$$\begin{aligned}\inf_{y\in D_{R_0}(\mathsf{p}_i)} |\det K_{yy}(y)| &= \inf_{y\in D_{R_0}(\mathsf{p}_i)} \left|\det\left(K_{yy}(\mathsf{p}_i)\{\mathbb{1}_d - (\mathbb{1}_d - T(\mathsf{p}_i)K_{yy}(y))\}\right)\right| \\ &= \inf_{y\in D_{R_0}(\mathsf{p}_i)} |\det K_{yy}(\mathsf{p}_i)| \left|\det\left(\mathbb{1}_d - (\mathbb{1}_d - T(\mathsf{p}_i)K_{yy}(y))\right)\right| \\ &\overset{(6.4.22)+(6.4.18)}{\geqslant} |\det K_{yy}(\mathsf{p}_i)| \left(1-\frac{1}{2}\right)^d \\ &\geqslant \frac{\varrho}{2^d} > 0 \ . \end{aligned} \tag{6.4.23}$$

The estimates (6.4.17) and (6.4.16) then follow easily. For instance, let us treat the second case *i.e.* $R_* = R_0'$. The case $R_* = R_0$ is proved in a similar way by firstly using Lemma 6.4.3, with $R = R_0$; then setting $\mathsf{p}_i = p_i$, $n_* = N$ and thus applying Theorem 6.4.1 to each H^i.
Let then $R_* = R_0'$. Thus, we can apply Theorem 6.4.1 to H^i. Hence, there exists a *Kolmogorov set*

$$\mathscr{K}^i \subset F_i^0 \times \mathbb{T}^d \ , \tag{6.4.24}$$

associated to H^i, with all the desired properties and satisfying[135]

$$\text{meas}\,(F_i^0 \times \mathbb{T}^d \backslash \mathscr{K}_i) \leqslant \mathsf{C}\,(12\pi)^d\,\frac{\vartheta\,\mathsf{T}}{\sigma_0}\,R_0^{d-1}\alpha \ . \tag{6.4.25}$$

[135] Where, ϱ_0 is replaced by $\varrho/2^d$, thanks to (6.4.23); T_0 and K_0 by T and K respectively.

Therefore

$$
\begin{aligned}
\mathrm{meas}\,(\mathfrak{D} \times \mathbb{T}^d \backslash \mathscr{K}) &\overset{(6.4.7)}{\leqslant} \mathrm{meas}\left(\left(\bigcup_{i=1}^{n_0} F_i^0 \times \mathbb{T}^d \right) \backslash \left(\bigcup_{i=1}^{n_0} \mathscr{K}^i \right) \right) \\
&\overset{(6.4.24)}{\leqslant} \sum_{i=1}^{n_0} \mathrm{meas}\, \left(F_i^0 \times \mathbb{T}^d \backslash \mathscr{K}^i \right) \\
&\overset{(6.4.25)}{\leqslant} \sum_{i=1}^{n_0} \mathsf{C}\ (6\pi)^d\ \frac{\vartheta\ \mathsf{T}}{\sigma_0}\ R_0'^{d-1} \alpha \\
&= \mathsf{C}\ (12\pi)^d\ n_0^{\frac{1}{d}}\ \frac{\vartheta\ \mathsf{T}}{\sigma_0} \left(n_0\ R_0'^d \right)^{\frac{d-1}{d}}\ \alpha \\
&\overset{(6.4.9)}{\leqslant} \mathsf{C}\ (12\pi)^d\ n_0^{\frac{1}{d}}\ \frac{\vartheta\ \mathsf{T}}{\sigma_0} \left(\mathrm{meas}\, \mathfrak{D} \right)^{\frac{d-1}{d}}\ \alpha\ .
\end{aligned}
$$

∎

Appendices

A On the initial order of truncation κ_0 of the Fourier series in Theorem 2.1.4

Let

$$\Theta > 0,\ 0 < \vartheta < 1,\ 0 < \sigma \leqslant \frac{1}{20},\ \nu > \bar{\nu} > d \geqslant 2,\ \beta := 1 - \frac{1}{\bar{\nu}} + \frac{1}{\nu},\ 0 < \widetilde{c} \leqslant (1-\beta)\, e,$$

with[136]

$$\vartheta \leqslant \frac{\bar{\nu}}{\widetilde{c}(d-1)^{\beta}},$$

$$\begin{aligned}
\widetilde{\kappa}_0 &= \left[-\frac{\log \Theta}{(1-\vartheta)\sigma} \right], \\
\widetilde{C}_{11} &= \exp\left((1-\vartheta) \left(\left(\frac{\bar{\nu}}{\widetilde{c}\vartheta} \right)^{1/\beta} + \frac{1}{20} \right) \right), \\
\widetilde{C}_{12} &= \left(\frac{e^{-\frac{1-\vartheta}{20}}}{2C_6} \left((1-\vartheta)\widetilde{c} \right)^{\bar{\nu}} \right)^{-\nu/\bar{\nu}}, \\
\widetilde{C}_{13} &= \exp\left((1-\vartheta) \left(\left(\frac{2C_0 C_4}{C_5} \right)^{1/\bar{\nu}} + \frac{1}{20} \right) \right).
\end{aligned}$$

Then

Lemma A.1

[136]Notice that

$$\frac{\bar{\nu}}{\widetilde{c}(d-1)^{\beta}} \geqslant \frac{\nu\bar{\nu}}{(\nu-\bar{\nu})\, e} \cdot \frac{\bar{\nu}}{(d-1)^{\beta}} > \frac{\bar{\nu}^2}{(\bar{\nu}-1)^{\beta}\, e} > \frac{\bar{\nu}}{e} > \frac{2}{e} > \frac{1}{2},$$

so that one can choose $\vartheta = \frac{1}{2}$ and in that case, if one chooses in addition $\widetilde{c} = \bar{c}$, then $\widetilde{\kappa}_0 = \kappa_0$, $\widetilde{C}_{11} = C_{11}$, $\widetilde{C}_{12} = C_{12}$ and $\widetilde{C}_{13} = C_{13}$, with $\bar{c}$, C_{11}, C_{12} and C_{13} as in §2.1.3.1 and §**??**.

(i) If $\Theta < \min(20^\nu \widetilde{C}_{11}^{-1}, \widetilde{C}_{12}^{-1})\sigma^\nu$ *then*

$$\kappa_0^{\bar{\nu}}\sigma^{\bar{\nu}}\, e^{-\kappa_0\sigma} \leqslant \Theta\, e^{\frac{1-\vartheta}{20}} < \frac{1}{2C_6\kappa_0^{\bar{\nu}}}, \quad \kappa_0\sigma > d-1. \tag{A.1}$$

(ii) If $\Theta \leqslant 20^\nu \widetilde{C}_{13}^{-1}\sigma^\nu$ *and* $2h_0\kappa_0^{\bar{\nu}} \leqslant \alpha$ *then*

$$\frac{4C_4h_0}{\alpha\sigma^{\bar{\nu}}} \leqslant \frac{C_5}{C_0}. \tag{A.2}$$

Proof Above all, notice that (for any $0 < \beta < 1$)

$$\forall\, t > 1, \quad \frac{t}{\log t} \geqslant (1-\beta)\, et^\beta \geqslant \tilde{c}t^\beta. \tag{A.3}$$

Let $t := \kappa_0\sigma$.
Let's prove (i). Assume that $\Theta < \min(20^\nu \widetilde{C}_{11}^{-1}, \widetilde{C}_{12}^{-1})\sigma^\nu$. Then

$$\begin{aligned} t > -\frac{\log\Theta}{1-\vartheta} - \sigma > -\frac{\log\left((20\sigma)^\nu \widetilde{C}_{11}^{-1}\right)}{1-\vartheta} - \frac{1}{20} \geqslant -\frac{\log(\widetilde{C}_{11}^{-1})}{1-\vartheta} - \frac{1}{20} \\ = \left(\frac{\bar{\nu}}{\tilde{c}\vartheta}\right)^{1/\beta} \geqslant d-1 \geqslant 1. \end{aligned}$$

Therefore $t > 1$ and $t > \left(\frac{\bar{\nu}}{\tilde{c}\vartheta}\right)^{1/\beta}$, so that

$$\begin{aligned} \frac{t}{\log t} \overset{(A.3)}{\geqslant} \tilde{c}t^\beta \geqslant \frac{\bar{\nu}}{\vartheta} \quad &\Longrightarrow \quad t^{\bar{\nu}} \leqslant e^{\vartheta t} \\ &\Longrightarrow \quad t^{\bar{\nu}}\, e^{-t} \leqslant e^{-(1-\vartheta)t} \leqslant e^{\log\Theta + (1-\vartheta)\sigma} \leqslant \Theta\, e^{\frac{1-\vartheta}{20}}. \end{aligned}$$

On the other hand, since $\Theta \leqslant (20\sigma)^\nu \widetilde{C}_{11}^{-1} \leqslant \widetilde{C}_{11}^{-1} < 1$ then

$$\begin{aligned} \Theta\kappa_0^{\bar{\nu}} &\leqslant \Theta\left(\frac{\log\Theta^{-1}}{(1-\vartheta)\sigma}\right)^{\bar{\nu}} \overset{(A.3)}{\leqslant} \frac{\Theta^{1-\bar{\nu}(1-\beta)}}{((1-\vartheta)\tilde{c}\sigma)^{\bar{\nu}}} = \frac{\Theta^{\bar{\nu}/\nu}}{((1-\vartheta)\tilde{c}\sigma)^{\bar{\nu}}} \\ &< \frac{(\widetilde{C}_{12}^{-1}\sigma^\nu)^{\bar{\nu}/\nu}}{((1-\vartheta)\tilde{c}\sigma)^{\bar{\nu}}} = \frac{e^{-\frac{1-\vartheta}{20}}}{2C_6}. \end{aligned}$$

Finally we prove (ii). If $\quad \Theta \leqslant 20^\nu \widetilde{C}_{13}^{-1}\sigma^\nu \quad$ and $\quad 2h_0\kappa_0^{\bar\nu} \leqslant \alpha \quad$ then

$$\begin{aligned}
\frac{4C_4 h_0}{\alpha\sigma^{\bar\nu}} &\leqslant \frac{2C_4}{t^{\bar\nu}}\\
&< 2C_4\left(-\frac{\log\Theta}{1-\vartheta}-\sigma\right)^{-\bar\nu}\\
&< 2C_4\left(-\frac{\log\left((20\sigma)^\nu \widetilde{C}_{13}^{-1}\right)}{1-\vartheta}-\frac{1}{20}\right)^{-\bar\nu}\\
&\leqslant 2C_4\left(-\frac{\log\left(\widetilde{C}_{13}^{-1}\right)}{1-\vartheta}-\frac{1}{20}\right)^{-\bar\nu} = \frac{C_5}{C_0}.
\end{aligned}$$

$\blacksquare$

B Smooth contraction mapping Lemma

Let r, s, σ, δ, $\mathrm{L} > 0$. Let $u \in C^\infty(\mathbb{R}^d \times \mathbb{T}^d, \mathbb{T}^d)$, with $x \mapsto x + u(y,x)$ holomorphic on $\mathbb{T}^d_{s+\delta}$ for any given $y \in \mathbb{R}^d$. Assume,

$$\frac{1}{\sigma}\|u\|_{0,s+\delta},\ \|u_x\|_{0,s+\delta} \leqslant \mathrm{L} \leqslant \delta \leqslant \frac{1}{2}\ , \tag{B.1}$$

where

$$\|\cdot\|_{0,s+\delta} := \sup_{\mathbb{R}^d\times\mathbb{T}^d_{s+\delta}} |\cdot|\ .$$

Assume also that for any $n \in \mathbb{N}$ there exists a constant $C_n > 0$ with the following property: for any $\beta_1, \beta_2 \in \mathbb{N}_0^d$ with $|\beta_1|_1 + |\beta_2|_1 \leqslant n$,

$$r^{|\beta_1|_1}\sigma^{|\beta_2|_1-1}\|\partial_y^{\beta_1}\partial_x^{\beta_2}u\|_{0,s} \leqslant C_{u,n}\mathrm{L}\ ,. \tag{B.2}$$

Lemma B.1 *Under the above assumptions, there exists a unique map $v \in C^\infty(\mathbb{R}^d \times \mathbb{T}^d, \mathbb{T}^d)$, with $x \mapsto x + u(y,x)$ holomorphic on $\mathbb{T}^d_s$ such that for any given $y \in \mathbb{R}^d$, the map $x \mapsto x + v(y,x)$ is the inverse of $x \mapsto x + u(y,x)$. Moreover, for any $n \in \mathbb{N}$ there a constant $\widehat{C}_n > 0$ such that for any $\beta_1, \beta_2 \in \mathbb{N}_0^d$ with $|\beta_1|_1 + |\beta_2|_1 \leqslant n$,*

$$r^{|\beta_1|_1}\sigma^{|\beta_2|_1-1}\|\partial_y^{\beta_1}\partial_x^{\beta_2}v\|_{0,s} \leqslant \widehat{C}_n\mathrm{L} \tag{B.3}$$

and

$$\|v\|_{0,s} \leqslant \|u\|_{0,s+\delta} \ , \qquad \|v_x\|_{0,s} \leqslant \frac{\|u_x\|_{0,s+\delta}}{1-\delta} \ . \tag{B.4}$$

Furthermore, $\widehat{C}_n$ can be expressed in term of C_n, for any $n \in \mathbb{N}_0$.

Proof Let $\mathcal{F}$ be the set of $w \in C^0(\mathbb{R}^d \times \mathbb{T}^d_s, \mathbb{C}^d)$ such that

$$\|w\|_{0,s} \leqslant \delta \ .$$

Then, $(\mathcal{F}, \ \|\cdot\|_{0,s})$ is a Banach space and for any $w \in \mathcal{F}$,

$$\|\operatorname{Im}(x + w(y,x))\|_{0,s} < s + \|w\|_{0,s} \leqslant s + \delta.$$

Hence the map

$$F \colon \mathcal{F} \ni w \mapsto -u(\pi_1, \pi_2 + w) \in \mathcal{F}$$

is well–defined. Notice that

$$x + v(y,x) + u(y, x + v(y,x)) = x \iff v(y,x) = -u(y, x + v(y,x)) \iff v = F(v).$$

Hence, we have to show that F admits a unique fixed point. But

$$\|F(w_1) - F(w_2)\|_{0,s} \leqslant \|u_x\|_{0,s+\delta} \|w_1 - w_2\|_{0,s} \leqslant \delta \|w_1 - w_2\|_{0,s} \ , \quad \forall \ w_1, w_2 \in \mathcal{F}$$

i.e. , F is a contraction. Therefore, by the Banach's Fixed Point (or contraction mapping) Theorem, F admits a unique fixed point, say v, and v is obtained as the uniform limit of the sequence $(F^n(0))_n$. Thus, by Weierstrass's Theorem, $x \mapsto x + v(y,x)$ is holomorphic on $\mathbb{T}^d_s$, for each $y \in \mathbb{R}^d$. Moreover

$$\|v\|_{0,s} = \|F(v)\|_{0,s} \leqslant \|u\|_{0,s+\delta}$$

and, by differentiating $v = F(v)$ *w.r.t* x, we get

$$v_x = -(\mathbb{1}_d + u_x)^{-1} u_x = -\left(\sum_{n=0}^{\infty} (-u_x)^n \right) u_x$$

so that

$$\|v_x\|_{0,s} \leqslant \left(\sum_{n=0}^{\infty} \|u_x\|^n_{0,s+\delta} \right) \|u_x\|_{0,s+\delta} \leqslant \frac{\|u_x\|_{0,s+\delta}}{1-\delta} \ ,$$

which conclude the proof of (B.4). Next, we shall proceed inductively for the remainder of the proof. We have

$$\frac{1}{\sigma}\|v\|_{0,s} \leqslant \frac{1}{\sigma}\|u\|_{0,s+\delta} \leqslant \mathrm{L} \ ,$$

which proofs (B.3)$_{n=0}$. Set $w(y,x) = (x + v(y,x)$ *i.e.* $w = \pi_2 + v$. Now, fix $m \in \mathbb{N}_0$ and assume that $v \in C^m(\mathbb{R}^d \times \mathbb{T}^d_s, \mathbb{C}^d)$ and (B.3)$_n$ holds, for any $0 \leqslant n \leqslant m$. Then, using the multivariate Fàa Di Bruno's formula (see [CS96], Theorem 2.1) to differentiate $v = F(v)$, for any $\beta = (\beta_1, \beta_2) \in \mathbb{N}_0^d \times \mathbb{N}_0^d$, with $|\beta|_1 = m + 1$, we have[137]

$$
\begin{aligned}
-\frac{1}{\sigma}\Lambda^\beta \partial^\beta v &= -\frac{1}{\sigma}\Lambda^\beta \partial_y^{\beta_1}\partial_x^{\beta_2} v \\
&= -\frac{1}{\sigma}\Lambda^\beta \partial_y^{\beta_1}\partial_x^{\beta_2} F(v) \\
&= \frac{1}{\sigma}\sum_{\substack{\lambda_1\in\mathbb{N}_0^d \\ \lambda_1\leqslant\beta_1}} \Lambda^{(\lambda_1,0)} \sum_{\substack{\lambda_2\in\mathbb{N}_0^d \\ 1\leqslant|\lambda_2|_1\leqslant|(\beta_1-\lambda_1,\beta_2)|_1}} \Lambda^{(0,\lambda_2)}\, \partial_x^{\lambda_2}\partial_y^{\lambda_1} u \sum_{j=1}^{|(\beta_1-\lambda_1,\beta_2)|_1} \sum_{(k,l)\in\mathscr{S}(j,(\beta_1-\lambda_1,\beta_2),\lambda_2)} \\
&\quad \Lambda^{(\beta_1-\lambda_1,\beta_2-\lambda_2)}\, ((\beta_1-\lambda_1,\beta_2))! \prod_{i=1}^{j} \frac{(\partial^{l_i} w)^{k_i}}{k_i!(l_i!)^{|k_i|_1}} \\
&= \sum_{\substack{\lambda=(\lambda_1,\lambda_2)\in\mathbb{N}_0^d\times\mathbb{N}_0^d \\ \lambda_1\leqslant\beta_1 \\ 1\leqslant|\lambda|_1\leqslant m+1}} \frac{1}{\sigma}\Lambda^\lambda\, \partial^\lambda u \sum_{j=1}^{m+1-|\lambda_1|_1} \sum_{(k,l)\in\mathscr{S}(j,\beta-(\lambda_1,0),\lambda_2)} (\beta_1-\lambda_1)!\beta_2! \prod_{i=1}^{j} \frac{(\frac{1}{\sigma}\Lambda^{l_i}\partial^{l_i} w)^{k_i}}{k_i!(l_i!)^{|k_i|_1}} \\
&= \sum_{\substack{\lambda=(\lambda_1,\lambda_2)\in\mathbb{N}_0^d\times\mathbb{N}_0^d \\ \lambda_1\leqslant\beta_1 \\ \lambda_1\neq 0 \ \text{ or } |\lambda_2|_1\neq 1 \\ 1\leqslant|\lambda|_1\leqslant m+1}} \frac{1}{\sigma}\Lambda^\lambda\, \partial^\lambda u \sum_{j=1}^{m+1-|\lambda_1|_1} \sum_{(k,l)\in\mathscr{S}(j,\beta-(\lambda_1,0),\lambda_2)} (\beta_1-\lambda_1)!\beta_2! \prod_{i=1}^{j} \frac{(\frac{1}{\sigma}\Lambda^{l_i}\partial^{l_i} w)^{k_i}}{k_i!(l_i!)^{|k_i|_1}} + \\
&\quad + \sum_{\substack{(\lambda_1,\lambda_2)\in\mathbb{N}_0^d\times\mathbb{N}_0^d \\ \lambda_1=0,\ |\lambda_2|_1=1}} \partial^\lambda u \sum_{j=1}^{1} \sum_{(k,l)=(\lambda_2,\beta)} \beta! \prod_{i=1}^{1} \frac{(\frac{1}{\sigma}\Lambda^{l_i}\partial^{l_i} w)^{k_i}}{k_i!(l_i!)^{|k_i|_1}} \\
&= \sum_{\substack{\lambda=(\lambda_1,\lambda_2)\in\mathbb{N}_0^d\times\mathbb{N}_0^d \\ \lambda_1\leqslant\beta_1 \\ \lambda_1\neq 0 \ \text{ or } |\lambda_2|_1\neq 1 \\ 1\leqslant|\lambda|_1\leqslant m+1}} \frac{1}{\sigma}\Lambda^\lambda\, \partial^\lambda u \sum_{j=1}^{m+1-|\lambda_1|_1} \sum_{(k,l)\in\mathscr{S}(j,\beta-(\lambda_1,0),\lambda_2)} (\beta_1-\lambda_1)!\beta_2! \prod_{i=1}^{j} \frac{(\frac{1}{\sigma}\Lambda^{l_i}\partial^{l_i} w)^{k_i}}{k_i!(l_i!)^{|k_i|_1}} + \\
&\quad + u_x \cdot \frac{1}{\sigma}\Lambda^\beta \partial^\beta(\pi_2 + v)
\end{aligned}
$$

[137] With the convention $0^0 = 1$.

i.e. ,

$$\frac{1}{\sigma}\Lambda^\beta\partial^\beta v = -(\mathbb{1}_d+u_x)^{-1}\left(\frac{1}{\sigma}\Lambda^\beta u_x\,\partial^\beta\pi_2 + \sum_{\substack{\lambda=(\lambda_1,\lambda_2)\in\mathbb{N}_0^d\times\mathbb{N}_0^d\\ \lambda_1\leqslant\beta_1\\ \lambda_1\neq 0\ \text{ or }|\lambda_2|_1\neq 1\\ 1\leqslant|\lambda|_1\leqslant m+1}}\frac{1}{\sigma}\Lambda^\lambda\,\partial^\lambda u\sum_{j=1}^{m+1-|\lambda_1|_1}\sum_{(k,l)\in\mathscr{S}(j,\beta-(\lambda_1,0),\lambda_2)}\right.$$

$$\left.(\beta_1-\lambda_1)!\beta_2!\prod_{i=1}^{j}\frac{(\frac{1}{\sigma}\Lambda^{l_i}\partial^{l_i}(\pi_2+v))^{k_i}}{k_i!(l_i!)^{|k_i|_1}}\right),$$

where

- $\Lambda := (\underbrace{r,\cdots,r}_{d},\underbrace{\sigma,\cdots,\sigma}_{d})$,
- $\mathscr{S}(j,\beta-(\lambda_1,0),\lambda_2) := \Big\{(k,l)=(k_1,\cdots,k_j,l_1,\cdots,l_j)\in\left(\mathbb{N}_0^d\right)^j\times\left(\mathbb{N}_0^{2d}\right)^j : \prod_{i=1}^{j}|k_i|_1>0,$ $0\prec l_1\prec\cdots\prec l_j\,,\quad\sum_{i=1}^{j}k_i=\lambda_2\quad\text{and}\quad\sum_{i=1}^{j}|k_i|_1 l_i=\beta-(\lambda_1,0)\Big\}$,
- $\forall\; k\in\mathbb{N},\ (a,b)\in\mathbb{N}_0^k\times\mathbb{N}_0^k\,,(a\preccurlyeq b\Longleftrightarrow a_j\leqslant b_j\ ,\forall 1\leqslant j\leqslant k)$,

and, for all $k\in\mathbb{N},\ (a,b)\in\mathbb{N}_0^k\times\mathbb{N}_0^k\,, a\prec b$ if and only if one of the following holds

(*i*) $|a|_1<|b|_1$ or

(*ii*) $|a|_1=|b|_1$ and there exists $1<j\leqslant k$ such that $a_i=b_i$ for all $1\leqslant i<j-1$ and $a_j<b_j$.

Therefore, $v\in C^{m+1}(\mathbb{R}^d\times\mathbb{T}_s^d,\mathbb{C}^d)$. Moreover, since $\|(\mathbb{1}_d+u_x)^{-1}\|\leqslant\frac{1}{1-\delta}\leqslant 2$, by the inductive hypothesis, (B.1) and (B.2), we have

$$\begin{aligned}\frac{1}{\sigma}\Lambda^{\beta}\|\partial^{\beta}v\|_{0,s} &\leqslant 2\left(\|u_x\|_{0,s+\delta}+\sum_{\substack{\lambda=(\lambda_1,\lambda_2)\in\mathbb{N}_0^d\times\mathbb{N}_0^d\\ \lambda_1\leqslant\beta_1\\ \lambda_1\neq 0\quad\text{or }|\lambda_2|_1\neq 1\\ 1\leqslant|\lambda|_1\leqslant m+1}} C_{|\lambda|_1}\mathrm{L}\sum_{j=1}^{m+1-|\lambda_1|_1}\sum_{(k,l)\in\mathscr{S}(j,\beta-(\lambda_1,0),\lambda_2)}\right.\\
&\qquad\left.(\beta_1-\lambda_1)!\beta_2!\prod_{i=1}^{j}\frac{(1+\widehat{C}_{|l_i|_1}\mathrm{L})^{|k_i|_1}}{k_i!(l_i!)^{|k_i|_1}}\right)\\
&\leqslant 2\left(\mathrm{L}+\sum_{\substack{\lambda=(\lambda_1,\lambda_2)\in\mathbb{N}_0^d\times\mathbb{N}_0^d\\ \lambda_1\leqslant\beta_1\\ \lambda_1\neq 0\quad\text{or }|\lambda_2|_1\neq 1\\ 1\leqslant|\lambda|_1\leqslant m+1}} C_{|\lambda|_1}\mathrm{L}\sum_{j=1}^{m+1-|\lambda_1|_1}\sum_{(k,l)\in\mathscr{S}(j,\beta-(\lambda_1,0),\lambda_2)}(\beta_1-\lambda_1)!\beta_2!(1+\widehat{C}_m)^{|\lambda_2|_1}\right)\\
&\leqslant \widehat{C}_{m+1}\mathrm{L}\ ,\end{aligned}$$

where $\widehat{C}_{m+1}>0$ is an universal constant, independent upon β. Finally, notice that $\widehat{C}_{m+1}$ can be expressed in term of C_{m+1} if $\widehat{C}_m$ can be expressed in term of C_m. These concludes the proof of the Lemma. ∎

C Extension of Lipschitz–Hölder continuous functions with control on the sup–norm

We aim to recall here a very deep Extension Theorem for Lipschitz–Hölder continuous function, following closely [Min70].[138]

Theorem C.1 (G. J. Minty[Min70]) *Let $(V, \langle \cdot\, , \cdot \rangle)$ be a separable inner product space, $\varnothing \neq A \subseteq V$, $b > 0$, $0 < a \leqslant 1$ and $g \colon A \to \mathbb{R}^d$ a (a, b)–Lipschitz–Hölder continuous function on A i.e.* [139]

$$|g(x_1) - g(x_2)|_2 \leqslant b \, \|x_1 - x_2\|^a \,, \qquad \forall\, x_1, x_2 \in A \,. \tag{C.1}$$

Then, there exists a global (a, b)–Lipschitz–Hölder continuous function[140] *$G \colon V \to \mathbb{R}^d$ such that $G|_A = g$. Futhermore, G can be chosen in such away that $G(V)$ is contained in the closed convex hull of $g(A)$. Hence, in particular,*

$$\sup_{x \in V} \|G(x)\| = \sup_{x \in A} \|g(x)\| \quad \textit{and} \quad \sup_{x_1 \neq x_2 \in V} \frac{\|G(x_1) - G(x_2)\|}{\|x_1 - x_2\|} = \sup_{x_1 \neq x_2 \in A} \frac{\|g(x_1) - g(x_2)\|}{\|x_1 - x_2\|} \,. \tag{C.2}$$

We need some preliminaries to prove the Theorem. Given $n \in \mathbb{N}$, we shall denote

$$\Upsilon_n := \{\lambda = (\lambda_1, \cdots, \lambda_n) \in [0,1]^n \ : \ \lambda_1 + \cdots + \lambda_n = 1\} \,.$$

Definition C.2 (Kirszbraun function) *Let V_1 be a $\mathbb{R}$–vector space and X a non–empty set. A function $f \colon V_1 \times X \times X \to \mathbb{R}$ is called Kirszbraun function (K–function) if:*

(i) f is convex and for any $x_1, x_2 \in X$ and for any finite–dimensional subspace S of V_1, the function $f(\cdot, x_1, x_2) \colon S \ni y \to f(y, x_1, x_2)$ is Lower semicontinuous[141]*;*

(ii) for any $n \in \mathbb{N}$, for any $(y_1, x_1), \cdots, (y_n, x_n) \in V_1 \times X$, for any $x \in X$ and for any $(\lambda_1, \cdots, \lambda_n) \in \Upsilon_n$, the inequality

$$\sum_{1 \leqslant i,j \leqslant n} \lambda_i \lambda_j f(y_i - y_j, x_i, x_j) \geqslant 2 \sum_{i=1}^n \lambda_i f(y_i - y, x_i, x) \,, \qquad y := \sum_{j=1}^n \lambda_j y_j \tag{C.3}$$

holds.

[138] Recall that, Kirszbraun's Theorem (see [Fed], §2.10.43) asserts only that one can extend a Lipschitz continuous function without increasing the Lipschitz constant.

[139] Recall that $|\cdot|_2$ denotes the Euclidean norm on $\mathbb{R}^d$.

[140] *i.e.* satisfying (C.1) on V.

[141] *i.e.* for any $t \in \mathbb{R}$, the sublevel set $\{y \in S : f(y, x_1, x_2) \leqslant t\}$ is closed in S endowed with the canonical topology.

Then, the following holds.

Theorem C.3 (G. J. Minty[Min70]) *Let* $f\colon \mathbb{R}^d \times V \times V \to \mathbb{R}$ *be a K–function,* $n \in \mathbb{N}$, $(y_1, x_1), \cdots, (y_n, x_n) \in \mathbb{R}^d \times V$*. Assume that* f *is continuous and for any* $1 \leqslant i, j \leqslant n$,

$$f(y_i - y_j, x_i, x_j) \leqslant 0 \; . \tag{C.4}$$

Then, given any $x \in V$*, there exists* y *in the convex hull of* $\{y_1, \cdots, y_n\}$ *such that* $f(y_i - y, x_i, x) \leqslant 0$*, for any* $1 \leqslant i \leqslant n$*.*

Proof Consider the function

$$F\colon \Upsilon_n \times \Upsilon_n \ni (\lambda, \mu) \mapsto \sum_{i=1}^{n} \lambda_i f\left(y_i - \sum_{j=1}^{n} \mu_j y_j, x_i, x\right) \; .$$

Then, it is clear that F is convex and lower semicontinuous in μ, concave and upper semicontinuous in λ. Thus, since Υ_n is compact and thanks to the von Neumann's Minimax Theorem, there exists $(\lambda^0, \mu^0) \in \Upsilon_n \times \Upsilon_n$ such that

$$F(\lambda^0, \mu^0) \leqslant \max_{\lambda\in\Upsilon_n} F(\lambda, \mu^0) = \min_{\mu\in\Upsilon_n} \max_{\lambda\in\Upsilon_n} F(\lambda, \mu) = \max_{\lambda\in\Upsilon_n} \min_{\mu\in\Upsilon_n} F(\lambda, \mu) = \min_{\mu\in\Upsilon_n} F(\lambda^0, \mu) \leqslant F(\lambda^0, \mu^0) \; .$$

Hence,

$$F(\lambda, \mu^0) \leqslant F(\lambda^0, \mu^0) \leqslant F(\lambda^0, \mu) \; , \qquad \forall \; \lambda, \mu \in \Upsilon_n \; . \tag{C.5}$$

But,

$$2F(\lambda^0, \lambda^0) = 2\sum_{i=1}^{n} \lambda_i^0 f\left(y_i - \sum_{j=1}^{n} \lambda_j^0 y_j, x_i, x\right) \overset{(C.3)}{\leqslant} \sum_{1\leqslant i,j\leqslant n} \lambda_i \lambda_j f(y_i - y_j, x_i, x_j) \overset{(C.4)}{\leqslant} 0 \; .$$

Set

$$y^0 := \sum_{j=1}^{n} \mu_j^0 y_j \; .$$

Therefore, for any $1 \leqslant i \leqslant n$,

$$f(y_i - y^0, x_i, x) = F(\delta_i^i, \mu^0) \overset{(C.5)}{\leqslant} F(\lambda^0, \lambda^0) \leqslant 0 \; ,$$

where δ_j^i is the Kronecker delta: $\delta_j^i = 1$ if $i = j$ and 0 otherwise. ∎

We shall need also the following.

Lemma C.4 *Let* $x_1, \cdots, x_n \in \mathbb{R}^d$*. Then,* (*i*) *given any* $\beta, a_1, \cdots, a_n > 0$*, we have*[142]

$$\sum_{1\leqslant i,j\leqslant n} \frac{\langle x_i\ ,\ x_j\rangle}{(a_i+a_j)^\beta} = \frac{1}{\Gamma(\beta)}\int_0^\infty \left\| \sum_{i=1}^n e^{-a_i t} x_i \right\|^2 t^{\beta-1} dt \geqslant 0\ . \tag{C.6}$$

(*ii*) *given any* $(\lambda_1, \cdots, \lambda_n) \in \Upsilon_n$ *and any* $0 < a \leqslant 1$*,*

$$\sum_{1\leqslant i,j\leqslant n} \lambda_i\lambda_j |x_i - x_j|_2^{2a} \leqslant \sum_{1\leqslant i,j\leqslant n} \lambda_i\lambda_j (|x_i|_2^2 + |x_j|_2^2)^a \leqslant 2\sum_{i=1}^n \lambda_i\, |x_i|_2^{2a}\ . \tag{C.7}$$

Proof (*i*) is trivial. Let us prove (*ii*). Above all, we recall the Bernoulli inequality:

$$(1+x)^r \leqslant 1 + rx\ , \qquad \forall\ x \geqslant -1\ ,\ \forall\ 0 \leqslant r \leqslant 1\ . \tag{C.8}$$

The case $a = 1$ is obvious. Let then $0 < a < 1$. By the continuity of the norm, up to approximating the zero vector by a sequence of non–zero vectors, we can assume that

[142] Γ being the Euler's Gamma function.

each $x_i \neq 0$, $i = 1, \cdots, n$. Thus, we have

$$\begin{aligned}
\sum_{1\leqslant i,j\leqslant n} \lambda_i\lambda_j |x_i - x_j|_2^{2a} &= \sum_{1\leqslant i,j\leqslant n} \lambda_i\lambda_j \langle x_i - x_j, x_i - x_j\rangle^a \\
&= \sum_{1\leqslant i,j\leqslant n} \lambda_i\lambda_j (|x_i|_2^2 + |x_j|_2^2)^a \left(1 - \frac{2\langle x_i\ ,\ x_j\rangle}{|x_i|^2 + |x_j|_2^2}\right)^a \\
&\overset{(C.8)}{\leqslant} \sum_{1\leqslant i,j\leqslant n} \lambda_i\lambda_j (|x_i|_2^2 + |x_j|_2^2)^a \left(1 - \frac{2a\langle x_i\ ,\ x_j\rangle}{|x_i|_2^2 + |x_j|_2^2}\right) \\
&= \sum_{1\leqslant i,j\leqslant n} \lambda_i\lambda_j (|x_i|_2^2 + |x_j|_2^2)^a - 2a \sum_{1\leqslant i,j\leqslant n} \frac{\langle \lambda_i x_i\ , \lambda_j x_j\rangle}{(|x_i|_2^2 + |x_j|_2^2)^{1-a}} \\
&\overset{(C.6)}{\leqslant} \sum_{1\leqslant i,j\leqslant n} \lambda_i\lambda_j (|x_i|_2^2 + |x_j|_2^2)^a \\
&= \sum_{1\leqslant i,j\leqslant n} \lambda_i\lambda_j \max\{|x_i|_2, |x_j|_2\}^{2a} \left(1 + \frac{\max\{|x_i|_2, |x_j|_2\}^2}{\max\{|x_i|_2, |x_j|_2\}^2}\right)^a \\
&\overset{(C.8)}{\leqslant} \sum_{1\leqslant i,j\leqslant n} \lambda_i\lambda_j \max\{|x_i|_2, |x_j|_2\}^{2a} \left(1 + a\left(\frac{\min\{|x_i|_2, |x_j|_2\}}{\max\{|x_i|_2, |x_j|_2\}}\right)^2\right) \\
&\leqslant \sum_{1\leqslant i,j\leqslant n} \lambda_i\lambda_j \max\{|x_i|_2, |x_j|_2\}^{2a} \left(1 + \left(\frac{\min\{|x_i|_2, |x_j|_2\}}{\max\{|x_i|_2, |x_j|_2\}}\right)^{2a}\right) \\
&= \sum_{1\leqslant i,j\leqslant n} \lambda_i\lambda_j \left(|x_i|_2^{2a} + |x_j|_2^{2a}\right) \\
&= 2\sum_{i=1}^{n} \lambda_i\, |x_i|_2^{2a}\ .
\end{aligned}$$

∎

Now, we are in position to prove Theorem C.1.

Proof of Theorem C.1 The proof is divided into three steps.
Step 1 We show that

$$f : \mathbb{R}^d \times V \times V \ni (y, x_1, x_2) \mapsto |y|_2^2 - b^2 \|x_1 - x_2\|^{2a}$$

is a K–function. f is obviously continuous (actually, C^∞) and convex in y. Now, let $n \in \mathbb{N}$,

$(y_1, x_1), \cdots, (y_n, x_n) \in \mathbb{R}^d \times V$, $x \in V$ and $(\lambda_1, \cdots, \lambda_n) \in \Upsilon_n$, and set $y := \sum_{j=1}^{n} \lambda_j y_j$. Then

$$\begin{aligned}
\sum_{1\leqslant i,j\leqslant n} \lambda_i\lambda_j f(y_i - y_j, x_i, x_j) &= \sum_{1\leqslant i,j\leqslant n} \lambda_i\lambda_j |(y_i - y) - (y_j - y)|_2^2 - b^2 \sum_{1\leqslant i,j\leqslant n} \lambda_i\lambda_j \|(x_i - x) - (x_j - x)\|^{2a} \\
&= \sum_{1\leqslant i,j\leqslant n} \lambda_i\lambda_j (|y_i - y|_2^2 + \|y - y_j\|^2 + 2\langle y_i - y, y - y_j\rangle) - \\
&\quad - b^2 \sum_{1\leqslant i,j\leqslant n} \lambda_i\lambda_j \|x_i - x_j\|^{2a} \\
&= 2\sum_{i=1}^{n} \lambda_i |y_i - y|_2^2 + 2b^2 \left\langle \sum_{i=1}^{n} \lambda_i (y_i - y), \sum_{i=1}^{n} \lambda_j (y - y_j) \right\rangle - \\
&\quad - b^2 \sum_{1\leqslant i,j\leqslant n} \lambda_i\lambda_j |x_i - x_j|^{2a} \\
&= 2\sum_{i=1}^{n} \lambda_i |y_i - y|_2^2 - b^2 \sum_{1\leqslant i,j\leqslant n} \lambda_i\lambda_j \|(x_i - x) - (x_j - x)\|^{2a} \\
&\overset{(C.7)}{\geqslant} 2\sum_{i=1}^{n} \lambda_i |y_i - y|_2^2 - 2b^2 \sum_{i=1}^{n} \lambda_i \|x_i - x\|^{2a} \\
&= 2\sum_{i=1}^{n} \lambda_i f(y_i - y, x_i, x)\ .
\end{aligned}$$

Step 2 We want to show that we can extend g to $\underline{A \bigcup \{x_0\}}$ in such away that the image of x_0 by the extension lies in closed convex hull $\overline{\mathrm{conv}\,(g(A))}$ of $g(A)$, for any $x_0 \in V$. If $x_0 \in A$, there is nothing to do. Let then $x_0 \in V\backslash A$. Set[143]

$$\mathscr{C}(x) := \overline{\mathrm{conv}\,(g(A))} \bigcap \left\{y \in \mathbb{R}^d : f(g(x) - y, x, x_0) \leqslant 0\right\}\ , \qquad x \in A\ .$$

Then, for any $x \in A$, $\mathscr{C}(x)$ is a compact convex subset of $\mathbb{R}^d$. Now pick any $x_1, \cdots, x_{d+1} \in A$ and set $y_i := g(x_i)\ ,\ 1 \leqslant i \leqslant d+1$. Thus, (C.1) implies

$$f(y_i - y_j, x_i, x_j) \leqslant 0\ , \qquad \forall\ 1 \leqslant i \leqslant d+1\ .$$

Thanks to **Step 1**, we can apply Theorem C.3. Therefore, there exists y_0 in the convex hull of $\{y_1, \cdots, y_n\}$ such that $f(y_i - y_0, x_i, x_0) \leqslant 0$, for any $1 \leqslant i \leqslant d+1$. Hence,

$$\bigcap_{i=i}^{d+1} \mathscr{C}(x_i) \neq \varnothing \qquad \text{(since it contains } y_0\text{)}.$$

[143]Notice that, for any $x \in A$, $b\|x - x_0\|^a > 0$ and $\{y \in \mathbb{R}^d : f(g(x) - y, x, x_0) \leqslant 0\}$ is the closed ball (with respect to the Euclidean norm) centered at $g(x)$ with radius $b\|x - x_0\|^a$.

Thus, by Helly's Theorem[144], there exists

$$y_{x_0} \in \bigcap_{x \in A} \mathscr{C}(x) \ .$$

Consequently, the extension g_{x_0} of g to $A \bigcup \{x_0\}$ is obtained by setting $g_{x_0}(x_0) := y_{x_0}$.
Step 3 Pick any countable dense subset D of V. Then, by **Step 2**, we can extend g inductively to $A \bigcup D$. Denote by g_D such an extension and notice that $g_D(A \bigcup D) \subset \overline{\operatorname{conv}(g(A))}$ and satisfies (C.1) on $A \bigcup D$, by construction. Now, pick any $x^0, x^1 \in V \backslash A$ and sequences $\{x_n^i\} \subset D$ converging to x^i, $i = 0, 1$. Fix $i = 0, 1$. Then, for any $n, m \in \mathbb{N}$,

$$|g_D(x_n^i) - g_D(x_m^i)|_2 \leqslant b \|x_n^i - x_m^i\|^a \ .$$

Hence, the sequence $\{g_D(x_n^i)\} \subset \mathbb{R}^d$ is Cauchy and, therefore, converges to a $y^i \in \mathbb{R}^d$ and y^i does not depend upon the sequence chosen but only upon x^i. Now, by

$$\begin{aligned} &g_D(x_n^0),\ g_D(x_n^1) \in g_D(D) \subset \overline{\operatorname{conv}(g(A))} \ , \\ &|g_D(x_n^i) - g(x)|_2 \leqslant b \|x_n^i - x\|^a \ , \qquad \forall\, x \in A \end{aligned}$$

and

$$|g_D(x_n^0) - g_D(x_n^1)|_2 \leqslant b \|x_n^0 - x_n^1\|^a \ ,$$

for all $n \geqslant 0$, we get, by passing to the limit,

$$\begin{aligned} &y^0, y^1 \in \overline{\operatorname{conv}(g(A))} \ , \\ &|y^i - g(x)|_2 \leqslant b \|x^i - x\|^a \ , \qquad \forall\, x \in A \end{aligned}$$

and

$$|y^0 - y^1|_2 \leqslant b \|x^0 - x^1\|^a \ .$$

Then, a desired extension is obtained by just setting

$$f(x) := g(x) \ ,$$

for $x \in A$ and

$$f(x) := \lim g_D(x_n) \ ,$$

for $x \in V \backslash A$ and $\{x_n\} \subset D$ any sequence converging to x. ∎

[144] See [DGK21]

D Lebesgue measure and Lipschitz continuous map

Lemma D.1 *Let $\emptyset \neq A \subset \mathbb{R}^d$ be a Lebesgue–measurable set and $f\colon A \to \mathbb{R}^d$ be Lipschitz continuous with*

$$\|f - \mathrm{id}\|_{L,A} := \sup_{\substack{x,y\in A\\ x\neq y}} \frac{|f(x)-f(y)|}{|x-y|} \leqslant \delta \,. \tag{D.1}$$

Then

$$|\,\mathrm{meas}\,(f(A)) - \mathrm{meas}\,(A)| \leqslant ((1+\delta)^d - 1)\,\mathrm{meas}\,(A)\,. \tag{D.2}$$

Remark D.2 Notice that the inequality (D.2) is sharp as shown by the example $f = (1+\delta)\,\mathrm{id}$.

Proof By Theorem C.1 (see Appendix C), $f-\mathrm{id}$ can be extended to a Lipschitz continuous $g\colon \mathbb{R}^d \circlearrowleft$ with

$$\|g\|_{L,\mathbb{R}^d} = \|f-\mathrm{id}\|_{L,A} \leqslant \delta\,.$$

Now, by Rademacher's Theorem, there exists a set $N \subset \mathbb{R}^d$ with $\mathrm{meas}\,(N) = 0$ and such that g is differentiable on $\mathbb{R}^d\backslash N$. Then[145]

$$\|g_y\|_{\mathbb{R}^d\backslash N} = \|g\|_{L,\mathbb{R}^d\backslash N} \leqslant \|g\|_{L,\mathbb{R}^d} \leqslant \delta\,.$$

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

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

[145] Let's point out that $\mathbb{R}^d\backslash N$ is non–convex if N is non–empty. Netherless, one can just approximate a segment by curves contained in $\mathbb{R}^d\backslash N$ and with length arbitrarily close to the length of the segment.

Thus, by the change of variable (or area or coarea) formula[146], we have

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

∎

[146] See [EG15], §3.3

E Whitney's smoothness

Definition E.1 *Let $A \subset \mathbb{R}^d$ be non–empty and $n \in \mathbb{N}_0$, $m \in \mathbb{N}$. A function $f\colon A \to \mathbb{R}^m$ is said 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 $f \in C^n_W(A, \mathbb{R}^m)$, if for any $\varepsilon > 0$ and $y_0 \in A$, there exists $\delta > 0$ such that, for any $y, y' \in A \cap B_\delta(y_0)$ and $\nu \in \mathbb{N}_0^d$, with $|\nu|_1 \leqslant n$,*

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

The following is proven in [Chi86, §2.7, pg. 58] for $d = 1$.

Lemma E.2 *Let $A \subset \mathbb{R}^d$ 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*

$$a := \sum_{m=1}^{\infty} \|f_m\|_{r_m,A}\, r_m^{-n} < \infty, \qquad \|f_m\|_{r_m,A} := \sup_{B^d_{r_m}(A)} |f_m| . \tag{E.2}$$

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

Proof Let $\nu \in \mathbb{N}_0^d$, with $|\nu|_1 \leqslant n$. We start showing that

$$f_\nu = \sum_{m=1}^{\infty} \partial_y^\nu f_m$$

is well–defined on A. Indeed, for any $m \geqslant 1$, $f_m \in C^\infty(A)$ and, by Cauchy's estimate,

$$\| \sum_{m=1}^{\infty} \partial_y^\nu f_m \|_A \leqslant \sum_{m=1}^{\infty} \|\partial_y^\nu f_m\|_{\frac{r_m}{2},A} \leqslant 2^n \sum_{m=1}^{\infty} \|f_m\|_{r_m,A}\, r_m^{-|\nu|_1} \leqslant 2^n r_1^{n-|\nu|_1} \sum_{m=1}^{\infty} \|f_m\|_{r_m,A}\, r_m^{-n} \overset{(E.2)}{<} \infty ,$$

where

$$\| \cdot \|_A := \sup_A |\cdot| .$$

Finally, we show that $(\partial_y^\nu f)_{\nu\in\mathbb{N}_0^d, |\nu|_1\leqslant n}$ are the Whitney's derivatives of f. Fix then $y_0 \in A$, $0 < \varepsilon < a$ and $\nu \in \mathbb{N}_0^d$, with $|\nu|_1 \leqslant n$. Set

$$b := 2^n\, a \sum_{\substack{\mu\in\mathbb{N}_0^d \\ |\mu|_1=n+1}} 1 .$$

Let $m_1 \in \mathbb{N}$ such that[147]

$$2^n \sum_{m=m_1}^{\infty} \|f_m\|_{r_m,A}\, r_m^{-n} \leqslant \frac{\varepsilon^{n+1}}{(2b)^n} \qquad (< \varepsilon)\ . \tag{E.3}$$

Let[148]

$$\delta := \frac{\varepsilon}{4b}\, r_{m_1}$$

and

$$f^{[k]} := \sum_{m=1}^{k} f_m\ , \qquad k \geqslant 1\ .$$

Now, pick $y, y' \in A \cap B_\delta(y_0)$, with $y \neq y'$. Let then[149] $m_2 \geqslant m_1$ such that

$$\frac{\varepsilon}{2b}\, r_{m_2+1} \leqslant |y' - y| < \frac{\varepsilon}{2b}\, r_{m_2}\ . \tag{E.4}$$

Notice that $f^{[m_2]}$ is holomorphic on $D_{r_{m_2}}(A)$ and

$$0 < r := |y' - y| < \frac{\varepsilon}{2b}\, r_{m_2} \leqslant \frac{\varepsilon}{2a}\, r_{m_2} < \frac{r_{m_2}}{2}\ .$$

Moreover, for any $1 \leqslant m \leqslant m_2$,

$$\begin{cases} r_m - |y' - y| \geqslant r_{m_2} - |y' - y| \overset{(E.4)}{>} \dfrac{b}{\varepsilon}|y' - y| + \left(\dfrac{b}{\varepsilon} - 1\right)|y' - y| > \dfrac{b}{\varepsilon}|y' - y|\ , \\ r_m - |y' - y| = \left(\dfrac{r_m}{2} - |y' - y|\right) + \dfrac{r_m}{2} \geqslant \left(\dfrac{\varepsilon}{2b}\, r_{m_2} - |y' - y|\right) + \dfrac{r_m}{2} \overset{(E.4)}{>} \dfrac{r_m}{2}\ . \end{cases} \tag{E.5}$$

Therefore, by Taylor–Lagrange's formula and Cauchy's estimates, we have (for some $0 <$

[147] Such a m_1 exists by (E.2).

[148] Let us point out that δ does not depend upon ν. These is crucial! Actually, δ does not even depend upon y_0.

[149] Notice that such a m_2 exists since $|y' - y| \leqslant |y' - y_0| + |y_0 - y| < 2\delta = \frac{\varepsilon}{2b} r_{m_1}$ and the sequence $(r_m)_m$ is strictly decreasing.

$t < 1$)

$$\left| f_\nu^{[m_2]}(y') - \sum_{\substack{\mu\in\mathbb{N}_0^d \\ |\mu|_1 \leqslant n-|\nu|_1}} \frac{1}{\mu!} f_{\nu+\mu}^{[m_2]}(y)(y'-y)^\mu \right| = \left| \sum_{\substack{\mu\in\mathbb{N}_0^d \\ |\mu|_1 = n-|\nu|_1+1}} \frac{1}{\mu!} f_{\nu+\mu}^{[m_2]}(y+t(y'-y))(y'-y)^\mu \right|$$
$$\leqslant \sum_{\substack{\mu\in\mathbb{N}_0^d \\ |\mu|_1 = n-|\nu|_1+1}} \sum_{m=1}^{m_2} \|\partial_y^{\nu+\mu} f_m\|_{r,A}\, r^{|\mu|_1}$$
$$\leqslant r^{n-|\nu|_1+1} \sum_{\substack{\mu\in\mathbb{N}_0^d \\ |\mu|_1 = n-|\nu|_1+1}} \sum_{m=1}^{m_2} \|f_m\|_{r_m,A}\, (r_m - r)^{-(|\nu|_1+|\mu|_1)}$$
$$\leqslant \frac{b}{2^n a} r^{n-|\nu|_1+1} \sum_{m=1}^{m_2} \|f_m\|_{r_m,A}\, (r_m - r)^{-(n+1)}$$
$$\overset{(E.5)}{\leqslant} \frac{b}{2^n a} r^{n-|\nu|_1+1} \sum_{m=1}^{m_2} \|f_m\|_{r_m,A} \left(\frac{r_m}{2}\right)^{-n} \left(\frac{b}{\varepsilon}|y'-y|\right)^{-1}$$
$$= \varepsilon\, r^{n-|\nu|_1} \frac{1}{a} \sum_{m=1}^{m_2} \|f_m\|_{r_m,A} {r_m}^{-n}$$
$$\leqslant \varepsilon\, r^{n-|\nu|_1}\ . \tag{E.6}$$

Furthermore, for any $\mu \in \mathbb{N}_0^d$, with $|\mu|_1 \leqslant n$,

$$\sum_{m>m_2} \|\partial_\mu f_m\|_A \leqslant \sum_{m>m_2} \|\partial_\mu f_m\|_{\frac{\varepsilon}{4b} r_m, A}$$
$$\leqslant \sum_{m>m_2} \|f_m\|_{\frac{\varepsilon}{2b} r_m, A} \left(\frac{\varepsilon}{4b} r_m\right)^{-|\mu|_1}$$
$$= \sum_{m>m_2} \|f_m\|_{\frac{\varepsilon}{2b} r_m, A} \left(\frac{\varepsilon}{4b} r_m\right)^{-n} \left(\frac{\varepsilon}{4b} r_m\right)^{n-|\mu|_1}$$
$$\leqslant \left(\frac{\varepsilon}{4b} r_{m_2+1}\right)^{n-|\mu|_1} \left(\frac{2b}{\varepsilon}\right)^n 2^n \sum_{m>m_1} \|f_m\|_{r_m,A}\, r_m^{-n}$$
$$\overset{(E.4)+(E.3)}{\leqslant} r^{n-|\mu|_1} \left(\frac{2b}{\varepsilon}\right)^n \frac{\varepsilon^{n+1}}{(2b)^n}$$
$$= \varepsilon\, r^{n-|\mu|_1}\ . \tag{E.7}$$

Thus,

$$\begin{aligned}
\left| f_\nu(y') - \sum_{\substack{\mu\in\mathbb{N}_0^d\\ |\mu|_1\leqslant n-|\nu|_1}} \frac{1}{\mu!} f_{\nu+\mu}(y)(y'-y)^\mu \right| &\leqslant |f_\nu(y') - f_\nu^{[m_2]}(y')| + \\
&+ \left| f_\nu^{[m_2]}(y') - \sum_{\substack{\mu\in\mathbb{N}_0^d\\ |\mu|_1\leqslant n-|\nu|_1}} \frac{1}{\mu!} f_{\nu+\mu}^{[m_2]}(y)(y'-y)^\mu \right| + \\
&+ \left| \sum_{\substack{\mu\in\mathbb{N}_0^d\\ |\mu|_1\leqslant n-|\nu|_1}} \frac{1}{\mu!} \left(f_{\nu+\mu}^{[m_2]}(y) - f_{\nu+\mu}(y) \right) (y'-y)^\mu \right| \\
&\overset{(E.6)}{\leqslant} \sum_{m>m_2} \|\partial_\nu f_m\|_A + \\
&+ \varepsilon\, r^{n-|\nu|_1} + \\
&+ \sum_{\substack{\mu\in\mathbb{N}_0^d\\ |\mu|_1\leqslant n-|\nu|_1}} \sum_{m>m_2} \|\partial_{\nu+\mu} f_m\|_A\, r^{|\mu|_1} \\
&\overset{(E.7)}{\leqslant} \varepsilon\, r^{n-|\nu|_1} + \\
&+ \varepsilon\, r^{n-|\nu|_1} + \\
&+ \varepsilon\, r^{n-|\nu|_1} \sum_{\substack{\mu\in\mathbb{N}_0^d\\ |\mu|_1\leqslant n-|\nu|_1}} 1 \\
&\leqslant (2+(n+1)^d)\, \varepsilon\, |y'-y|^{n-|\nu|_1} ,
\end{aligned}$$

which conclude the proof, by the arbitrariness of $0 < \varepsilon < a$. ∎

Remark E.3 1. Actually, we proved something stronger. Namely, for any $\varepsilon > 0$, there exists $\delta > 0$ such that, for any $y, y' \in A$ and $\nu \in \mathbb{N}_0^d$, with $|y'-y| < \delta$ and $|\nu|_1 \leqslant n$,

$$\left| f_\nu(y') - \sum_{\substack{\mu\in\mathbb{N}_0^d\\ |\mu|_1\leqslant n-|\nu|_1}} \frac{1}{\mu!} f_{\nu+\mu}(y)(y'-y)^\mu \right| \leqslant \varepsilon\, a\, |y'-y|^{n-|\nu|_1} . \tag{E.8}$$

2. In fact, f satisfies the following "uniform" Whitney's condition, provided $n \geqslant 1$: for any $y, y' \in A$ and $\nu \in \mathbb{N}_0^d$, with $|y'-y| \leqslant r_0$ and $0 \leqslant |\nu|_1 \leqslant n-1$,

$$\left| f_\nu(y') - \sum_{\substack{\mu\in\mathbb{N}_0^d \\ |\mu|_1\leqslant n-1-|\nu|_1}} \frac{1}{\mu!} f_{\nu+\mu}(y)(y'-y)^\mu \right| \leqslant a\left(2^n + 2\,e^d\right)|y'-y|^{n-|\nu|_1}\,. \tag{E.9}$$

Indeed, for such given y, y', ν, let $m \in \mathbb{N}_0$ such that $r_{m_3+1} < |y'-y| \leqslant r_{m_3}$. Then, by similar computations as above, we have

$$\begin{aligned}
\left| f_\nu(y') - \sum_{\substack{\mu\in\mathbb{N}_0^d \\ |\mu|_1\leqslant n-1-|\nu|_1}} \frac{1}{\mu!} f_{\nu+\mu}(y)(y'-y)^\mu \right| &\leqslant |f_\nu(y') - f_\nu^{[m_3]}(y')| + \\
&+ \left| f_\nu^{[m_3]}(y') - \sum_{\substack{\mu\in\mathbb{N}_0^d \\ |\mu|_1\leqslant n-1-|\nu|_1}} \frac{1}{\mu!} f_{\nu+\mu}^{[m_3]}(y)(y'-y)^\mu \right| + \\
&+ \left| \sum_{\substack{\mu\in\mathbb{N}_0^d \\ |\mu|_1\leqslant n-1-|\nu|_1}} \frac{1}{\mu!}\left(f_{\nu+\mu}^{[m_3]}(y) - f_{\nu+\mu}(y)\right)(y'-y)^\mu \right| \\
&\leqslant \frac{2^n\,a}{(n-m)!}\,|y'-y|^{n-|\nu|_1} + \\
&+ 2 \sum_{\substack{\mu\in\mathbb{N}_0^d \\ |\mu|_1\leqslant n-1-|\nu|_1}} \frac{1}{\mu!} \sum_{m>m_3} \|\partial_{\nu+\mu} f_m\|_A\,|y'-y|^{|\mu|_1} \\
&\leqslant \left(\frac{2^n\,a}{(n-m)!} + 2\,a \sum_{\mu\in\mathbb{N}_0^d} \frac{1}{\mu!}\right)|y'-y|^{n-|\nu|_1} \\
&= a\,\left(2^n + 2\,e^2\right)\,|y'-y|^{n-|\nu|_1}\,.
\end{aligned}$$

Finally, we recall the very deep Whitney's extension theorem.

Theorem E.4 (Whitney [Whi34]) *Let $A \subset \mathbb{R}^d$ and $f \in C_W^n(A,\mathbb{R})$, $n \in \mathbb{N}_0$. If A is closed, then there exists $\bar f \in C^n(\mathbb{R}^d,\mathbb{R})$, real–analytic on $\mathbb{R}^d \backslash A$ and such that $\bar f_\nu = f_\nu$ on A, for any $\nu \in \mathbb{N}_0^d$, with $|\nu|_1 \leqslant n$.*

F Generalized Steiner's formula

We aim here to recall the generalized Steiner's formula to compute the volume of the two halves tubes that composes a (uniform) tubular neighborhood of an embedded hypersurface without boundary of $\mathbb{R}^d$.[150]
Let $\mathfrak{S}$ be a smooth, bounded, orientable hypersurface without boundary of $\mathbb{R}^d$ (equipped with the Euclidean metric). Fix the orientation given by a smooth unit normal vector field of $\mathfrak{S}$

$$\mathbf{n} = (\mathbf{n}_1, \cdots, \mathbf{n}_d) \colon \mathfrak{S} \to N\mathfrak{S} = (T\mathfrak{S})^{\perp} , \quad |\mathbf{n}|_2 = \mathbf{n}_1^2 + \cdots + \mathbf{n}_d^2 = 1 .$$

Let $dy := dy_1 \wedge \cdots \wedge dy_d$ be the volume form on $\mathbb{R}^d$ (which induces the Lebesgue–measure meas on $\mathbb{R}^d$) and ∇ be the Levi–Civita connection on $\mathbb{R}^d$. Then, let $d\mathfrak{S}$ be the induced area–form on $\mathfrak{S}$, defined by

$$d\mathfrak{S}(X_1, \cdots, X_{d-1}) = dy(X_1, \cdots, X_{d-1}, \mathbf{n}) ,$$

for any $X_1, \cdots, X_{d-1} \in \Gamma(\mathfrak{S})$, where $\Gamma(\mathfrak{S})$ denotes the Lie algebra of smooth vector fields on $\mathfrak{S}$. Define the shape operator $\mathrm{S} \colon \Gamma(\mathfrak{S}) \to \Gamma(\mathfrak{S})$ by

$$\mathrm{S}X = -\nabla_X \mathbf{n} .$$

Define the map $e_c \colon \{(y, u) : y \in \mathfrak{S} , u = \pm\mathbf{n}(y)\} \to [0, \infty]$ by[151]

$$e_c(y, u) := \sup\{t > 0 : \operatorname{dist}(y + tu, \mathfrak{S}) = t\} .$$

Then, define the *minimal focal distance*

$$\operatorname{minfoc}(\mathfrak{S}) := \inf\{e_c(y, u) : y \in \mathfrak{S} , u = \pm\mathbf{n}(y)\} .$$

Given $\delta > 0$, define the two half–tubes about $\mathfrak{S}$

$$\mathscr{T}^{\pm}(\mathfrak{S}, \delta) := \{y \pm t\mathbf{n}(y) : y \in \mathfrak{S}, \, 0 \leqslant t \leqslant \delta\}$$

and the δ–tubular neighborhood of $\mathfrak{S}$

$$\mathscr{T}(\mathfrak{S}, \delta) := \mathscr{T}^{+}(\mathfrak{S}, \delta) \bigcup \mathscr{T}^{-}(\mathfrak{S}, \delta) .$$

150 For a genralization to a non–uniform tubular neighborhood, see [Roc13].
151 $e_c(y, u)$ is the distance from $y \in \mathfrak{S}$ to its *cut-focal point* in the direction u if such a cut–focal point exists; otherwise $e_c(y, u) = \infty$.

The contraction operators C^j on the space of double forms of type[152] (p,q) are defined inductively as follows: $\mathrm{C}^0(\Lambda)=\Lambda$ and, for $j\geqslant 1$,

$$\mathrm{C}^j(\Lambda)(X_1,\cdots,X_{p-j})(Y_1,\cdots,Y_{q-j})=\sum_{i=0}^{d-1}\mathrm{C}^{j-1}(\Lambda)(X_1,\cdots,X_{p-j},E_i)(Y_1,\cdots,Y_{q-j},E_i)\;,$$

where $\{E_1,\cdots,E_{d-1}\}$ is any orthonormal frame field of $\mathfrak{S}$. Let $\mathbf{R}^{\mathfrak{S}}$ be the *curvature tensor* of $\mathfrak{S}$. $\mathbf{R}^{\mathfrak{S}}$ being a double form of type $(2,2)$, one can then take the wedge product of $\mathbf{R}^{\mathfrak{S}}$ with itself j times to get the double form $(\mathbf{R}^{\mathfrak{S}})^j$ of type $(2j,2j)$. Set $\mathrm{C}^0((\mathbf{R}^{\mathfrak{S}})^0))=1$ and define the $(2j)$–th and $(2j+1)$–th integrated mean curvatures of $\mathfrak{S}$ in $\mathbb{R}^d$ as follows $(j\geqslant 0)$:

$$\begin{aligned}\mathbf{k}_{2j}(\mathbf{R}^{\mathfrak{S}})&\coloneqq\frac{1}{j!(2j)!}\int_{\mathfrak{S}}\mathrm{C}^{2j}((\mathbf{R}^{\mathfrak{S}})^j)\,d\mathfrak{S}\;,\\ \mathbf{k}_{2j+1}(\mathbf{R}^{\mathfrak{S}},\mathrm{S})&\coloneqq\frac{1}{j!(2j)!}\int_{\mathfrak{S}}\left\{\mathrm{tr}\,(\mathrm{S})\,\mathrm{C}^{2j}((\mathbf{R}^{\mathfrak{S}})^j)-2j\,\mathrm{tr}\,\left(\mathrm{S}\mathrm{C}^{2j-1}((\mathbf{R}^{\mathfrak{S}})^j)\right)\right\}d\mathfrak{S}\;.\end{aligned}$$

Thus, the foolowing holds.

Theorem F.1 ([Gra12], pg. 224)

$$\mathrm{meas}\,\left(\mathscr{T}^{\pm}(\mathfrak{S},\delta)\right)=\sum_{j=0}^{\left[\frac{d-1}{2}\right]}\frac{\mathbf{k}_{2j}(\mathbf{R}^{\mathfrak{S}})\,\delta^{2j+1}}{1\cdot 3\cdots(2j+1)}\mp\sum_{j=0}^{\left[\frac{d}{2}-1\right]}\frac{\mathbf{k}_{2j+1}(\mathbf{R}^{\mathfrak{S}},\mathrm{S})\,\delta^{2j+2}}{1\cdot 3\cdots(2j+1)(2j+2)}\;,\tag{F.1}$$

for any $0\leqslant\delta\leqslant\mathrm{minfoc}\,(\mathfrak{S})$.

[152] A double form of type (p,q) is a $\mathfrak{F}(\mathfrak{S})$–linear map $\Lambda\colon\Gamma(\mathfrak{S})^p\times\Gamma(\mathfrak{S})^q\to\mathfrak{F}(\mathfrak{S})$, which is antisymmetric in the first p variables and in the last q as well, where $\mathfrak{F}(\mathfrak{S})$ denotes the space of smooth functions on $\mathfrak{S}$.

G Some others facts on Lipschitz continuous functions

In the following, we prove that any set is contained in some enlargement of itself through any contracting mapping which is bounded on the former set.

Lemma G.1 *Let $g\colon \mathbb{C}^d \to \mathbb{C}^d$ be Lipschitz continuous function. Assume that*

$$\delta := \sup_{\mathbb{R}^d} |g - \mathrm{id}| < \infty \,, \tag{G.1}$$

$$\|g - \mathrm{id}\|_{L,\mathbb{R}^d} < 1. \tag{G.2}$$

Then, for any $\varnothing \neq A \subset \mathbb{C}^d$,[153]

$$A \subset g\left(\overline{D_\delta(A)}\right) \,.$$

Proof Set $f := g - \mathrm{id}$ and 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

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

But, for any $y \in \overline{D_\delta(0)}$,

$$|h(y)| = |f(y + \bar{y})| \leqslant \|f\|_{\mathbb{R}^d} \overset{(G.1)}{\leqslant} \delta \,,$$

i.e. $h\colon \overline{D_\delta(0)} \to \overline{D_\delta(0)}$. Moreover, h is a contraction since $\|h\|_{L,\overline{D_\delta(0)}} \leqslant \|f\|_{L,\mathbb{R}^d} \overset{(G.2)}{<} 1$. Thus, we can apply the Banach's fixed point Theorem to complete the proof. ∎

[153] Where $\overline{D_\delta(A)}$ denotes the closed δ–neighborhood of A in $\mathbb{C}^d$.

Bibliography

[AKN06] V. I. Arnold, V.V. Kozlov, and A. I. Neishtadt. Mathematical aspects of classical and celestial mechanics, 2006.

[AMM78] Ralph Abraham, Jerrold E Marsden, and Jerrold E Marsden. *Foundations of mechanics*, volume 36. Benjamin/Cummings Publishing Company Reading, Massachusetts, 1978.

[Arn63a] V.I. Arnold. Proof of A.N. Kolmogorov's theorem on the conservation of conditionally periodic motions with a small variation in the Hamiltonian. *Russian Math. Surv*, 18(9), 1963.

[Arn63b] Vladimir I Arnold. Small denominators and problems of stability of motion in classical and celestial mechanics, usp. math. nauk 18 n. 6, 91. *Russ. Math. Surv*, 18, 1963.

[Arn63c] Vladimir Igorevich Arnol'd. Small denominators and problems of stability of motion in classical and celestial mechanics. *Russian Mathematical Surveys*, 18(6):85–191, 1963.

[BC18] L. Biasco and L. Chierchia. Explicit estimates on the measure of primary KAM tori. *Annali di Matematica Pura ed Applicata (1923-)*, 197(1):261–281, 2018.

[Cas57] JW Cassells. An introduction to diophantine approximation. *Cambridge tract*, (45), 1957.

[CC95] A. Celletti and L. Chierchia. A constructive theory of Lagrangian tori and computer-assisted applications. In *Dynamics reported*, pages 60–129. Springer, 1995.

[CC97] A. Celletti and L. Chierchia. On the stability of realistic three-body problems. *Communications in mathematical physics*, 186(2):413–449, 1997.

[CC06] A. Celletti and L. Chierchia. KAM tori for N–body problems: a brief history. In *Periodic, Quasi-Periodic and Chaotic Motions in Celestial Mechanics: Theory and Applications*, pages 117–139. Springer, 2006.

[Chi86] L. Chierchia. Quasi-periodic Schroedinger operators in one dimension, absolutely continuous spectra, bloch waves, and integrable hamiltonian systems. Technical report, New York Univ., NY (USA), 1986.

[Chi90] L. Chierchia. A direct method for constructing solutions of the Hamilton–Jacobi equation. *Meccanica*, 25(4):246–252, 1990.

[Chi08] L. Chierchia. Kolmogorov's 1954 paper on nearly-integrable Hamiltonian systems. *Regular and Chaotic Dynamics*, 13(2):130–139, 2008.

[Chi12] L. Chierchia. Kolmogorov–Arnold–Moser (KAM) Theory. In *Mathematics of Complexity and Dynamical Systems*, pages 810–836. Springer, 2012.

[CP10] Luigi Chierchia and Gabriella Pinzari. Properly–degenerate KAM Theory (following V.I. Arnold). *Discrete Contin. Dyn. Syst. Ser. S*, 3(4):545–578, 2010.

[CS96] G. Constantine and T. Savits. A multivariate Faa di Bruno formula with applications. *Transactions of the American Mathematical Society*, 348(2):503–520, 1996.

[DGK21] Ludwig Danzer, Branko Grunbaum, and Victor Klee. Helly's Theorem and its relatives1. *Monatsh. Math*, 31:60–97, 1921.

[DS01] A. C. Da Silva. *Lectures on symplectic geometry*, volume 3575. Springer, 2001.

[Dug57] Rene Dugas. A history of mechanics, trans. *JR Maddox (Switzerland: Editions Du Griffon, 1955)*, pages 124–127, 1957.

[Dum14] H Scott Dumas. *The KAM Story: A Friendly Introduction to the Content, History, and Significance of Classical Kolmogorov–Arnold–Moser Theory*. World Scientific Publishing Company, 2014.

[EG15] L. C. Evans and R. F. Gariepy. *Measure Theory and fine properties of functions*. CRC press, 2015.

[Fed] H. Federer. Geometric Measure Theory. 1969.

[Fer23a] Enrico Fermi. Beweis dass ein mechanisches normalsystem im allgemeinen quasi-ergodisch ist. *Physikalische Zeitschrift*, 24:261–265, 1923.

[Fer23b] Enrico Fermi. Dimostrazione che in generale un sistema meccanico normale è quasi—ergodico. *Il Nuovo Cimento (1911-1923)*, 25(1):267–269, 1923.

[Fer23c] Enrico Fermi. Generalizzazione del teorema di incrè sopra la non esistenza di integeali uniformi di un sistema di equazioni canoniche normali. *Il Nuovo Cimento (1911-1923)*, 26(1):105–114, 1923.

[Fer24] Enrico Fermi. Úber die existenz quasi-ergodischer systeme. *Phys. Z.*, 25:166, 1924.

[Gra12] Alfred Gray. *Tubes*, volume 221. Birkhäuser, 2012.

[KF99] A. N. Kolmogorov and S. V. Fomin. *Elements of the theory of functions and functional analysis*, volume 1. Courier Corporation, 1999.

[Kol54a] A. N. Kolmogorov. On preservation of conditionally periodic motions under a small change in the Hamiltonian function. In *Dokl. Akad. Nauk SSSR*, volume 98, pages 527–530, 1954.

[Kol54b] André Kolmogorov. Théorie générale des systèmes dynamiques de la mécanique classique. In *Proceedings of the International Congress of Mathematicians, Amsterdam, 1954, Vol. 1, pp. 315–333*, volume 1, pages 315—-333. Erven P. Noordhoff N.V., Groningen; North-Holland Publishing Co., Amsterdam, 1954.

[Min70] G. J. Minty. On the extension of Lipschitz, Lipschitz–Hölder continuous, and monotone functions. *Bulletin of the American Mathematical Society*, 76(2):334–339, 1970.

[Mos62] J Moser. On invariant curves of area-preserving mappings of an annulus. *Nachr. Akad. Wiss. Göttingen, II*, pages 1–20, 1962.

[Mos66a] Jürgen Moser. A rapidly convergent iteration method and non-linear differential equations–ii. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 20(3):499–535, 1966.

[Mos66b] Jürgen Moser. A rapidly convergent iteration method and non-linear partial differential equations–i. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 20(2):265–315, 1966.

[Mos67] J. Moser. Convergent series expansions for quasi-periodic motions. *Mathematische Annalen*, 169(1):136–176, 1967.

[Mos73] Jurgen Moser. *Stable and Random Motions in Dynamical Systems: With Special Emphasis on Celestial Mechanics; Hermann Weyl Lectures*. Princeton University Press, 1973.

[Mou02] Forest Ray Moulton. *An introduction to celestial mechanics*. Macmillan, New York, 1902.

[MZ05] Jürgen Moser and Eduard Zehnder. *Notes on dynamical systems*, volume 12. American Mathematical Soc., 2005.

[Nei81] AI Neishtadt. Estimates in the kolmogorov theorem on conservation of conditionally periodic motions. *Journal of Applied Mathematics and Mechanics*, 45(6):766–772, 1981.

[Poi90] Henri Poincaré. Sur le problème des trois corps et les équations de la dynamique. *Acta mathematica*, 13(1):A3–A270, 1890.

[Poi99] Henri Poincaré. Les méthodes nouvelles de la mécanique céleste (vol 1-3), 1892/1893/1899.

[Pös82] J Pöschel. Integrability of hamiltonian systems on cantor sets. *Communications on Pure and Applied Mathematics*, 35(5):653–696, 1982.

[Pös01] J. Pöschel. A Lecture on the Classical KAM Theorem. In *Proc. Symp. Pure Math*, volume 69, pages 707–732, 2001.

[Roc13] Raymond Roccaforte. The volume of non-uniform tubular neighborhoods and an application to the n-dimensional Szegö theorem. *Journal of Mathematical Analysis and Applications*, 398(1):61–67, 2013.

[Rüs75] H. Rüssmann. On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus. *Dynamical systems, theory and applications*, pages 598–624, 1975.

[Sch66] Wolfgang M Schmidt. On badly approximable numbers and certain games. *Transactions of the American Mathematical Society*, 123(1):178–199, 1966.

[Sch69] Wolfgang M Schmidt. Badly approximable systems of linear forms. *Journal of Number Theory*, 1(2):139–154, 1969.

[Vil08] Jordi Villanueva. Kolmogorov theorem revisited. *Journal of Differential Equations*, 244(9):2251–2276, 2008.

[Vil18] Jordi Villanueva. A parameterization method for lagrangian tori of exact symplectic maps of $\mathbb{R}^{2r}$. 2018.

[Whi34] H. Whitney. Analytic extensions of differentiable functions defined in closed sets. *Transactions of the American Mathematical Society*, 36(1):63–89, 1934.

[Zeh10] E. Zehnder. *Lectures on dynamical systems: Hamiltonian vector fields and symplectic capacities*, volume 11. European Mathematical Society, 2010.