Proof of Kolmogorov's 1954 theorem on persistence of quasi-periodic motions

Fabio Pusateri

Academic Year 2005-2006

1 Introduction

In 1954 A.N. Kolmogorov showed evidence in [Kol54] of the following theorem:

Theorem 1.1 (Kolmogorov). *Let H be an Hamiltonian in the form $H(y,x) = E + \omega \cdot y + Q(y,x) + \epsilon P(y,x)$ where Q and P are real-analytic functions over $B^d \times \mathbb{T}^d$ (here B^d is an euclidean ball in $\mathbb{R}^d$) with $\partial_y^\alpha Q(0,x) = 0$ for $|\alpha| \leq 1$, $\omega \in \mathbb{R}^d$, $E \in \mathbb{R}$.*
Assume that

$$\det\langle Q_{yy}(0,\cdot)\rangle = \det \int_{\mathbb{T}^d} Q_{yy}(0,x) \, \frac{dx}{(2\pi)^d} \neq 0$$

then for almost all $\omega \in \mathbb{R}^d$ there exists ϵ_0 such that for all $|\epsilon| \leq \epsilon_0$ there exists Φ_ϵ symplectic diffeomorphism which maps H into the Hamiltonian $N_\epsilon = E_\epsilon + \omega \cdot y' + Q_\epsilon(y', x')$, with $\partial_{y'}^\alpha Q_\epsilon(0,x') = 0$ for $|\alpha| \leq 1$ and where we have denoted $(x,y) = \Phi_\epsilon(y',x')$.
Moreover we have $|E_\epsilon - E|$, $\|Q_\epsilon - Q\|_{C^1}$ and $\|\Phi_\epsilon - \, id\,\|_{C^1}$ are all $O(\epsilon)$ (where $\|f\|_{C^1} := \sup|f| + \sup|f'|$).

Our aim is to give a proof of this theorem following the original ideas gave by Kolmogorov itself and focusing our attention on the estimate, in terms of some constants depending on different parameters, of the size of ϵ_0. We are interested in particular in the dependence of ϵ_0 from the diophantine constant γ because it is strictly related to the dimension of invariant tori in the phase space for the perturbed Hamiltonian H. For an elegant and extremely authoritative proof performed adopting a slightly different scheme refer to [Arn63]; our proof is instead inspired by the original scheme suggested by Kolmogorov and is based on [Chi05].

In order to explain how we are going to proceed, we want now to give an equivalent, but in some way more "quantitative" version of Kolmogorov's theorem.

First we set some notations. Let $\Omega \in \mathbb{C}^d$ we define the following sets:

$$\begin{aligned}
\Omega_r &:= \bigcup_{x_0 \in \Omega} \{x \in \mathbb{C}^d : |x - x_0| < r\}\,, \\
\mathbb{T}^d_\sigma &:= \{x \in \mathbb{C}^d : |\text{ Im } x_j| < \sigma,\ \text{Re } x_j \in \mathbb{T}\ \ \forall\, j = 1 \ldots d\}\,, \\
\mathcal{D}^d_{\gamma,\tau} &:= \{\omega \in \mathbb{R}^d : |\omega \cdot n| > \frac{\gamma}{|n|^\tau},\ \ \forall\, n \in \mathbb{Z}\}\,;
\end{aligned}$$

we shall refer to an element $\omega \in \mathcal{D}^d_{\gamma,\tau}$ as a Diophantine-(γ, τ) vector. Let $f : \Omega \to \mathbb{R}$ be a real-analytic function on an open set $\Omega \subseteq \mathbb{R}^d$ with analytic complex extension on

$$\Omega_r = \bigcup_{x_0 \in \Omega} \{x \in \mathbb{C}^d : |x - x_0| < r\}$$

we put

$$|f|_r = \sup_{\Omega_r} |f|;$$

if $f : \mathbb{T}^d \to \mathbb{R}$ is real-analytic with complex extension on $\mathbb{T}^d_\sigma$ we define

$$|f|_\sigma = \sup_{\mathbb{T}^d_\sigma} |f|;$$

if $f : \Omega \times \mathbb{T}^d \to \mathbb{R}$ is real-analytic with complex extension on the cartesian product $\Omega_r \times \mathbb{T}^d_\sigma$ we naturally put

$$|f|_{r,\sigma} = \sup_{\Omega_r \times \mathbb{T}^d_\sigma} |f|.$$

The same definitions can be obviously given if f is a function whose analytic extension assumes values in $\mathbb{C}^n$ or $\text{mat}_{\mathbb{C}}(n \times n)$, where in this case $|\cdot|$ is some appropriate norm in the space considered. The theorem we are going to prove is the following:

Theorem 1.2. *Let $H(y, x) = E + \omega \cdot y + Q(y, x) + \epsilon P(y, x)$ be a real-analytic Hamiltonian over $B^d \times \mathbb{T}^d$ with analytic extension for P and Q on the complex domain $B^d_r \times \mathbb{T}^d_\sigma$, for some $r > 0$ and $0 < \sigma \leq 1$ and $\omega \in \mathcal{D}^d_{\gamma,\tau}$. Suppose $Q(0, x) = \partial_y Q(0, x) = 0$ and*

$$\det\langle Q_{yy}(0, \cdot)\rangle \neq 0\ .$$

Let $\sigma_\infty < \sigma, r_\infty < r$ take

$$
\begin{aligned}
\mu &= |P|_{r,\sigma} \\
M &= \max\left\{\frac{1}{r}|Q|_{r,\sigma}, |Q_y|_{r,\sigma}, r|Q_{yy}|_{r,\sigma}\right\} \\
\lambda &= \max\left\{\frac{1}{\sigma_\infty}, \frac{1}{\sigma-\sigma_\infty}\right\} \\
\nu &= \max\left\{\frac{r}{r_\infty}, \frac{r}{r-r_\infty}\right\} \\
S &= \frac{1}{r}|\langle Q_{yy}(0,\cdot)\rangle^{-1}| \\
Z &= |\omega| \; ;
\end{aligned}
$$

and define

$$
\begin{aligned}
\Gamma_1 &= \max\left\{M\gamma^{-1}, 1\right\} \\
\Gamma_2 &= \max\{MS, 1\} \\
\Gamma_3 &= \max\{ZS, 1\} \\
\Gamma_4 &= \max\left\{M^{-1}, S\right\} \; ;
\end{aligned}
$$

there exists a positive constant $c(\tau, d) \geq 1$ such that if

$$\epsilon CD\mu < 1$$

where

$$C := c\nu^{14}\lambda^{4(\tau+d)}r^{-1}\Gamma_1^4\Gamma_2^4\Gamma_3\Gamma_4$$

and $D := 2^{12(\tau+d)+30}$, then there exists a symplectic diffeomorphism

$$\Phi : (y', x') \in B^d_{r_\infty} \times \mathbb{T}^d_{\sigma_\infty} \to (y, x) \in B^d_r \times \mathbb{T}^d_\sigma$$

which puts the Hamiltonian H into Kolmogorov's normal form

$$N'(y', x'; \epsilon) = E'(\epsilon) + \omega \cdot y' + Q'(y', x'; \epsilon) = H \circ \Phi \, ;$$

we also have that $|E' - E|$, $\|Q' - Q\|_{C^1} \leq \epsilon CD\mu M r$ and $\|\Phi - \mathrm{id}\|_{C^1} \leq \epsilon CD\mu r$.

1.1 Some useful estimates

Define $\mathcal{H}(\Omega_r), \mathcal{H}(\mathbb{T}^d_\sigma), \mathcal{H}(\Omega_r \times \mathbb{T}^d_\sigma)$ as the spaces of real-analytic functions having holomorphic extension on the prescribed domain and finite norm (respectively $|f|_r, |f|_\sigma$ or $|f|_{r,\sigma} < \infty$).

Lemma 1.1 (Cauchy's estimate)**.** *Let $f \in \mathcal{H}(\Omega_r)$, for all $p \in \mathbb{N}^d$ and $\forall\ 0 < \rho < r$ we have:*

$$|\partial_y^p f(y)|_\rho \leq \frac{p\,!}{(r-\rho)^{|p|_1}} |f|_r$$

Proof The proof of this lemma can be easily obtained by Cauchy's integral formula for analytic functions □

Of course, lemma 1.1 can be easily generalized to $f \in \mathcal{H}(\mathbb{T}^d_\sigma)$ or $\mathcal{H}(\Omega_r \times \mathbb{T}^d_\sigma)$.

Now define, for $\omega \in \mathcal{D}^d_{\gamma,\tau}$, the operator

$$\mathcal{D}_\omega = \sum_{i=1}^{d} \omega_i \partial_{x_i} .$$

Then, let $f \in \mathcal{H}(\mathbb{T}^d_\sigma)$, we are interested in solving the equation

$$\mathcal{D}_\omega u = f . \tag{1.1}$$

First recall that $f(y,x) \in C(B^p(\bar{y}) \times \mathbb{R}^d, \mathbb{R}^m)$, 2π-periodic in the second set of variables, is analytic if and only if there exist positive numbers M, r and ξ such that its Fourier's coefficients $\hat{f}_{k,n}$ satisfy

$$\|\hat{f}_{k,n}\|_\infty \leq M r^{-|k|_1} e^{-|n|_1 \xi} . \tag{1.2}$$

Now observe that if $u(x) = \sum_{n \in \mathbb{Z}^d} \hat{u}_n e^{in\cdot x}$ is the Fourier series for u, then

$$\mathcal{D}_\omega u = \sum_{n \in \mathbb{Z}^d} in \cdot \omega \hat{u}_n e^{in\cdot x}$$

so it is easily verified that $\langle \mathcal{D}_\omega u \rangle = 0$ (the Fourier coefficient corresponding to $n = 0$ is zero). Therefore to solve equation (1.1) we must necessarily require $\langle f \rangle = 0$. Expanding f in its Fourier series we obtain $f(x) = \sum_{n \in \mathbb{Z}^d \setminus \{0\}} \hat{f}_n e^{in\cdot x}$ so that equation (1.1) becomes

$$\sum_{n \in \mathbb{Z}^d \setminus \{0\}} in \cdot \omega \hat{u}_n e^{in\cdot x} = \sum_{n \in \mathbb{Z}^d \setminus \{0\}} \hat{f}_n e^{in\cdot x}$$

and hence

$$\hat{u}_n = \frac{\hat{f}_n}{in \cdot \omega} .$$

Finally observe that

$$u(x) = \sum_{n \in \mathbb{Z}^d \setminus \{0\}} \frac{\hat{f}_n}{in \cdot \omega} \; e^{in \cdot x}$$

converges absolutely by means of (1.2) (here $p = 0$ so that k does not appear) and the diophantine estimate satisfied by ω. We can now state

Lemma 1.2. *Let $f \in \mathcal{H}(\mathbb{T}^d_\sigma)$ and $\omega \in \mathcal{D}^d_{\gamma,\tau}$; if u is the only solution to $\mathcal{D}_\omega u = f$ with $\langle u \rangle = 0$, then there exists $c = c(\tau, d)$ such that*

$$|u|_{\sigma-\delta} \leq \frac{c}{\gamma} \frac{|f|_\sigma}{\delta^{d+\tau}}$$

Proof We have the following inequalities:

$$\begin{aligned} |u|_{\sigma-\delta} \; &\leq \; \left| \sum_{n \neq 0} \frac{\hat{f}_n}{in \cdot \omega} \; e^{in \cdot x} \right|_{\sigma-\delta} \leq \sum_{n \neq 0} \frac{|f|_\sigma}{|n \cdot \omega|} e^{-|n|\sigma} |e^{in \cdot x}|_{\sigma-\delta} \\ &\leq \; \sum_{n \neq 0} \frac{|n|^\tau}{\gamma} |f|_\sigma e^{-|n|\sigma} e^{|n|(\sigma-\delta)} = \frac{|f|_\sigma}{\gamma} \sum_{n \neq 0} e^{-|n|\delta} |n|^\tau \end{aligned}$$

where we have used equation (1.2) for f with $p = 0$ and $\xi = \delta$, while it effectively results by calculus that we can choose $M = |f|_\sigma$. We want now to estimate $\sum_{n \neq 0} e^{-|n|\delta} |n|^\tau$. Approximating the sum with an integral we have

$$\begin{aligned} \sum_{n \neq 0} e^{-|n|\delta} |n|^\tau \; &= \; c' \int_{\mathbb{R}^d} e^{-|x|\delta} |x|^\tau \; dx = \frac{c'}{\delta^\tau} \int_{\mathbb{R}^d} e^{-|\delta x|} |\delta x|^\tau \; dx = \\ &= \; \frac{c'}{\delta^{\tau+d}} \int_{\mathbb{R}^d} e^{-|y|} |y|^\tau \; dy = \frac{c(\tau, d)}{\delta^{\tau+d}} \end{aligned}$$

and the lemma is proved □

Combining this two preceding lemmata and simultaneously generalizing the result in Lemma 1.2 to further inversions of the operator $\mathcal{D}_\omega$, we get

Lemma 1.3. *Let $f \in \mathcal{H}(\mathbb{T}^d_\sigma)$ with $\langle f \rangle = 0$; for every choice of $\alpha \in \mathbb{N}^d$ and $p \in \mathbb{N}$ we have:*

$$|\partial^\alpha \mathcal{D}_\omega{}^{-p} f|_{\sigma-\delta} \leq c(\tau, d, p, \alpha) \frac{|f|_\sigma}{\gamma^p \delta^{p\tau+d+|\alpha|_1}} \; .$$

1.2 Diffeomorphisms on $\mathbb{T}^d$

Consider $a \in \mathcal{H}(\mathbb{T}^d_\sigma)$ and the following analytic function on $\mathbb{T}^d$:

$$\phi : x \in \mathbb{T}^d \to \phi(x) = x + a(x) \in \mathbb{T}^d \, ;$$

we want to give sufficient conditions on a in order to obtain that ϕ is an analytic diffeomorphism on $\mathbb{T}^d$. Our aim is to provide an inverse analytic function for ϕ, that is to say, $\tilde{\phi}(x') = x' + \tilde{a}(x')$ such that $\phi \circ \tilde{\phi} = \text{ id } = \tilde{\phi} \circ \phi$. Let's see what does this mean in terms of a and $\tilde{a}$:

$$\phi \circ \tilde{\phi}(x') = x' \iff \tilde{\phi}(x') + a(\tilde{\phi}(x')) = x' \iff$$

$$\iff x' + \tilde{a}(x') + a(\tilde{\phi}(x')) = x' \iff \tilde{a}(x') = -a(x' + \tilde{a}(x'))\, .$$

We now state the following lemma:

Lemma 1.4. *Let $a \in \mathcal{H}(\mathbb{T}^d_\xi)$ and take $\xi' < \xi$ such that $|a|_\xi \le \xi - \xi'$ and $|a_x|_\xi < 1$; then there exists a unique $\tilde{a} \in \mathcal{H}(\mathbb{T}^d_{\xi'})$ with $|\tilde{a}|_{\xi'} \le \xi - \xi'$ such that:*

$$-a(x' + \tilde{a}) = \tilde{a}.$$

Proof We initially define the following space

$$\mathcal{X} := \{b \in \mathcal{H}(\mathbb{T}^d_{\xi'}) : |b|_{\xi'} \le \xi - \xi'\} \; ;$$

$\mathcal{X}$ is a closed non-empty subset of the Banach space $\mathcal{H}(\mathbb{T}^d_{\xi'})$ and therefore is a Banach space itself. Let $\Phi : b(x') \in \mathcal{X} \to -a(x' + b(x')) \in \mathcal{H}(\mathbb{T}^d_{\xi'})$ we state that Φ is a contraction in $\mathcal{X}$; in fact for every choice of b and $c \in \mathcal{X}$ we have:

1. $|\text{ Im } x' + b(x')| \le |\text{ Im } x'| + |b(x')| < \xi' + |b(x')| \le \xi$ and this implies, by the hypotheses done on a, that $|\Phi(b)|_{\xi'} = |a(x' + b(x'))|_{\xi'} < \xi - \xi'$.
2. $|\Phi(b) - \Phi(c)|_{\xi'} = |a(x' + b(x')) - a(x' + c(x'))|_{\xi'} \le |a_x|_\xi |b - c|_{\xi'} <$ $< |b - c|_{\xi'}$ by Lagrange's theorem applied on a .

Thesis follows from Banach fixed point Theorem □

Observe that for any $a \in \mathcal{H}(\mathbb{T}^d_{\bar{\xi}})$, with $\bar{\xi} > \xi$, by Lemma 1.1 we can estimate $|a_x|_\xi$ as follows:

$$|a_x|_\xi \;=\; \sup_{\mathbb{T}^d_\xi} |a_x| = \sup_{\mathbb{T}^d_\xi} \sup_i \sum_{j=1}^d \left| \frac{\partial a_i}{\partial x_j} \right| \le \sup_i \sum_{j=1}^d \frac{|a|_{\bar{\xi}}}{\bar{\xi} - \xi} = d \frac{|a|_{\bar{\xi}}}{\bar{\xi} - \xi}.$$

Now combining this last estimate and Lemma 1.4 taking $\xi = \bar{\xi}$, we have

Proposition 1.1. *Let $a \in \mathcal{H}(\mathbb{T}^d_\xi)$ and let $\xi' < \xi$ such that $|a|_\xi < \frac{\xi - \xi'}{d+1}$ then there exists a unique $\tilde{a} \in \mathcal{H}(\mathbb{T}^d_{\xi'})$ with $|\tilde{a}|_{\xi'} \le |a|_\xi$ and $-a(x' + \tilde{a}) = \tilde{a}$; therefore $\phi(x) = x + a(x)$ is an analytic diffeomorphism on $\mathbb{T}^d$.*

2 Kolmogorov's idea and first step of the proof

2.1 Reduction of the perturbation to order ϵ^2

Let $H^{(0)} = N^{(0)} + \epsilon P^{(0)} = E + \omega \cdot y + Q^{(0)}(y,x) + \epsilon P^{(0)}$ be the real-analytic Hamiltonian in Kolmogorov's theorem defined on the phase space $\mathcal{U} := B^d \times \mathbb{T}^d$, endowed with the standard symplectic form

$$dy \wedge dx := \sum_{i=1}^{d} dy_i \wedge dx_i$$

(that is to say that Hamilton's equations are $\dot{x} = H_y, \dot{y} = -H_x$). Recall that $\omega \in \mathcal{D}^d_{\gamma,\tau}$ and $P, Q \in \mathcal{H}(B^d_r \times \mathbb{T}^d_\sigma)$ with Q quadratic in y. The first step (and main idea) of the proof, is to find a symplectic transformation Φ which maps $H^{(0)}$ into an Hamiltonian $H^{(1)}$ having the same form but whose perturbative part is order of ϵ^2.

Proposition 2.1. *Consider $H^{(0)}$ as previously defined and suppose to have*

$$\det\langle Q_{yy}(0,\cdot)\rangle \neq 0\,. \tag{2.1}$$

Then there exists a symplectic transformation $\Phi : (y', x') \to (y, x)$ generated by the second species function $F(y', x) = y' \cdot x + \epsilon g(y', x)$ where

$$g(y', x) = b \cdot x + s(x) + a(x) \cdot y'$$

for some $b \in \mathbb{R}^d$, $s : \mathbb{T}^d \to \mathbb{R}$ and $a : \mathbb{T}^d \to \mathbb{R}^d$ both analytic functions, such that

$$H^{(0)} \circ \Phi = H^{(1)} = E^{(1)} + \omega \cdot y' + Q^{(1)}(y', x') + \epsilon^2 P^{(1)}(y', x')$$

with $Q^{(1)}$ quadratic in y' and $Q^{(1)}, P^{(1)}$ real-analytic functions.

Proof By the definition of $F(y', x)$ we have the implicit definition of Φ given by:

$$\begin{cases} x' &= \frac{\partial F}{\partial y'} &= x + \epsilon a(x) \\ \\ y &= \frac{\partial F}{\partial x} &= y' + \epsilon(b + s_x(x) + (a_x(x))^T \cdot y') \end{cases}$$

Assume that $\varphi(x) = x' = x + \epsilon a(x)$ is a diffeomorphism on $\mathbb{T}^d$ with inverse $\tilde{\varphi}(x') = x = x' + \epsilon \tilde{a}(x')$. Following the Hamilton-Jacobi proceeding we aim to express $H^{(0)}(y, x)$ in the new variables (y', x'); notice that we will often leave x

instead of $\tilde{\varphi}(x')$ for simplicity, and we will not sometime use the apex 0 since there's no ambiguity for the moment. By Taylor's formula we have:

$$\begin{aligned} H(y,x) &= H(y'+\epsilon g_x, x) = H(y',x) + \epsilon H_y(y',x)\cdot g_x + \epsilon^2 \tilde{P}_1(y',x) = \\ &= H(y',x) + \epsilon\left[N_y(y',x) + \epsilon P_y(y',x)\right]\cdot g_x + \epsilon^2 \tilde{P}_1(y',x) = \\ &= H(y',x) + \epsilon N_y(y',x)\cdot g_x + \epsilon^2 \tilde{P}_2(y',x) = \\ &= N(y',x) + \epsilon\left[P(y',x) + N_y(y',x)\cdot g_x\right] + \epsilon^2 \tilde{P}_2(y',x) \end{aligned} \tag{2.2}$$

where we have put $\tilde{P}_1(y',x) = \int_0^1 (1-t) H_{yy}(y'+t\epsilon g_x, x)\langle g_x, g_x\rangle\, dt$ and obviously $\tilde{P}_2(y',x) = P_y(y',x)\cdot g_x + \tilde{P}_1(y',x)$. We now focus our attention on $H(y',x) + \epsilon N_y(y',x)\cdot g_x$ in order to put it into the desired Kolmogorov's normal form with at least a perturbative part of order ϵ^2. Recalling that for an analytic function f we have $\mathcal{D}_\omega f = \omega \cdot f_x$ we obtain:

$$\begin{aligned} &N_y(y',x)\cdot g_x = (\omega + Q_y)\cdot(b + s_x + (a_x)^T \cdot y') = \\ =\ &\omega\cdot b + \omega\cdot s_x + \omega\cdot(a_x)^T\cdot y' + Q_y\cdot(b+s_x) + Q_y\cdot(a_x)^T\cdot y' = \\ =\ &\omega\cdot b + \mathcal{D}_\omega s + \mathcal{D}_\omega a\cdot y' + Q_y\cdot(b+s_x) + \tilde{Q}_1(y',x)\,. \end{aligned}$$

with

$$\tilde{Q}_1(y',x) = Q_y\cdot(a_x)^T\cdot y'\,.$$

Now by Taylor's formula applied on $Q_y(y',x)$, and recalling that $Q_y(0,x) = 0$, we have

$$N_y(y,x)\cdot g_x = \omega\cdot b + \mathcal{D}_\omega s + \mathcal{D}_\omega a\cdot y' + Q_{yy}(0,x)\cdot y'\cdot(b+s_x) + \tilde{Q}_2(y',x)$$

where we have naturally put

$$\tilde{Q}_2(y',x) = \tilde{Q}_1(y',x) + \left(\int_0^1 (1-t) Q_{yyy}(ty',x)\, dt\right)\langle y', y', b+s_x\rangle\,.$$

Combining the expression found for $N_y(y,x)\cdot g_x$ and equation (2.2), reorganizing the terms and applying Taylor's formula on $P(y',x)$, we obtain:

$$H(y,x) = N(y',x) + \epsilon[P(y',x) + \omega\cdot b + \mathcal{D}_\omega s + \mathcal{D}_\omega a\cdot y' +$$

$$+Q_{yy}(0,x)\cdot y'\cdot(b+s_x) + \tilde{Q}_2(y',x)] + \epsilon^2\tilde{P}_2(y',x) =$$

$$= E + \epsilon(\omega\cdot b) + \omega\cdot y' + Q(y',x) + \epsilon[P(0,x) + P_y(0,x)\cdot y' +$$

$$+\tilde{Q}_3(y',x) + \mathcal{D}_\omega s + \mathcal{D}_\omega a\cdot y' + Q_{yy}(0,x)\cdot y'\cdot(b+s_x)]$$

$$+\epsilon^2\tilde{P}_2(y',x) \tag{2.3}$$

having defined

$$\tilde{Q}_3(y',x) = \tilde{Q}_2(y',x) + \left(\int_0^1 (1-t)P_{yy}(ty',x)\,dt\right)\langle y',y'\rangle\,.$$

Starting from the equation (2.3) we want now to determine b , s and a. Observe that since $\tilde{Q}_3(0,x)=0$ we have:

$$[\ldots]_{y'=0} = P(0,x) + \mathcal{D}_\omega s = (P(0,x) - \langle P(0,x)\rangle + \mathcal{D}_\omega s) + \langle P(0,x)\rangle$$

so taking

$$s(x) = \mathcal{D}_\omega{}^{-1}(P(0,x) - \langle P(0,x)\rangle) \tag{2.4}$$

it results $[\ldots]_{y'=0} = \langle P(0,x)\rangle$.

For what concerns the linear part in y' we want to maintain the same frequency ω of $H^{(0)}$. Since the term $\omega\cdot y'$ is already given by $N(y',x)$ we have to require

$$P_y(0,x) + \mathcal{D}_\omega a + Q_{yy}(0,x)\cdot(b+s_x) = 0\,. \tag{2.5}$$

By averaging we have

$$\langle P_y(0,\cdot)\rangle + \langle Q_{yy}(0,\cdot)\cdot b\rangle + \langle Q_{yy}(0,\cdot)\cdot s_x(\cdot)\rangle = 0$$

and by hypotheses $\langle Q_{yy}(0,\cdot)\rangle$ is invertible so that we can take

$$b = -\langle Q_{yy}(0,\cdot)\rangle^{-1}\langle P_y(0,\cdot) + Q_{yy}(0,\cdot)\cdot s_x(\cdot)\rangle \tag{2.6}$$

in order to have the average of the left member in (2.5) to be 0. We are now able to solve equation (2.5) taking

$$a = -\mathcal{D}_\omega{}^{-1}\left[P_y(0,x) + Q_{yy}(0,x)\cdot(b+s_x)\right] \tag{2.7}$$

In conclusion by (2.3) , (2.4) , (2.6) and (2.7) we have:

$$H(y,x) = H\circ\Phi(y',x') = H^{(1)}(y',\tilde\varphi(x')) = N^{(1)}(y',\tilde\varphi(x'))$$

$$+\,\epsilon^2 P^{(1)}(y',\tilde\varphi(x')) = E^{(1)} + \omega\cdot y' + Q^{(1)}(y',\tilde\varphi(x')) + \epsilon^2 P^{(1)}(y',\tilde\varphi(x'))\,.$$

where:

$$E^{(1)} = E + \epsilon(\omega\cdot b + \langle P(0,\cdot)\rangle)\,; \tag{2.8}$$

$$Q^{(1)}(y',\tilde\varphi(x')) = Q(y',\tilde\varphi(x')) + \epsilon\tilde Q_3(y',\tilde\varphi(x'))\,; \tag{2.9}$$

$$P^{(1)}(y',\tilde\varphi(x')) = \tilde P_2(y',\tilde\varphi(x')) = P_y(y',\tilde\varphi(x'))\cdot g_x(\tilde\varphi(x')) +$$

$$+\quad \int_0^1 (1-t)H_{yy}(y'+t\epsilon g_x,\tilde\varphi(x'))\langle g_x(\tilde\varphi(x')), g_x(\tilde\varphi(x'))\rangle\,dt\,. \tag{2.10}$$

More expressly we recall that $\tilde Q_3 = Q_1 + Q_2 + Q_3$ with

$$\begin{aligned}
Q_1(y',\tilde\varphi(x')) &= \tilde Q_1(y',\tilde\varphi(x')) = Q_y(y',\tilde\varphi(x'))\cdot(a_x)^T(\tilde\varphi(x'))\cdot y' \\
Q_2(y',\tilde\varphi(x')) &= \left(\int_0^1 (1-t)Q_{yyy}(ty',\tilde\varphi(x'))\,dt\right)\langle y',y',b+s_x(\tilde\varphi(x'))\rangle \\
Q_3(y',\tilde\varphi(x')) &= \left(\int_0^1 (1-t)P_{yy}(ty',\tilde\varphi(x'))\,dt\right)\langle y',y'\rangle.
\end{aligned} \tag{2.11}$$

To end the proof we observe that $Q^{(1)}(y',\tilde\varphi(x'))$ is quadratic in y' so that $N^{(1)}$ is effectively in the desired Kolmogorov's normal form $\square$

Lemma 2.1. *The non-degeneracy condition holds for $N^{(1)}(y',\tilde\varphi(x'))$ as found in proposition 2.1, that is:*

$$\det\langle Q^{(1)}_{yy}(0,\cdot)\rangle \neq 0$$

Proof $Q^{(1)}(y',\tilde\varphi(x')) = Q(y',\tilde\varphi(x')) + \epsilon\tilde Q_3(y',\tilde\varphi(x'))$ so by derivation and averaging we have

$$\langle Q^{(1)}_{yy}(y',\cdot)\rangle = \langle Q_{yy}(y',\cdot)\rangle + \epsilon\langle\partial^2_y\tilde Q_3(y',\cdot)\rangle = \langle Q_{yy}(y',\cdot)\rangle + o(\epsilon)$$

Thesis follows for small enough ϵ , since $\det\langle Q_{yy}(y',\cdot)\rangle \neq 0$ by hypotheses. We postpone for the moment the discussion with full details on the estimate of how small must ϵ be in order to have $\langle Q^{(1)}(y',\tilde\varphi(x'))\rangle$ invertible $\square$

2.2 Control on the domain of Φ

Recall that $H^{(0)} = N^{(0)} + \epsilon P^{(0)} = E + \omega \cdot y + Q^{(0)} + \epsilon P^{(0)}$ with $\omega \in \mathcal{D}^d_{\gamma,\tau}$ for some fixed $\gamma \in \mathbb{R}$, and $P, Q \in \mathcal{H}(B^d_r \times \mathbb{T}^d_\sigma)$. Let $\sigma_\infty < \sigma < 1$ and $0 < r_\infty < r$ we define

$$
\begin{aligned}
M &:= \max\left\{\frac{1}{r}|Q|_{r,\sigma},\ |Q_y|_{r,\sigma},\ r|Q_{yy}|_{r,\sigma}\right\} \\
S &:= \frac{1}{r}\left|\langle Q_{yy}(0,\cdot)^{-1}\rangle\right| \\
\lambda &:= \max\left\{\frac{1}{\sigma_\infty}\ ,\ \frac{1}{\sigma-\sigma_\infty}\right\} \\
\nu &:= \max\left\{\frac{r}{r_\infty}\ ,\ \frac{r}{r-r_\infty}\right\} \\
Z &:= |\omega| \\
\mu &:= |P|_{r,\sigma}
\end{aligned}
$$

We want now to give estimates on $|g|$ in order to apply proposition 1.1 to $g(y', x) = b \cdot x + s(x) + a(x) \cdot y'$ obtaining that the application

$$\varphi : x \longmapsto \frac{\partial F}{\partial y'} = x + \epsilon a(x) = x'$$

is effectively a diffeomorphism on $\mathbb{T}^d$ and by consequence so is $\tilde{\varphi} : x' \longmapsto x' + \epsilon\tilde{a}(x') = x$, i.e. the first component of Φ. Recall that we have $b \in \mathbb{R}^d$ and by definition of s and a in equations (2.4) , (2.6) , (2.7) and lemma 1.2 there exists $0 < \delta < \sigma$ such that $s \in \mathcal{H}(\mathbb{T}^d_{\sigma-\delta})$ and $a \in \mathcal{H}(\mathbb{T}^d_{\sigma-2\delta})$; here δ is the loss of analycity due to the inversion of the operator $\mathcal{D}_\omega$.

Remark 2.1. *Let $\rho < r$ and $\delta < \sigma$ be respectively the losses of analycity in y and x; combining lemmata 1.1 and 1.2, for any $f \in \mathcal{H}(\mathbb{T}^d_\sigma)$ or $\mathcal{H}(B^d_r \times \mathbb{T}^d_\sigma)$ and $l \in \mathbb{N}^d$ this two estimates hold:*

$$\left|\partial^l_x \mathcal{D}_\omega{}^{-1} f(x)\right|_{\sigma-\delta} \leq \frac{c}{\gamma}\frac{|f|_\sigma}{\delta^{q+|l|_1}} \tag{2.12}$$

$$\left|\partial^p_y \partial^l_x f(y,x)\right|_{r-\rho,\sigma-\delta} \leq c\,\frac{|f|_{r,\sigma}}{\rho^{|p|_1}\delta^{|l|_1}} \tag{2.13}$$

where we take the same constant $c \geq 1$ for both inequalities and for any f scalar or vectorial function, matrix or tensor and where $q = d + \tau$.

Lemma 2.2. *There exists a constant $c_1 \geq 1$ depending on $q = \tau + d$, and $B_1 \geq 1$ depending on M, S, γ, μ and r such that for all $0 < \delta < \sigma - \sigma_\infty$*

$$\max\left\{|s|_{\sigma-\frac{\delta}{2}}, |s_x|_{\sigma-\frac{\delta}{2}}, |b|, |a|_{\sigma-\delta}\, r, |a_x|_{\sigma-\delta}\, r, |g_x|_{r,\sigma-\delta}\right\} \leq c_1 B_1 \delta^{-2q} r\,.$$

Proof Using inequalities (2.12) and (2.13) and recalling the definitions of s, b and a in (2.4), (2.6) and (2.7), we estimate separately each quantity. First of all we have

$$|s|_{\sigma-\frac{\delta}{2}}\,,\ |s_x|_{\sigma-\frac{\delta}{2}} \leq \frac{c}{\gamma} 2^q \delta^{-q} |P(0,x) - \langle P(0,x)\rangle|_\sigma \leq$$

$$\leq \frac{c}{\gamma} 2^{q+1} \delta^{-q}\, |P(0,x)|_\sigma \leq \frac{\bar{c}}{\gamma} \delta^{-q} |P(y,x)|_{r,\sigma} \leq \bar{c} \delta^{-q} \mu \gamma^{-1}$$

with $\bar{c} = c 2^{q+1}$.

Furthermore we may estimate

$$|b| = |\langle Q_{yy}(0,\cdot)\rangle^{-1} \langle P_y(0,\cdot) + Q_{yy}(0,\cdot) \cdot s_x(\cdot)\rangle| \leq$$

$$\leq Sr \sup_{\mathbb{T}^d} \left(|P_y(0,x)| + |Q_{yy}(0,x) \cdot s_x(x)|\right) \leq$$

$$\leq Sr \left(\sup_{B^d \times \mathbb{T}^d_\sigma} |P_y(y,x)| + \sup_{B^d \times \mathbb{T}^d_\sigma} |Q_{yy}(y,x)| |s_x(x)|_{\sigma-\frac{\delta}{2}} \right) \leq$$

$$\leq Sr \left(c\, \frac{\mu}{r} + c \frac{M}{r}\, \bar{c} \delta^{-q} \mu \gamma^{-1} \right) \leq c\bar{c} S \mu r^{-1} \left(1 + M r^{-1}\, \delta^{-q} \gamma^{-1}\right) \leq$$

$$\leq c\bar{c} S \mu \delta^{-q} \left(1 + M\gamma^{-1}\right) \leq c' S \mu \delta^{-q} A_1$$

where we define the first auxiliary constant

$$A_1 := \max\left\{M\gamma^{-1},\, 1\right\}$$

and $c' := 2c\bar{c} = c^2 2^{q+2}$.

Using (2.12) and (2.13) once again, from the definition of a in (2.7) we get

$$|a|_{\sigma-\delta}, |a_x|_{\sigma-\delta} \leq \frac{c}{\gamma}\frac{2^q}{\delta^q}|P_y(0,x) + Q_{yy}(0,x)\cdot(b+s_x)|_{\sigma-\frac{\delta}{2}} \leq$$

$$\leq \frac{\bar{c}}{\gamma}\delta^{-q}\left[\sup_{B^d\times\mathbb{T}^d_\sigma}|P_y(y,x)| + \sup_{B^d\times\mathbb{T}^d_\sigma}|Q_{yy}(y,x)|\left(|b| + |s_x(x)|_{\sigma-\frac{\delta}{2}}\right)\right] \leq$$

$$\leq \frac{\bar{c}}{\gamma}\delta^{-q}\left[c\frac{\mu}{r} + c\frac{M}{r}\left(|b| + |s_x|_{\sigma-\frac{\delta}{2}}\right)\right] \leq$$

$$\leq \frac{\bar{c}}{\gamma}\delta^{-q}\left[c\mu r^{-1} + cMr^{-1}\left(c'S\mu\delta^{-q}A_1 + \bar{c}\delta^{-q}\mu\gamma^{-1}\right)\right] \leq$$

$$\leq c\bar{c}c'\delta^{-2q}\gamma^{-1}\mu r^{-1}\left[1 + \left(MSA_1 + M\gamma^{-1}\right)\right] \leq$$

$$\leq c\bar{c}c'\delta^{-2q}\gamma^{-1}\mu r^{-1}\left[1 + A_1\left(MS+1\right)\right] \leq$$

$$\leq \hat{c}\delta^{-2q}\gamma^{-1}\mu r^{-1}A_1A_2$$

where

$$A_2 := \max\{MS,\, 1\}$$

and we take $\hat{c} := 2c\bar{c}c' = c^3 2^{2q+4}$. By using the preceding estimates we have

$$|g_x(y',x)|_{r,\sigma-\delta} = |b + s_x(x) + a_x(x)^T\cdot y'|_{r,\sigma-\delta} \leq$$

$$\leq |b| + |s_x|_{\sigma-\delta} + |a_x|_{\sigma-\delta}|y'| \leq$$

$$\leq c'S\mu\delta^{-q}A_1 * \bar{c}\delta^{-q}\mu\gamma^{-1} + \hat{c}\delta^{-2q}\gamma^{-1}\mu r^{-1}A_1A_2 r \leq$$

$$\leq \hat{c}\delta^{-2q}\left[S\mu A_1 + \mu\gamma^{-1} + \mu\gamma^{-1}A_1A_2\right] \leq$$

$$\leq 2\hat{c}\delta^{-2q}rA_1A_2\left[S\mu r^{-1} + \mu r^{-1}\gamma^{-1}\right] \leq \hat{c}\delta^{-2q}rA_1A_2A_3$$

where

$$A_3 = \mu r^{-1}\max\left\{S,\, \gamma^{-1}\right\};$$

observe that A_3 is linear in μ and so is the final estimate that proves the lemma with $c_1 = 4\hat{c} = c^3 2^{2q+6}$ and $B_1 = A_1A_2A_3$ □

With these estimates we can now obtain the following

Proposition 2.2. *There exists $c_2 \geq c_1$ such that if*

$$\epsilon c_2 B_1 \rho^{-1} \delta^{-2q} r < 1 \tag{2.14}$$

then

1. *$\varphi(x) = x + \epsilon a(x)$, with a as in (2.7), is an analytic diffeomorphism on $\mathbb{T}^d$.*
2. *If $\tilde{\varphi}(x') = x' + \epsilon \tilde{a}(x'; \epsilon)$ is its inverse, we have*
$$|\tilde{a}|_{\sigma - \frac{3}{2}\delta} \leq |a|_{\sigma-\delta} \leq c_1 \delta^{-2q} B_1 \, .$$
3. *$\tilde{\varphi} : \mathbb{T}^d_{\sigma-\frac{3}{2}\delta} \longmapsto \mathbb{T}^d_{\sigma-\delta}$, $\varphi : \mathbb{T}^d_{\sigma-2\delta} \longmapsto \mathbb{T}^d_{\sigma-\frac{3}{2}\delta}$ and $\tilde{\varphi} \circ \varphi = \ id \ = \varphi \circ \tilde{\varphi}$ on $\mathbb{T}^d_{\sigma-2\delta}$.*
4. *Let $\rho < r$ then $\forall \ y' \in B_{r-\rho}$, $x \in \mathbb{T}^d_{\sigma-2\delta}$ we have $y' + t\epsilon g_x(y', x) \in B_{r-\frac{\rho}{2}}$, $\forall \ t \in [0,1]$; in particular $y = y' + \epsilon g_x(y', x) \in B_{r-\frac{\rho}{2}}$*

Proof The first three statements follow directly from proposition 1.1 with $\xi = \sigma - \delta$, $\xi' = \sigma - \frac{3}{2}\delta$ and taking $c_2 = 2c_1(d+1)$ so that condition $|a|_{\sigma-\delta} < \frac{\xi-\xi'}{d+1} = \frac{\delta}{2(d+1)}$ holds by the estimate in lemma 2.2. Again by lemma 2.2 and by hypotheses we obtain $\epsilon |g_x(y', x)|_{r-\rho, \sigma-2\delta} \leq \frac{\rho}{2}$ so that the last statement is also proved $\square$

With this proposition we are now able to control domain and codomain of Φ; for instance we may use the following kind of estimates:

$$|P^{(1)} \circ \Phi|_{r-\rho, \sigma-2\delta} \leq |P|_{r-\frac{\rho}{2}, \sigma-\delta}$$

$$|Q_i \circ \Phi|_{r-\rho, \sigma-2\delta} \leq |Q_i|_{r-\frac{\rho}{2}, \sigma-\delta} \qquad \text{for } \ i = 1, 2, 3$$

2.3 Estimates on $E^{(1)} - E^{(0)}$, $Q^{(1)} - Q^{(0)}$ and $P^{(1)}$

To complete the first step of the proof of Kolmogorov's theorem we want now to estimate the difference between the energies and the quadratic parts of $N^{(0)}$ and $N^{(1)}$, and the size of the new perturbation $P^{(1)}$.

Lemma 2.3. *There exists $c_3 \geq c_2$ constant depending on $q = \tau + d$, and $B_2 \geq B_1$ depending on M, S, μ, Z, γ and r such that:*

$$\max \Big\{ \big| E^{(1)} - E^{(0)} \big| \, , \epsilon \big| P^{(1)}(y', \tilde{\varphi}(x')) \big|_{r-\rho, \sigma-2\delta} \, ,$$

$$(\rho/2)^{|\alpha|_1} \big| \partial^\alpha_y \left(Q^{(1)}(y', \tilde{\varphi}(x')) - Q^{(0)}(y', \tilde{\varphi}(x') \right) \big|_{r-\frac{3}{2}\rho, \sigma-2\delta} \Big\} \leq$$

$$\leq \epsilon c_3 \rho^{-3} \delta^{-4q} r^3 B_2 \mu$$

for any $|\alpha|_1 \leq 2$.

Proof Identity (2.8) and lemma 2.2 yield:

$$\left|E^{(1)} - E^{(0)}\right| = \epsilon|\omega \cdot b + \langle P(0,\cdot)\rangle| \leq \epsilon(|\omega|\,|b| + |P|_{r.\sigma}) \leq$$

$$\leq \epsilon(Zc'S\delta^{-q}\mu A_1 + \mu) \leq \epsilon c'\delta^{-q}\mu A_1\,(ZS+1) \leq \epsilon c'\delta^{-q}\mu A_1 A_4$$

with

$$A_4 := \max\{ZS,\, 1\}$$

Moreover, by identity (2.9) we have $Q^{(1)} - Q^{(0)} = \epsilon\tilde{Q}^{(3)} = \epsilon(Q_1 + Q_2 + Q_3)$; thus, we may estimate separately the three terms using definitions in (2.11) and the estimates prove in lemma 2.2; it result

$$|Q_1(y', \tilde{\varphi}(x'))|_{r-\frac{3}{2}\rho,\sigma-2\delta} \leq |Q_1(y', x))|_{r-\frac{3}{2}\rho,\sigma-\delta} \leq$$

$$\leq |Q_y(y', x)|_{r-\rho,\sigma-\delta}|a_x(x)|_{\sigma-\delta}|y'| \leq cc_1\rho^{-1}\delta^{-2q}\gamma^{-1}\mu M A_1 A_2 r$$

and

$$|Q_2(y', \tilde{\varphi}(x'))|_{r-\frac{3}{2}\rho,\sigma-2\delta} \leq |Q_2(y', x)|_{r-\frac{3}{2}\rho,\sigma-\delta} \leq$$

$$\leq c\frac{8}{27}M\rho^{-3}r^3(|b| + |s_x|_{\sigma-\delta}) \leq$$

$$\leq cM\rho^{-3}r^3\left(c_1 S\mu\delta^{-q}A_1 + c_1\delta^{-q}\mu\gamma^{-1}\right) \leq$$

$$\leq cc_1\rho^{-3}\delta^{-2q}r^3 M\left(S\mu A_1 + \mu\gamma^{-1}\right) \leq cc_1\rho^{-3}\delta^{-2q}MA_1A_3r^4\,;$$

analogously, for what concerns Q_3 we have:

$$|Q_3(y', \tilde{\varphi}(x'))|_{r-\frac{3}{2}\rho,\sigma-2\delta} \leq |Q_3(y', x)|_{r-\frac{3}{2}\rho,\sigma-\delta} \leq$$

$$\leq |P_y(y', x)|_{r-\rho,\sigma}|y'|^2 \leq c\mu\rho^{-2}r^2\,.$$

Now recall that $A_3 = \mu r^{-1}\max\{S, \gamma^{-1}\}$ and then

$$MA_3 = \mu r^{-1}\max\left\{MS, M\gamma^{-1}\right\} \leq \mu r^{-1}\max\{A_2, A_1\} \leq \mu r^{-1}A_1A_2\,;$$

besides observe that obviously $\rho r^{-1} < 1$ and therefore we have:

$$(\rho/2)^{|\alpha|_1}\left|\partial_y^\alpha\left(Q^{(1)}(y',\tilde{\varphi}(x')) - Q^{(0)}(y',\tilde{\varphi}(x'))\right)\right|_{r-2\rho,\sigma-2\delta} \leq$$

$$\leq \epsilon\left(|Q_1|_{r-\frac{3}{2}\rho,\sigma-2\delta} + |Q_2|_{r-\frac{3}{2}\rho,\sigma-2\delta} + |Q_3|_{r-\frac{3}{2}\rho,\sigma-2\delta}\right) \leq$$

$$\leq \epsilon\left(cc_1\rho^{-1}\delta^{-2q}\gamma^{-1}\mu M A_1 A_2 r + cc_1\rho^{-3}\delta^{-2q}MA_1A_3r^4 + c\mu\rho^{-2}r^2\right) \leq$$

$$\leq \epsilon cc_1\rho^{-3}\delta^{-2q}r^3\left(\gamma^{-1}\mu M A_1 A_2 + MA_1A_3r + \mu\right) \leq$$

$$\leq \epsilon cc_1\rho^{-3}\delta^{-2q}r^3\left(\mu {A_1}^2 A_2 + \mu {A_1}^2 A_2 + \mu\right) \leq$$

$$\leq \epsilon cc_1\rho^{-3}\delta^{-2q}r^3 {A_1}^2 {A_2}^2\mu$$

It remains now to be proved the estimate for P; by identity (2.10) we have

$$|P^{(1)}(y',\tilde{\varphi}(x'))|_{r-\rho,\sigma-2\delta} \leq |P^{(1)}(y',x)|_{r-\rho,\sigma-\delta} \leq$$

$$\leq \left|P_y^{(0)}(y',x')\right|_{r-\rho,\sigma-\delta}|g_x(x)|_{\sigma-\delta} + \left|H_{yy}^{(0)}(y',x)\right|_{r-\frac{\delta}{2},\sigma-\delta}|g_x(x)|^2_{\sigma-\delta} \leq$$

$$\leq c\rho^{-1}\mu c_1\delta^{-2q}B_1 r + \left(\left|Q_{yy}^{(0)}(y',x)\right|_{r-\frac{\rho}{2},\sigma-\delta}\right.$$

$$\left. + \epsilon\left|P_{yy}^{(0)}(y',x)\right|_{r-\frac{\rho}{2},\sigma-\delta}\right)\ |g_x(x)|^2_{\sigma-\delta} \leq$$

$$\leq cc_1\rho^{-1}\delta^{-2q}\mu B_1 r + \left(4c\rho^{-2}Mr + \epsilon 4c\rho^{-2}\mu\right)\left(c_1\delta^{-2q}B_1 r\right)^2 \leq$$

$$\leq cc_1\rho^{-1}\delta^{-2q}\mu B_1 r + 4cc_1^2\rho^{-2}\delta^{-4q}{B_1}^2 r^3 M\left(1 + \epsilon\frac{\mu}{Mr}\right) \leq$$

$$\leq 4cc_1^2\rho^{-2}\delta^{-4q}r^2 B_1\left[\mu + MrB_1\left(1 + \epsilon\frac{\mu}{Mr}\right)\right] \leq$$

$$\leq 4cc_1^2\rho^{-2}\delta^{-4q}r^2 B_1\mu\left[1 + {A_1}^2{A_2}^2\left(1 + \epsilon\frac{\mu}{Mr}\right)\right] \leq$$

$$\leq 12cc_1^2\rho^{-2}\delta^{-4q}r^2{A_1}^2{A_2}^2 B_1\mu$$

if we impose on ϵ the condition

$$\epsilon\frac{\mu}{Mr} \leq 1$$

(note that this condition will be automatically satisfied by stronger conditions we will impose later). The lemma is so proved taking

$$c_3 = 12cc_1^2 = 3c^7 2^{4q+14} \tag{2.15}$$

and

$$B_2 = {A_1}^3 {A_2}^3 A_3 A_4 \quad \square \tag{2.16}$$

3 Iteration and conclusion

3.1 Inductive step and convergence of the scheme

In lemma 2.1 we have proved that Kolmogorov's non-degeneracy condition holds for $Q^{(1)} = Q^{(0)} \circ \Phi$ and hence we can iterate proposition 2.1 obtaining via consecutive symplectic transformations the following scheme:

$$\begin{aligned} H &= H^{(0)} = N^{(0)} + \epsilon P^{(0)} \overset{\Phi^{(0)}}{\longmapsto} H^{(1)} = N^{(1)} + \epsilon^2 P^{(1)} \overset{\Phi^{(1)}}{\longmapsto} H^{(2)} = \\ &= N^{(2)} + \epsilon^4 P^{(2)} \dots \overset{\Phi^{(j-1)}}{\longmapsto} H^{(j)} = N^{(j)} + \epsilon^{2^j} P^{(j)} \dots \end{aligned} \tag{3.1}$$

(notice that here $\Phi^{(0)} = \Phi$ in proposition 2.1); to prove theorem 1.2 we must therefore provide in some way the convergence of the scheme.

With proposition 2.1 we have reduced the analycity domain from $B^d_r \times \mathbb{T}^d_\sigma$ to $B^d_{r-2\rho} \times \mathbb{T}^d_{\sigma-2\delta}$, where this loss is due to the inversion of the operator $\mathcal{D}_\omega$ and to the necessity of estimating the derivatives of some analytic functions (see lemmata 1.1 and 1.2) . Let r_j and δ_j be the losses of analycity at each step and $B^d_{r_j} \times \mathbb{T}^d_{\sigma_j}$ the analycity domain after j iterations; in order to be able to iterate infinitely many times the proceeding shown, obtaining a non-empty analycity domain, we must then require that the sequences $\sigma_0 = \sigma$, $\sigma_1 = \sigma_0 - 2\delta_0$, $\sigma_2 = \sigma_1 - 2\delta_1 \dots \sigma_{j+1} = \sigma_j - 2\delta_j = \sigma_0 - 2\sum_{k=1}^{j} \delta_k$ and $r_0 = r$, $r_1 = r_0 - 2\rho_0$, $r_2 = r_1 - 2\rho_1 \dots r_{j+1} = r_j - 2\rho_j = r_0 - 2\sum_{k=1}^{j} \rho_k$ admit a strictly positive limit. For any $\sigma_\infty < \sigma_0$ and $r_\infty < r_0$ we put

$$\delta_j = \frac{1}{2^j} \frac{\sigma_0 - \sigma_\infty}{2} \qquad \rho_j = \frac{1}{2^j} \frac{r_0 - r_\infty}{2} \tag{3.2}$$

in order to have a final analycity domain $B^d_{r_\infty} \times \mathbb{T}^d_{\sigma_\infty}$.

Recall that in lemmata 2.2 and 2.3 we defined

$$\begin{aligned} A_1 &= \max\left\{M\gamma^{-1},\, 1\right\} \\ A_2 &= \max\{MS,\, 1\} \\ A_3 &= \mu \max\left\{S,\, \gamma^{-1}\right\} := \mu \hat{A}_3 \\ A_4 &= \max\{ZS,\, 1\} \end{aligned}$$

and took $B_1 = A_1A_2A_3$ and $B_2 = {A_1}^3{A_2}^3A_3A_4$. We now define inductively the following quantities

$$M_j := \frac{1}{r_j}|Q^{(j)}|_{r_j,\sigma_j}\ , \quad S_j := \frac{1}{r_j}\left|\langle Q^{(j)}_{yy}(0,\cdot)^{-1}\rangle\right|\ , \quad \mu_j := |P^{(j)}|_{r_j,\sigma_j}$$

and the following real numbers

$$\lambda_j := \max\left\{\tfrac{1}{\sigma_\infty}\ ,\ \tfrac{1}{\sigma_j-\sigma_\infty}\right\} \qquad \nu_j := \max\left\{\tfrac{r}{r_\infty}\ ,\ \tfrac{r}{r_j-r_\infty}\right\}$$

$$A_1^{(j)} := \max\{M_j\gamma^{-1},\, 1\} \qquad A_2^{(j)} := \max\{M_jS_j,\, 1\}$$

$$A_3^{(j)} := \mu_j\hat{A}_3^{(j)} = \mu_j r_j^{-1}\max\{S_j,\, \gamma^{-1}\} \quad A_4^{(j)} := \max\{ZS_j,\, 1\}$$

$$B_1^{(j)} := A_1^{(j)}A_2^{(j)}A_3^{(j)} \qquad B_2^{(j)} := {A_1^{(j)}}^3{A_2^{(j)}}^3A_3^{(j)}A_4^{(j)}$$

with the notation $M_0 = M$, $S_0 = S$, $\lambda_0 = \lambda$, $\nu_0 = \nu$, $\mu_0 = \mu$. We are now ready to state

Lemma 3.1. *There exist positive constants $c_4 \geq c_3$ and q_4, depending on $q = \tau + d$, such that if*

$$\epsilon CD\mu < 1 \tag{3.3}$$

with $C = c_4\nu^{14}\lambda^{4q}r^{-1}A_1^4A_2^4A_4\max\{M^{-1},S\}$, $D = 2^{q_4}$, then it is possible to define iteratively (by the scheme described) Hamiltonians $H^{(j)} = N^{(j)} + \epsilon^{2^j}P^{(j)}$ real-analytic on $B^d_{r_j}\times\mathbb{T}^d_{\sigma_j}$ and symplectic transformations $\Phi^{(j)}$ such that $H^{(j+1)} = H^{(j)}\circ\Phi^{(j)}$.

Moreover, referring to the quantities previously defined, for every $j\in\mathbb{N}$ we have

$$Mr \leq M_jr_j \leq 2Mr \tag{3.4}$$

$$S_jr_j \leq 2Sr \tag{3.5}$$

$$\epsilon^{2^j}\mu_j \leq \frac{(\epsilon CD\mu)^{2^j}}{CD^{j+1}} \tag{3.6}$$

and by mere consequence

$$\begin{aligned} A_1^{(j)} &\leq 2A_1\nu \\ A_2^{(j)} &\leq 4A_2\nu^2 \\ \hat{A}_3^{(j)} &\leq 2\hat{A}_3\nu^2 \\ \epsilon^{2^j}A_3^{(j)} &\leq A_3\nu^2 \\ A_4^{(j)} &\leq 2A_4\nu \\ \epsilon^{2^j}\frac{\mu_jr_j}{M_j} &\leq 1 \end{aligned}$$

and

$$\begin{aligned}\epsilon^{2^j} B_1^{(j)} &\leq B_1 \\ \epsilon^{2^j} B_2^{(j)} &\leq B_2\,.\end{aligned}$$

Furthermore, it results that the symplectic transformation $\Phi^{(j)} : B^d_{r_{j+1}} \times \mathbb{T}^d_{\sigma_{j+1}} \to B^d_{r_j} \times \mathbb{T}^d_{\sigma_j}$ *generated by* $F_j(y', x) = y' \cdot x + \epsilon^{2^j} g_j(y', x)$ *(we denote* $F_0 = F$ *), where* $g_j(y', x) = b_j \cdot x + s_j(x) + a_j(x) \cdot y'$ *, is a symplectic diffeomorphism since*

$$\epsilon^{2^j} c_2 B_1^{(j)} {\rho_j}^{-1} {\delta_j}^{-2q} r_j < 1 \tag{3.7}$$

for all $j \in \mathbb{N}$.

Proof We want now to prove by induction inequalities (3.4) to (3.7) . For $j = 0$ condition (3.6) is trivial and (3.4) and (3.5) are obviously satisfied. For what concerns (3.7) observe that

$$\begin{aligned}{\delta_0}^{-m} &= \left(\frac{\sigma_0 - \sigma_\infty}{2}\right)^{-m} \leq 2^m \lambda^m \\ {\rho_0}^{-m} &= \left(\frac{r_0 - r_\infty}{2}\right)^{-m} \leq 2^m \left(\frac{\nu}{r}\right)^m\end{aligned}$$

and therefore we have

$$\epsilon c_2 B_1 \rho_0^{-1} \delta_0^{-2q} r \leq \epsilon c_2 A_1 A_2 A_3 2\nu r^{-1} 2^{2q} \lambda^{2q} r =$$

$$= \quad \epsilon \mu c_2 2^{2q+1} M^{-1} A_1^2 A_2^2 \nu \lambda^{2q} r^{-1} \leq \epsilon C D \mu < 1$$

by hypotheses, taking $c_4 \geq c_2$, $C \geq c_4 M^{-1} A_1^2 A_2^2 \nu \lambda^{2q} r^{-1}$ and $q_4 \geq 2q + 1$ so that (3.7) holds for $j = 0$. During the proof we will come across several lower bounds on c_4, q_4 and C and in the end we will take the worst in order to have all conditions required satisfied simultaneously.

Assume now by induction that conditions from (3.4) to (3.7) hold for $i = 0 \ldots j-1$. Recall that by consequence of lemma 2.3 we have for all $|p| \leq 1$

$$\left|P^{(j+1)}\right|_{r_j, \sigma_j} \leq c_3 \rho_j^{-2} \delta_j^{-4q} r_j^2 {A_1^{(j)}}^3 {A_2^{(j)}}^3 {\hat{A}_3}^{(j)} A_4^{(j)} \mu_j^2 := \mu_{j+1}$$

$$\left|E^{(j+1)} - E^{(j)}\right| \leq \epsilon^{2^j} \mu_{j+1}$$

$$(r - r_j)^{|p|_1} \left|\partial_y^p \left(Q^{(j+1)} - Q^{(j)}\right)\right|_{r_j, \sigma_j} \leq \epsilon^{2^j} \mu_{j+1}$$

where we have denoted $\mu_0 = \mu$ and

$$\mu_1 = c_3 \rho^{-2} \delta^{-4q} r^2 A_1^3 A_2^3 \hat{A}_3 A_4 \mu^2 \ .$$

We now verify (3.6): for all $1 \leq i \leq j$ it results

$$\mu_i = c_3 \rho_{i-1}^{-2} \delta_{i-1}^{-4q} r_{i-1}^2 A_1^{(i-1)^3} A_2^{(i-1)^3} \hat{A}_3^{\ (i-1)} A_4^{(i-1)} \mu_{i-1}^2 \leq$$

$$\leq c_3 \left(\frac{r_o - r_\infty}{2^i} \right)^{-2} \left(\frac{\sigma_o - \sigma_\infty}{2^i} \right)^{-4q} r^2 2^{11} \nu^{12} A_1^3 A_2^3 \hat{A}_3 A_4 \mu_{i-1}^2 \leq$$

$$\leq c_3 2^{4q+13} 2^{(4q+2)(i-1)} \nu^{14} \lambda^{4q} A_1^4 A_2^4 A_4 M^{-1} r^{-1} \mu_{i-1}^2 \leq C_0 D_0^{i-1} \mu_{i-1}^2$$

where this last inequality is obtained taking $C_0 \geq c_4 \nu^{14} \lambda^{4q} A_1^4 A_2^4 A_4 M^{-1} r^{-1}$ with $c_4 \geq c_3 2^{4q+13}$ and $D_0 \geq 2^{4q+2}$ (that is $q_4 \geq 4q + 2$). Now let $\hat{\mu}_i = C_0 D_0^{i+1} \mu_i$ we have

$$\hat{\mu}_i \leq \left(C_0 D_0^{i+1} \right) \left(C_0 D_0^{i-1} \mu_{i-1}^2 \right) = C_0^2 D_0^{2i} \mu_{i-1}^2 = \hat{\mu}_{i-1}^2 \,;$$

therefore iterating we obtain

$$\hat{\mu}_i \leq \hat{\mu}_0^{2^i}$$

for all $i \leq j$ that is, for all $C \geq C_0$ and $D \geq D_0$ it results (taking $i = j$)

$$C D^{j+1} \mu_j \leq (C D \mu)^{2^j} \qquad \Rightarrow \qquad \epsilon^{2^j} \mu_j \leq \frac{(\epsilon C D \mu)^{2^j}}{C D^{j+1}} \,;$$

thus, condition (3.6) holds for every $j \in \mathbb{N}$.

Using (3.6) and hypothesis (3.3) we can obtain

$$\begin{aligned}
|Q^{(j)}|_{r_j, \sigma_j} &= \left| Q^{(0)} + \sum_{i=1}^{j} Q^{(i)} - Q^{(i-1)} \right|_{r_j, \sigma_j} \leq \\
&\leq |Q^{(0)}|_{r_1, \sigma_1} + \sum_{i=1}^{j} |Q^{(i)} - Q^{(i-1)}|_{r_i, \sigma_i} \leq \\
&\leq rM + \sum_{i=1}^{j} \epsilon^{2^{i-1}} \mu_{i-1} \leq rM + \sum_{i=1}^{j} \frac{(\epsilon C D \mu)^{2^{i-1}}}{C D^i} \leq \\
&\leq rM + \sum_{i=1}^{j} \frac{1}{C D^i} \leq rM + \frac{1}{C} \sum_{i=1}^{+\infty} D^{-i} = rM + \frac{1}{C(D-1)} \leq \\
&\leq rM + rM \leq 2rM \ .
\end{aligned}$$

since $C^{-1} \le C_0^{-1} \le Mr$, so that (3.4) is verified.

Let us verify (3.5). Let $B_i := \langle Q_{yy}^{(j)}(0,\cdot)\rangle$ for $i = 0 \dots j-1$, we want to prove $|B_j^{-1}| \le 2S$. Recall that if $A \in \mathrm{mat}(d \times d)$ then $(I+A)^{-1} = \sum_{k=0}^{\infty}(-1)^k A^k$ and $|(I+A)^{-1}| \le \frac{1}{1-|A|}$. Now

$$B_j = B_0 + \sum_{i=1}^{j} B_i - B_{i-1} = B_0 + \hat{B} = B_0 \left(I + B_0^{-1}\hat{B} \right)$$

where obviously we took $\hat{B} = \sum_{i=1}^{j}(B_i - B_{i-1})$. By hypothesis B_0 is invertible, such that to invert B_j we have to invert $I + B_0^{-1}\hat{B}$, that is we want to prove $|B_0^{-1}\hat{B}| < 1$:

$$\begin{aligned}
|B_0^{-1}\hat{B}| &\le |B_0^{-1}| \sum_{i=1}^{j} |B_i - B_{i-1}| \le \\
&\le Sr \sum_{i=1}^{j} \left| \langle \partial_y^2 Q^{(i)}(0,\,\cdot) - Q^{(i-1)}(0,\,\cdot)\rangle \right|_{r_i,\sigma_i} \le \\
&\le Sr \sum_{i=1}^{j} \frac{c}{(r-r_i)^2} \left| Q^{(i)} - Q^{(i-1)} \right|_{r_i,\sigma_i} \le \\
&\le cS \frac{r}{(r-r_i)^2} \sum_{i=1}^{j} \epsilon^{2^i} \mu_{i-1} \le Sc\frac{\nu}{r} \sum_{i=1}^{j} \frac{(\epsilon C D_0 \mu)^{2^i}}{C D_0^{i+1}} \le \\
&\le \frac{cS\nu}{rC} \sum_{i=1}^{\infty} \frac{1}{D_0^{i+1}} = \frac{cS\nu}{rCD_0(D_0-1)} \le \frac{cS\nu}{rC} \le \frac{1}{2}\ .
\end{aligned}$$

if we assume $C \ge 2cS\nu r^{-1}$. The new condition on C_0 is now

$$C_0 \ge c_4 \nu^{14} \lambda^{4q} A_1^4 A_2^4 A_4 r^{-1} \max\{M^{-1}, S\}\ .$$

Also, we have just proved that B_j is invertible and

$$|B_j^{-1}| = r_j S_j = |B_0^{-1}| \left| \left(I + B_0^{-1}\hat{B} \right)^{-1} \right| \le |B_0^{-1}| \frac{1}{1 - |B_0^{-1}\hat{B}|} \le 2Sr.$$

so that $S_j r_j \le 2Sr$ for every $j \in \mathbb{N}$.

To end the proof of this lemma we still need to verify (3.7) for $i = j$. Using

(3.4) to (3.6) and hypothesis (3.3) we have

$$\epsilon^{2^j} c_2 B_1^{(j)} \rho_j{}^{-1} \delta_j{}^{-2q} r_j = c_2 \epsilon^{2^j} \mu_j A_1^{(j)} A_2^{(j)} \hat{A}_3{}^{(j)} \frac{2^{j+1}}{r_o - r_\infty} \left(\frac{2^{j+1}}{\sigma_o - \sigma_\infty} \right)^{2q} r_j \leq$$

$$\leq c_2 \epsilon^{2^j} \mu_j 2^4 A_1 A_2 \hat{A}_3 2^{(2q+1)(j+1)} \nu^6 \lambda^{2q} \leq$$

$$\leq c_2 \frac{(\epsilon C D \mu)^{2^j}}{C D^{j+1}} 2^4 A_1 A_2 \hat{A}_3 2^{(2q+1)(j+1)} \nu^6 \lambda^{2q} \leq$$

$$\leq \frac{c_2}{C D^{j+1}} 2^4 M^{-1} r^{-1} A_1^2 A_2^2 2^{(2q+1)(j+1)} \nu^6 \lambda^{2q} < 1$$

if we take $C \geq c_4 M^{-1} r^{-1} A_1^2 A_2^2 \nu^6 \lambda^{2q}$ with $c_4 \geq c_2 2^4$ and $D \geq 2^{2q+1}$. This lemma is proved by taking

$$\begin{aligned} c_4 &= c_3 2^{4q+13} \\ q_4 &= 4q+2 \qquad \square \end{aligned} \tag{3.8}$$

We are now ready to prove the convergence of the scheme described in (3.1) with the following

Proposition 3.1. *Let $\Phi = \Phi^{(0)}, \Phi^{(1)} \dots \Phi^{(j)}$ the sequence of symplectic diffeomorphisms obtained iterating lemma 2.1; if we define*

$$\Psi^{(j)} = \Phi^{(0)} \circ \Phi^{(1)} \circ \cdots \circ \Phi^{(j)} : B^d_{r_{j+1}} \times \mathbb{T}^d_{\sigma_{j+1}} \to B^d_r \times \mathbb{T}^d_\sigma$$

then the sequence $\Psi^{(j)}$ converges (uniformly) to a symplectic diffeomorphism $\Psi := \lim_{j \to \infty} \Psi^{(j)}$ such that

1. $\Psi = \mathrm{id} + O(\epsilon)$
2. $H^{(0)} \circ \Psi = N^{(\infty)} = E^{(\infty)} + \omega \cdot y' + Q^{(\infty)}(y', x')$

with $N^{(\infty)}$ (that is N' in theorem 1.2) analytic on $B^d_{r_\infty} \times \mathbb{T}^d_{\sigma_\infty}$.

Proof We prove uniform convergence of $\Psi^{(j)}$ which also guarantees the analycity of $N^{(\infty)}$. Let's write $\Psi^{(j)}$ through a telescopic series:

$$\Psi^{(j)} = \Psi^{(0)} + \sum_{i=1}^{j} \Psi^{(i)} - \Psi^{(i-1)} = \Phi + \sum_{i=1}^{j} \Psi^{(i)} - \Psi^{(i-1)} .$$

In lemma 2.2 we obtained that

$$|\Phi - \mathrm{id}|_{r_1, \sigma_1} \leq \epsilon c_2 B_1 \delta_0^{-2q} r$$

since

$$(\Phi - \mathrm{id}\,)(y', x') = \epsilon\left(b + s_x(x) + a_x^T(x) \cdot y',\ \tilde{a}(x')\right)_{x=\tilde{\varphi}(x')}$$

and each term was estimated with $c_1 B_1 \delta^{-2q} r$ and $c_2 \geq 4c_1$. By induction we can therefore assume

$$|\Phi^{(i)} - \mathrm{id}\,|_{r_i,\sigma_i} \leq \epsilon^{2^i} c_2 B_1^{(j)} \delta_j^{-2q} r_j$$

which implies, together with lemma 3.1 ,

$$\left|\Psi^{(i)} - \Psi^{(i-1)}\right|_{r_{i+1},\sigma_{i+1}} = \left|\Phi^{(i)} \circ \Psi^{(i-1)} - \Psi^{(i-1)}\right|_{r_{i+1},\sigma_{i+1}} \leq$$

$$\leq \epsilon^{2^i} c_2 B_1^{(j)} \delta_j^{-2q} r_j = c_2 \epsilon^{2^i} \mu_j A_1^{(j)} A_2^{(j)} \hat{A}_3{}^{(j)} \delta_j^{-2q} r_j \leq$$

$$\leq c_2 \epsilon^{2^i} \mu_j 2^4 \nu^5 A_1 A_2 \hat{A}_3 \lambda^{2q} 2^{2q(j+1)} r \leq$$

$$\leq c_2 \frac{(\epsilon C_0 D_0 \mu)^{2^i}}{C_0 D_0^{i+1}} 2^4 \nu^5 A_1^2 A_2^2 M^{-1} r^{-1} \lambda^{2q} 2^{2q(i+1)} r \leq (\epsilon C_0 D_0 \mu)^{2^i} r$$

since in lemma 3.1 we took $C_0 \geq c_4 \nu^{16} \lambda^{4q} A_1^4 A_2^4 A_4 M^{-1} r^{-1}$ and $D_0 \geq 2^{4q+2}$ (notice that $\nu > 1$) . Therefore we can estimate $|\Psi - \mathrm{id}\,|$ as follows :

$$|\Psi - \mathrm{id}\,|_{r_\infty,\sigma_\infty} \leq |\Phi - \mathrm{id}\,|_{r_\infty,\sigma_\infty} + \sum_{i=1}^{\infty} \left|\Psi^{(i)} - \Psi^{(i-1)}\right|_{r_\infty,\sigma_\infty} \leq$$

$$\leq |\Phi - \mathrm{id}\,|_{r_1,\sigma_1} + \sum_{i=1}^{\infty} \left|\Psi^{(i)} - \Psi^{(i-1)}\right|_{r_{i+1},\sigma_{i+1}} \leq$$

$$\leq \epsilon c_2 B_1 \delta_0^{-2q} r + \sum_{i=1}^{\infty} (\epsilon C_0 D_0 \mu)^{2^i} r \leq$$

$$\leq \epsilon \mu c_2 A_1 A_2 \hat{A}_3 \delta_0^{-2q} r + \sum_{i=1}^{\infty} (\epsilon C_0 D_0 \mu)^{2^i} r \leq$$

$$\leq \epsilon \mu c_2 M^{-1} r^{-1} A_1^2 A_2^2 2^{2q(i+1)} \lambda^{2q} r + \sum_{i=2}^{\infty} (\epsilon C_0 D_0 \mu)^i r \leq$$

$$\leq \epsilon C_0 D_0 \mu r + \sum_{i=2}^{\infty} (\epsilon C_0 D_0 \mu)^i r$$

since, always by lemma 3.1, it results $C_0 \geq c_2 M^{-1} r^{-1} \lambda^{2q}$ and $D_0 \geq 2^{2q}$. Then, taking $D \geq 2D_0$, that is to say the new hypothesis is $\epsilon C_0 D \mu < 1$ and hence

$\epsilon C_0 D_0 \mu < \frac{1}{2}$, we obtain

$$|\Psi - \mathrm{id}|_{r_\infty, \sigma_\infty} \leq \epsilon C_0 D_0 \mu r + \sum_{i=2}^{\infty} (\epsilon C_0 D_0 \mu)^i r \leq$$

$$\leq \epsilon C_0 D_0 \mu r + \frac{(\epsilon C_0 D_0 \mu)^2}{1 - \epsilon C_0 D_0 \mu} r \leq \epsilon C_0 D_0 \mu r + 2 (\epsilon C_0 D_0 \mu)^2 r \leq \epsilon C_0 D \mu r.$$

Thus $\Psi^{(j)}$ converges uniformly to Ψ and $N^{(\infty)} = H^{(0)} \circ \Psi$ is analytic. To conclude we trivially observe that

$$\epsilon^{2^j} \left| P^{(j)} \right|_{\sigma_j,\, \sigma_j} \leq \epsilon^{2^j} \mu_j \leq (\epsilon C D \mu)^{2^j} \overset{j \to \infty}{\longrightarrow} 0$$

so that $N^{(\infty)}$ is effectively in Kolmogorov's normal form $\square$

3.2 Final estimates

To completely prove theorem 1.2 we still need to estimate $|E^{(\infty)} - E^{(0)}|$ and $\left\| Q^{(\infty)} - Q^{(0)} \right\|_{C^1}$. Recall first that in order to have all inductive conditions satisfied we must take $\epsilon C D \mu < 1$ for any

$$C \geq C_0 = c_4 \nu^{14} \lambda^{4q} r^{-1} \max\{M^{-1}, S\} A_1^4 A_2^4 A_4 \tag{3.9}$$

$$D \geq D_0 = 2^{4q+2} \tag{3.10}$$

with $c_4 = c_3 2^{4q+13}$. Now using an estimate done in the proof of lemma 2.3 we have

$$\left| E^{(1)} - E^{(0)} \right| \leq c_1 \epsilon \delta_0^{-q} \mu A_1 A_2 \,;$$

therefore, by inductive hypotheses and lemma 3.1 we obtain

$$\left| E^{(j+1)} - E^{(j)} \right| \leq c_1 \epsilon^{2^j} \delta_j^{-2q} \mu_j A_1^{(j)} A_2^{(j)} \leq$$

$$\leq c_1 \epsilon^{2^j} \mu_j \lambda^{2q} 2^{2q(j+1)} 2^3 \nu^3 A_1 A_2 \leq$$

$$\leq \frac{(\epsilon C D_0 \mu)^{2^j}}{C D_0^{j+1}} 2^{2q(j+1)} c_1 2^3 \nu^3 \lambda^{2q} A_1 A_2 \leq (\epsilon C D_0 \mu)^{2^j} M r$$

for any $C \geq C_0$. Now, writing $E^{(\infty)}$ as a telescopic series and taking $D \geq 2D_0$, in order to have $\epsilon C D_0 \mu < \frac{1}{2}$, it results

$$\begin{aligned}
&\left|E^{(\infty)} - E^{(0)}\right| \leq \left|E^{(1)} - E^{(0)}\right| + \sum_{j=1}^{\infty} \left|E^{(j+1)} - E^{(j)}\right| \leq \\
&\leq c_1 \epsilon \delta_0^{-q} \mu A_1 A_2 + Mr \sum_{j=1}^{\infty} \left(\epsilon C D_0 \mu\right)^{2^j} \leq \\
&\leq c_1 \epsilon \lambda^q 2^q \mu A_1 A_2 + Mr \sum_{j=2}^{\infty} \left(\epsilon C D_0 \mu\right)^j \leq \\
&\leq \epsilon C D_0 \mu M r + Mr \frac{\left(\epsilon C D_0 \mu\right)^2}{1 - \epsilon C D_0 \mu} \leq \epsilon C D \mu M r \ .
\end{aligned}$$

In a completely analogous way we can estimate $\left|\partial_y^p (Q^{(\infty)} - Q^{(0)})\right|_{r_\infty, \sigma_\infty}$; in lemma 2.3 we obtained

$$\rho_0^{|p|_1} \left|\partial_y^p (Q^{(1)} - Q^{(0)})\right|_{r_1, \sigma_1} \leq c_3 \epsilon \rho_0^{-3} \delta_0^{-2q} r^3 A_1^2 A_2^2 \mu \,.$$

Thus, by induction

$$\begin{aligned}
&(r - r_{j+1})^{|p|_1} \left|\partial_y^p (Q^{(j+1)} - Q^{(j)})\right|_{r_{j+1}, \sigma_{j+1}} \leq \\
&\leq c_3 \epsilon^{2^j} \rho_j^{-3} \delta_j^{-2q} r_j^3 {A_1^{(j)}}^2 {A_2^{(j)}}^2 \mu_j \leq c_3 \epsilon^{2^j} \mu_j 2^{(3+2q)(j+1)} \nu^3 \lambda^{2q} 2^6 \nu^6 A_1^2 A_2^2 \leq \\
&\leq \frac{\left(\epsilon C D_0 \mu\right)^{2^j}}{C D_0^{j+1}} 2^{(3+2q)(j+1)} \nu^9 \lambda^{2q} c_3 2^6 A_1^2 A_2^2 \leq \left(\epsilon C D_0 \mu\right)^{2^j} Mr \,;
\end{aligned}$$

writing as usual $Q^{(\infty)}$ as a telescopic series we obtain for $|p|_1 \leq 2$

$$(r - r_\infty)^{|p|_1} \left|\partial_y^p (Q^{(\infty)} - Q^{(0)})\right|_{r_\infty, \sigma_\infty} \leq (r - r_1)^{|p|_1} \left|\partial_y^p (Q^{(1)} - Q^{(0)})\right|_{r_1, \sigma_1} +$$

$$+ \sum_{j=1}^{\infty} (r - r_{j+1})^{|p|_1} \left|\partial_y^p (Q^{(j+1)} - Q^{(j)})\right|_{r_{j+1}, \sigma_{j+1}} \leq$$

$$\leq c_3 \epsilon \rho_0^{-3} \delta_0^{-2q} r^3 A_1^2 A_2^2 \mu + Mr \sum_{j=1}^{\infty} (\epsilon C D_0 \mu)^{2^j} \leq$$

$$\leq \epsilon c_3 2^{3+2q} \nu^3 \lambda^{2q} A_1^2 A_2^2 \mu + Mr \sum_{j=2}^{\infty} (\epsilon C D_0 \mu)^j \leq$$

$$\leq \epsilon C D_0 \mu M r + Mr \frac{(\epsilon C D_0 \mu)^2}{1 - \epsilon C D_0 \mu} \leq \epsilon C D \mu M r$$

having imposed the same previous condition $D \geq 2D_0$.

We now conclude remarking that by the estimates done we can take $\epsilon C D \mu < 1$ with (see (3.9), (3.10), (2.15), (3.8))

$$\begin{aligned} c_4 &= 3c^7 2^{8(\tau+d)+27} \\ C &= c_4 \nu^{14} \lambda^{4(\tau+d)} A_1^4 A_2^4 A_4 r^{-1} \max\{M^{-1}, S\} \\ D &= 2^{4(\tau+d)+3} \end{aligned}$$

where $c = c(\tau, d)$ is defined in lemma 1.2); this condition is equivalent to

$$\epsilon < \epsilon_0 := \frac{r}{3\mu} c^{-7} 2^{-(12(\tau+d)+30)} \nu^{-3} \lambda^{-4(\tau+d)} (A_1 A_2)^{-4} A_4^{-1} \min\{M, S^{-1}\}\,. \quad (3.11)$$

References

[Arn63] V. I. Arnol'd. Proof of a theorem by A. N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian. *Russian Math. Survey*, 18:13–40, 1963.

[Chi05] L. Chierchia. Il Problema dei 3 Corpi Ristretto Circolare Piano, 2004-2005. Lecture notes from the course AM8. Università degli Studi Roma Tre.

[Kol54] A.N. Kolmogorov. On the Conservation of Conditionally Periodic Motions under Small Perturbation of the Hamiltonian. *Dokl. Akad. Nauk SSR*, 98:527–530, 1954.