EXPONENTIALLY STABLE MANIFOLDS IN THE NEIGHBOURHOOD OF ELLIPTIC EQUILIBRIA

ANTONIO GIORGILLI
Dipartimento di Matematica e Applicazioni, Via degli Arcimboldi 8, 20126 — Milano, Italy.

DANIELE MURARO
Università di Milano Bicocca, Piazza dell'Ateneo Nuovo 1, 20126 — Milano, Italy.

Abstract. We consider a Hamiltonian system in a neighbourhood of an elliptic equilibrium which is a minimum for the Hamiltonian. With appropriate non–resonance conditions we prove that in the neighbourhood of the equilibrium there exist low dimensional manifolds that are exponentially stable in Nekhoroshev's sense. This generalizes the theorem of Lyapounov on the existence of periodic orbits. The result may be meaningful for, e.g., the dynamics of non–linear chains of FPU type.

1. Introduction

In the celebrated report published in 1955[8] Fermi, Pasta and Ulam pointed out the phenomenon of localization of energy in low frequency modes, or freezing of high frequencies, in a discretized model of a string. The phenomenon has been later investigated by many authors (see, e.g., [6], [2], [3], [12], [13], [11], [15], [9], [4], [7] and the references therein). In particular in a recent paper the existence of the so called "natural packets of modes" has been discussed on the basis of a numerical experiment[1].

The present paper is devoted to a theoretical investigation of that phenomenon, with the aim of understanding the underlying dynamics. This is performed by generalizing a theorem of Lyapounov on continuation of periodic orbits combined with ideas coming from Nekhoroshev's theory on exponential stability.

The phenomenon of localization of energy is the following. One considers a non-linear chain of N particles with a nearest neighbours interaction as described by the Hamiltonian

$$H(q,p) = \sum_{j=1}^{N} \frac{p_j^2}{2} + \sum_{j=0}^{N} V(q_{j+1} - q_j) \ , \quad q_0 = q_{N+1} = 0 \ ,$$

where

$$V(r) = \frac{1}{2} r^2 + \frac{\alpha}{3} r^3 + \frac{\beta}{4} r^4 \ .$$

Here, q and p are the displacements of the particles from equilibrium and the momenta of the particles, respectively, and α and β are parameters. It is well known that the model above represents in fact a system of N harmonic oscillators with frequency spectrum

$$\omega_k = 2\sin\frac{k\pi}{2(N+1)} \ ,$$

and with a nonlinear coupling due to the cubic and quartic terms of the interaction potential.

Starting (as Fermi, Pasta and Ulam did) with initial conditions with the whole energy concentrated on the lowest frequency mode, one observes that the energy flows in a short time only to a few modes of low frequency, that we call the "natural packet", while the high frequencies remain frozen. A further sharing of energy with the high frequency modes takes a longer time that grows very fast when the energy E decreases to zero. In the recent paper [1] the size of the natural packet as a function of the specific energy E/N and of the number N of particles has been investigated in detail, showing that the phenomenon of localization of energy may persist in the so called thermodynamic limit, i.e., when the number of particles grows very large. This fact may be very relevant for the foundations of Statistical Mechanics.

From the mathematical viewpoint we are actually dealing with a Hamiltonian system with n degrees of freedom in a neighbourhood of an elliptic equilibrium, with Hamiltonian

$$H(x,y) = H_0(x,y) + H_1(x,y) + \ldots \ , \tag{1}$$

where

$$H_0(x,y) = \frac{1}{2}\sum_{l=1}^{n}\omega_l(x_l^2 + y_l^2) \ ,$$

$\omega \in \mathbf{R}^n$ is the real vector of the frequencies, $\omega_1, \ldots, \omega_n$ are all positive, and $H_s(x,y)$ for $s \geq 1$ is a homogeneous polynomial of degree $s+2$. We assume also that the equilibrium is a minimum for the Hamiltonian, so that $\omega_1, \ldots, \omega_n$ are all positive.

A possible interpretation of the phenomenon of localization of energy in the FPU model is the following. Let us consider for a moment only the quadratic part of the Hamiltonian, i.e., let $H_1 = \ldots = 0$. A packet of $m < n$ modes is easily constructed by considering initial condition with $x_{m+1} = y_{m+1} = \ldots = x_n = y_n = 0$, and with arbitrary values for $x_1, y_1, \ldots, x_m, y_m$. In such a case, all orbits lie on an invariant subspace of dimension $2m$, defined by $x_{m+1} = y_{m+1} = \ldots = x_n = y_n = 0$, which is foliated in invariant tori of dimension m. The problem is whether something similar may happen for the complete nonlinear system. For $m = 1$ a positive answer is given by Lyapounov's theorem on the continuation of periodic orbits that reduce to the normal modes when the perturbation decreases to zero. For $m > 1$ one is tempted to generalize the theorem of Lyapounov by proving the existence of an invariant manifold of dimension $2m$ that reduces to the invariant plane for vanishing perturbation. However, such an attempt will likely fail for two reasons. The first one is that the construction of such an invariant manifold may be performed in a formal sense, as we shall do below, but the resulting series are likely to be divergent, as typically happens for perturbation expansions. The second reason is that, according to the results of the numerical exploration, the internal dynamics of the natural packet appears to exhibit a chaotic behaviour, which is not compatible with the plain foliation in invariant tori of the corresponding unperturbed case.

Our suggestion is to overcome the two difficulties above as follows. First, we look for an approximate normal form that possesses an invariant manifold, but with no restrictions on the dynamics on that manifold. Second, we look for a quasi–invariance of that manifold in the sense of Nekhoroshev's theory on exponential stability, i.e., we prove that the orbit is confined in a very small neighbourhood of the quasi–invariant manifold for an exponentially long time.

We give now a formal statement. Let $\Delta_\varrho \subset \mathbf{R}^{2n}$ be a polydisk of radius ϱ centered at the origin. Our main result is the following

Theorem: *Let the Hamiltonian (1) be analytic in a neighbourhood of the origin, and assume that the frequencies $\omega_1,\ldots,\omega_n$ are all positive and satisfy for any $\mu,\nu \in \{m+1,\ldots,n\}$ the diophantine conditions*

$$
\begin{aligned}
|k_1\omega_1+\ldots+k_m\omega_m \pm \omega_\nu| &\geq \gamma|k|^{-\tau}, \\
|k_1\omega_1+\ldots+k_m\omega_m \pm 2\omega_\nu| &\geq \gamma|k|^{-\tau}, \\
|k_1\omega_1+\ldots+k_m\omega_m \pm \omega_\nu \pm \omega_\mu| &\geq \gamma|k|^{-\tau} \qquad \text{for } \mu\neq\nu
\end{aligned}
\tag{2}
$$

where γ and τ are positive constants and $m<n$ is a positive integer. Then there exists a positive ϱ^* and a local analytic manifold V of dimension $2m$, tangent at the origin to the subspace $x_{m+1}=y_{m+1}=\ldots=x_n=y_n=0$, such that the following holds true: let $\varrho<\varrho^*/3^{\tau+2}$, and let $\sigma=\exp\left(-\frac{1}{2a}\left(\frac{\varrho^*}{\varrho}\right)^a\right)$, with $a=1/(\tau+2)$; then there exists a positive $\bar\varrho<\varrho$ such that for every orbit with initial point $(x_0,y_0)\in\Delta_{\bar\varrho}$ satisfying $\mathrm{dist}\big((x_0,y_0),V\big)<\sigma/2$ one has $(x_t,y_t)\in\Delta_\varrho$ and $\mathrm{dist}\big((x_t,y_t),V\big)<\sigma$ for

$$
|t| < T_* \exp\left(\frac{1}{2a}\left(\frac{\varrho^*}{\varrho}\right)^a\right),
$$

with a constant T_*.

Remark. By a standard argument from diophantine theory one may choose $\tau \geq m-1$.

The idea of looking for a quasi–invariant manifold has been suggested to us by the papers of Cherry [5] and Moser [14], who in the case $m=2$ prove the existence of an invariant manifold, but with the assumption that the frequencies ω_1,ω_2 are independent over the reals. This result does not apply to the case of an elliptic equilibrium. However, the technique that has been used in [10] may be adapted in order to prove our theorem.

The paper is organized as follows. In sect. 2 we characterize the normal form of the Hamiltonian and recall the formal algorithm. In sect. 3 we give the quantitative estimates for the normal form and the remainder. In sect. 4 we show how to use the normal form in order to complete the proof of the theorem.

In order to make the paper self–contained, some technical details are deferred to two appendixes.

2. Formal algorithm

We first discuss the formal algorithm that gives the Hamiltonian a normal form up to a finite order r. The general method based on composition of canonical transformations via Lie series is rather well known. Thus we recall only the relevant formulae. For completeness sake we include a short deduction in Appendix A. Our aim is to write the Hamiltonian after r normalization steps as

$$H^{(r)}(x,y) = H_0(x,y) + Z_1(x,y) + \ldots + Z_r(x,y) + \sum_{s>r} H_s^{(r)}(x,y) \ , \tag{3}$$

where $Z_1(x,y),\ldots,Z_r(x,y)$ satisfy the condition of being in "normal form", in a sense to be made precise. The original Hamiltonian $H_0 \equiv H$ is given in this form with $r=0$, so that it contains no function Z. The normalization procedure is worked out step by step, using a generating sequence $\{\chi_1,\ldots,\chi_r\}$. Precisely, we perform a sequence of r canonical transformations, r being arbitrary, by recursively determining the new Hamiltonian as

$$H^{(r)} = \exp{(L_{\chi_r})} H^{(r-1)},$$

where $L_\varphi \cdot := \{\cdot,\varphi\}$, the Poisson bracket with φ, and

$$\exp{(L_\varphi)} = 1 + L_\varphi + \frac{1}{2!}L_\varphi^2 + \ldots = \sum_{s\geq 0}\frac{1}{s!}L_\varphi^s.$$

The generating function is determined by solving with respect to χ_r and Z_r the equation

$$L_{H_0}\chi_r + Z_r = H_r^{(r-1)} \ . \tag{4}$$

Assuming for a moment that a solution of the latter equation has been found, the transformed Hamiltonian is determined as

$$\begin{aligned} H_{sr+m}^{(r)} &= \frac{1}{s!}L_{\chi_r}^s Z_m + \sum_{p=0}^{s-1}\frac{1}{p!}L_{\chi_r}^p H_{(s-p)r+m}^{(r-1)} \qquad \text{for } r\geq 2\ ,\ s\geq 1 \text{ and } 1\leq m<r\ ,\\ H_{sr}^{(r)} &= \frac{1}{(s-1)!}L_{\chi_r}^{s-1}\left(\frac{1}{s}Z_r + \frac{s-1}{s}H_r^{(r-1)}\right) + \sum_{p=0}^{s-2}\frac{1}{p!}L_{\chi_r}^p H_{(s-p)r}^{(r-1)}\\ &\qquad\qquad \text{for } r\geq 1 \text{ and } s\geq 2\ . \end{aligned} \tag{5}$$

Let us now discuss how to solve equation (4). The usual remark is that L_{H_0} is a linear operator mapping the space $\mathcal{P}_s$ of homogeneous polynomials of degree s in x,y into itself. Moreover, L_{H_0} may be diagonalized via the canonical transformation to complex variables ξ,η

$$x_l = \frac{\xi_l + i\eta_l}{\sqrt{2}},\quad y_l = \frac{i(\xi_l - i\eta_l)}{\sqrt{2}}. \tag{6}$$

For, in the new variables one has $H_0 = i\sum_{l=1}^n \omega_l\xi_l\eta_l$, and one checks immediately that

$$L_{H_0}\xi^j\eta^k = i\langle j-k,\omega\rangle\xi^j\eta^k \tag{7}$$

for any monomial $\xi^j\eta^k = \xi_1^{j_1}\cdot\ldots\cdot\xi_n^{j_n}\cdot\eta_1^{k_1}\ldots\eta_n^{k_n}$. Denoting by $\mathcal{P}$ the linear space of all polynomials, the kernel $\mathcal{N}$ and the range $\mathcal{C}$ of L_{H_0} are introduced in the usual way, namely $\mathcal{N} = L_{H_0}^{-1}(0)$, the inverse image of the null vector in $\mathcal{P}$, and $\mathcal{C} = L_{H_0}(\mathcal{P})$. Both $\mathcal{N}$ and $\mathcal{C}$ are subspaces of $\mathcal{P}$, and in view of L_{H_0} being diagonalizable we also have $\mathcal{N}\cap\mathcal{C} = \{0\}$ and $\mathcal{N}\oplus\mathcal{C} = \mathcal{P}$. A consequence of these properties is that L_{H_0} restricted to the subspace $\mathcal{C}$ is uniquely inverted, i.e., the equation $L_{H_0}\chi = \psi$ with $\psi\in\mathcal{C}$ admits a unique solution χ satisfying $\chi\in\mathcal{C}$.

We now come to make precise the meaning of "normal form" in our case. Having fixed the positive integer $m < n$ consider the disjoint subsets of $\mathbf{Z}^n$.

$$
\begin{aligned}
\mathcal{K}^\sharp &= \{k\in\mathbf{Z}^n \ :\ k_{m+1} = \ldots = k_n = 0\}\ ,\\
\mathcal{K}^\natural &= \{k\in\mathbf{Z}^n \ :\ 1\le |k_{m+1}|+\ldots+|k_n|\le 2\}\ ,\\
\mathcal{K}^\flat &= \{k\in\mathbf{Z}^n \ :\ |k_{m+1}|+\ldots+|k_n| > 2\}\ .
\end{aligned}
\tag{8}
$$

One has $\mathbf{Z}^n = \mathcal{K}^\sharp\cup\mathcal{K}^\natural\cup\mathcal{K}^\flat$, of course. Considering only integer vectors j,k with non–negative components, introduce the subspaces of $\mathcal{P}$

$$
\begin{aligned}
\mathcal{P}^\sharp &= \mathrm{span}\left\{x^j y^k \ :\ j+k\in\mathcal{K}^\sharp\right\}\\
\mathcal{P}^\natural &= \mathrm{span}\left\{x^j y^k \ :\ j+k\in\mathcal{K}^\natural\right\}\\
\mathcal{P}^\flat &= \mathrm{span}\left\{x^j y^k \ :\ j+k\in\mathcal{K}^\flat\right\}
\end{aligned}
\tag{9}
$$

These subspaces are clearly disjoint, and moreover one has $\mathcal{P} = \mathcal{P}^\sharp\oplus\mathcal{P}^\natural\oplus\mathcal{P}^\flat$. Let us now define the subspace

$$\mathcal{Z}^\natural = \{f\in\mathcal{P}^\natural : \{f, I_\nu\} = 0,\ \nu = m+1,\ldots,n\}$$

and split $\mathcal{P}^\natural = \mathcal{Z}^\natural\oplus\mathcal{W}^\natural$ where $\mathcal{W}^\natural$ is the complementary subspace to $\mathcal{Z}^\natural$ in $\mathcal{P}^\natural$. Moreover, let us consider the subspaces $\mathcal{Z}$ and $\mathcal{W}$ of $\mathcal{P}$ defined as

$$\mathcal{Z} = \mathcal{P}^\sharp\oplus\mathcal{Z}^\natural\oplus\mathcal{P}^\flat,\quad \mathcal{W}\equiv\mathcal{W}^\natural.$$

One clearly has $\mathcal{Z}\cap\mathcal{W} = \{0\}$ and $\mathcal{Z}\oplus\mathcal{W} = \mathcal{P}$, and moreover we also have, by construction, $\mathcal{W}\subset\mathcal{R}$ (see below). With this setting the equation

$$L_{H_0}\chi + Z = \psi, \tag{10}$$

with known ψ, admits a simple solution. Split $\psi = \psi_{\mathcal{Z}}+\psi_{\mathcal{W}}$ with $\psi_{\mathcal{Z}}\in\mathcal{Z}$ and $\psi_{\mathcal{W}}\in\mathcal{W}$; such a decomposition exists and is unique, since $\mathcal{Z}$ and $\mathcal{W}$ are complementary subspaces. Then set $Z = \psi_{\mathcal{Z}}$ and determine $\chi = L_{H_0}^{-1}\psi_{\mathcal{W}}$ as the solution of (10) made unique with the prescription $\chi\in\mathcal{R}$.

It is also interesting to characterize $\mathcal{P}^\natural$ and $\mathcal{Z}^\natural$ using the complex coordinates, and to investigate its relations with the null space $\mathcal{N}$. By the way, this will allow us to check in a direct way that $\mathcal{W}\subset\mathcal{R}$, which is essential in order to assure the consistency of the whole construction. Let us consider $\xi^j\eta^k\in\mathcal{P}^\natural$. Then at most two of the factors ξ_ν,η_ν with $\nu = m+1,\ldots,n$ may appear. Recall that in complex variables $I_\nu = i\xi_\nu\eta_\nu$, we have

$$\{\xi^j\eta^k, I_\nu\} = i(j_\nu - k_\nu)\xi^j\eta^k\ .$$

Thus, a basis for $\mathcal{Z}^\natural$ is the set of monomials $\xi_\nu\eta_\nu \cdot \varphi(\xi_1,\ldots,\xi_m,\eta_1,\ldots,\eta_m)$, where φ is any monomial. This fully characterizes $\mathcal{Z}^\natural$. Let us now characterize the subspace $\mathcal{N}\cap\mathcal{P}^\natural$. This is easily done by looking at (7). In view of the diagonal form of L_{H_0} it is easily seen that a basis of $\mathcal{N}$ is made by all monomials $\xi^j\eta^k$ such that $\langle j-k,\omega\rangle = 0$. Thus, considering again a monomial $\xi^j\eta^k \in \mathcal{P}^\natural$ we have either

$$\langle j-k,\omega\rangle = (j_1-k_1)\omega_1 + \ldots + (j_m-k_m)\omega_m + \lambda_\nu\omega_\nu \quad \text{with } \lambda_\nu = \pm 1,$$

or

$$\langle j-k,\omega\rangle = (j_1-k_1)\omega_1 + \ldots + (j_m-k_m)\omega_m + \lambda_\nu\omega_\nu + \lambda^{'}_\mu\omega_\mu \quad \text{with } \lambda_\nu, \lambda^{'}_\mu = \pm 1,$$

the case $\mu=\nu$ being allowed. In the first case $\langle j-k,\omega\rangle$ can not vanish in view of the first condition (2); in the second case, in view of the second and third condition (2) $\langle j-k,\omega\rangle$ may vanish only if $\mu=\nu$ and $\lambda_\nu+\lambda^{'}_\mu = 0$, i.e., if $\xi^j\eta^k \in \mathcal{Z}^\natural$. Thus we have $\mathcal{N}\cap\mathcal{P}^\natural \subset \mathcal{Z}^\natural$, and in view of $\mathcal{W}=\mathcal{W}^\natural$ we conclude $\mathcal{W}\subset\mathcal{R}$, as claimed.

3. Quantitative estimates for normal form and remainder

We first introduce a norm on the space $\mathcal{P}_s$ of homogeneous polynomials of degree s. Pick a real vector $R\in\mathbf{R}^n$ with positive components and consider the domain

$$\Delta_R = \{(x,y)\in\mathbf{C}^n \ :\ |x_j|\le R_j\ ,\ |y_j|\le R_j\ ,\ 1\le j\le n\}\ , \tag{11}$$

namely a polydisk which is the product of disks of radii $R_1,\ldots,R_n$ in the planes of the complex coordinates $(x_1,\ldots,x_n)$ and $(y_1,\ldots,y_n)$, respectively. Let also

$$\Lambda = \min_{1\le j\le n} R_j\ . \tag{12}$$

The norm $\|f\|_R$ of the function $f=\sum_{j,k} f_{j,k}x^jy^k$ in the polydisk Δ_R is defined as

$$\|f\|_R = \sum_{j,k} |f_{j,k}|R^{j+k}\ . \tag{13}$$

A family of polydisks $\Delta_{\delta R}$ of radii δR, with $0<\delta\le 1$ will be considered below. With a minor abuse the simplified notation $\|\cdot\|_\delta$ in place of $\|\cdot\|_{\delta R}$ will be used.

The main technical tool is the following inequality, that resembles the Cauchy's estimate for derivatives of functions: assume that $\|\chi\|_{1-\delta'}$ and $\|f\|_{1-\delta'}$ are known for some non negative $\delta'<1$. Then for any positive $1>\delta>\delta'$ and for integer $p\ge 1$ we have

$$\|L^p_\chi f\|_{1-\delta} \le \frac{p!}{e^2}\left(\frac{e}{d\Lambda}\right)^{2p} \|\chi\|^p_{1-\delta'}\|f\|_{1-\delta'} \tag{14}$$

where $d:=\delta-\delta'$. The proof of this inequality is given in appendix B.

We look now for quantitative estimates of all functions entering the construction of the normal form according to the algorithm in section 2. We proceed as follows:

(i) we transform the Hamiltonian (1) to complex variables ξ,η via the transformation (6);

(ii) we perform an arbitrary number of normalization steps, thus getting the Hamiltonian $H^{(r)}$ in the form (3), in complex variables;

(iii) we transform back to real variables by applying the inverse of (6) to $H^{(r)}$.

We remark that the transformations to complex variables and back are harmless for what concerns the norms of the functions. For, if $f \in \mathcal{P}_s$ is given in real (resp. complex) variables and $f' \in \mathcal{P}_s$ is the transformed function in complex (resp. real) variables, then one has

$$\|f'\| \le 2^{s/2}\|f\|.$$

Let the non increasing sequence $\{\alpha_r\}_{r\ge1}$ be defined as

$$\alpha_r = \min_{0<|k|\le r+2} |\langle k,\omega\rangle| \qquad \text{for } k \in \mathcal{K}^\natural$$

By the hypotheses of the Theorem we have $\alpha_r > 0$ for all $r \ge 1$. We consider the solution of the equation (4) for the generating function. This is better done in complex variables ξ,η defined by (6), because the operator L_{H_0} takes a diagonal form. We immediately get

$$\|\chi_r\|_{1-\delta} \le \frac{\|H_r^{(r-1)}\|_{1-\delta}}{\alpha_r}, \qquad \|Z_r\|_{1-\delta} \le \|H_r^{(r-1)}\|_{1-\delta}, \tag{15}$$

where we have assumed that $\|H_r^{(r-1)}\|_{1-\delta}$ is known for some positive $\delta < 1$. We look now for recursive estimates generated by the formal algorithm. To this end, let $\{d_r\}_{r\ge1}$ be a non increasing sequence such that $\sum_r d_r = d < 1$, and let $\delta_0 = 0$ and $\delta_r = d_1 + \ldots + d_r$. Let us define the sequence $\{\eta_{r,s}\}_{r\ge0,s\ge0}$ as

$$\begin{aligned} \eta_{0,s} &= C^{s-1} \\ \eta_{r,s} &= C^{s-1}\mu_{r,s}\frac{1}{\alpha_1 d_1^2 \ldots \alpha_r d_r^2}\cdot\left(\frac{1}{\alpha_r d_r^2}\right)^{s-r-1} \end{aligned} \tag{16}$$

where

$$C = \frac{e^2E}{\Lambda^2} + h$$

and the sequence $\{\mu_{r,s}\}_{r\ge0,s\ge0}$ is recursively defined as

$$\begin{aligned} \mu_{0,0} &= 0, \qquad \mu_{0,s} = 1 \qquad \text{for } s > 0, \\ \mu_{r,sr+m} &= \sum_{p=0}^{s} \mu_{r-1,r}^{p}\mu_{r-1,(s-p)r+m} \qquad \text{for } s \ge 0,\ 0 \le m < r. \end{aligned}$$

We also consider the sequence of boxed domains $\Delta_{(1-\delta_r)R}$. The recursive estimates are collected in the following

Lemma 1: *If the Hamiltonian (1) satisfies (in complex variables)*

$$\|H_s\|_1 \le h^{s-1}E \qquad \textit{for } s \ge 1$$

with constants $h \geq 0$ and $E > 0$, then the following estimates hold true:

$$\|\chi_r\|_{1-\delta_{r-1}} \leq \frac{\eta_{r-1,r}E}{\alpha_r} \tag{17}$$

$$\|Z_r\|_{1-\delta_{r-1}} \leq \eta_{r-1,r}E \tag{18}$$

$$\|H_s^{(r)}\|_{1-\delta_r} \leq \eta_{r,s}E \tag{19}$$

Proof. By induction. For $r = 0$ only the estimate (19) is meaningful and is trivially true in view of the definition of $\eta_{0,s}$. In order to perform the induction, first remark that if (19) holds true up to $r-1$, then (17) and (18) are true for r; this is a straightforward consequence of the inequalities (15), with $\delta = \delta_{r-1}$. Thus, it remains to prove only that (19) hold true for r as a consequence of (17) and (18) and of the recursive definitions (5). Remarking that Z_r and $H_r^{(r-1)}$ are estimated by exactly the same quantity it is safe to estimate $\|H_{sr}^{(r)}\|$ by replacing Z_r with $H_r^{(r-1)}$ in the second of (5). This is tantamount to extending the sum in the second of (5) to $p = s-1$ and making it identical with the sum in the first of (5), with $m = 0$. Thus, assuming that $H^{(r-1)}$ is known, by the general estimate (14) (with $\delta = \delta_r,\ \delta' = \delta_{r-1}$ and $d = d_r$) one has

$$\begin{aligned}\|H_{sr+m}^{(r)}\|_{1-\delta_r} \leq \frac{1}{e^2}\left(\frac{e}{d_r\Lambda}\right)^{2s} \|\chi_r\|_{1-\delta_{r-1}}^s \|Z_m\|_{1-\delta_{m-1}} \\ +\frac{1}{e^2}\sum_{p=0}^{s-1}\left(\frac{e}{d_r\Lambda}\right)^{2p} \|\chi_r\|_{1-\delta_{r-1}}^p \|H_{(s-p)r+m}^{(r-1)}\|_{1-\delta_{r-1}} \\ \text{for } r \geq 2,\ s \geq 1,\ 1 \leq m < r\end{aligned}$$

$$\|H_{sr}^{(r)}\|_{1-\delta_r} \leq \frac{1}{e^2}\sum_{p=0}^{s-1}\left(\frac{e}{d_r\Lambda}\right)^{2p} \|\chi_r\|_{1-\delta_{r-1}}^p \|H_{(s-p)r}^{(r-1)}\|_{1-\delta_{r-1}} \qquad \text{for } r \geq 1,\ s \geq 2$$

Let us first consider the estimate of $\|H_{sr+m}^{(r)}\|_{1-\delta_r}$. By the inductive hypothesis and by (17) and (18) one has

$$\begin{aligned}\|H_{sr+m}^{(r)}\|_{1-\delta_r} &= \frac{1}{e^2}\left(\frac{e}{d_r\Lambda}\right)^{2s}\left(\frac{\eta_{r-1,r}E}{\alpha_r}\right)^s \eta_{m-1,m}E \\ &\quad + \frac{1}{e^2}\sum_{p=0}^{s-1}\left(\frac{e}{d_r\Lambda}\right)^{2p}\left(\frac{\eta_{r-1,r}E}{\alpha_r}\right)^p \eta_{r-1,(s-p)r+m}E \\ &\leq C^s\left(\frac{1}{\alpha_r d_r^2}\right)^s \eta_{r-1,r}^s\eta_{m-1,m}E + \sum_{p=0}^{s-1} C^p\left(\frac{1}{\alpha_r d_r^2}\right)^p \eta_{r-1,r}^p\eta_{r-1,(s-p)r+m}E\end{aligned}$$

Replacing $\eta_{r,s}$ in the last estimate one gets

$$\|H_{sr+m}^{(r)}\|_{1-\delta_r} \leq C^{sr+m-1}\frac{1}{\alpha_1 d_1^2\ldots\alpha_r d_r^2}\cdot\left[\left(\frac{1}{\alpha_1 d_1^2\ldots\alpha_r d_r^2}\right)^{s-1}\left(\frac{1}{\alpha_1 d_1^2\ldots\alpha_{m-1} d_{m-1}^2}\right)\mu_{r-1,r}^s\mu_{m-1,m}\right.$$

$$+\sum_{p=0}^{s-1}\alpha_r d_r^2\left(\frac{1}{\alpha_1 d_1^2\ldots\alpha_r d_r^2}\right)^p\left(\frac{1}{\alpha_{r-1}d_{r-1}^2}\right)^{(s-p-1)r+m}\mu_{r-1,r}^p\mu_{r-1,(s-p)r+m}\Bigg]E$$

Recalling that $\{\alpha_r\}_{r\geq 1}$ and $\{d_r\}_{r\geq 1}$ are non increasing sequences, we obtain

$$\|H_{sr+m}^{(r)}\|_{1-\delta_r}\leq C^{sr+m-1}\frac{1}{\alpha_1 d_1^2\ldots\alpha_r d_r^2}\left(\frac{1}{\alpha_r d_r^2}\right)^{(s-1)r+m-1}\mu_{r,sr+m}E=\eta_{r,sr+m}E$$

Similarly one gets

$$\|H_{sr}^{(r)}\|_{1-\delta_r}\leq\frac{1}{e^2}\sum_{p=0}^{s-1}\left(\frac{e}{d_r\Lambda}\right)^{2p}\left(\frac{\eta_{r-1,r}E}{\alpha_r}\right)^p\eta_{r-1,(s-p)r}E\leq\eta_{r,sr}E$$

so that (19) follows. This concludes the proof of lemma 1. *Q.E.D.*

The estimates of lemma 1 hold true for any r, and depend on the arbitrary choice of $d_1,\ldots,d_r$ and on the quantities $\alpha_1,\ldots,\alpha_r$, which in turn depend on the frequencies $\omega_1,\ldots,\omega_n$. However, it is immediate to realize that, e.g., $\|\chi_r\|$ grow at least as $r!$. For, by the diophantine theory we know that the best estimate for the divisors α_r is

$$\alpha_r\geq\gamma(r+2)^{-\tau} \tag{20}$$

with $\gamma>0$ and $\tau>1$; and the product of divisors in the definition of $\eta_{r,r-1}$ generates the factorial. This prevents the possibility of proving the convergence of the whole procedure for $r\to\infty$, with our estimates.

Thus, we simply stop the construction of the normal form at some finite r, that we keep fixed although still arbitrary. Moreover we replace the sequence α_r with the diophantine estimate (20), i.e. we set $\alpha_r=\gamma(r+2)^{-\tau}$ with constants $\gamma>0$, $\tau\geq m-1$, (see the remark after the Theorem). Finally we make the choice $d_1=\ldots=d_r=d/r$ with $0<d<1$. With this setting we have

Lemma 2: *For a fixed r the normal form is estimated by*

$$\|Z_s\|_{1-d}\leq C_1^{s-1}(s+1)!^\tau r^{2(s-1)}E \qquad \text{for } s\leq r \tag{21}$$

$$\|H_s^{(r)}\|_{1-d}\leq C_2C_1^{s-1}(r+2)!^\tau(r+2)^{(\tau+2)s-\tau(r+1)-1}E \qquad \text{for } s>r \tag{22}$$

where C_1,C_2 are positive constants and $d<1$ is arbitrary.

The proof is just matter of replacing the diophantine estimate for α_r and the choice for d_r in the definition (16) of $\eta_{r,s}$, and performing some elementary manipulations. The sequence $\{\mu_{r,s}\}_{s\geq r\geq 0}$ is estimated by

$$\mu_{r-1,r}<\beta^{r-1},\qquad \mu_{r,s}<r(\beta+1)^s \qquad \text{for } s>r \tag{23}$$

with a positive constant β. The latter estimates are proven in Appendix B.

4. Use of the normal form and proof of the Theorem

So far we may write the Hamiltonian in normal form up to order r as

$$H^{(r)}=H_0^\sharp+H_0^\natural+Z^\sharp+Z^\natural+Z^\flat+\mathcal{R}^\sharp+\mathcal{R}^\natural+\mathcal{R}^\flat \tag{24}$$

where

$$
\begin{aligned}
H_0^\sharp &= \sum_{j=1}^{m} \frac{\omega_j}{2}(x_j^2+y_j^2), \qquad H_0^\natural = \sum_{\nu=m+1}^{n} \frac{\omega_\nu}{2}(x_\nu^2+y_\nu^2), \\
Z^\sharp &= Z_1^\sharp + \ldots + Z_r^\sharp, \qquad \mathcal{R}^\sharp = \sum_{s>r} (H_s^{(r)})^\sharp,
\end{aligned} \tag{25}
$$

and similar expressions for $Z^\natural, Z^\flat, \mathcal{R}^\natural$ and $\mathcal{R}^\flat$, where the superscript indicates the projection of the function on the subspaces $\mathcal{P}^\sharp, \mathcal{P}^\natural$ and $\mathcal{P}^\flat$ defined by (9). Consider the harmonic actions of the oscillators, namely

$$
\begin{aligned}
I_j &= \frac{x_j^2+y_j^2}{2}, \qquad j=1,\ldots,m, \\
I_\nu &= \frac{x_\nu^2+y_\nu^2}{2}, \qquad \nu=m+1,\ldots,n;
\end{aligned}
$$

we use different subscripts j and ν in order to identify the first m actions and the remaining ones, respectively. The crucial point is that we have

$$
\begin{aligned}
\dot I_j &= \{I_j, Z^\sharp\} + \{I_j, Z^\natural\} + \{I_j, Z^\flat\} + \{I_j, \mathcal{R}\} \\
\dot I_\nu &= \{I_\nu, Z^\flat\} + \{I_\nu, \mathcal{R}^\natural + \mathcal{R}^\flat\},
\end{aligned}
$$

where $\mathcal{R} = \mathcal{R}^\sharp + \mathcal{R}^\natural + \mathcal{R}^\flat$. For, the Poisson brackets between any of the actions $I_1,\ldots,I_n$ and H_0 vanish because H_0 is a function of $I_1,\ldots,I_n$ only. Moreover, $\{I_\nu, Z^\sharp\}$ and $\{I_\nu, \mathcal{R}^\sharp\}$ do vanish because $Z^\sharp$ and $\mathcal{R}^\sharp$ depend on x_j, y_j with $1 \le j < m$, while I_ν depends on x_ν, y_ν with $\nu > m$. Finally, $\{I_\nu, Z^\natural\}$ vanishes in view of the definition of the subspace $\mathcal{Z}^\natural$, to which $Z^\natural$ belongs.

Let us now, for a moment, forget the remainder $\mathcal{R}$ which is not in normal form. Then we have to consider only the equations

$$
\begin{aligned}
\dot I_j &= \{I_j, Z^\sharp\} + \{I_j, Z^\natural\} + \{I_j, Z^\flat\}, \\
\dot I_\nu &= \{I_\nu, Z^\flat\} \in \mathcal{P}^\flat.
\end{aligned}
$$

If we choose initial conditions such that $I_{m+1}(0) = \ldots = I_n(0) = 0$, which implies $x_\nu = y_\nu = 0$ for $\nu = m+1,\ldots,n$, and with $I_1(0),\ldots,I_m(0)$ arbitrary, then we have $I_{m+1}(t) = \ldots = I_n(t) = 0$ for all t. This because the r.h.s. of the equation for I_ν contains at least one factor $x_{m+1}, y_{m+1},\ldots,x_n,y_n$, which is zero. Thus, the $2m-$dimensional subspace $I_{m+1}(0) = \ldots = I_n(0) = 0$ is invariant. Taking into account the near the identity transformation of coordinates to the normal form, it turns out immediately that this subspace is a local analytic manifold for the original system. Concerning instead the dynamics on the manifold, it is immediate to see that nothing can be said a priori. For, the dynamics is clearly generated by the reduced Hamiltonian $\bar H(x_1,\ldots,x_m,y_1,\ldots,y_m) = H_0^\sharp + Z^\sharp$, with no restrictions on the nonlinear term $Z^\sharp$.

The manifold $x_\nu = y_\nu = 0$, $\nu = m+1,\ldots,n$ is no more invariant if one takes into account the unnormalized remainder $\mathcal{R}$. In this case we proceed in the spirit of Nekhoroshev's theory, looking for a long time quasi-invariance in the neighbourhood of that manifold.

From now on we fix the free parameters R and d by setting $R_1 = \ldots = R_N = 1$, and $d = 1/4$. Then we proceed in two steps. First we prove that for small energy there is a neighbourhood of the origin that remains invariant. Next we shall deduce exponential stability estimates for the quasi-invariant manifold.

The first step is a standard matter, due to the quadratic form of H_0, and in view of the hypothesis that all the frequencies $\omega_1, \ldots, \omega_n$ are positive. For, the Hamiltonian has a minimum of zero energy at the origin, so that for sufficiently small energy E there are two polydisks Δ^n_{ϱ} and $\Delta^n_{\overline{\varrho}}$ with $\overline{\varrho} < \varrho = O(E^{1/2})$, such that $\Delta^n_{\overline{\varrho}} \subset \mathcal{D}_E \subset \Delta^n_{\varrho}$, where $\mathcal{D}_E$ is a connected component of the invariant subset of phase space defined by $H^{(r)}(x, y) \le E$.

Thus, we proceed with the second step using the estimates (21) and (22). Let us consider a domain $\Delta^m_{\varrho} \times \Delta^{n-m}_{\sigma}$ which is the product of polydisks

$$\begin{aligned} \Delta^m_{\varrho} &= \{x_j^2 + y_j^2 \le \varrho^2,\ j = 1, \ldots, m\} \\ \Delta^{n-m}_{\sigma} &= \{x_\nu^2 + y_\nu^2 \le \sigma^2,\ \nu = m+1, \ldots, n\}, \end{aligned}$$

where ϱ and σ are positive, and $\sigma \ll \varrho$; a choice for σ will be made below. Our aim is to estimate $\dot{I}_\nu(x, y)$ in $\Delta^m_{\varrho} \times \Delta^{n-m}_{\sigma}$. Recalling that

$$\dot{I}_\nu = \{I_\nu, Z^\flat\} + \{I_\nu, \mathcal{R}^\natural + \mathcal{R}^\flat\}$$

and using the more explicit expression (25) for $Z^\flat, \mathcal{R}^\natural$ and $\mathcal{R}^\flat$ we have

$$\begin{aligned} \{I_\nu, Z^\flat\} &= \sum_{s=1}^{r} \{I_\nu, Z^\flat_s\} \\ \{I_\nu, \mathcal{R}^\natural + \mathcal{R}^\flat\} &= \sum_{s>r} \{I_\nu, (H^{(r)}_s)^\natural + (H^{(r)}_s)^\flat\}. \end{aligned}$$

By the general estimate (29) of the Poisson bracket, and by (21) and (22) we have

$$\begin{aligned} \|\{I_\nu, Z^\flat_s\}\|_{1-2d} &\le \frac{C_1^{s-1}}{2d^2}(s+1)!^\tau r^{2(s-1)} E, \\ \|\{I_\nu, (H^{(r)}_s)^\natural + (H^{(r)}_s)^\flat\}\|_{1-2d} &\le \frac{C_2 C_1^{s-1}}{2d^2}(r+2)!^\tau (r+2)^{(\tau+2)s - \tau(r+1) - 1} E. \end{aligned}$$

The inequalities $\|Z^\flat_s\|_{1-d} \le \|Z_s\|_{1-d}$ and $\|(H^{(r)}_s)^\natural + (H^{(r)}_s)^\flat\|_{1-d} \le \|H^{(r)}_s\|_{1-d}$ have been used here. Remark now that $\{I_\nu, Z^\flat_s\} \in \mathcal{P}^\flat$ is a homogeneous polynomial of degree $s+2$ which contains at least three factors x_ν, y_ν with $m+1 \le \nu \le n$; similarly, $\{I_\nu, (H^{(r)}_s)^\natural + (H^{(r)}_s)^\flat\}$ contains at least one such factor. Thus for all $(x, y) \in \Delta^m_{\varrho} \times \Delta^{n-m}_{\sigma}$, we have (recall that $d = 1/4$)

$$\begin{aligned} \bigl|\{I_\nu, Z^\flat_s\}(x, y)\bigr| &\le \bigl\|\{I_\nu, Z^\flat_s\}(x, y)\bigr\|_{\frac{1}{2}} 2^{s+2} \sigma^3 \varrho^{s-1} \quad \text{for } 1 \le s \le r\ , \\ \bigl|\{I_\nu, (H^{(r)}_s)^\natural + (H^{(r)}_s)^\flat\}(x, y)\bigr| &\le \bigl\|\{I_\nu, (H^{(r)}_s)^\natural + (H^{(r)}_s)^\flat\}\bigr\|_{\frac{1}{2}} 2^{s+2} \sigma \varrho^{s+1}, \quad \text{for } s > r\ . \end{aligned}$$

Adding up all contributions we get

$$
\begin{aligned}
|\dot{I}_\nu(x,y)| \le 64\cdot 12^\tau E\sigma^3 &\sum_{s=1}^{r}[2(s+1)^\tau C_1 r^2\varrho]^{s-1}\\
&+32C_2E\sigma\frac{(r+2)!^\tau}{(r+2)^{\tau(r+1)+1}}\sum_{s\ge 0}[2C_1(r+2)^{\tau+2}\varrho]^s C_1^r(r+2)^{(\tau+2)(r+1)}\varrho^{r+2}
\end{aligned}
$$

Assuming

$$
\varrho < \frac{1}{4C_1(r+2)^{\tau+2}} \tag{26}
$$

and with a trivial use of the sum of the geometric series we readily conclude

$$
|\dot{I}_\nu(x,y)| \le 128\cdot 12^\tau E\sigma^3 + 64C_2E(C_1e^2)^r(r+2)!^{\tau+2}\sigma\varrho^{r+2} \qquad \text{for } \nu = m+1,\ldots,n.
$$

We now follow the usual procedure for Nekhoroshev's estimates. Having fixed ϱ we make the optimal choice for r

$$
r+2 = \left(\frac{\varrho^*}{\varrho}\right)^a, \qquad \varrho^* = \frac{1}{C_1e^2}, \qquad a = \frac{1}{\tau+2}, \tag{27}
$$

which clearly satisfies (26). With a trivial calculation using Stirling's formula we get

$$
(C_1e^2\varrho)^{r+2}(r+2)!^{\tau+2} \simeq \exp\left(-\frac{1}{a}\left(\frac{\varrho^*}{\varrho}\right)^a\right).
$$

Remark that (27) is sense only if $r \ge 1$, which imposes the smallness condition on ϱ

$$
\varrho < \frac{\varrho^*}{3^{\tau+2}}.
$$

This allows us to remove the arbitrary normalization order from our estimates, thus getting

$$
|\dot{I}_\nu(x,y)| < \frac{D}{2}\left[\sigma^3 + \sigma\exp\left(-\frac{1}{a}\left(\frac{\varrho^*}{\varrho}\right)^a\right)\right]
$$

where D is a positive constant independent of ϱ and σ. Finally, we impose the initial point to be very close to the subspace $I_{m+1} = \ldots = I_n = 0$ by setting

$$
\sigma^2 = \exp\left(-\frac{1}{a}\left(\frac{\varrho^*}{\varrho}\right)^a\right).
$$

This gives the final estimate

$$
|\dot{I}_\nu(x,y)| < D\sigma^3. \tag{28}
$$

Now, assume that $(x(0),y(0)) \in \Delta_\varrho^m \times \Delta_{\frac{\sigma}{2}}^{n-m}$ and consider the inequality

$$
|I_\nu(t) - I_\nu(0)| \le |t|D\sigma^3
$$

which holds true for $|t| < \tau$, where τ is defined as

$$\tau = \sup\left\{ t > 0 : |I_\nu(s)| < \frac{\sigma^2}{2} \ \forall |s| < t \right\}$$

In order that $(x(t), y(t)) \in \Delta_\varrho^m \times \Delta_\sigma^{n-m}$, we must have

$$|I_\nu(t) - I_\nu(0)| < \frac{\sigma^2}{4}$$

By (28) this is certainly satisfied if

$$|t| < \frac{1}{4D\sigma} = \frac{1}{4D} \exp\left(\frac{1}{2a} \left(\frac{\varrho^*}{\varrho} \right)^a \right)$$

Setting $T_* = 1/4D$ the proof of the Theorem is concluded.

A. Deduction of equation (5)

A justification for equation (5) can be found as follows. Let us rearrange the terms in the expansion of $\exp(L_{\chi_r})H^{(r-1)}$. Considering first H_0 and $H_r^{(r-1)}$ together, one has

$$\begin{aligned} \exp(L_{\chi_r})\big(H_0 + H_r^{(r-1)}\big) = H_0 &+ L_{\chi_r} H_0 + \sum_{s\geq 2} \frac{1}{s!} L_{\chi_r}^s H_0 \\ &+ H_r^{(r-1)} + \sum_{s\geq 1} \frac{1}{s!} L_{\chi_r}^s H_r^{(r-1)} \ . \end{aligned}$$

where H_0 is the first term in the transformed Hamiltonian $H^{(r)}$ in (3). By equation (4) one can replace the unwanted term $H_r^{(r-1)}$ with the normalized one Z_r. The two sums may be collected and simplified by calculating

$$\begin{aligned} &\sum_{s\geq 2} \frac{1}{s!} L_{\chi_r}^s H_0 + \sum_{s\geq 1} \frac{1}{s!} L_{\chi_r}^s H_r^{(r-1)} \\ &\qquad = \sum_{s\geq 2} \frac{1}{(s-1)!} L_{\chi_r}^{s-1} \left[\frac{1}{s} \left(L_{\chi_r} H_0 + H_r^{(r-1)} \right) + \frac{s-1}{s} H_r^{(r-1)} \right] \\ &\qquad = \sum_{s\geq 2} \frac{1}{(s-1)!} L_{\chi_r}^{s-1} \left(\frac{1}{s} Z_r + \frac{s-1}{s} H_r^{(r-1)} \right) \ . \end{aligned}$$

where both $L_{\chi_r}^{s-1} Z_r$ and $L_{\chi_r}^{s-1} H_r^{(r-1)}$ are homogeneous polynomials of degree $sr + 2$, that are added to $H_{sr}^{(r)}$ in the second of (5).

Let us now calculate the transformation for the functions $Z_1, \ldots, Z_{r-1}$ that are already in normal form. Recalling that no such term exists for $r = 1$, for $r > 1$ one has

$$\exp(L_{\chi_r}) Z_m = Z_m + \sum_{s\geq 1} \frac{1}{s!} L_{\chi_r}^s Z_m \ , \quad \text{for } 1 \leq m < r \ .$$

The term Z_m is copied into $H^{(r)}$ in (3), while the term $L^s_{\chi_r} Z_m$ is a homogeneous polynomial of degree $sr+m+2$ that is added to $H^{(r)}_{sr+m}$ in the first of (5).

Finally, consider all terms $H^{(r-1)}_s$ with $s > r$, that may be written as $H^{(r-1)}_{lr+m}$ with $l \geq 1$ and $0 \leq m < r$, the case $l = 1$, $m = 0$ being excluded. One gets

$$\exp(L_{\chi_r}) H^{(r-1)}_{lr+m} = \sum_{p\geq 0} \frac{1}{p!} L^p_{\chi_r} H^{(r-1)}_{lr+m}$$

where $L^p_{\chi_r} H^{(r-1)}_{lr+m}$ is a homogeneous polynomial of degree $(p+l)r+m+2$. Collecting all homogeneous terms with $m = 0$, $l \geq 2$ and $p+l = s \geq 2$ one gets $\sum_{p=0}^{s-2} \frac{1}{p!} L^p_{\chi_r} H^{(r-1)}_{(s-p)r}$, that is added to $H^{(r)}_{sr}$ in the second of (5). Similarly, collecting all homogeneous terms with $0 < m < r$, $l \geq 1$ and $p+l = s \geq 1$ one gets $\sum_{p=0}^{s-1} \frac{1}{p!} L^p_{\chi_r} H^{(r-1)}_{(s-p)r+m}$, that is added to $H^{(r)}_{sr+m}$ in the first of (5). The latter case does not occur for $r = 1$. This completes the justification of equation (5).

B. Technical estimates

This appendix is devoted to the proof of the estimates (14) and (23). Let $0 \leq \max(\delta', \delta'') < \delta < 1$. We first prove a general estimate for the Lie derivative of a generic function f in a domain $\Delta_{(1-\delta)R}$, namely

$$\left\|L_\chi f\right\|_{1-\delta} \leq \frac{1}{(\delta-\delta')(\delta-\delta'')\Lambda^2} \left\|\chi\right\|_{1-\delta'} \left\|f\right\|_{1-\delta''} \tag{29}$$

with Λ as in (12). A proof of this estimate can be obtained as follows. Write, generically, $\chi = \sum_{j,k} c_{j,k} x^j y^k$ and $f = \sum_{j,k} f_{j,k} x^j y^k$. Then compute

$$L_\chi f = \sum_{j,k,j',k'} \sum_{l=1}^{n} \frac{j'_l k_l - j_l k'_l}{x_l y_l} c_{j,k} f_{j',k'} x^{j+j'} y^{k+k'} \ . \tag{30}$$

Using the definition of norm evaluate

$$\begin{aligned}
\|L_\chi f\|_{1-\delta} &\leq \sum_{j,k,j',k'} \sum_{l=1}^{n} \frac{|j'_l k_l - j_l k'_l|}{R_l^2} |c_{j,k}|\, |f_{j',k'}| (1-\delta)^{|j+k|+|j'+k'|-2} R^{j+k+j'+k'} \\
&\leq \frac{1}{\Lambda^2} \sum_{j,k} \sum_{j',k'} \sum_{l=1}^{n} |j'_l k_l - j_l k'_l|\ |c_{j,k}| \big((1-\delta') - (\delta-\delta')\big)^{|j+k|-1} R^{j+k} \\
&\qquad \times |f_{j',k'}| \big((1-\delta'') - (\delta-\delta'')\big)^{|j'+k'|-1} R^{j'+k'} \ .
\end{aligned}$$

If f is a generic function, then in view of $j_l \leq |j+k|$ and $k_l \leq |j+k|$ one has

$$\sum_{l=1}^{n} |j'_l k_l - j_l k'_l| < |j+k| \sum_{l=1}^{n} |j'_l + k'_l| = |j+k| \cdot |j'+k'| \ . \tag{31}$$

Replacing in the estimate above and using the elementary inequality[†]

$$m(\lambda-x)^{m-1}<\frac{\lambda^m}{x}\quad \text{for } 0<x<\lambda \text{ and } m\ge 1 \tag{32}$$

one gets

$$\begin{aligned}\|L_\chi f\|_{1-\delta} &\le \frac{1}{\Lambda^2}\sum_{j,k}|c_{j,k}|\,|j+k|\bigl((1-\delta')-(\delta-\delta')\bigr)^{|j+k|-1}R^{j+k}\\ &\qquad\times\sum_{j',k'}|f_{j',k'}|\,|j'+k'|\bigl((1-\delta'')-(\delta-\delta'')\bigr)^{|j'+k'|-1}R^{j'+k'}\\ &\le \frac{1}{(\delta-\delta')(\delta-\delta'')\Lambda^2}\sum_{j,k}|c_{j,k}|\,(1-\delta')^{|j+k|}R^{j+k}\\ &\qquad\times\sum_{j',k'}|f_{j',k'}|(1-\delta'')^{|j'+k'|}R^{j'+k'}\ ,\end{aligned}\tag{33}$$

from which (29) immediately follows in view of the definition of the norm.

Let now $p>1$, and assume that $\|\chi\|_{1-\delta'}$ and $\|f\|_{1-\delta'}$ are known for some non negative δ'. We look for an estimate of $\|L_\chi^p f\|_{1-\delta}$ with $\delta'<\delta<1$. Let $d=\delta-\delta'$ and $\tilde d=d/p$. Then by a straightforward application of (29) we have the recursive estimate

$$\|L_\chi^s f\|_{1-\delta'-s\tilde d}\le\frac{1}{s\tilde d^2\Lambda^2}\|\chi\|_{1-\delta'}\|\|_{1-\delta'-(s-1)\tilde d}.$$

By recursive application of the latter formula we get

$$\|L_\chi^p f\|_{1-\delta}\le\frac{p^{2p}}{p!}\frac{1}{(d\Lambda)^{2p}}\|\chi\|_{1-\delta'}^p\|f\|_{1-\delta'},$$

and (14) follows in view of the trivial inequality $p^p\le e^{p-1}p!$.

We come now to estimating the sequences $\mu_{r,r}$ and $\mu_{s,r}$ with $s\ge r\ge 2$. Start by proving the following properties:

$$\mu_{0,r}<\ldots<\mu_{r-1,r}=\mu_{r,r}=\mu_{r+1,r}=\ldots\quad\text{for } r\ge 1\ , \tag{34}$$

$$\mu_{1,s}=s\quad\text{for } s\ge 0\ , \tag{35}$$

$$\mu_{r,s}=\mu_{r-1,s}+\mu_{r-1,r}\mu_{r,s-r}\le\mu_{r-1,s}+\mu_{r,r}\mu_{s-r,s-r}\quad\text{for } s\ge r\ge 2\ , \tag{36}$$

$$\mu_{r,r}\le r+\sum_{j=2}^{r-1}\mu_{j,j}\mu_{r-j,r-j}\quad\text{for } r\ge 3\ . \tag{37}$$

Properties (34) and (35) follow from the definition. To check (36) let us write $s=kr+m$

[†] The function $x(\lambda-x)^{m-1}$ in the interval $0\le x\le\lambda$ has a maximum for $x=\lambda/m$, and the inequality follows from $x(\lambda-x)^{m-1}\le\frac{\lambda}{m}\left(\lambda-\frac{\lambda}{m}\right)^{m-1}<\frac{\lambda^m}{m}$.

with $0 \le m < r$ and calculate

$$\begin{aligned} \mu_{r,kr+m} &= \mu_{r-1,kr+m} + \sum_{p=1}^{k} \mu_{r-1,r}^{p} \mu_{r-1,(k-p)r+m} \\ &= \mu_{r-1,kr+m} + \mu_{r-1,r} \underbrace{\sum_{p=0}^{k-1} \mu_{r-1,r}^{p} \mu_{r-1,(k-1-p)r+m}}_{=\mu_{r,(k-1)r+m}} \ ; \end{aligned}$$

this gives the equality, while the inequality follows from (34). A proof of (37) is obtained by a repeated application of (36), followed by use of (35). For, calculate

$$\begin{aligned} \mu_{r,r} &= \mu_{r-1,r} \\ &\le \mu_{r-2,r} + \mu_{r-1,r-1}\mu_{1,1} \\ &\le \mu_{r-3,r} + \mu_{r-2,r-2}\mu_{2,2} + \mu_{r-1,r-1}\mu_{1,1} \\ &\dots \\ &\le \mu_{1,r} + \mu_{2,2}\mu_{r-2,r-2} + \dots \mu_{r-1,r-1}\mu_{1,1} \ ; \end{aligned}$$

this is nothing but (37) since $\mu_{1,r} = r$. Thus, the sequence $\{\nu_r\}_{r\ge1}$ defined as

$$\begin{aligned} &\nu_1 = 1 \ , \quad \nu_2 = 2 \ , \\ &\nu_r = r + \sum_{j=2}^{r-1} \nu_j \nu_{r-j} \quad \text{for } r \ge 3 \end{aligned} \tag{38}$$

is a majorant of $\{\mu_{r,r}\}_{r\ge1}$. An easy remark is that $\{\nu_r\}_{r\ge1}$ resembles the well known sequence $\{\lambda_r\}_{r\ge1}$ defined as

$$\begin{aligned} &\lambda_1 = 1 \ , \\ &\lambda_r = \sum_{j=1}^{r-1} \lambda_j \lambda_{r-j} \quad \text{for } r \ge 2 \ . \end{aligned} \tag{39}$$

Indeed it is an easy matter to see that $\nu_r \le a^{r-1}\lambda_r$ for some positive constant a (e.g., $a = 2$). Hence it is enough to study the latter sequence. It is known that

$$\lambda_r = \frac{2^{r-1}(2r-3)!!}{r!} \le 4^{r-1} \ , \tag{40}$$

where the standard notation $(2n+1)!! = 1 \cdot 3 \cdot \ldots \cdot (2n+1)$ has been used. As a check, let the function $g(z)$ be defined as $g(z) = \sum_{r\ge1} \lambda_r z^r$, so that $\lambda_r = g^{(r)}(0)/r!$. Then it is immediate to check that the recursive definition (39) is equivalent to the equation $g = z + g^2$. By repeated differentiation of the latter equation one readily finds

$$g' = \frac{1}{1-2g} \ , \ \dots \ , \ g^{(r)} = \frac{2^{r-1}(2r-3)!!}{(1-2g)^{2r-1}}$$

(check by induction). From this, (40) follows. Thus, $\mu_{r,r} \le \beta^{r-1}$ for some positive β.

By equations (34), (35), (36) and the estimate for $\mu_{r,r}$ one also has

$$\mu_{r,s} \le \mu_{r-1,s} + \beta^{s-2} \le \ldots \le s + r\beta^{s-2} \le r(\beta+1)^s$$

with $\beta \ge 1$ and $s \ge r \ge 2$

References

[1] Berchialla, L., Galgani L. and Giorgilli, A.: *Localization of energy in FPU chains*, preprint (2002).

[2] Bocchieri, P., Scotti, A., Bearzi, B. and Loinger, A.: *Anharmonic chain with Lennard–Jones interaction*, Phys. Rew. A **2**, 2013–2019 (1970).

[3] Carotta, M.C., Ferrario, C., Lo Vecchio, G., Carazza, B. and Galgani, L.: *New phenomenon in the stochastic transition of coupled oscillators*, Phys. Rev A **17**, 786 (1978).

[4] Casetti, L., Cerruti–Sola, M., Pettini, M. and Cohen, E.G.D.: *The Fermi–Pasta–Ulam problem revisited: Stochasticity thresholds in nonlinear Hamiltonian systems*, Phys. Rev. E **55**, 6566–6574 (1997).

[5] Cherry, T. M.: *On the solutions of Hamiltonian systems in the neighbourhood of a singular point*, Proc. London Math. Soc., Ser. 2, **27**, 151–170 (1926).

[6] Izrailev, F.M. and Chirikov, B.V.: *Stochasticity of the simplest dynamical model with divided phase space* Dokl. Akad. Nauk. SSSR **166**, 57 (1966); Sov. Phys. Dokl. **11**, 30 (1966).

[7] De Luca, J., Lichtenberg, A.J. and Ruffo, S.: *Finite time to equipartition in the thermodynamic limit*, Phys. Rev E **60**, 3781–3786 (1999).

[8] Fermi, E., Pasta, J. and Ulam, S.: *Studies of nonlinear problems*, Los Alamos document LA–1940 (1955).

[9] Galgani, L., Giorgilli, A., Martinoli, A. and Vanzini, S.: *On the problem of energy equipartition for large systems of the Fermi–Pasta–Ulam type: analytical and numerical estimates*, Physica D **59**, 334–348 (1992).

[10] Giorgilli, A.: *Unstable equilibria of Hamiltonian systems*, Disc. and Cont. Dynamical Systems, Vol. 7, N. 4, 855–871 (2001).

[11] Kantz, H., Livi, R. and Ruffo, S.: *Equipartition thresholds in chains of anharmonic oscillators*, J. Stat. Phys. **76**, 627–643 (1994).

[12] Livi, R., Pettini, M., Ruffo, S., Sparpaglione, M. and Vulpiani, A.: *Relaxation to different stationary states in the Fermi–Pasta–Ulam model*, Phys. Rev A **28**, 3544–3552 (1983).

[13] Livi, R., Pettini, M., Ruffo, S., Sparpaglione, M. and Vulpiani, A.: *Equipartition threshold in nonlinear large Hamiltonian systems: the Fermi–Pasta–Ulam model*, Phys. Rev. A **31**, 1039–1045 (1985).

[14] Moser, J.: *On the generalization of a theorem of A. Liapounoff*, Comm. Pure Appl. Math. **11**, 257–271 (1958).

[15] Pettini, M. and Landolfi, M.: *Relaxation properties and ergodicity breaking in nonlinear Hamiltonian dynamics*, Phys. Rev A **41**, 768–783 (1990).