# Elliptic tori in FPU non-linear chains with a small number of nodes* 

CHIARA CARACCIOLO<br>Dipartimento di Matematica, Università degli Studi di Roma "Tor Vergata", via della Ricerca Scientifica 1, 00133 - Roma (Italy).<br>UGO LOCATELLI<br>Dipartimento di Matematica, Università degli Studi di Roma "Tor Vergata", via della Ricerca Scientifica 1, 00133 - Roma (Italy). e-mails: caraccio@mat.uniroma2.it, locatell@mat.uniroma2.it,

12th February 2021


#### Abstract

We revisit an algorithm constructing elliptic tori, that was originally designed for applications to planetary hamiltonian systems. The scheme is adapted to properly work with models of chains of $N+1$ particles interacting via anharmonic potentials, thus covering also the case of FPU chains. After having preliminarily settled the Hamiltonian in a suitable way, we perform a sequence of canonical transformations removing the undesired perturbative terms by an iterative procedure. This is done by using the Lie series approach, that is explicitly implemented in a programming code with the help of a software package, which is especially designed for computer algebra manipulations. In the cases of FPU chains with $N=4,8$, we successfully apply our new algorithm to the construction of elliptic tori for wide sets of the parameter ruling the size of the perturbation, i.e., the total energy of the system. Moreover, we explore the stability regions surrounding 1D elliptic tori. We compare our semi-analytical results with those provided by numerical explorations of the FPU-model dynamics, where the latter ones are obtained by using techniques based on the so called frequency analysis. We find that our procedure works up to values of the total energy that are of the same order of magnitude with respect to the maximal ones, for which elliptic tori are detected by numerical methods.


[^0]C. Caracciolo, U. Locatelli

## 1 Introduction

In the 50 's, Fermi, Pasta and Ulam designed one of the very first numerical experiments in the history of science (see [15]), realized thanks to the remarkable implementation of M. Tsingou on the computer Maniac I in the Los Alamos laboratories. It is well known that the authors were surprised by the results. In the so called FPU non-linear chains, when just a few slow frequency normal modes are initially excited with relatively low amplitudes of oscillation, the system does not seem to tend toward the equipartition of the energy among all the normal modes themselves. Such a behavior was observed to persist also for very long times of numerical integration and was unexpected, because it is in contrast with the predictions of the Statistical Mechanics. This phenomenon has been explained in terms of metastability (see, e.g., the reviews [1] and [2]). Moreover, the thermalization was observed for larger values of the initial amplitudes of oscillation. The complete understanding of the eventual transition to ergodicity is still an open problem, in particular for what concerns the "thermodinamic limit", i.e., by keeping fixed the value of the specific energy $E_{S}=E / N$ with $N \rightarrow \infty, E$ being the total energy of the system and $N+1$ the number of nodes making part of the non-linear FPU chain (see, e.g., those sections in [3] that are devoted to review part of the wide scientific literature on this topic).

The present work is in the spirit of the approach described in [10] and [11, where the persistence of (meta)stable configurations has been interpreted in terms of the so called $q$-tori. They are invariant tori of lower dimension, i.e., smaller than the maximal one that is equal to the number of degrees of freedom $N-1$. In the limit of the total energy going to zero, the motion on $q$-tori is given by the composition of small oscillations for a reduced number of normal modes. For relatively low values of the initial oscillations amplitudes, the sharing of the energy with the remaining modes is very limited. For what concerns such a phenomenon of (so called) localization, the behavior of the motions on $q$-tori is very similar to those considered in the pioneering numerical simulations of the FPU chains. Therefore, a special attention is devoted to the linear stability of $q$-tori, because the concept of similarity between those orbits would be obviously reinforced, if the initial conditions with just a few modes excited would be in a sort of stability region around the $q$-tori. In [10] and [11, $q$-tori are determined by using the method of Poincaré-Lindstedt series, that is able to compute the shape of invariant manifolds and the motion on them; however, in order to have information about the local dynamics in the neighborhood of those tori, some additional work is required (see., e.g., [8]). Let us also recall that $q$-tori are a generalization of $q$-breathers, that are periodic orbits which can be obtained as a continuation of the small oscillations involving one single normal mode. Their role in the dynamics of FPU chains has been widely investigated starting by a sequence of articles written by S. Flach and some of his coworkers (see [44, [17] and therein references).

In KAM theory, linearly stable $q$-tori are commonly known as elliptic tori, that are compact invariant manifold of dimension smaller than the number of degrees of freedom. Moreover, they are such that the transverse motions are given by a superposition of small oscillations (as in the case of the elliptic equilibrium points). Our main goal is to
introduce a suitable normal form approach explicitly constructing elliptic tori that are invariant with respect to the flow induced by the Hamiltonian of the FPU model. In the present work, we refer to a theoretical background that is going back to Melnikov and the first complete proofs of existence of elliptic tori (see [38], [14] and [45]). In particular, we adapt to the present context the algorithm designed to construct specific elliptic tori for planetary Hamiltonian systems, that are peculiar because the transverse dynamics is slow with respect to the one on the torus, which is related to the orbital motions of revolution around the central star (see [49] and [26]).

In the scientific literature, a special care is devoted to the study of the motions starting from initial conditions of the same type as those considered in the original numerical experiments on FPU chains. In this paper, we focus on a case that is widely debated, although it is even more specific: all the energy is initially given just to the first mode, that is related to the largest wavelength. If the nonlinearity were neglected in the Hamilton equations, the oscillations of the chain (in a plane where the equilibrium positions and the corresponding displacements with respect to equilibrium of all the particles are plotted) would have the profile similar to an arc, that is the discretization of a half of a graph drawing the sine function. Therefore, hereafter we will refer to these initial conditions by calling them of "semi-sinusoidal" type. We aim to establish a relation between the eventual quasi-periodicity of some motions starting from this kind of initial conditions and the stability region surrounding elliptic tori. For such a purpose, we find that 1D invariant manifold are extremely effective. With respect to the terminology previously introduced, they can be seen as linearly stable $q$-breathers.

Let us emphasize that our approach, being based on a normal form method, is in a very good position for this type of investigation, because it directly gives the equations of motion ruling the dynamics in the neighborhood of the elliptic tori without any additional work. In the present paper we mainly focus on the comparisons between the results provided by the explicit construction of normal forms with those given by a suitable study of the numerical integrations. The accuracy of the former computational method (that is often said of semi-analytic type) is strongly limited when the number of degrees of freedom is increased. This is the reason why we prefer to consider non-linear chains with small numbers of nodes, in order to obtain sharp results eventually giving a clear validation of our approach. In this respect, we have to admit that the spirit of our work is evidently far from the original motivations making so interesting the study of the FPU model, because they mainly concern with the pathway to the statistical equilibrium and the eventual occurrence of obstacles represented by metastable configurations.

The paper is organised as follows. In section 2, we introduce the basic settings in a Hamiltonian framework that are needed to properly define the FPU model. The algorithm constructing the normal form for elliptic tori is described in section 3. As a further difference with respect to the works on $q$-tori, where the numerical integrations were analyzed by using the Generalized ALignment Indices method (GALI for short, see [50]), here the results provided by the semi-analytic method are compared with those obtained by applying Frequency Analysis (see, e.g., [35]). The adaptation of this latter technique of numerical investigation to the present context is described in section 4. Most of the results testing both the validity of our approach and its performances are collected in section 5 .

Section 6 is devoted to the introduction of suitable Birkhoff normal forms centered around the invariant elliptic tori; this allows to study the dynamics in the vicinity of them. Both the conclusions and the main perspectives are discussed in section 7 .

## 2 Settings for the definition of the Hamiltonian model

The Hamiltonian describing a FPU chain of $N+1$ particles is the following:

$$
\begin{equation*}
\left.\mathcal{H}(\boldsymbol{y}, \boldsymbol{x})=\frac{1}{2} \sum_{j=0}^{N-1}\left[y_{j}^{2}+\left(x_{j+1}-x_{j}\right)^{2}\right)\right]+\frac{\alpha}{3} \sum_{j=0}^{N-1}\left(x_{j+1}-x_{j}\right)^{3}+\frac{\beta}{4} \sum_{j=0}^{N-1}\left(x_{j+1}-x_{j}\right)^{4} \tag{1}
\end{equation*}
$$

with $\alpha, \beta \in \mathbb{R}$, where $x_{k}$ is the displacement of the $k$-th particle with respect to the equilibrium position and $y_{k}$ its canonically conjugate momentum. Usually, $N=2^{p}$ and boundary conditions are fixed so that $x_{0}=x_{N}=0, y_{0}=y_{N}=0$. The case $\alpha \neq 0, \beta=0$ is called $\alpha$-model; when $\alpha=0, \beta \neq 0$ it is called $\beta$-model.
Using the variables $(\boldsymbol{Y}, \boldsymbol{X})$, that are related to the normal modes of oscillation and are defined by the following canonical transformation $\forall l=1, \ldots, N-1$ :

$$
\begin{equation*}
x_{l}=\sqrt{\frac{2}{N}} \sum_{j=1}^{N-1} \frac{X_{j}}{\sqrt{\nu_{j}}} \sin \left(\frac{j l \pi}{N}\right), \quad y_{l}=\sqrt{\frac{2}{N}} \sum_{j=1}^{N-1} Y_{j} \sqrt{\nu_{j}} \sin \left(\frac{j l \pi}{N}\right) \tag{2}
\end{equation*}
$$

with

$$
\begin{equation*}
\nu_{j}=2 \sin \left(\frac{j \pi}{2 N}\right) \tag{3}
\end{equation*}
$$

the Hamiltonian can be rewritten as follows:

$$
\begin{equation*}
\mathcal{H}(\boldsymbol{Y}, \boldsymbol{X})=\frac{1}{2} \sum_{j=1}^{N-1}\left[\nu_{j}\left(X_{j}^{2}+Y_{j}^{2}\right)\right]+\mathcal{H}^{*}(\boldsymbol{X}) \tag{4}
\end{equation*}
$$

with $\mathcal{H}^{*}$ containing only cubic and quartic terms in $\boldsymbol{X}$.
Now, it is convenient to temporarily introduce action-angle coordinates adapted to a fixed number of $n_{1} \leq N-1$ harmonic oscillators, by means of the canonical transformations

$$
Y_{j}=\sqrt{2 I_{j}} \cos \varphi_{j}, \quad X_{j}=\sqrt{2 I_{j}} \sin \varphi_{j}
$$

$\forall j=1, \ldots, n_{1}$. This leads Hamiltonian (4) to the following form:

$$
\begin{equation*}
\mathcal{H}(\boldsymbol{I}, \boldsymbol{\xi}, \boldsymbol{\varphi}, \boldsymbol{\eta})=\sum_{j=1}^{n_{1}} \omega_{j} I_{j}+\sum_{j=1}^{n_{2}} \Omega_{j} \frac{\left(\xi_{j}^{2}+\eta_{j}^{2}\right)}{2}+\mathcal{H}^{*}(\boldsymbol{I}, \boldsymbol{\xi}, \boldsymbol{\varphi}, \boldsymbol{\eta}), \tag{5}
\end{equation*}
$$

where we have split $\boldsymbol{\nu}$ as $\boldsymbol{\nu}=(\boldsymbol{\omega}, \boldsymbol{\Omega}) \in \mathbb{R}^{n_{1}} \times \mathbb{R}^{n_{2}}$, with $n_{2}=N-1-n_{1}$, and we have introduced $\xi_{j}=Y_{j+n_{1}}, \eta_{j}=X_{j+n_{1}}, \forall j=1, \ldots, n_{2}$.

We proceed now by expanding this Hamiltonian so as to approach a normal form adapted to an elliptic torus. For this purpose, we pick a fixed valud $\boldsymbol{I}^{\star}$ of the $n_{1}-$ dimensional action vector $\boldsymbol{I}$ and then we perform a translation, by defining

$$
\begin{equation*}
p_{j}=I_{j}-I_{j}^{\star} \quad \forall j=1, \ldots, n_{1}, \tag{6}
\end{equation*}
$$

and leaving the other coordinates unchanged. Finally, we expand the Hamiltonian in power series of $\boldsymbol{p}, \boldsymbol{\xi}, \boldsymbol{\eta}$ and, renaming the angles $\boldsymbol{\varphi}$ as $\boldsymbol{q}$, in Fourier series in the angles $\boldsymbol{q}$, so that the Hamiltonian will assume the following form:

$$
\begin{equation*}
\mathcal{H}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta})=\boldsymbol{\omega}^{(0)} \cdot \boldsymbol{p}+\sum_{j=1}^{n_{2}} \frac{\Omega_{j}^{(0)}}{2}\left(\xi_{j}^{2}+\eta_{j}^{2}\right)+\mathcal{F}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta}) . \tag{7}
\end{equation*}
$$

Let us stress that in the expression of the Hamiltonian above, we have omitted the parametric dependency with respect to the initial translation vector $\boldsymbol{I}^{\star}$, in spite of its crucial importance, that we are going to discuss in this final part of the present section. As a first remark, let us recall that in the phase space the distance from the equilibrium point corresponding to the origin of the canonical coordinates $\boldsymbol{I}, \boldsymbol{\xi}$ and $\boldsymbol{\eta}$ is essentially governed by the initial translation vector $\boldsymbol{I}^{\star}$ and so also for what concerns the width of the oscillations in the normal mode variables $(\boldsymbol{Y}, \boldsymbol{X})$. From an heuristic point of view, this allows us to see that the size of the perturbations is ruled by $\boldsymbol{I}^{\star}$.

Hereafter, in the notation we will stress the parametric dependency on the initial translation $\boldsymbol{I}^{\star}$ (actually, on the angular velocity vector $\boldsymbol{\omega}^{(0)}$ that can be put in a bijective correspondence with $\boldsymbol{I}^{\star}$ ) just when it is really necessary for the understanding. In particular, this will be mandatory in section 3.4, where the formal statement of theorem 3.1 is reported. In fact, that rigorous result claims that our algorithm constructing elliptic tori (which is described in the next section) is converging just in a subset of an open ball of the initial vectors $\boldsymbol{I}^{\star}\left(\boldsymbol{\omega}^{(0)}\right)$. Under suitable conditions, such a subset has positive Lebesgue measure which is larger and larger (in a probabilistic sense) for perturbations shrinking to zero.

## 3 Construction of invariant elliptic tori by a normal form algorithm

We present the formal algorithm making reference to a generic Hamiltonian with $n_{1}+n_{2}$ degrees of freedom, where the canonical coordinates $(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta})$ can be naturally split in two parts, that are $(\boldsymbol{p}, \boldsymbol{q}) \in \mathbb{R}^{n_{1}} \times \mathbb{T}^{n_{1}}$, action-angle variables in the neighborhood of the wanted torus, and $(\boldsymbol{\xi}, \boldsymbol{\eta}) \in \mathbb{R}^{n_{2}} \times \mathbb{R}^{n_{2}}$, in the neighborhood of the origin.

Let us introduce some notation. For some fixed positive integer $K$ we introduce the distinguished classes of functions $\widehat{\mathcal{P}} \widehat{\hat{m}, \hat{\ell}}$, with integers $\hat{m}, \hat{\ell}, s \geq 0$; any generic function

[^1]$g \in \widehat{\mathcal{P}}_{\hat{m}, \hat{\ell}}^{s K}$ can be written as
\[

$$
\begin{equation*}
g(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta})=\sum_{\substack{\boldsymbol{m} \in \mathbb{N}^{n_{1}} \\|\boldsymbol{m}|=\tilde{m}}} \sum_{\substack{(\ell, \bar{\ell}) \in \mathbb{N}^{2 n_{2}} \\|\ell|+|\bar{\ell}|=\grave{\ell}}} \sum_{\substack{k \in \mathbb{Z}^{n_{1}} \\|\boldsymbol{k}| \leq s K}} \boldsymbol{p}^{\boldsymbol{m}} \boldsymbol{\xi}^{\ell} \boldsymbol{\eta}^{\bar{\ell}}\left[c_{\boldsymbol{m}, \ell, \bar{\ell}, \boldsymbol{k}} \cos (\boldsymbol{k} \cdot \boldsymbol{q})+d_{\boldsymbol{m}, \ell, \bar{\ell}, \boldsymbol{k}} \sin (\boldsymbol{k} \cdot \boldsymbol{q})\right], \tag{8}
\end{equation*}
$$

\]

with the coefficients $c_{\boldsymbol{m}, \ell, \bar{\ell}, \boldsymbol{k}}, d_{\boldsymbol{m}, \ell, \bar{\ell}, \boldsymbol{k}} \in \mathbb{R}$; in the previous formula, we have introduced the symbol $|\cdot|$ to denote the $\ell_{1}$-norm and we have adopted the multi-index notation, i.e., $\boldsymbol{p}^{\boldsymbol{m}}=\prod_{j=1}^{n_{1}} p_{j}^{m_{j}}$. We stress that the positive integer $K$ is chosen in such a way to exploit the Fourier decay of the coefficients in the analytic functions. As it is detailed in subsection 4.4 of [24], this is made in order to split all the Hamiltonian in blocks of different orders of magnitude with respect the natural small parameter related to the quasi-integrable problem that is the object of our study. To fix the ideas, a suitable choice of the parameter $K$ allows to obtain initial estimates of the type described into the fundamental assumption (c) of theorem 3.1. We say that $g \in \mathcal{P}_{\ell}^{s K}$ if

$$
\begin{equation*}
g \in \bigcup_{\substack{\hat{m} \geq 0, \hat{\ell} \geq 0 \\ 2 \hat{m}+\hat{\ell}=\ell}} \widehat{\mathcal{P}}_{\hat{m}, \hat{\ell}}^{s K} \tag{9}
\end{equation*}
$$

Hereafter, we will omit the dependence of the function from the variables, unless it has some special meaning. Moreover, we will adopt the usual notation for the average of a function $g$ with respect to the generic angles $\boldsymbol{\vartheta} \in \mathbb{T}^{n}$, i.e., $\langle g\rangle_{\vartheta}=\int_{\mathbb{T}^{n}} \mathrm{~d} \vartheta_{1} \ldots \mathrm{~d} \vartheta_{n} g /(2 \pi)^{n}$.

Let us consider a Hamiltonian that can be written in the following way:

$$
\begin{align*}
H^{(0)}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta})= & \boldsymbol{\omega}^{(0)} \cdot \boldsymbol{p}+\sum_{j=1}^{n_{2}} \frac{\Omega_{j}^{(0)}}{2}\left(\xi_{j}^{2}+\eta_{j}^{2}\right)+\sum_{s \geq 0} \sum_{\ell \geq 3} f_{\ell}^{(0, s)}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta})  \tag{10}\\
& +\sum_{s \geq 1} \sum_{\ell=0}^{2} f_{\ell}^{(0, s)}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta})
\end{align*}
$$

where $f_{\ell}^{(0, s)} \in \mathcal{P}_{\ell}^{s K}$, while the first upper index is related to the normalization step. Starting from the Hamiltonian described in (1) and following the prescriptions given in the previous section, one can easily check that a FPU chain of $N+1$ particles can be brought to the form above. In other words, the Hamiltonian in formula (7) can be expanded as $H^{(0)}$ in (10), where $f_{\ell}^{(0, s)}=0$ when $s \geq 2$ and $f_{\ell}^{(0,1)} \in \mathcal{P}_{\ell}^{K}, f_{\ell}^{(0,2)} \in \mathcal{P}_{\ell}^{2 K}$ $\forall \ell \geq 0$, with ${ }^{2} K=2$. This holds true, both for the $\alpha-$ model and the $\beta$ one.

Our main purpose is to eliminate from the Hamiltonian all the terms having total degree less than three in the square root of the actions; by referring to the paradigmatic

[^2]form described in (10), the unwanted terms are appearing in its last row. Actually, such a goal can be achieved by performing an infinite sequence of canonical transformations, so as to bring the Hamiltonian to the following final normal form:
\[

$$
\begin{equation*}
H^{(\infty)}(\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{\Xi}, \boldsymbol{\Theta})=\boldsymbol{\omega}^{(\infty)} \cdot \boldsymbol{P}+\sum_{j=1}^{n_{2}} \frac{\Omega_{j}^{(\infty)}}{2}\left(\Xi_{j}^{2}+\Theta_{j}^{2}\right)+\sum_{s \geq 0} \sum_{\ell \geq 3} f_{\ell}^{(\infty, s)}(\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{\Xi}, \boldsymbol{\Theta})+\mathcal{E}^{(\infty)}, \tag{11}
\end{equation*}
$$

\]

with $f_{\ell}^{(\infty, s)} \in \mathcal{P}_{\ell}^{s K}$ and $\mathcal{E}^{(\infty)} \in \mathcal{P}_{0}^{0}$, i.e., it is a constant. It is easy to check that

$$
\begin{equation*}
(\boldsymbol{P}(t), \boldsymbol{Q}(t), \boldsymbol{\Xi}(t), \boldsymbol{\Theta}(t))=\left(\mathbf{0}, \boldsymbol{Q}_{0}+\boldsymbol{\omega}^{(\infty)} t, \mathbf{0}, \mathbf{0}\right) \tag{12}
\end{equation*}
$$

is a solution of Hamilton's equations related to (11), since that Hamiltonian, except for its first part, contains terms of type $\mathcal{O}\left(\|\boldsymbol{P}\|^{2}\right), \mathcal{O}(\|\boldsymbol{P}\|\|(\boldsymbol{\Xi}, \boldsymbol{\Theta})\|)$ and $\mathcal{O}\left(\|(\boldsymbol{\Xi}, \boldsymbol{\Theta})\|^{3}\right)$ only. Therefore, the motion law (12), that is generated by the initial condition ( $\mathbf{0}, \boldsymbol{Q}_{0}, \mathbf{0}, \mathbf{0}$ ), is quasi-periodic with an angular velocity vector equal to $\boldsymbol{\omega}^{(\infty)}$ and the corresponding orbit lies on the $n_{1}$-dimensional invariant torus $\boldsymbol{P}=\mathbf{0}, \boldsymbol{\Xi}=\boldsymbol{\Theta}=\mathbf{0}$. The energy level of such a manifold is $H^{(\infty)}(\mathbf{0}, \boldsymbol{Q}, \mathbf{0}, \mathbf{0})=\mathcal{E}^{(\infty)}$. Moreover, it is elliptic in the sense that the transverse dynamics in a neighborhood of the invariant torus itself is given by oscillations whose corresponding angular velocity vector is approaching $\boldsymbol{\Omega}^{(\infty)}$ in the limit of $\|(\boldsymbol{\Xi}, \boldsymbol{\Theta})\|$ going to zero.

The formal algorithm for the construction of the normal form is composed by a sequence of canonical transformations, defined using the formalism of Lie series. We can summarize the $r$-th normalization step, by giving the formula defining the canonical change of coordinates that transforms the intermediate Hamiltonian $H^{(r-1)}$ into the subsequent $H^{(r)}$, i.e.,

$$
\begin{equation*}
\exp \left(\mathcal{L}_{\chi_{0}^{(r)}}\right) \circ \exp \left(\mathcal{L}_{\chi_{1}^{(r)}}\right) \circ \exp \left(\mathcal{L}_{\chi_{2}^{(r)}}\right) \circ \mathfrak{D}^{(r)} \tag{13}
\end{equation*}
$$

where $\mathcal{L}_{f} \cdot=\{\cdot, f\}$ is the Lie derivative operator related to the Poisson bracket $\{\cdot, \cdot\}$ and the Lie series $3^{3}$ operator

$$
\begin{equation*}
\exp \left(\mathcal{L}_{\chi_{j}^{(r)}}\right) \cdot=\sum_{i=0}^{\infty} \frac{1}{i!} \mathcal{L}_{\chi_{j}^{(r)}}^{i} \tag{14}
\end{equation*}
$$

removes the Hamiltonian terms with total degree in the square root of the actions equal to $j$ and with trigonometric degree in the angles $\boldsymbol{q}$ up to $r K$. Moreover, by a linear canonical transformation $\mathfrak{D}^{(r)}$, the terms that are quadratic in $(\boldsymbol{\xi}, \boldsymbol{\eta})$ and do not depend on the angles $\boldsymbol{q}$ are brought to a diagonal form. At the end of this $r$-th normalization

[^3]step, the ineliminable terms that are independent on the angles $\boldsymbol{q}$ and either linear in $\boldsymbol{p}$ or quadratic in $(\boldsymbol{\xi}, \boldsymbol{\eta})$ are added to the normal form part. This requires to update the angular velocities from $\left(\boldsymbol{\omega}^{(r-1)}, \boldsymbol{\Omega}^{(r-1)}\right)$ to $\left(\boldsymbol{\omega}^{(r)}, \boldsymbol{\Omega}^{(r)}\right)$, that is why in (11) the Hamiltonian in normal form has new frequency vectors $\boldsymbol{\omega}^{(\infty)}$ and $\boldsymbol{\Omega}^{(\infty)}$.

Now, let us suppose that the Hamiltonian at step $r-1$ can be expanded as follows:

$$
\begin{align*}
H^{(r-1)}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta})= & \boldsymbol{\omega}^{(r-1)} \cdot \boldsymbol{p}+\sum_{j=1}^{n_{2}} \frac{\Omega_{j}^{(r-1)}}{2}\left(\xi_{j}^{2}+\eta_{j}^{2}\right)+\sum_{s \geq r} \sum_{\ell=0}^{2} f_{\ell}^{(r-1, s)}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta})  \tag{15}\\
& +\sum_{s \geq 0} \sum_{\ell \geq 3} f_{\ell}^{(r-1, s)}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta})+\mathcal{E}^{(r-1)}
\end{align*}
$$

where $f_{\ell}^{(r-1, s)} \in \mathcal{P}_{\ell}^{s K}$ and $\mathcal{E}^{(r-1)} \in \mathcal{P}_{0}^{0}$, i.e., it is a constant gathering all the terms which cannot be eliminated by the first homological equations that have already been solved for all the normalization steps $1, \ldots, r-1$, that have been already performed. Let us emphasize that the starting Hamiltonian $H^{(0)}$ written in formula (10) is in form (15), with $r=1$ and $\mathcal{E}^{(0)}=0$. In the following subsections we are going to detail how the algorithm actually works.

### 3.1 First stage of the $r$-th normalization step

In the context of the $r$-th normalization step, the first stage aims to remove the terms depending just on the angles $\boldsymbol{q}$ up to the trigonometrical degree $r K$, i.e., the terms collected in $f_{0}^{(r-1, r)}$. We determine the generating function $\chi_{0}^{(r)}$ by solving the homological equation

$$
\begin{equation*}
\left\{\chi_{0}^{(r)}, \boldsymbol{\omega}^{(r-1)} \cdot \boldsymbol{p}\right\}+f_{0}^{(r-1, r)}(\boldsymbol{q})-\left\langle f_{0}^{(r-1, r)}(\boldsymbol{q})\right\rangle_{\boldsymbol{q}}=0 \tag{16}
\end{equation*}
$$

Let us remark that, since $f_{0}^{(r-1, r)}$ depends on $\boldsymbol{q}$ only, $\left\langle f_{0}^{(r-1, r)}\right\rangle_{\boldsymbol{q}}$ is a constant term that gives its contribution to the energy level of the final invariant elliptic torus; thus, we introduce

$$
\begin{equation*}
\mathcal{E}^{(r)}=\mathcal{E}^{(r-1)}+\left\langle f_{0}^{(r-1, r)}\right\rangle_{\boldsymbol{q}} . \tag{17}
\end{equation*}
$$

Starting from the Fourier expansion

$$
\begin{equation*}
f_{0}^{(r-1, r)}(\boldsymbol{q})=\sum_{|\boldsymbol{k}| \leq r K}\left[c_{\boldsymbol{k}} \cos (\boldsymbol{k} \cdot \boldsymbol{q})+d_{\boldsymbol{k}} \sin (\boldsymbol{k} \cdot \boldsymbol{q})\right] \tag{18}
\end{equation*}
$$

with $c_{\boldsymbol{k}}$, $d_{\boldsymbol{k}}$ real known coefficients, we readily get the following expression for the generating function

$$
\begin{equation*}
\chi_{0}^{(r)}(\boldsymbol{q})=\sum_{0<|\boldsymbol{k}| \leq r K}\left[-\frac{c_{\boldsymbol{k}} \sin (\boldsymbol{k} \cdot \boldsymbol{q})}{\boldsymbol{k} \cdot \boldsymbol{\omega}^{(r-1)}}+\frac{d_{\boldsymbol{k}} \cos (\boldsymbol{k} \cdot \boldsymbol{q})}{\boldsymbol{k} \cdot \boldsymbol{\omega}^{(r-1)}}\right] \tag{19}
\end{equation*}
$$

Therefore, the homological equation (16) admits a solution for the generating function $\chi_{0}^{(r)}$, provided that the frequency vector $\boldsymbol{\omega}^{(r-1)}$ is non-resonant up to the trigonometric degree $r K$. To fix the ideas, we assume that it fulfills the following Diophantine condition:

$$
\begin{equation*}
\min _{0<|\boldsymbol{k}| \leq r K}\left|\boldsymbol{k} \cdot \boldsymbol{\omega}^{(r-1)}\right| \geq \frac{\gamma}{(r K)^{\tau}} . \tag{20}
\end{equation*}
$$

Now, when we apply the canonical transformation $\exp \mathcal{L}_{\chi_{0}^{(r)}}$ to the Hamiltonian at step $r-1$, renaming the new variables as the old ones (by abuse of notation), the transformed Hamiltonian reads as

$$
\begin{align*}
H^{(\mathrm{I} ; r)} & =\exp \left(\mathcal{L}_{\chi_{0}^{(r)}}\right) H^{(r-1)} \\
& =\boldsymbol{\omega}^{(r-1)} \cdot \boldsymbol{p}+\sum_{j=1}^{n_{2}} \Omega_{j}^{(r-1)} \frac{\left(\xi_{j}^{2}+\eta_{j}^{2}\right)}{2}+\sum_{s \geq r} \sum_{\ell=0}^{2} f_{\ell}^{(\mathrm{I} ; r, s)}+\sum_{s \geq 0} \sum_{\ell \geq 3} f_{\ell}^{(\mathrm{I} ; r, s)}+\mathcal{E}^{(r)}, \tag{21}
\end{align*}
$$

where the functions $f_{\ell}^{(\mathrm{I} ; r, s)} \in \mathcal{P}_{\ell}^{s K}$ contributing to the definition of the new Hamiltonian are recursively defined as follows:

$$
\begin{array}{rlrl}
f_{0}^{(\mathrm{I} ; r, r)} & =0, & & \\
f_{0}^{(\mathrm{I} ; r, r+m)} & =f_{0}^{(r-1, r+m)} & & \text { for } 0<m<r, \\
f_{\ell}^{(\mathrm{I} ; r, s)} & =\sum_{j=0}^{\lfloor s / r\rfloor} \frac{1}{i!} \mathcal{L}_{\chi_{0}^{(r)}}^{i} f_{\ell+2 i}^{(r-1, s-i r)} & & \text { for } \ell=0, s \geq 2 r \text { or } \ell=1,2, s \geq r  \tag{22}\\
& \quad \text { or } \ell \geq 3, s \geq 0 .
\end{array}
$$

Indeed, in order to practically implement the definitions included in formula (22), we have found convenient to proceed in the following way. First, we define the new terms as the old ones, i.e., $f_{\ell}^{(\mathrm{I} ; r, s)}=f_{\ell}^{(r-1, s)}$; then, each term generated by Lie derivatives with respect to the generating function is added to the corresponding class. This is made by a sequence of redefinitions, for each of them an abuse of notation has to be tolerated. Since the Lie derivative of the generating function $\chi_{0}^{(r)}$ decreases by 1 the degree in $\boldsymbol{p}$, while the trigonometrical degree in the angles $\boldsymbol{q}$ is increased by $r K$, we set

$$
\begin{equation*}
f_{\ell-2 i}^{(\mathrm{I} ; r, s+i r)} \hookleftarrow \frac{1}{i!} \mathcal{L}_{\chi_{0}^{(r)}}^{i} f_{\ell}^{(r-1, s)} \quad \forall \ell \geq 0,1 \leq i \leq\lfloor\ell / 2\rfloor, s \geq 0 \tag{23}
\end{equation*}
$$

where with the notation $a \hookleftarrow b$ we mean that the quantity $a$ is redefined so as to be equal to $a+b$. Let us recall that $f_{0}^{(\mathrm{I} ; r, r)}=0$, because of the homological equation (16).

### 3.2 Second stage of the $r$-th normalization step

The second stage of the $r$-th normalization step acts on the Hamiltonian that is initially expanded as in (21), in order to remove the perturbing terms linear in $(\boldsymbol{\xi}, \boldsymbol{\eta})$ and independent of $\boldsymbol{p}$, that are those collected in $f_{1}^{(\mathrm{I} ; r, r)}$. Thus, we have to solve the following homological equation:

$$
\begin{equation*}
\left\{\chi_{1}^{(r)}, \boldsymbol{\omega}^{(r-1)} \cdot \boldsymbol{p}+\sum_{j=1}^{n_{2}} \frac{\Omega_{j}^{(r-1)}}{2}\left(\xi_{j}^{2}+\eta_{j}^{2}\right)\right\}+f_{1}^{(\mathrm{I} ; r, r)}(\boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta})=0 . \tag{24}
\end{equation*}
$$

In order to solve such an equation, it is convenient to temporarily replace the transverse variables $(\boldsymbol{\xi}, \boldsymbol{\eta})$ with action-angle canonical coordinates, by defining $\xi_{j}=\sqrt{2 J_{j}} \cos \left(\varphi_{j}\right)$
and $\eta_{j}=\sqrt{2 J_{j}} \sin \left(\varphi_{j}\right)$. Let us write the expansion of $f_{1}^{(\mathrm{I} ; r, r)}(\boldsymbol{q}, \boldsymbol{J}, \boldsymbol{\varphi})$ as follows:

$$
\begin{equation*}
f_{1}^{(\mathbb{I} r, r)}(\boldsymbol{q}, \boldsymbol{J}, \boldsymbol{\varphi})=\sum_{0 \leq \boldsymbol{k} \leq r K} \sum_{j=1}^{n_{2}} \sqrt{2 J_{j}}\left[c_{\boldsymbol{k}, j}^{( \pm)} \cos \left(\boldsymbol{k} \cdot \boldsymbol{q} \pm \varphi_{j}\right)+d_{\boldsymbol{k}, j}^{( \pm)} \sin \left(\boldsymbol{k} \cdot \boldsymbol{q} \pm \varphi_{j}\right)\right] \tag{25}
\end{equation*}
$$

where $c_{\boldsymbol{k}, j}^{( \pm)}$and $d_{\boldsymbol{k}, j}^{( \pm)}$are some real coefficients. Therefore, the generating function $\chi_{1}^{(r)}$ solving equation (24) is determined in such a way that

$$
\begin{equation*}
\chi_{1}^{(r)}(\boldsymbol{q}, \boldsymbol{J}, \boldsymbol{\varphi})=\sum_{0 \leq \boldsymbol{k} \leq r K} \sum_{j=1}^{n_{2}} \sqrt{2 J_{j}}\left[-\frac{c_{\boldsymbol{k}, j}^{( \pm)} \sin \left(\boldsymbol{k} \cdot \boldsymbol{q} \pm \varphi_{j}\right)}{\boldsymbol{k} \cdot \boldsymbol{\omega}^{(r-1)} \pm \Omega_{j}^{(r-1)}}+\frac{d_{\boldsymbol{k}, j}^{( \pm)} \cos \left(\boldsymbol{k} \cdot \boldsymbol{q} \pm \varphi_{j}\right)}{\boldsymbol{k} \cdot \boldsymbol{\omega}^{(r-1)} \pm \Omega_{j}^{(r-1)}}\right] \tag{26}
\end{equation*}
$$

This expression is well-defined, provided that the frequency vector $\boldsymbol{\omega}^{(r-1)}$ satisfies the so-called first Melnikov non-resonance condition up to order $r K$ (see [38]), i.e.,

$$
\begin{equation*}
\min _{\substack{0<\boldsymbol{k}|\leq r K\\| \boldsymbol{\ell} \mid=1}}\left|\boldsymbol{k} \cdot \boldsymbol{\omega}^{(r-1)}+\boldsymbol{\ell} \cdot \boldsymbol{\Omega}^{(r-1)}\right| \geq \frac{\gamma}{(r K)^{\tau}} \quad \text { and } \quad \min _{|\boldsymbol{\ell}|=1}\left|\boldsymbol{\ell} \cdot \boldsymbol{\Omega}^{(r-1)}\right| \geq \gamma, \tag{27}
\end{equation*}
$$

for some fixed values of both $\gamma>0$ and $\tau>n_{1}-1$. By applying the Lie series $\exp \left(\mathcal{L}_{\chi_{1}^{(r)}}\right)$ to the old Hamiltonian $H^{(\mathrm{I} ; r)}$, we have a new one, i.e., $H^{(\mathrm{II} ; r)}=\exp \left(\mathcal{L}_{\chi_{1}^{(r)}}\right) H^{(\mathrm{I} ; r)}$, where now the generating function $\chi_{1}^{(r)}$ is reconsidered to be dependent on the variables $(\boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta})$. The expansion of the new Hamiltonian will have exactly the same structure as that described in (21), with $f_{\ell}^{(\mathrm{II} ; r, s)}$ in place of $f_{\ell}^{(\mathrm{I} ; r, s)}$ and $f_{1}^{(\mathrm{II} ; r, r)}=0$. The new terms $f_{\ell}^{(\mathrm{II} ; r, s)}$ that compose the Hamiltonian can be determined with calculations similar to those listed during the description of the first stage of normalization. For what concerns the explicit computer algebra manipulations, it is enough to remark that the Lie derivative with respect to $\chi_{1}^{(r)}$ decreases by 1 the total degree in the square root of the actions. This explains why it is convenient to set the sequence of redefinitions of the new terms $f_{\ell}^{(\mathrm{II} ; r, s)}$ in such a way that

$$
\begin{equation*}
f_{\ell-i}^{(\mathrm{II} ;, s+i r)} \hookleftarrow \frac{1}{i!} \mathcal{L}_{\chi_{1}^{(r)}}^{i} f_{\ell}^{(\mathrm{I} ; r, s)} \quad \forall \ell \geq 0,1 \leq i \leq \ell, s \geq 0 \tag{28}
\end{equation*}
$$

This also means that $f_{1}^{(\mathrm{II} ; r, r)}=0$, due to the homological equation (24).

### 3.3 Third stage of the $r-$ th normalization step

The third and last stage of normalization is more elaborated. It aims to remove terms belonging to two different classes: first, those linear in $\boldsymbol{p}$ and independent of $(\boldsymbol{\xi}, \boldsymbol{\eta})$, moreover, other terms that are quadratic in $(\boldsymbol{\xi}, \boldsymbol{\eta})$ and independent of $\boldsymbol{p}$. Such a part of the perturbation is removed by the composition of two canonical transformations expressed by Lie series, with generating functions $X_{2}^{(r)}(\boldsymbol{p}, \boldsymbol{q}) \in \widehat{\mathcal{P}}_{1,0}^{r K}$ and $Y_{2}^{(r)}(\boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta}) \in \widehat{\mathcal{P}}_{0,2}^{r K}$, respectively. Moreover, the third stage is ended by a linear canonical transformation $\mathfrak{D}_{2}^{(r)}$ that leaves the pair $(\boldsymbol{p}, \boldsymbol{q})$ unchanged and it aims to diagonalize the terms that are quadratic in $(\boldsymbol{\xi}, \boldsymbol{\eta})$
and independent of the angles $\boldsymbol{q}$. Let us detail all these changes of coordinates, so that the algorithm is unambiguously defined.

The generating functions $X_{2}^{(r)}$ is in charge to remove terms that are linear in $\boldsymbol{p}$ and do depend on the angles $\boldsymbol{q}$ up to the trigonometric degree $r K$. Therefore, it is a solution of the following homological equation:

$$
\begin{equation*}
\left\{X_{2}^{(r)}, \boldsymbol{\omega}^{(r-1)} \cdot \boldsymbol{p}\right\}+f_{2}^{(\mathrm{II} ; r, r)}(\boldsymbol{p}, \boldsymbol{q})-\left\langle f_{2}^{(\mathrm{II} ; r, r)}(\boldsymbol{p}, \boldsymbol{q})\right\rangle_{\boldsymbol{q}}=0 \tag{29}
\end{equation*}
$$

Let us recall that $f_{2}^{(\mathrm{II} ; r, r)} \in \mathcal{P}_{2}^{r K}=\widehat{\mathcal{P}}_{1,0}^{r K} \cup \widehat{\mathcal{P}}_{0,2}^{r K}$; in general, such a function really depends on all the canonical variables, i.e., $f_{2}^{(\mathrm{II} ; r, r)}=f_{2}^{(\mathrm{II} ; r, r)}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta})$. Therefore, we denote with $f_{2}^{(\mathrm{II} ; r, r)}(\boldsymbol{p}, \boldsymbol{q})$ the subpart of $f_{2}^{(\mathrm{II} ; r, r)}$ that is depending just on $(\boldsymbol{p}, \boldsymbol{q})$. Analogously, in the following $f_{2}^{(I I ; r, r)}(\boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta})$ will denote the subpart of $f_{2}^{(I I ; r, r)}$ that does depend on all the canonical variables but the actions $\boldsymbol{p}$ and so on also for what concerns $f_{2}^{(\mathrm{II} ; r, r)}(\boldsymbol{\xi}, \boldsymbol{\eta})$. This highly non-standard notation will be adopted all along the present subsection. Let us here emphasize that the term $\left\langle f_{2}^{(I I ; r, r)}(\boldsymbol{p}, \boldsymbol{q})\right\rangle_{\boldsymbol{q}}$ will be added to the part in normal form, with a change of the angular velocity vector $\boldsymbol{\omega}^{(r-1)}$. We can write the expansion of $f_{2}^{(I I ; r, r)}(\boldsymbol{p}, \boldsymbol{q})$ as

$$
\begin{equation*}
f_{2}^{(\mathrm{II} ; r, r)}(\boldsymbol{p}, \boldsymbol{q})=\sum_{0<\boldsymbol{k} \leq r K} \sum_{j=1}^{n_{1}} p_{j}\left[c_{\boldsymbol{k}, j} \cos (\boldsymbol{k} \cdot \boldsymbol{q})+d_{\boldsymbol{k}, j} \sin (\boldsymbol{k} \cdot \boldsymbol{q})\right], \tag{30}
\end{equation*}
$$

thus, the corresponding solution of the homological equation (29) is given by

$$
\begin{equation*}
X_{2}^{(r)}(\boldsymbol{p}, \boldsymbol{q})=\sum_{0<\boldsymbol{k} \leq r K} \sum_{j=1}^{n_{1}} p_{j}\left[-\frac{c_{\boldsymbol{k}, j} \sin (\boldsymbol{k} \cdot \boldsymbol{q})}{\boldsymbol{k} \cdot \boldsymbol{\omega}^{(r-1)}}+\frac{d_{\boldsymbol{k}, j} \cos (\boldsymbol{k} \cdot \boldsymbol{q})}{\boldsymbol{k} \cdot \boldsymbol{\omega}^{(r-1)}}\right] \tag{31}
\end{equation*}
$$

Let us remark that $X_{2}^{(r)}$ is well-defined provided that the frequency vector $\boldsymbol{\omega}^{(r-1)}$ satisfies the non-resonance condition already reported in (20).

The generating function $Y_{2}^{(r)}$ aims to remove the part of the term of $f_{2}^{(\mathrm{II} ; r, r)}$ that is quadratic in $(\boldsymbol{\xi}, \boldsymbol{\eta})$ and does depend on the angles $\boldsymbol{q}$. Therefore, $Y_{2}^{(r)}$ has to solve the following homological equation:

$$
\begin{equation*}
\left\{Y_{2}^{(r)}, \boldsymbol{\omega}^{(r-1)} \cdot \boldsymbol{p}+\sum_{j=1}^{n_{2}} \frac{\Omega_{j}^{(r)}}{2}\left(\xi_{j}^{2}+\eta_{j}^{2}\right)\right\}+f_{2}^{(\mathrm{II} ; r, r)}(\boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta})-\left\langle f_{2}^{(\mathrm{II} ; r, r)}(\boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta})\right\rangle_{\boldsymbol{q}}=0 \tag{32}
\end{equation*}
$$

It is convenient to defer the discussion about the treatment of the term $\left\langle f_{2}^{(I I ; r, r)}(\boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta})\right\rangle_{\boldsymbol{q}}$, because it is rather peculiar. Now, let us temporarily replace the transverse canonical variables $(\boldsymbol{\xi}, \boldsymbol{\eta})$ with the action-angle coordinates $(\boldsymbol{J}, \boldsymbol{\varphi})$ (exactly in the same way as in the previous section (3.2), so as to write the expansion of the perturbing term we now want to remove in the following way:

$$
\begin{align*}
f_{2}^{(\mathrm{II} ; r, r)}(\boldsymbol{q}, \boldsymbol{J}, \boldsymbol{\varphi})=\sum_{0<\boldsymbol{k} \leq r K} \sum_{i, j=1}^{n_{2}} 2 \sqrt{J_{i} J_{j}} & {\left[c_{\boldsymbol{k}, i, j}^{( \pm, \pm)} \cos \left(\boldsymbol{k} \cdot \boldsymbol{q} \pm \varphi_{i} \pm \varphi_{j}\right)\right.}  \tag{33}\\
& \left.+d_{\boldsymbol{k}, i, j}^{( \pm, \pm)} \sin \left(\boldsymbol{k} \cdot \boldsymbol{q} \pm \varphi_{i} \pm \varphi_{j}\right)\right]
\end{align*}
$$

Thus, the generating function $Y_{2}^{(r)}$ is determined by equation (32) in such a way that

$$
\begin{align*}
Y_{2}^{(r)}(\boldsymbol{q}, \boldsymbol{J}, \boldsymbol{\varphi})=\sum_{0<\boldsymbol{k} \leq r K} \sum_{i, j=1}^{n_{2}} 2 \sqrt{J_{i} J_{j}}[ & -\frac{c_{\boldsymbol{k}, i, j}^{( \pm, \pm)} \sin \left(\boldsymbol{k} \cdot \boldsymbol{q} \pm \varphi_{i} \pm \varphi_{j}\right)}{\boldsymbol{k} \cdot \boldsymbol{\omega}^{(r-1)} \pm \Omega_{i}^{(r-1)} \pm \Omega_{j}^{(r-1)}} \\
& \left.+\frac{d_{\boldsymbol{k}, i, j}^{( \pm, \pm)} \cos \left(\boldsymbol{k} \cdot \boldsymbol{q} \pm \varphi_{i} \pm \varphi_{j}\right)}{\boldsymbol{k} \cdot \boldsymbol{\omega}^{(r-1)} \pm \Omega_{i}^{(r-1)} \pm \Omega_{j}^{(r-1)}}\right] \tag{34}
\end{align*}
$$

that is well defined provided that the angular velocity vector $\boldsymbol{\omega}^{(r-1)}$ satisfies both the already mentioned Diophantine inequality (20) and the so-called second Melnikov nonresonance condition up to order $r K$ (see [38]), i.e.,

$$
\begin{equation*}
\min _{\substack{0<k|k| r K, \\ \text { k| } \mid=2}}\left|\boldsymbol{k} \cdot \boldsymbol{\omega}^{(r-1)}+\boldsymbol{\ell} \cdot \boldsymbol{\Omega}^{(r-1)}\right| \geq \frac{\gamma}{(r K)^{\tau}} \tag{35}
\end{equation*}
$$

with fixed values of both the parameters $\gamma>0$ and $\tau>n_{1}-1$.
After having performed these two changes of coordinates, we still may have terms that are linear in $\boldsymbol{p}$ or quadratic in $(\boldsymbol{\xi}, \boldsymbol{\eta})$ and do not depend on $\boldsymbol{q}$. The former ones can be directly added to the part in normal form, whereas the latter have to be preliminarily put in diagonal form. This can be done by a linear canonical transformation $\mathfrak{D}^{(r)}$, that can be computed in a direct way ${ }^{4}$ so that

$$
\begin{equation*}
\left.\left(\sum_{j=1}^{n_{2}} \frac{\Omega_{j}^{(r-1)}\left(\xi_{j}^{2}+\eta_{j}^{2}\right)}{2}+f_{2}^{(\mathrm{II} ; r, r)}(\boldsymbol{\xi}, \boldsymbol{\eta})\right)\right|_{(\boldsymbol{\xi}, \boldsymbol{\eta})=\mathfrak{D}^{(r)}(\overline{\boldsymbol{\xi}}, \bar{\eta})}=\sum_{j=1}^{n_{2}} \frac{\Omega_{j}^{(r)}\left(\bar{\xi}_{j}^{2}+\bar{\eta}_{j}^{2}\right)}{2} \tag{36}
\end{equation*}
$$

as it is explained, e.g., in section 7 of [23]. Such a direct calculation can be performed provided that

$$
\begin{equation*}
\min _{|\ell|=2}\left|\boldsymbol{\ell} \cdot \boldsymbol{\Omega}^{(r-1)}\right| \geq \gamma \tag{37}
\end{equation*}
$$

Finally, we need to understand how these generating functions affect the Hamiltonian terms. Let us recall that the Poisson brackets with the generating functions $X_{2}^{(r)}$ and $Y_{2}^{(r)}$ do not change the total degree in the square root of the actions. In order to describe the definitions of the new Hamiltonian terms it is convenient to introduce the intermediate functions $g_{\ell}^{(r, s)}, g_{\ell}^{\prime(r, s)}$ in the following way. First, we define $g_{\ell}^{(r, s)}=f_{\ell}^{(\mathrm{II} ; r, s)}$ for all nonnegative values of the indexes $\ell, r$ and $s$; then, we consider the effects induced by the application of the Lie series with generating function $X_{2}^{(r)}$ to the Hamiltonian. In order to do that, we gather the new terms in $g_{\ell}^{(r, s)}$ according to both their total degree in the

[^4]square root of the actions and the trigonometric degree in the angles, this means that we perform sequences of redefinitions so that
\[

$$
\begin{equation*}
g_{\ell}^{(r, s+i r)} \hookleftarrow \frac{1}{i!} \mathcal{L}_{X_{2}^{(r)}}^{i} f_{\ell}^{(\mathrm{II} r, r, s)} \quad \forall i \geq 1, \ell \geq 0, s \geq 0 \tag{38}
\end{equation*}
$$

\]

In the same way, we gather in $g_{\ell}^{(r, s)}$ all the Hamiltonian terms created by the application of the Lie series having $Y_{2}^{(r)}$ as generating function. This is summarized by the following redefinitions: $g_{\ell}^{(r, s)}=g_{\ell}^{(r, s)}$ and

$$
\begin{equation*}
g_{\ell}^{\prime(r, s+i r)} \hookleftarrow \frac{1}{i!} \mathcal{L}_{Y_{2}^{(r)}}^{i} g_{\ell}^{(r, s)} \quad \forall i \geq 1, \ell \geq 0, s \geq 0 \tag{39}
\end{equation*}
$$

Let us remark that each class of type $\widehat{\mathcal{P}}_{\ell_{1}, \ell_{2}}^{s}$ or $\mathcal{P}_{\ell}^{s}$ is preserved by the diagonalization linear operator $\mathfrak{D}^{(r)}$, for all non-negative values of the indexes $\ell_{1}, \ell_{2}, \ell$ and $s$. Therefore, it is natural to put

$$
\begin{equation*}
f_{\ell}^{(r, s)}=g_{\ell}^{\prime(r, s)} \circ \mathfrak{D}^{(r)} \tag{40}
\end{equation*}
$$

for all indexes $\ell \geq 0$ and $s \geq 0$.
At the end of the $r$-th normalization step, it is convenient that the terms linearly depending just on $\boldsymbol{p}$ or $\boldsymbol{J}$ are included in the main part of the Hamiltonian, because all of them belong to the same class of functions, i.e. $\mathcal{P}_{2}^{0}$. For this purpose, we introduce the new angular velocity vector $\boldsymbol{\omega}^{(r)}$, in such a way that

$$
\begin{equation*}
\boldsymbol{\omega}^{(r)} \cdot \boldsymbol{p}=\boldsymbol{\omega}^{(r-1)} \cdot \boldsymbol{p}+f_{2}^{(\mathrm{II} ; r, 0)}(\boldsymbol{p}) \tag{41}
\end{equation*}
$$

while the new values of the components of $\boldsymbol{\Omega}^{(r)}$ are defined by equation (36), that also allows us to put $f_{2}^{(r, r)}=0$.

The Hamiltonian at the end of the $r$-th normalization step $H^{(r)}$ has the same structure of $H^{(r-1)}$ in (15), but the new perturbative terms $f_{\ell}^{(r, s)}$ with $\ell=0,1,2$ are expected to be smaller with respect to the previous ones, because of the Fourier decay of the coefficients jointly with the fact that we removed the part of perturbation up to the trigonometric degree $r K$.

### 3.4 On the rigorous analysis of the convergence of the algorithm

The non-resonance conditions we have assumed in (20), (27), (35) and (37) can be summarized in the following way:

$$
\begin{equation*}
\min _{\substack{0<\boldsymbol{k}|\leq r K \\ 0 \leq|\ell| \leq 2}}\left|\boldsymbol{k} \cdot \boldsymbol{\omega}^{(r-1)}+\boldsymbol{\ell} \cdot \boldsymbol{\Omega}^{(r-1)}\right| \geq \frac{\gamma}{(r K)^{\tau}} \quad \text { and } \quad \min _{0<|\ell| \leq 2}\left|\boldsymbol{\ell} \cdot \boldsymbol{\Omega}^{(r-1)}\right| \geq \gamma, \tag{42}
\end{equation*}
$$

with $\gamma>0$ and $\tau>n_{1}-1$. Let us here resume the parametric dependence of all the Hamiltonian terms on the initial value of the angular velocity vector $\boldsymbol{\omega}^{(0)}$, that is in a bijective correspondence with the initial translation vector $\boldsymbol{I}^{\star}$, as it has been discussed at the end of sect. 2. In particular, in the Diophantine inequalities reported in (42)
the angular velocity vectors at the $r$-th normalization step are functions of $\boldsymbol{\omega}^{(0)}$, i.e., $\boldsymbol{\omega}^{(r-1)}=\boldsymbol{\omega}^{(r-1)}\left(\boldsymbol{\omega}^{(0)}\right)$ and $\boldsymbol{\Omega}^{(r-1)}=\boldsymbol{\Omega}^{(r-1)}\left(\boldsymbol{\omega}^{(0)}\right)$. Let us recall that we do not try to keep a full control on the way $\boldsymbol{\omega}^{(r-1)}\left(\boldsymbol{\omega}^{(0)}\right)$ and $\boldsymbol{\Omega}^{(r-1)}\left(\boldsymbol{\omega}^{(0)}\right)$ are modified passing from the $r$-th normalization step to the next one. Such an approach is in contrast with what is usually done to construct the Komogorov normal form for maximal invariant tori, where the angular velocities are kept fixed (see [33] or, e.g., [24]), but it is somehow unavoidable because of the occurrence of the transversal angular velocities $\boldsymbol{\Omega}^{(r-1)}\left(\boldsymbol{\omega}^{(0)}\right)$ that in general cannot remain constant along the normalization procedure. This seems to prevent the complete construction of the normal form and, therefore, the proof of the existence of an elliptic torus. Nevertheless, following the approach designed by Pöschel in [45], it can be proved that the Lebesgue measure of the resonant regions where the Melnikov conditions are not satisfied shrinks to zero with the size of the perturbation. Therefore, an analysis made with a suitable scheme of estimates allows to prove the convergence of the algorithm, as it is summarized by the statement below.
Theorem 3.1 Consider the following family of real Hamiltonians:

$$
\begin{align*}
& H^{(0)}\left(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta} ; \boldsymbol{\omega}^{(0)}\right)=\boldsymbol{\omega}^{(0)} \cdot \boldsymbol{p}+\sum_{j=1}^{n_{2}} \frac{\Omega_{j}^{(0)}\left(\boldsymbol{\omega}^{(0)}\right)}{2}\left(\xi_{j}^{2}+\eta_{j}^{2}\right)+  \tag{43}\\
&+\sum_{s \geq 0} \sum_{\ell \geq 3} f_{\ell}^{(0, s)}\left(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta} ; \boldsymbol{\omega}^{(0)}\right)+\sum_{s \geq 1} \sum_{\ell=0}^{2} f_{\ell}^{(0, s)}\left(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta} ; \boldsymbol{\omega}^{(0)}\right)
\end{align*}
$$

where $(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta}) \in \mathcal{O}_{1} \times \mathbb{T}^{n_{1}} \times \mathcal{O}_{2}$ with $\mathcal{O}_{1}$ and $\mathcal{O}_{2}$ open neighborhoods of the origin in $\mathbb{R}^{n_{1}}$ and $\mathbb{R}^{2 n_{2}}$, respectively, while $\boldsymbol{\omega}^{(0)} \in \mathcal{U}, \mathcal{U}$ being an open subset of $\mathbb{R}^{n_{1}}$, and $f_{\ell}^{(0, s)} \in \mathcal{P}_{\ell}^{s K}$. Moreover, we assume that:
(a) all the functions $\Omega_{j}: \mathcal{U} \mapsto \mathbb{R}$ and $f_{\ell}^{(0, s)}: \mathcal{O}_{1} \times \mathbb{T}^{n_{1}} \times \mathcal{O}_{2} \times \mathcal{U} \mapsto \mathbb{R}$, appearing in (43), are analytic functions with respect to $\boldsymbol{\omega}^{(0)} \in \mathcal{U}$;
(b) $\Omega_{i}^{(0)}\left(\boldsymbol{\omega}^{(0)}\right) \neq \Omega_{j}^{(0)}\left(\boldsymbol{\omega}^{(0)}\right)$ and $\Omega_{i_{2}}^{(0)}\left(\boldsymbol{\omega}^{(0)}\right) \neq 0$ for $\boldsymbol{\omega}^{(0)} \in \mathcal{U}$ and $1 \leq i<j \leq n_{2}$, $1 \leq i_{2} \leq n_{2}$;
(c) for some fixed and positive values of $\varepsilon$ and $E$, one has

$$
\begin{equation*}
\sup _{\left(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta} ; \boldsymbol{\omega}^{(0)}\right) \in \mathcal{O}_{1} \times \mathbb{T}^{n_{1}} \times \mathcal{O}_{2} \times \mathcal{U}}\left|f_{\ell}^{(0, s)}\left(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta} ; \boldsymbol{\omega}^{(0)}\right)\right| \leq \varepsilon^{s} E \tag{44}
\end{equation*}
$$

$\forall s \geq 1, \ell \geq 0$ and $\forall \ell \geq 3$ when $s=0$.
Then, there is a positive $\varepsilon^{\star}$ such that for $0 \leq \varepsilon<\varepsilon^{\star}$ the following statement holds true: there exists a non-resonant set $\mathcal{U}^{(\infty)} \subset \mathcal{U}$ of positive Lebesgue measure and with the measure of $\mathcal{U} \backslash \mathcal{U}^{(\infty)}$ tending to zero for $\varepsilon \rightarrow 0$ for bounded $\mathcal{U}$, such that for each $\boldsymbol{\omega}^{(0)} \in$ $\mathcal{U}^{(\infty)}$ there exists an analytic canonical transformation $(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\xi}, \boldsymbol{\eta})=\psi_{\boldsymbol{\omega}^{(0)}}^{(\infty)}(\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{\Xi}, \boldsymbol{\Theta})$ leading the Hamiltonian to the normal form

$$
\begin{align*}
H^{(\infty)}\left(\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{\Xi}, \boldsymbol{\Theta} ; \boldsymbol{\omega}^{(0)}\right)= & \boldsymbol{\omega}^{(\infty)} \cdot \boldsymbol{P}+\sum_{j=1}^{n_{2}} \frac{\Omega_{j}^{(\infty)}\left(\Xi_{j}^{2}+\Theta_{j}^{2}\right)}{2}  \tag{45}\\
& +\sum_{s \geq 0} \sum_{\ell \geq 3} f_{\ell}^{(\infty, s)}\left(\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{\Xi}, \boldsymbol{\Theta} ; \boldsymbol{\omega}^{(0)}\right)+\mathcal{E}^{(\infty)}
\end{align*}
$$

where $\boldsymbol{\omega}^{(\infty)}=\boldsymbol{\omega}^{(\infty)}\left(\boldsymbol{\omega}^{(0)}\right), \boldsymbol{\Omega}^{(\infty)}=\boldsymbol{\Omega}^{(\infty)}\left(\boldsymbol{\omega}^{(0)}\right), f_{\ell}^{(\infty, s)} \in \mathcal{P}_{\ell}^{s K}$ and $\mathcal{E}^{(\infty)}=\mathcal{E}^{(\infty)}\left(\boldsymbol{\omega}^{(0)}\right)$ is a finite real value fixing the constant energy level that corresponds to the invariant torus $\left\{\left(\boldsymbol{P}=\mathbf{0}, \boldsymbol{Q} \in \mathbb{T}^{n_{1}}, \boldsymbol{\Xi}=\mathbf{0}, \boldsymbol{\Theta}=\mathbf{0}\right)\right\}$.

Actually, the statement above does not add anything new with respect to that reported in [45]. In spite of the fact that, here, we do not aim to deal with another purely analytical demonstration of the existence of elliptic tori, what really matters is in the proof of theorem 3.1, because it is based on a formal algorithm that is substantially the same (apart very minor differences) with respect to that described in subsections 3.1] 3.3. Therefore, the theorem above rigorously ensures the convergence of the whole procedure, provided the perturbation is small enough. Such an approach, that goes back to the proof of KAM theorem using classical series expansions, has been showed to be in a very good position for the applications to realistic models (see, e.g., [24], [19], [27]), while the proof adopted in 45] is based on a fast convergence scheme of quadratic type (a so called Newton-like method).

Hereafter, we abandon mainly theoretical arguments; in particular, in section 5, our normal form algorithm will be used in practice, for explicitly constructing elliptic tori in the FPU model. In particular, since we are not going to apply the theorem, we do not care about the threshold value $\varepsilon^{\star}$, that is extremely low. Actually, it is well known that computer assisted proofs can ensure the existence of invariant KAM surfaces for values of the small parameter that are very close to its upper limit; moreover, there are versions of this kind of rigorous results that are suitable for elliptic tori (see [16] and [37]). Since we just want to extract from the proof of theorem 3.1 a key element, i.e., the widely described constructive algorithm, and use it in practice, we will limit ourselves to check for its convergence with a numerical criterion that will be shown later.

## 4 Frequency Analysis and the search for elliptic tori

Several powerful numerical approaches have been designed in order to explore the dynamics of Hamiltonian systems. In the present context, we think that the so called Frequency Analysis Method (hereafter FA, see [35] for an introduction) is in a better position, because it can be efficiently adapted to the search for elliptic tori (see, e.g., [42] and [12] for applications to problems of interest in Astronomy).

### 4.1 FA as a tool to investigate the dynamics of FPU chains

Following the usual prescriptions provided in the framework of FA, we study complex signals, that are formed by pairs of real canonical variables of cartesian type, namely we consider the motion laws $t \mapsto Y_{j}(t)+\mathrm{i} X_{j}(t)$, where $j=1, \ldots, N-1$ and variables $(\boldsymbol{Y}, \boldsymbol{X})$ are related to the normal modes of oscillation (recall formulæ (2)-(4)). Actually, we focus

[^5]on the following decomposition of the signals:
\[

$$
\begin{equation*}
Y_{j}(t)+\mathrm{i} X_{j}(t) \simeq \sum_{s=1}^{\mathcal{N}_{C}} A_{j, s} e^{\mathrm{i}\left(v_{j, s} t+\phi_{j, s}\right)} \tag{46}
\end{equation*}
$$

\]

where $\mathcal{N}_{C}$ is the number of the components we want to consider, while $A_{j, s} \in \mathbf{R}_{+}$and $\phi_{j, s} \in[0,2 \pi)$ are the amplitude and the initial phase of the $s$-th component of the $j$-th signal, respectively. Moreover, $v_{j, s}$ is a local maximum point of the "tuning function" related to the fundamental integral of the frequency analysis method, i.e.

$$
\begin{equation*}
v \mapsto \mathcal{T}_{j}(v)=\frac{1}{T}\left|\int_{0}^{T} \mathrm{~d} t\left[Y_{j}(t)+\mathrm{i} X_{j}(t)\right] e^{-\mathrm{i} v t} \mathcal{W}(t)\right| \tag{47}
\end{equation*}
$$

Here, $\mathcal{W}$ is a suitable weight function such that $\int_{0}^{T} \mathrm{~d} t \mathcal{W}(t)=T$. In all the computations, we have adopted a so-called "Hanning filter" adapted to the interval of time $[0, T]$, i.e., we have set $\mathcal{W}(t)=1+\cos [\pi(2 t / T-1)]$. Since we do not know exactly the motion laws $t \mapsto Y_{j}(t)+\mathrm{i} X_{j}(t)$, in practice we obviously limit ourselves to deal with a discretization of the signal made by sets of finite points computed on a regular grid in the interval $[0, T]$, i.e., for $t=i \Delta$, where $i=0, \ldots, \mathcal{N}_{P}$ and the sampling time is $\Delta=T / \mathcal{N}_{P}$.

In order to produce the discretized signals $\left\{\left(Y_{j}(i \Delta), X_{j}(i \Delta)\right)\right\}_{i=0}^{\mathcal{N}_{P}}$ to be analysed, we have preliminarly used the symplectic integrator of type $\mathcal{S B A} \mathcal{B}_{3 C}$, which is described in [36]. The splitting of the initial Hamiltonian (1) in two integrable parts $\mathcal{A}$ and $\mathcal{B}$ (that is requested for implementing such a kind of numerical integrator) is made so that the quadratic part is stored in $\mathcal{A}$, i.e., $\left.\mathcal{A}=\frac{1}{2} \sum_{j=0}^{N-1}\left[y_{j}^{2}+\left(x_{j+1}-x_{j}\right)^{2}\right)\right]$, and $\mathcal{B}=\mathcal{H}-\mathcal{A}$. In order to reduce the effects due to the accumulation of the round-off errors, we have found convenient to perform the numerical integrations by implementing the long double floating-point arithmetics in our $\mathbf{C}$ code, while we used the standard double precision variables in all the $\mathbf{C}$ programs devoted to the FA.

As a first application of FA, let us consider a very classical type of numerical exploration about the dynamical behavior of the FPU chains. Indeed, we are going to consider initial conditions where just the first oscillation mode is excited, in agreement with what was done in the original experiments about the FPU model; therefore, we study the chaoticity of the motions starting from those initial data. Actually, we set

$$
\begin{equation*}
x_{l}(0)=\sqrt{\frac{2}{N}} \frac{A}{\sqrt{\nu_{1}}} \sin \left(\frac{l \pi}{N}\right), \quad y_{l}(0)=0 \tag{48}
\end{equation*}
$$

$\forall l=1, \ldots, N-1$. Thus, the corresponding energy level $E=\mathcal{H}(\boldsymbol{y}(0), \boldsymbol{x}(0))$ depends uniquely on the amplitude $A$ of the initial excitation of the first mode, since we keep $\alpha$ and $\beta$ as fixed parameters (recall definition (11). We numerically determine the global maximum point $\omega_{f ; 1}$ of the function $v \mapsto \mathcal{T}_{1}(v)$ defined in (47). Of course, the calculation of the integral in (47) is made by a quadrature method that is applied after having replaced the function $t \mapsto\left(Y_{1}(t), X_{1}(t)\right)$ (where $\left.t \in[0, t]\right)$ with the finite sequence $\left\{\left(Y_{1}(i \Delta), X_{1}(i \Delta)\right)\right\}_{i=0}^{T / \Delta}$. Therefore, this same procedure is repeated so as to determine


Figure 1: Variation of the first angular velocity as a function of the specific energy in FPU chains. All the considered motions start from initial conditions where just the first normal mode is excited. The results are plotted in red, green and blue when they refer to $N=4,8$ and 16 , resp. On the left, the $\alpha$-model is studied with $\alpha=1 / 4$; on the right, the $\beta$-model is analysed with $\beta=1 / 4$.
the global maximum point of the function defined in (47) when the integral appearing in that formula is performed on the interval $[T, 2 T]$ instead of $[0, T]$. In Figure [1, the variation $\left|\Delta \omega_{f ; 1}\right|$ of the maximum points over those two subsequent time windows is reported as a function of the specific energy $E_{S}=E / N$ for different number of particles, namely $N=4,8,16$, for both a $\alpha$-model and a $\beta$ one (left and right boxes, resp.). All the signals analysed to produce the plots in Figure 1 are such that the sampling time is $\Delta=0.5$, while the total extent of the time interval covered by the signals is $2 T=2^{17}=131072$. All the preliminary numerical integrations (that have been necessary to create the signals) have been performed by setting the timestep at 0.125 .

Following the usual prescriptions about FA, we can assume that the dynamics is stable and the orbits are on invariant tori, if the variations of the angular velocities are at most a few order of magnitude larger than the round-off error, else the motions have to be considered as chaotic. Let us focus on the plots in the left box of Figure 1. First, we remark that it was well expected that the values of $\left|\Delta \omega_{f ; 1}\right|$ are increasing with respect the number $N$ of nodes: on one hand, it is known that the variation of the local maxima of the function (47) is $\mathcal{O}\left(1 / T^{4}\right)$ when quasi-periodic motions are studied and the Hanning filter is adopted (see [35]), on the other hand, $\omega_{f ; 1} \simeq \nu_{1}$ that is defined in (3), then the fundamental period corresponding to the main oscillation (of the first normal mode) is $\mathcal{O}(N)$. Therefore, in relative terms, when $N$ is increased, larger times are needed to let the angular velocities relax to their final values; this explains why the variation of $\left|\Delta \omega_{f ; 1}\right|$ gets bigger with $N$, by keeping fixed the total time $2 T$. By comparison with Figure 6 of [7], we see that our results are in good agreement with those obtained by computing the maximal Lyapunov exponent for the $\alpha$-model: the transition to the instability occurs for values of the specific energy $E_{S}$ larger than 0.2 and 0.02 , when $N=8$ and $N=16$, respectively. However, the FA highlights also the weak chaoticity regions likely due to the crossing of resonances. This occurs for values of the specific energy that are smaller
than the previous ones and, then, such that the diffusion in the phase space is not very remarkable. By comparing the plots in the right box of Figure 1 with those in the left one, it is quite evident that both the crossing of resonances and the ultimate transition to chaos is sharper in the case of the $\beta$-model. This was easy to expect, due to the fact that the perturbing terms are quartic instead of being cubic.

### 4.2 FA to locate invariant elliptic tori

Computational algorithms that are commonly used into the framework of FA can be efficiently adapted for the localization of elliptic tori. For such a purpose, we basically follow the approach described in [43], where an attempt to provide a theoretical foundation of such a method is also made. The technique that iteratively refines the approximation of a generic $n_{1}$-dimensional elliptic torus is the keystone of this procedure. It can be summarized as follows, by referring to the specific case of the FPU chains, just to fix the ideas. Before entering the details of the procedure, we emphasize that it is tailored for lower dimensional tori with a purely elliptic character in their transverse oscillations; the method is expected to not be working in the case of hyperbolic dynamics in at least one degree of freedom.
(a) First, let us consider an initial condition $(\boldsymbol{Y}(0), \boldsymbol{X}(0))$ eventually close enough to an invariant elliptic torus.
(b) By using a suitable numerical integrator, we produce $N-1$ discretized signals $\left\{\left(Y_{j}(i \Delta), X_{j}(i \Delta)\right)\right\}_{i=0}^{T / \Delta}, N$ being the number of nodes, $\Delta$ the sampling time, $T$ the width of the total time interval of integration and $j=1, \ldots, N-1$.
(c) For all signals, we perform the quasi-periodic decompositions described in the previous subsections, i.e., $Y_{j}(t)+\mathrm{i} X_{j}(t) \simeq \sum_{s=1}^{\mathcal{N}_{C}} A_{j, s} e^{\mathrm{i}\left(v_{j, s} t+\phi_{j, s}\right)}, \mathcal{N}_{C}$ being the fixed number of summands. By eventually reordering the components, we can assume that the amplitudes are decreasing with respect to the index $s$, i.e., $A_{j, 1} \geq A_{j, 2} \geq \ldots \geq A_{j, \mathcal{N}_{C}}$ $\forall j=1, \ldots, N-1$.
(d) By proceeding as prescribed at points (d1)-(d4) below, we determine a $n_{1}$-dimensional vector $\boldsymbol{\omega}_{f}$, whose components are the so called fundamental angular velocities.
(d1) As in the previous subsection, we set $\omega_{f ; 1}=v_{1,1}$ that is the global maximum point of the function $\mathcal{T}_{1}(v)$ defined in (47).
(d2) For each $j=2, \ldots, N-1$, we select the largest summand (corresponding to a value, say, $\bar{s}_{j}$ of the index $s$ ) that cannot be explained as a linear combination with integer coefficients of the current components of the fundamental angular velocities. This means that $\forall 1 \leq s \leq \bar{s}_{j}-1$ there is a $\overline{\boldsymbol{k}} \in \mathbb{Z}^{j-1}$ such that $\left|v_{j, s}-\sum_{i=1}^{j-1} \bar{k}_{i} \omega_{i}\right| \leq \varepsilon_{\text {tol }}$ with $|\overline{\boldsymbol{k}}| \leq K_{M}$, while $\left|v_{j, \bar{s}_{j}}-\sum_{i=1}^{j-1} k_{i} \omega_{i}\right|>\varepsilon_{\text {tol }}$ $\forall \boldsymbol{k} \in \mathbb{Z}^{j-1}$ such that $|\boldsymbol{k}| \leq K_{M}$. The parameters $\varepsilon_{\text {tol }}$ and $K_{M}$ are conveniently fixed so as to let the whole procedure converge (see the discussion at the end of the present section).
(d3) For each $j=2, \ldots, N-1$, we extract the new $j$-th component of the fundamental angular velocities vector from the value $v_{j, \bar{s}_{j}}$ selected as explained at the previous point (d2).
(d4) When each $j$-th component of the fundamental angular velocities is expected to be approximately equal to a value $\nu_{j}$ that is known a priori (for instance, refer to equation (3) in the case of FPU chains), then the choice can be adapted on such a purpose. In such a situation, first, $\forall j=1, \ldots, N-1$ it is convenient to look for the integer coefficients $\bar{k}_{1}, \ldots, \bar{k}_{j}$ such that

$$
\left|\bar{k}_{j} v_{j, \bar{s}_{j}}+\sum_{i=1}^{j-1} \bar{k}_{i} \omega_{f ; i}-\nu_{j}\right|=\min _{\substack{k \in \mathbb{Z} j \\|k| \leq K_{M}, k_{j} \neq 0}}\left|k_{j} v_{j, \bar{s}_{j}}+\sum_{i=1}^{j-1} k_{i} \omega_{f ; i}-\nu_{j}\right|,
$$

therefore, we put $\omega_{f ; j}=\bar{k}_{j} v_{j, \bar{s}_{j}}+\sum_{i=1}^{j-1} \bar{k}_{i} \omega_{f ; i}$. In words, by linear combinations with integer coefficents, we iteratively define a set of fundamental angular velocities that are close as much as possible to the expected values.
(e) We modify the initial conditions as follows. Let us introduce $(\tilde{\boldsymbol{Y}}(0), \tilde{\boldsymbol{X}}(0))$ so that $\tilde{Y}_{j}(0)+\mathrm{i} \tilde{X}_{j}(0)=\sum_{s=1}^{\mathcal{N}_{C}} \tilde{A}_{j, s} e^{\mathrm{i} \phi_{j, s}} \forall j=1, \ldots, N-1$, where $\tilde{A}_{j, s}=0$ when $\mid v_{j, s}-$ $\boldsymbol{k} \cdot \boldsymbol{\omega}_{f} \mid>\varepsilon_{\text {tol }} \forall \boldsymbol{k} \in \mathbb{Z}^{n_{1}}$ such that $|\boldsymbol{k}| \leq K_{M}$, otherwise we simply keep $\tilde{A}_{j, s}=A_{j, s}$. Once again, the previous formulæ can be rephrased in a more clear way as follows: the initial conditions are chosen in order to suppress all the components related to angular velocities that are not a linear combination (by integer coefficients) of the fundamental vector $\boldsymbol{\omega}_{f}$, which has been determined as explained at the previous points (d1)-(d4).
(f) We redefine the initial conditions so that $(\boldsymbol{Y}(0), \boldsymbol{X}(0))=(\tilde{\boldsymbol{Y}}(0), \tilde{\boldsymbol{X}}(0))$. The execution of the algorithm is positively concluded when the signals are close enough to their quasi-periodic approximations based on the fundamental vector of angular velocities $\boldsymbol{\omega}_{f}$, i.e., we check if the following condition is satisfied $\forall j=1, \ldots, N-1$ :

$$
\frac{\max _{\left\{t_{i}: t_{i}=i \Delta, \forall 0 \leq i \leq T / \Delta\right\}}\left|Y_{j}\left(t_{i}\right)+\mathrm{i} X_{j}\left(t_{i}\right)-\sum_{s=1}^{\mathcal{N}_{C}} \tilde{A}_{j, s} e^{\mathrm{i}\left(v_{j, s} t_{i}+\phi_{j, s}\right)}\right|}{\left|A_{1,1}\right|} \leq \mu_{t o l}
$$

the execution of the algorithm is ended by setting a flag variable on an error status if we reached a prefixed maximum $\sqrt{6}$ value on the number of times the points (b)-(f) are iterated; otherwise, we restart from point (b).

A good choice of the parameters $\varepsilon_{t o l}, \mu_{t o l}, K_{M}$ and $\mathcal{N}_{C}$ is crucial for avoiding infinite loops. On one hand, by adopting a small value of the tolerance errors $\varepsilon_{t o l}$ and $\mu_{t o l}$, the algorithm is forced to look for very accurate quasi-periodic approximations. On the other hand, there are so many calculations involved (during both the numerical integrations

[^6]and the FA) that the accumulation of the round-off errors is relevant; therefore, if we ask for too much precision, for instance, the stop condition described at point (f) above could never be fulfilled. We are obviously interested in keeping a large enough value of the upper bound $K_{M}$ on the $l_{1}-$ norm of the Fourier harmonics $\boldsymbol{k}$. However, because of the Fourier decay of the coefficients in the Hamiltonian expansions, terms related to harmonics having a big norm $|\boldsymbol{k}|$ are expected to play a negligible role; therefore, it is essential to adopt a small enough value for the upper bound $K_{M}$ in order to prevent the wrong choice of spurious terms in the definitions introduced at the previous points (d2), (d4) and (e). The pros and cons concerning the choice of the number of components are similar to those regarding the tolerance errors: small values of $\mathcal{N}_{C}$ decrease the accuracy of the selected approximations at the end of the whole procedure, but too large values of $\mathcal{N}_{C}$ increase dramatically the computational time and could prevent its convergence because of the fact that components related to very small amplitudes are strongly affected by numerical errors.

The applications of this method, based on FA and designed for the localization of elliptic tori, are discussed in the next section. For all those numerical computations, we have found that it is convenient to set $\varepsilon_{t o l}=10^{-12}, \mu_{\text {tol }}=2 \times 10^{-6}, K_{M}=20$ and $\mathcal{N}_{C}=25$. Moreover, we will keep fixed the parameters concerning also the preliminary symplectic integration (that are the timestep, $\Delta$ and $T$ ), whose values are reported in the comments to Figure 1 in the previous subsection 4.1.

## 5 Semi-analytic approach vs. numerical explorations

The approach described in section 3 can be implemented in a programming code. For such a purpose, we have used X@óvoৎ, a software package especially designed for doing computer algebra manipulations into the framework of Hamiltonian perturbation theory (see [28] for an introduction to its main concepts). In this section we will discuss the results of the application of the algorithm constructing elliptic tori for the FPU chains; moreover, we will compare these results with the technique for doing numerical explorations that has been previously described in section 4 .

### 5.1 Applications to the case of two-dimensional elliptic tori

Let us start by studying a non-trivial case that is as simple as possible. We focus on a FPU chain having $N+1$ particles, with $N=4$. This means that the system has $N-1=3$ degrees of freedom, which is the minimum among the non-integrable systems we can consider, when ${ }^{7} N=2^{m}$ with $m \in \mathbb{N}$. First, we begin studying 2D elliptic tori.

As it has been widely stressed in section 2, all the procedure constructing the normal form for a $n_{1}$-dimensional elliptic torus basically depends on the preliminar translation vector $\boldsymbol{I}^{\star} \in \mathbb{R}_{+}^{n_{1}}$, which also rules the size of the perturbation. Moreover, looking at the

[^7]

Figure 2: Numerical check of the convergence of the algorithm constructing the normal form related to an elliptic torus for the FPU model: we plot the norm of the generating functions at each normalization step $r$. The symbols $+, \times, \triangle, *$ and $\square$ refer to the norm of the generating functions $\chi_{0}^{(r)}, \chi_{1}^{(r)}, X_{2}^{(r)}, Y_{2}^{(r)}$ and the diagonalizing transformation $\mathfrak{D}^{(r)}$. The norm is calculated by adding up the absolute value of all the coefficients appearing in the expansion of each generating function or the linear change of coordinates defined by $\mathfrak{D}^{(r)}$, for what concerns the latter case.
description of the algorithm made in section 3, it is expected to converge if and only if the generating functions decrease in a geometrical way. Actually, an automatic criterion to decide about the practical usefulness of the normal form expansions can be given by monitoring the behavior of the norms of the generating functions (see, e.g., [51). In order to fix the ideas, it is convenient to see how such an approach can be adapted to the present framework.

The results summarized in Figure 2 refer to the $\alpha$-model with $\alpha=0.25$, where the components of the initial translation $\left(I_{1}^{\star}, I_{2}^{\star}\right)$ are fixed as $(0.0001,0.0001)$ and $(1,1)$ for what concerns the left box and the right one, respectively. Actually, the construction of the normal forms related to both those translation vectors are reported also in Figure 3: they correspond to the symbols plotted in the centre of the left box and in its topright corner, resp. For the sake of simplicity, we have decided to truncate the initial expansions of $H^{(0)}$ in (10) so that the total degree $\ell$ of the classes of functions $\mathcal{P}_{\ell}^{s K}$ is not greater than 8. Such an upper bound is automatically satisfied by any Hamiltonian $H^{(r)}$, because the Poisson brackets with the generating functions do not increase the total degree in the square root of the actions. In order to deal with all the computations required by the algorithm in a reasonably short amount of time, we have limited ourselves to perform just the first 12 normalization steps. With these parameters, the computation of a single 2D elliptic torus requires 77 seconds of CPU time on a computer equipped with a Xeon 18-Core 6154 (3.0 Ghz) processor. Accordingly, we have decided to truncate the Fourier expansions so as to contain all the terms of the final generating functions $X_{2}^{(12)}, Y_{2}^{(12)} \in \mathcal{P}_{2}^{12 K}$. With these last settings, the computational algorithm described in section 3 is completely defined. In the semi-log plot reported in the left box

[^8]

Figure 3: On the left, success/failure of our procedure constructing 2-dimensional elliptic tori for different translation vectors, the symbol is plotted when the algorithm is considered to be convergent, because it satisfies the rules (A)-(B), the $\triangle$ otherwise. On the right, plot of the ratio between the energy of the first two normal modes (see the text for a more precise definition) for motions of the same type as those studied in the first work about FPU chains. The curve appearing in the right box has been reported also on the left as a solid line.
of Figure 2, the decrease of the norms of $\chi_{0}^{(r)}, \chi_{1}^{(r)}, Y_{2}^{(r)}, X_{2}^{(r)}$ is pleasantly regular apart a few unavoidable fluctuations. Such generating functions are listed according to their size in an increasing order, that is coherent with the expected accumulation of small divisors, as it has been described in [26] and in the reference reported in footnote ${ }^{5}$. Since the linear canonical transformation $\mathfrak{D}^{(r)}$ is determined with the aim to remove the same kind of terms for any normalization step $r$, there is not any small divisor entering the solution of equation (36). Therefore, it is natural to expect that the norm corresponding to $\mathfrak{D}^{(r)}$ is smaller than any other. This explains also why its plot in the left box of Figure 2 is the most irregular, since it is very exposed to the numerical effects due to round-off errors, because of its smallness. In short, the generating functions introduced at the 12th normalization step are 6 or 7 order of magnitude smaller than those defined at the beginning of the algorithm and their decrease is rather regular. This strongly suggests that the sequence of the corresponding canonical transformations is more and more close to the identity and their composition is converging to the change of coordinates bringing the Hamiltonian to the wanted normal form. On the other hand, the behavior of the norms appearing in the right box of Figure 2 is not converging at all. Looking at the different values of the vectors $\boldsymbol{I}^{\star}$ that are corresponding to the left and right boxes, respectively, and keeping in mind the discussion in section 2, the following explanation of the observed phenomena looks quite natural: in the case of the right box, the initial translation is so large that the induced perturbation prevent the existence of the wanted elliptic torus. This does not occur when $\boldsymbol{I}^{\star}$ and the corresponding perturbation are small enough, as for the plot in the left box.

We are now ready to formulate some practical rules, in order to establish when we can assume that the procedure is properly working or not. We consider that the normalization algorithm is converging if both the following conditions are satisfied:
(A) $\frac{\left\|X_{2}^{(r)}\right\|}{\left\|X_{2}^{(1)}\right\|+\left\|X_{2}^{(2)}\right\|}<(0.95)^{r-1}$ for $r>2$;
(B) $\frac{\left\|X_{2}^{(\bar{r})}\right\|}{\left\|X_{2}^{(1)}\right\|}<\frac{1}{1000}$, where $\bar{r}=12$ is the last normalization step performed.

Let us stress that the geometrical decrease of the generating functions (when it occurs) is usually so sharp that a reasonable modification of the values appearing in the r.h.s. of the inequalities included in the rules (A)-(B) above does not considerably affect the results. We have decided to refer the tests to the sequence $X_{2}^{(r)}$ because their norms are usually greater than the corresponding ones for all other generating functions, as already remarked in the comments about Figure 2, Let us recall that the functional norm is given by the sum of the absolute values of all the coefficients appearing in the (finite) expansion. We have considered many different initial translations, so that the components of every vector $\boldsymbol{I}^{\star}$ are located in correspondence with the points of a regular 2D grid of $\log -\log$ type. For each initial translation $\boldsymbol{I}^{\star}$ making part of such a finite set, we have performed the normalization procedure, at the end of which we have applied the practical criterion described above so as to discriminate between the convergence and the divergence of that single application of the constructive algorithm. The successes and failures of this explicit method are summarized in the left box of Figure 3. Actually, the location of the points making part of the regular grid is rescaled in terms of the energy of the quadratic approximation based on the normal modes oscillations; this means that we put the values of $\nu_{1} I_{1}^{\star}$ and $\nu_{2} I_{2}^{\star}$ in abscissas and ordinates, resp. This has been done for an easier comparison with the results reported in the right box of the same figure.

The plot in the left box of Figure 3 has been realized by considering values of the energy of the first two normal modes in a wide range from $10^{-8}$ to 5 ; the normalization algorithm looks to be convergent often, even when both those values are approximately equal to 1 . This is certainly a remarkably good performance, because it is not far from the upper bound on the specific energy over the which quasi-periodic motions are no more detectable, by exploring the dynamics with the FA method (see Figure (1). However, a more careful analysis of the results highlights that the motions studied in the present section are quite different with respect to those of the previous one. Let us now focus on the same initial conditions of "semi-sinusoidal" type (48) that have been considered in [15]; then, we can numerically integrate the equations of motion and calculate the quasi-periodic decomposition (46). In the right box of Figure 3, we report the ratio $\left(\nu_{2,1} A_{2,1}^{2}\right) /\left(\nu_{1,1} A_{1,1}^{2}\right)$ of the energies of the maximal quasiperiodic components related to the second and the first normal mode of oscillation, respectively. The plot of such a quantity, as a function of the total energy, clearly shows that the oscillations with angular velocities close to that of the second normal mode get considerably excited just for large enough initial amplitudes of the first mode. On the other hand, the more the energy is low, the more the sharing between normal modes is inhibited (as it is well known, see, e.g., (4). For values of the total energy that are relatively small, it is reasonable to assume that the two different ways we adopted to define the energies related to the normal modes are approximately equal. Therefore, it makes sense to draw the same curve that appears in the right box of Figure 3 also on the left, after having reported the numerical data from the former to the latter in the corresponding units. Such a graphical operation makes
evident that the motions generated by initial conditions of "semi-sinusoidal" type are not in the domain of applicability of our algorithm when it is used to explicitly construct 2D elliptic tori. Furthermore, the curve we have drawn in the left box looks parallel and well separated with respect to a sort of natural border of the domain of applicability of the method. This remark and the inhibition effect on the sharing of the energy (that is initially exciting just the first mode) strongly suggest that the kind of motions originally studied in the FPU chains are substantially more similar to those on 1D elliptic tori. This is the reason why, hereafter, such invariant objects will play a prominent role in our investigations, in spite of the fact that they are simply linearly stable periodic orbits.

### 5.2 Applications to the case of one-dimensional elliptic tori

One-dimensional elliptic tori in FPU chains can be found, by using a continuation method based on the algorithm explained in subsection 4.2. Let us recall that in the scientific literature one can find several different numerical techniques that have been shown to be very efficient in searching for periodic orbits. However, we prefer a method based on FA that is designed to work properly just for (quasi)periodic motions lying on lower dimensional tori having an elliptic character in the transverse degrees of freedom. This is perfectly coherent with the type of normal forms we aim to construct. In this context, let us also mention that there are also methods that are mainly analytic by their nature. For instance, an approach based on the Reeb's theorem requires to perform a preliminary averaging procedure on the angular part of the canonical coordinates; then, non-degenerate equilibrium points of the reduced (and averaged) system are proved to correspond to periodic solutions of the original one. Moreover, they also inherit their eventual linear stability from the corresponding equilibrium points (see [46] and [34] for a recent application to rotating Hénon-Heiles systems).
In order to describe in a more detailed way the continuation method based on FA, the procedure can be summarized as it follows.

- The numerical procedure of subsection 4.2 is performed from point (a) to (f), starting from an initial condition $(\boldsymbol{Y}(0), \boldsymbol{X}(0))$ of type (48), where $A$ is so small that such an algorithm is able to easily find the corresponding 1D elliptic torus. At the end of this stage, we compute the value of the energy $\tilde{E}=\mathcal{H}(\tilde{\boldsymbol{Y}}(0), \tilde{\boldsymbol{X}}(0))$, where the Hamiltonian $\mathcal{H}$ is defined in (4) and $(\tilde{\boldsymbol{Y}}(0), \tilde{\boldsymbol{X}}(0))$ is defined at the last execution of point (e).
- Every time the conclusion at point (f) is positive, the corresponding value of the energy $\tilde{E}$ is updated in the same way we described just above. Therefore, the whole procedure of subsection 4.2 is restarted from point (a), right after having redefined $(\underset{\boldsymbol{Y}}{\boldsymbol{Y}}(0), \boldsymbol{X}(0))=(1+\zeta)(\tilde{\boldsymbol{Y}}(0), \tilde{\boldsymbol{X}}(0))$, where $\zeta$ is a small parameter (see below) and $(\tilde{\boldsymbol{Y}}(0), \tilde{\boldsymbol{X}}(0))$ is the initial conditions vector corresponding to the last computation of $\tilde{E}$, which is also the last numerical determination of the elliptic torus.
- If the prescription at point (f) returns a negative conclusion and $\zeta$ is still larger than a prefixed minimum value $\bar{\zeta}$, then $\zeta$ is reduced by half. Therefore, the procedure of
subsection 4.2 is restarted from point (a), after having redefined the initial conditions in the same way we have described just above.
- If the small parameter $\zeta \leq \bar{\zeta}$, then the whole continuation method is finally stopped.


Figure 4: Comparisons between 1D tori constructed by our normal form algorithm (dots denoted by $\bullet$ ) and those detected by using the continuation method (solid lines), for both the $\alpha$-model and the $\beta$ one (on the left / right, resp.) in the cases with $N=4,8$ (top / bottom, resp.).

The solid lines appearing in Figure 4 are obtained by applying this continuation method. In such numerical experiments, the value of $\zeta$ has been initially set at 0.1 , while we put the threshold value $\bar{\zeta}=0.00025$. Let us remark that in abscissas we report the values of the specific energy $E_{S}$ instead of the total one. By comparing Figures 4 and 1, one can appreciate that for the $\beta$-model the breakdown threshold with respect to $E_{s}$ for the invariant manifold corresponding to 1 D elliptic tori is approximately the same as that for the motions starting from the same type of initial conditions originally considered in [15] and for which the transition to chaoticity looks very sharp. For what concerns the $\alpha$-model, 1D elliptic tori appear to be more robust also because they seem not being affected from the transitions through resonances, that apparently exerts a strong impact in the numerical experiments summarized in Figure 1 .


Figure 5: Study of the transition to chaoticity in the $\beta$-model of the FPU chain with $N=4$ (and $\beta=1 / 4$ ). On the left, behavior of the eigenvalues $\lambda_{ \pm j}=e^{ \pm i \theta_{j}}$ related to any monodromy matrix, that has been computed for each of the 1D elliptic tori considered in the top-right box of Figure 4. More precisely, the values of $\pm \theta_{j}$ for $j=1,2,3$ are plotted as a function of the specific energy of the corresponding periodic orbit. On the right, the variation of the frequency is reported again as a function of the specific energy. Each plotted value refers to a motion starting from an initial condition that is slightly modified with respect to the corresponding periodic orbit considered in the left box. See the text for more details.

In Figure 4, there are also some points that have been plotted with bold dots, so that they can be easily seen. They refer to the applications of our normalization algorithm, which has been described in section 3 and has been adapted to construct 1D elliptic tori. Actually, we have selected a finite set of values for the initial translation $I_{1}^{\star} \in \mathbb{R}_{+}$, in such a way that the sequence $\log I_{1}^{\star}$ makes a regular grid. For each single element of such a set, we have performed our whole normalization procedure starting from that value of the initial translation $I_{1}^{\star}$. Let us now focus on the comparison between expansions (11) and (15). When the algorithm seems to be converging from a numerical point of view, the limit values $\omega_{1}^{(\infty)}$ and $\mathcal{E}^{(\infty)}$, which are quantities characterizing the wanted 1D elliptic torus, are expected to be well approximated by $\omega_{1}^{(\bar{r})}$ and $\mathcal{E}^{(\bar{r})}$, resp., which are explicitly computed during each construction of the normal form up to the step $\bar{r}$. Therefore, also our semi-analytic algorithm provides the value of both the angular velocity and the energy level which are directly comparable with those given by the continuation method based on FA, that has been previously described in this same section. Looking at Figure [4, one can appreciate that the agreement between these two computational methods is excellent, because the superposition between the plots is nearly perfect in all the four cases we have considered.

The extension of the range for relatively large values of the energy is definitely the hardest challenge for what concerns the applicability of our algorithm. However, let us check that the numerical method we used for locating elliptic tori does not introduce any artificial simplification. In other words, we want to ensure that the family of periodic curves cannot be easily prolonged by continuation over what we call breakdown threshold. This is because of a very concrete dynamical obstruction: the transition to a resonant
chaotic region. In Figure 5, we consider just the case of the $\beta$-model with $N=4$. Let us remark that the monodromy matrix related to a linearly stable periodic orbit can be easily calculated (see, e.g., [30] for a gentle introduction to these concepts). This can be done by referring to the corresponding normal form approximation of a 1D elliptic torus, that we can write as $\omega_{1} P_{1}+\sum_{j=1}^{2} \Omega_{j}\left(\Xi_{j}^{2}+\Theta_{j}^{2}\right) / 2$, in analogy with the statement of theorem 3.1. In fact, its six eigenvalues are such that $\lambda_{ \pm j}=\exp \left( \pm i 2 \pi \Omega_{j} / \omega_{1}\right)$ for $j=1,2$ and $\lambda_{ \pm 3}=1$, being the period equal to $2 \pi / \omega_{1}$. For what concerns the graphs reported on the left of Figure [5 let us emphasize that we focus on all the 1D elliptic tori that have been numerically found, in order to draw the solid curve appearing in the top-right box of Figure 4. One can easily deduce that the ratio $\Omega_{1} / \omega_{1}$ is going to converg $\epsilon^{9}$ to 2 in correspondence with the breakdown threshold, whose characteristic value can be expressed in terms of the specific energy as $E_{S} \simeq 1.25$. In the box on the right of Figure 5 we consider a slight modification of every initial condition related to a 1D elliptic torus, whose the corresponding monodromy matrix has been studied in the plot on the left. In fact, the values of the canonical coordinates are multiplied by a common factor 1.01 for all these initial conditions; then, the corresponding variation of the first angular velocity $\left|\Delta \omega_{f ; 1}\right|$ is computed (according to the description in subsection 4.1) and it is plotted in the graph included in the box on the right. The sudden jump by eight orders of magnitude that is corresponding to the last evaluation about the variation of the fundamental frequency, jointly with the comments concerning the eigenvalues of the monodromy matrix, allows us to draw the following conclusion. 1D elliptic tori are entering the chaotic zone surrounding the $2: 1$ resonance between the angular velocities $\omega_{1}$ and $\Omega_{1}$, when the specific energy is converging to the breakdown value $E_{s} \simeq 1.25$. This explains why the one-parameter family of 1D elliptic tori cannot be further prolonged for larger values of $E_{S}$, by preserving its continuity. The same situation has been observed in the cases with the $\alpha$-model instead of the $\beta$-one and/or $N=8$, where the patterns of the graphs representing the eigenvalues of the monodromy matrix are eventually more complicated. For the sake of brevity, the figures corresponding to these other cases are omitted.

Looking at the comparison between the boxes on top of Figure 4 with those on bottom, it is evident that our algorithm works better in the case $N=4$ with respect to $N=8$. In general, it is very natural to expect that, for our semi-analytic procedure, the situation is more difficult when the number of degrees of freedom is increased. In particular, we had to reduce a few parameters ruling the truncations of the Hamiltonians, in order to make reasonable both the computational time and the usage of memory in the case with $N=8$. In fact, while for $N=4$ these parameters are exactly the same as those described in subsection 5.1, when $N=8$ we have decided to truncate the expansions of $H^{(r)}$ so that the total degree in the square root of the actions is not greater than 4 , for

[^9]any normalization step $r$ ranging from 0 to $\bar{r}$, where this upper bound on $r$ has been decreased to $\bar{r}=8$ (in the case with $N=4$, it was fixed to $\bar{r}=12$ ). Since we have kept the same practical criterion to decide if the algorithm is converging or not, this means that rule (B) of subsection 5.1 has become more restrictive, because the decrease of the generating function must be faster to compensate the reduction on the number of normalization steps. For what concerns the CPU usage, the computation of a single 1D elliptic torus requires around 33 seconds when $N=4$ and 112 seconds in the case with $N=8$. Of course, such a comparison is rather rough, because the truncation parameters have been chosen in a different way with the aim of accomplishing all the computational tasks in a reasonable total time, as we have discussed above. Both the truncation rules and the computational resources are exactly the same as in the case of the 2D elliptic tori that has been described in the previous subsection. Here, the time needed by the computation of a single invariant torus is about $43 \%$ with respect to the amount that has been reported during the discussion of Figure 2. The reason of such a gain can be easily understood by looking at the expansion of a generating function; to fix the ideas, let us focus on equation (26). For instance, since the number of summands making part of the generating function $\chi_{1}^{(r)}$ is increased by a factor $r$ passing from $n_{1}=1$ and $n_{2}=2$ to $n_{1}=2$ and $n_{2}=1$ (let us recall that we fixed $K=2$ when dealing with FPU models), one immediately realizes that the whole normalization procedure is based on Poisson brackets that involve many more terms in the case of the explicit construction of 2D elliptic tori. This explains why more computational time is needed in that case.

Apparently, our approach works better when it is applied to the $\beta$-models instead of $\alpha$ ones; this is highlighted by the comparison between the boxes on the left of Figure 4 with those on the right. In our opinion, this phenomenon is due to two main reasons. First, the transition to the chaoticity is much sharper in the case of a quartic perturbation instead of a cubic one; therefore, it is natural to expect that a normal form method can approach better the breakdown threshold for any invariant manifold, when the $\beta$-model is considered. Moreover, there is an accurate integrable approximation of the $\alpha$-model: the Toda lattice, because their expansions coincide up to terms of third order (see [18] and, e.g., the introduction of [47] in [20]). We think that this fact should play an important role in order to stabilize the dynamics when relatively high energies are considered, but our approach is completely blind to such an accurate approximation provided by a Toda-like Hamiltonian. This further explains why we have achieved better results in the case of the $\beta$-model.

By comparing the cases reported in Figure 4 all together, the best performance is in the $\beta$-model with $N=4$. In that case, our algorithm is able to construct the normal form for a 1D elliptic torus, when the energy is about $93 \%$ of the value corresponding to its breakdown threshold $E_{S} \simeq 1.25$, as it has been numerically computed by using the continuation method based on FA. The results for the $\beta$-model are still good for $N=8$, where our algorithm fulfills the construction of tori up to $64 \%$ of the threshold energy. On the other hand, for the $\alpha$-model the performance is good only for $N=4$, since for $N=8$ we are really far from the numerical threshold.

## 6 Stability in the neighborhood of elliptic tori

According to the discussion at the end of subsection 5.1, the orbits generated by initial conditions of "semi-sinusoidal" type are close to be 1D elliptic tori. In the present section, we enforce this concept, trying to explain the apparent stability (i.e., quasi-periodicity) of the former motions in terms of their vicinity to the latter ones. We emphasize that 1D elliptic tori are expected to be quite robust, because either the construction of the corresponding normal form is stopped (due to the eventual occurrence of resonant terms) or, if it is converging, arbitrarily small divisors cannot be produced. In order to understand this argument, it is convenient to reconsider the non-resonance condition (42), in the case the sequence of the angular velocity vector $\boldsymbol{\omega}^{(r-1)}$ is one-dimensional. Here, we aim to describe how the elliptic tori influence the dynamics in their neighborhood, by implementing a suitable Birkhoff normalization algorithm. Thus, we basically follow the ideas described in [39] with an approach similar to that fully developed in [25] and [27], where the main perturbative terms are removed from a Hamiltonian in Kolmogorov's normal form by an iterative procedure. Moreover, for those dynamical systems such a method has been also complemented with suitable final estimates about the time of effective stability. We adapt that approach in such a way to remove the terms eventually depending on the angles which appear in the expansion of the normal form Hamiltonian (11) and are $o\left(\|\boldsymbol{P}\|+\|(\boldsymbol{\Xi}, \boldsymbol{\Theta})\|^{2}\right)$.

Let us first rewrite the expansion (11) of Hamiltonian $H^{(\infty)}$, that is the normal form corresponding to an elliptic torus, in the following way:

$$
\begin{equation*}
\mathcal{H}^{(0)}(\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{J}, \boldsymbol{\varphi})=\sum_{j=1}^{n_{1}} \omega_{j}^{(\infty)} P_{j}+\sum_{j=1}^{n_{2}} \Omega_{j}^{(\infty)} J_{j}+\sum_{\ell>2} f_{\ell}^{(0)}(\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{J}, \boldsymbol{\varphi}), \tag{49}
\end{equation*}
$$

where $(\boldsymbol{J}, \boldsymbol{\varphi})$ are the action-angle variables related to the canonical coordinates $(\boldsymbol{\Xi}, \boldsymbol{\Theta})$ that are transverse to the elliptic torus; moreover, the perturbing terms $f_{\ell}^{(0)}$ are of total order $\ell$ in the square root of the actions $(\boldsymbol{P}, \boldsymbol{J})$, while they are of trigonometrical degree not greater than $\ell$ with respect to the angles $\varphi$. Indeed, such functions $f_{\ell}^{(0)}$ are given by the series $\sum_{s \geq 0} f_{\ell}^{(\infty, s)}$, where the summands $f_{\ell}^{(\infty, s)}$ appear in (11). Therefore, the angular dependence of $f_{\ell}^{(0)}$ is analytic with respect to $\boldsymbol{Q}$.

By performing a finite sequence of canonical transformations (expressed, once again, in terms of the Lie series formalism), we aim to lead the Hamiltonian to the Birkhoff normal form up to order $r$, that is

$$
\begin{equation*}
\mathcal{H}^{(r)}(\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{J}, \boldsymbol{\varphi})=\sum_{s=0}^{r} Z_{s}(\boldsymbol{P}, \boldsymbol{J})+\sum_{\ell>r} f_{\ell}^{(r)}(\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{J}, \boldsymbol{\varphi}), \tag{50}
\end{equation*}
$$

where the first part $\mathcal{Z}^{(r)}=\sum_{s=0}^{r} Z_{s}$ is the integrable part, because it depends on the actions only, while $\mathcal{R}^{(r)}=\sum_{\ell>r} f_{\ell}^{(r)}=\mathcal{O}\left(\|(\boldsymbol{P}, \boldsymbol{J})\|^{(r+3) / 2}\right)$ is the remainder.

Let us briefly describe the $r$-th normalization step in a fully generic way. For such a purpose, it is convenient to recollect the action-angle variables in $(\boldsymbol{I}, \boldsymbol{\vartheta}) \in \mathbb{R}^{n} \times \mathbb{T}^{n}$, where $\boldsymbol{I}=(\boldsymbol{J}, \boldsymbol{P})$ and $\boldsymbol{\vartheta}=(\boldsymbol{Q}, \boldsymbol{\varphi})$. We focus on the expansion of the Hamiltonian produced by
the first $r-1$ normalization steps, that is $\mathcal{H}^{(r-1)}(\boldsymbol{I}, \boldsymbol{\vartheta})=\sum_{s=0}^{r-1} Z_{s}(\boldsymbol{I})+\sum_{\ell \geq r} f_{\ell}^{(r)}(\boldsymbol{I}, \boldsymbol{\vartheta})$, where the main integrable term $Z_{0}$ can be rewritten as $\boldsymbol{\nu} \cdot \boldsymbol{I}$ with $\boldsymbol{\nu}=\left(\boldsymbol{\omega}^{(\infty)}, \boldsymbol{\Omega}^{(\infty)}\right)$. We aim to remove those perturbative terms having a total degree in the square root of the actions that is equal to $r+2$. Therefore, the $r$-th generating function is determined as the solution of the homological equation

$$
\begin{equation*}
\left\{\chi^{(r)}, \boldsymbol{\nu} \cdot \boldsymbol{I}\right\}+f_{r}^{(r-1)}=Z_{r} . \tag{51}
\end{equation*}
$$

This means that the new summand making part of the normal form is given by $Z_{r}(\boldsymbol{I})=$ $\left\langle f_{r}^{(r-1)}\right\rangle_{\vartheta}$, while the new generating function can be written as

$$
\begin{equation*}
\chi^{(r)}(\boldsymbol{I}, \boldsymbol{\vartheta})=\sum_{\substack{|\boldsymbol{m}|=r+2}} \sum_{\substack{\boldsymbol{k}=\left(k_{1}, \boldsymbol{k}_{2}\right) \in \mathbb{Z}^{n_{1}} \times \mathbb{Z}^{n_{2}} \\ \boldsymbol{k} \neq \boldsymbol{o},\left|\boldsymbol{k}_{2}\right| \leq r+2}} \boldsymbol{I}^{\boldsymbol{m} / 2}\left[-\frac{c_{\boldsymbol{k}, \boldsymbol{m}} \sin (\boldsymbol{k} \cdot \boldsymbol{\vartheta})}{\boldsymbol{k} \cdot \boldsymbol{\nu}}+\frac{d_{\boldsymbol{k}, \boldsymbol{m}} \cos (\boldsymbol{k} \cdot \boldsymbol{\vartheta})}{\boldsymbol{k} \cdot \boldsymbol{\nu}}\right] \tag{52}
\end{equation*}
$$

where the coefficients $c_{\boldsymbol{k}, \boldsymbol{m}}$ and $d_{\boldsymbol{k}, \boldsymbol{m}}$ appear in the following expansion of the perturbative term to be averaged:

$$
\begin{equation*}
f_{r}^{(r-1)}(\boldsymbol{I}, \boldsymbol{\vartheta})=\sum_{|\boldsymbol{m}|=r+2} \sum_{\substack{\boldsymbol{k}=\left(\boldsymbol{k}_{1}, k_{2}\right) \in \mathbb{Z}^{n_{1}} \times \boldsymbol{k}_{2} \mid \leq r+2}} \boldsymbol{I}^{\boldsymbol{m} / 2}\left[c_{\boldsymbol{k}, \boldsymbol{m}} \cos (\boldsymbol{k} \cdot \boldsymbol{\vartheta})+d_{\boldsymbol{k}, \boldsymbol{m}} \sin (\boldsymbol{k} \cdot \boldsymbol{\vartheta})\right] \tag{53}
\end{equation*}
$$

As it is usual, the iterative definition of the functions $f_{\ell}^{(r)} \forall \ell>r$ appearing in formula (50) can be easily obtained. In fact, it is just matter to gather separately all the summands having a total degree in the square root of the actions equal to $\ell+2$ among those appearing in the expansion of the Lie series which defines the Hamiltonian in Birkhoff normal form up to order $r$, i.e.,

$$
\begin{equation*}
\mathcal{H}^{(r)}=\exp \left(\mathcal{L}_{\chi^{(r)}}\right) \mathcal{H}^{(r-1)} \tag{54}
\end{equation*}
$$

After having replaced $r$ with $r+1$, the next normalization step could be exhaustively defined by using again all the formulæ (51)-(54). This ends the description of the iterative algorithm constructing the Birkhoff normal form up to a finite order in the neighborhood of an elliptic torus.

In order to ensure that the homological equation (51) can be solved by the generating function $\chi^{(r)}$ written in (52), once again we have to assume a suitable non-resonance condition on the angular velocity vector, for instance, the Diophantine one, i.e.,

$$
\begin{equation*}
\min _{\substack{\boldsymbol{k}=\left(\boldsymbol{k}_{1}, \boldsymbol{k}_{1}\right) \geq \mathbb{Z}^{n_{1} \times \mathbb{Z}^{n_{2}}} \\ \boldsymbol{k \neq 0}\left|,\left|k_{2}\right| \leq r+2\right.}}|\boldsymbol{k} \cdot \boldsymbol{\nu}| \geq \frac{\gamma}{(r+2)^{\tau}}, \tag{55}
\end{equation*}
$$

for some fixed value of the parameters $\gamma>0$ and $\tau \geq n-1$, with $n=n_{1}+n_{2}$. Let us recall that the algorithm constructing the Birkhoff normal form is not convergent. This means that we cannot make the Hamiltonian integrable, by iterating the procedure up to infinity. Indeed, it is convenient to perform the $r$-th normalization step until the sup-norm of the remainder $\mathcal{R}^{(r)}=\sum_{\ell>r} f_{\ell}^{(r)}$ is decreasing (with respect to the previous one, that is $\mathcal{R}^{(r-1)}$ ) on the set $\left\{(\boldsymbol{I}, \boldsymbol{\vartheta}) \in \mathcal{B}_{\varrho}(\mathbf{0}) \times \mathbb{T}^{n}\right\}$, where $\varrho$ is the maximal distance in action from the elliptic torus in its neighborhood. The so called estimates à la Nekhoroshev
are based on a rather standard argument that can be summarized as follows: the size of the remainder can be minimized in correspondence with an optimal step $\bar{r} \propto \varrho^{-1 /[2(1+\tau)]}$; therefore, it is easy to prove that $\sup _{\mathcal{B}_{\ell}(\mathbf{0}) \times \mathbb{T}^{n}}\left|\mathcal{R}^{(\bar{r})}\right|$ is bounded from top by a quantity $\sim \exp \left(-\left(\varrho^{*} / \varrho\right)^{1 /[2(1+\tau)]}\right)$ that is exponentially small with respect to the inverse of a fractional power of the distance $\varrho$ from the torus, $\varrho^{*}$ being a suitable positive constant. In words, this phenomenon is due to the fact that the distance $\varrho$ from the torus rules the perturbation: in its vicinity, the integrable approximation is better and the perturbation (that is given by the remainder) is smaller; otherwise, the approximation provided by the Birkhoff normal form is not so accurate, the perturbative terms play a remarkable role and just a few normalization steps can be profitably performed. Much more details can be found in the introductory lectures [22], while a quantitative analysis in the context of some computer-assisted proofs is described in [5].

Birkhoff normal forms are commonly used in order to provide lower bounds on the diffusion time that is needed by a Hamiltonian system to escape from a local region of the phase space. For the sake of simplicity, let us rewrite the Birkhoff's normal form up to order $r$, that is described in (50), in the following more concise way: $\mathcal{H}^{(r)}(\boldsymbol{I}, \boldsymbol{\vartheta})=\mathcal{Z}^{(r)}(\boldsymbol{I})+$ $\mathcal{R}^{(r)}(\boldsymbol{I}, \boldsymbol{\vartheta})$. Since $\dot{\boldsymbol{I}}=\left\{\boldsymbol{I}, \mathcal{H}^{(r)}\right\}=\left\{\boldsymbol{I}, \mathcal{R}^{(r)}\right\}$, the smaller is the remainder, the longer will be the escaping time from the neighborhood $\mathcal{B}_{\varrho}(\mathbf{0}) \times \mathbb{T}^{n}$ (around the elliptic torus) for any motion starting from an initial condition with $\|\boldsymbol{I}(0)\| \leq \varrho_{0}<\varrho$. One can easily verify that also the lower bound for the diffusion time is of order $\sim \exp \left(\left(\varrho^{*} / \varrho\right)^{1 /[2(1+\tau)]}\right)$ for $\varrho \rightarrow 0_{+}$. Therefore, for some realistic models of physical interest, this can be used to ensure the effective stability, because the diffusion time can be made larger than the expected lifetime of the system, by choosing a value of $\varrho_{0}$ that is small enough. In the present work, we do not provide any evaluation of the drift speed $\dot{\boldsymbol{I}}$, also because there is not a natural lifetime of the FPU model to be compared with the estimates about the diffusion time. Therefore, we prefer to show the evidence that the quasi-periodic motions, observed in the numerical experiments reported in subsection 4.1, are generated by some initial conditions, which are so close to an elliptic torus that the remainder of the Birkhoff normal form can be made effectively small.

In order to produce the plots reported in Fig. 6, we reconsider all the normal forms related to the one-dimensional elliptic tori we have explicitly constructed for both the $\alpha$-model and the $\beta$ one in the cases with $N=4,8$. More precisely, we study separately every Kolmogorov-like normal form of type (11), which corresponds to a dot (denoted by •) that has been plotted in Fig. 4. Let us recall that we have preliminarly produced such a Hamiltonian by several algebraic manipulations (using the libraries of utilities provided by the software package X@óvoऽ). In order to not increase dramatically the computational complexity of the normalization algorithm, we prefer to avoid the search for the optimal step. Indeed, for each of those Kolmogorov-like normal forms, we perform a further Birkhoff normalization up to order $r=5$ for $N=4$ and $r=1$ for $N=8$. Moreover, $\forall s=0, \ldots, r$ we truncate the Hamiltonians $\mathcal{H}^{(s)}$, defined by the iterative procedure summarized by formulæ (51)-(54), up to the maximal trigonometric degree 24 (or 16) in the angles $\boldsymbol{\vartheta}$ and up to order 8 (or 4 ) in the square root of the actions $\boldsymbol{I}$ when $N=4(N=8$, respectively). Let us stress that these limits force us to manipulate several millions of terms in the case with $N=8$, because the Hamiltonians depend on
seven pairs of canonically conjugate variables. For each of the Birkhoff normal forms $\mathcal{H}^{(r)}(\boldsymbol{I}, \boldsymbol{\vartheta})=\mathcal{Z}^{(r)}(\boldsymbol{I})+\mathcal{R}^{(r)}(\boldsymbol{I}, \boldsymbol{\vartheta})$ expanded in the vicinity of a 1D elliptic torus, we select among the initial conditions of "semi-sinusoidal" type the configuration minimizing the absolute value of the remainder $10\left|\mathcal{R}^{(r)}\right|$. In view of the truncation parameters, such a remainder $\mathcal{R}^{(r)}=\sum_{\ell>r} f_{\ell}^{(r)}$ is of order 8 (or 4) with respect to $\|\boldsymbol{I}\|^{1 / 2}$. In Fig. 6, for each of the Birkhoff normal forms $\mathcal{H}^{(r)}$ we have computed, we plot the value of the specific energy for such an initial condition and the corresponding minimum of the absolute value for what concerns the remainder $\left|\mathcal{R}^{(r)}\right|$.

It is very meaningful to compare every case reported in Fig. 6 with the corresponding one in Fig. T. The simpler behavior is shown in the case of the $\beta$-model with $N=4$ : with the exception of some fluctuations for very small amplitudes of the initial conditions of "semi-sinusoidal" type that are probably due to round-off errors, the size of the remainder grows linearly as a function of the specific energy in the semi-log plot (i.e., with the prescribed polynomial law) up to values of $E_{S} \simeq 1$. This is in agreement with the corresponding plot of the variation of the first angular velocity in Fig. [1, that is of order of the roundoff error until a sharp transition to larger values (that we associate to a chaotic regime) occurs for $E_{S} \gtrsim 1$. In the case of the $\alpha-$ model with $N=4$, the chaotic regime can be appreciated in Fig. 1 in the region where $E_{S}>0.3$ and it looks rather intermittent up to values of the specific energy close to 1 . Once again, this is in rather good agreement with the behavior of the size of the remainder in the vicinity of 1D elliptic tori, whose corresponding values are reported in the graph appearing below in the top-left box of Fig. 6. In fact, the abscissa of the last plotted point corresponds to a value of the specific energy $E_{S}$ that is slightly larger than 1 . Moreover, the slope of the curve looks lower in the region with $E_{S} \in(0.3,1)$ with respect to the trend we can appreciate for smaller energies. Let us recall that such a change of the slope in the behavior of the reminder size is usually connected with a decrease of the optimal step, when the normalization algorithm à la Birkhoff is performed (see, e.g., the discussion in subsection 4.3 of [27]). This explains why such a procedure is less effective when $E_{S} \in(0.3,1)$. When the $\alpha-$ model is considered with $N=8$, in Fig. 1 a few of very narrow chaotic regions can be observed in the region with $E_{S} \in(0.05,0.2)$, while for greater values of the specific energy the prevalence of chaos is very definite. We think that the agreement with the plot in the bottom-left box of Fig. 6 is rather acceptable, because there is a sharp decrease of the slope, when $E_{S}$ is approximately in the interval $(0.05,0.2)$ and the graph stops there. In the last case, that is the $\beta$-model with $N=8$, the variation of the first angular velocity in Fig. 1 shows a sudden increase in the interval of values of $E_{S} \in(0.05,0.1)$; moreover, there is a sharp transition to the chaotic regime for $E_{S} \gtrsim 0.2$. There is a small track of the intermediate chaotic region in the behavior of the size of the remainder in the bottom-right box of Fig. 6 (see the rather irregular plot of three consecutive points with specific energy $E_{S} \lesssim 0.1$ ); moreover, also this graph stops at $E_{S} \simeq 0.2$. In short, the

[^10]

Figure 6: Energy level $\left|\mathcal{R}^{(r)}\right|$ for the remainder of the Birkhoff normal form around an elliptic torus (after having performed $r=5,1$ normalization steps in the cases with $N=4,8$, respectively). Such an evaluation is done for the initial conditions of "semi-sinusoidal" type minimizing the remainder itself. The corresponding value of the specific energy for such initial conditions is reported on the axis of abscissas. The $\alpha$-model and the $\beta$ one are considered on the left and on the right, resp.; the cases with $N=4,8$ are reported from top to bottom. In the top-left box, also the values for the remainder of the Birkhoff normal forms around 2D elliptic tori are reported and the corresponding dots (each of them marked with the symbol •) are connected by dashed segments. All the other graphs refer to 1D elliptic tori.
coherence between the data reported in Figs. 6 and 17 reinforces the idea that a motion starting from initial conditions of "semi-sinusoidal" type is quasi-periodic, when it is in the stability region surrounding an elliptic torus, where the size of the remainder is negligible with respect to the value of the total energy of the Hamiltonian.

In the top-left box of Fig. 6, the size of the remainder for the Birkhoff normal forms around 2D elliptic tori is also included (the corresponding plot is connected by dashed segments). In more detail, this further process performing the elimination of the perturbative terms starts from the normal forms represented by the dots appearing on the diagonal edge of the region marked with many symbols $\bullet$, that are plotted in the left box of Fig. 3, Let us recall that they are expected to correspond to the 2D elliptic tori that are the closest ones with respect to the initial conditions of "semi-sinusoidal" type. By
comparing those plots, one can appreciate that the remainders of Birkhoff normal forms centered about 2D elliptic tori are much larger than the corresponding ones that are related to linearly stable periodic orbits. This fact unambiguously shows that the latter are much more effective, for what concerns the approximation of the orbits originating by the initial conditions of "semi-sinusoidal" type. Such a remark is in agreement with the discussion at the end of subsection 5.1 and it can be seen as a natural completion of that.

## 7 Conclusions and perspectives

In this work we have shown the effectiveness of our procedure explicitly constructing the normal form for elliptic tori. It is based on an adaptation of the algorithm introduced in [49] and proved to be convergent in [26] under rather mild hypotheses. There is a major difference with respect to the framework which has been assumed in those works, because it is typical into the study of planetary systems. There, the dynamics that is transverse to the elliptic tori is secular with respect to the revolution periods of the orbits lying on the elliptic tori; here, there is not any separation between the degrees of freedom according to whether they are fast or slow. An exhaustive analysis of the convergence of our algorithm has been deferred to another work (see the reference reported in footnote ${ }^{5}$ ), while here we prefer to focus on the semi-analytic applications to FPU chains with a small number of nodes. This means that the formal procedure constructing the normal forms related to elliptic tori is explicitly implemented with the help of a specific computer algebra software, specialized in the implementation of perturbation methods based on canonical transformations expressed by Lie series. The results given by such a semianalytic approach are compared with the purely numerical ones, which are produced by other technical tools that are used often in the field of Astronomy and are commonly known with the name of frequency analysis (FA) methods. FA is used here also in a version that is extremely performing in the search of elliptic tori. We have devoted a special care in the comparisons between the semi-analytic results and the numerical ones, for what concerns the periodic orbits that are linearly stable in the transverse dynamics (describing the limit of small oscillations around the periodic orbits themselves). We have considered both the $\alpha$-model and the $\beta$ one in the cases with $N=4,8$. As it was rather natural to expect, we have obtained our best result when the perturbation is given by the quartic terms of the potential energy (i.e., the $\beta$-model) with the lowest significant number of particles $(N=4)$. In such a specific case, our semi-analytic algorithm has been able to construct 1-D elliptic tori up to values of the energy that are about $93 \%$ of their breakdown threshold, according to the numerical exploration performed by using FA, which has been widely described in subsection 5.2. Such a level of performance is comparable with that of the best computer-assisted proof (hereafter, CAP) for what concerns the existence of KAM tori for a Hamiltonian system, i.e., the so called forced pendulum problem (see [9]). Let us also recall that, in some special models described by symplectic maps, there are CAPs working up to values of the small parameter that are extremely close to the numerical breakdown threshold for the wanted invariant tori (see [16]). In this context, it is natural to raise the following couple of interesting questions.

Can our algorithm be translated in a fully rigorous CAP? Would its performances be better than those provided by an approach (see, e.g., [37]) that is not based on the construction of a normal form for elliptic tori? Our work is not focusing on CAPs, therefore, the answers are eventually deferred to a future research project. However, we think that in the present paper, we have settled an interesting way to numerically evaluate the breakdown threshold for elliptic tori, which is based on a technique using FA. In our opinion, the comparisons with the numerical experiments on FPU chains with a small number of nodes could be used as the stress-test to benchmark the CAPs about the persistence of elliptic tori, as well as this has been done in the recent past for the maximal KAM tori that are invariant with respect to the symplectic map or the forced pendulum.

In the present work, both the semi-analytic method and the numerical explorations carefully study some properties of the motions starting from the initial conditions of "semisinusoidal" type, that were first considered in the very pioneering simulations designed by Fermi, Pasta, Ulam and Tsingou. The comparison between these two different approaches is in support of the following picture. When the behavior of the motion is so far from the statistical equipartition of the energy (among the normal modes) to be quasi-periodic, then the orbit is close to an elliptic torus. The concept of vicinity is made meaningful with respect to a partial Birkhoff normal form that is constructed into the neighborhood of the nearest elliptic torus to the initial conditions: we evaluate the size of its remainder and we compare its behavior with that of a chaos indicator that is calculated by using a technique based on FA.

Although any discussion about the asymptotic properties for the number of nodes $N$ going to infinity is beyond the scopes of the present work, it would be interesting to extend our approach to values of $N$ larger than 4 or 8 . In order to do that, in our semianalytic method based on the construction of suitable normal forms we should include some substantial improvements; let us sketch a few of them. It is well known that when initial conditions of "semi-sinusoidal" type and rather low values of the total energy are considered, the oscillation amplitudes of the high frequency normal modes are extremely small (see, e.g., [4). Therefore, in our codes doing computer algebra manipulations we should introduce a hierarchy in the truncation rules of the expansions (that are now completely homogeneous), in such a way to approximate more carefully the dynamics of the low frequency normal modes with respect to those oscillating faster. Such a source of improvement is very technical, but it is essential in order to implement a normal form approach that is able to properly describe the behavior of the FPU chains also for a not so small number of nodes. Moreover, our semi-analytic method has shown better performances in the applications to $\beta$-models rather than to $\alpha$-ones. In order to compensate this imbalance, the introduction of action-angle variables well adapted to the Toda lattice is probably unavoidable. Producing the preliminary expansions of the FPU Hamiltonian in such coordinates is a task far from being trivial.

In a recent work, we have improved very significantly some previous results about the secular dynamics of the $v$-Andromedæ planetary system, by constructing the normal forms for an elliptic torus and, therefore, for the orbit lying on a (KAM) invariant manifold. In order to achieve such a goal, the application of the same algorithm described and tested in the present paper has played a prominent role. Such an implementation has been
so successful also because of the two following facts: in a purely secular approximation of a planetary system, there is not any more distinction between fast and slow dynamics (as in the framework considered here); moreover, the problem is described by a Hamiltonian depending on the canonical variables as an even polynomial (i.e., the same situation holding true for the $\beta$-model). We think that the back transfer of our approach into the field of Astronomy could provide interesting results. There are possible applications that look rather promising, for what concerns the partially unknown 3D architecture of the extrasolar planetary systems detected by the radial velocity method. This holds particularly true for the orbits of the exoplanets that are in a so called mean-motion resonance (see [48] as an example of a recent work about such a subject).

## Acknowledgments

This work was partially supported by the MIUR-PRIN project 20178CJA2B - "New Frontiers of Celestial Mechanics: theory and Applications", the "Beyond Borders" programme of the University of Rome Tor Vergata through the project ASTRID (CUP E84I19002250005) and by the "Progetto Giovani 2019" programme of the National Group of Mathematical Physics (GNFM-INdAM) through the project "Low-dimensional Invariant Tori in FPU-like Lattices via Normal Forms". The authors acknowledge also the MIUR Excellence Department Project awarded to the Department of Mathematics of the University of Rome "Tor Vergata" (CUP E83C18000100006).

## References

[1] D. Bambusi, A. Ponno: Resonance, Metastability and Blow up in FPU, in G. Gallavotti, man. ed., see [20], 191-205 (2008).
[2] G. Benettin, A. Carati, L. Galgani, A. Giorgilli: The Fermi-Pasta-Ulam Problem and the Metastability Perspective, in G. Gallavotti, man. ed., see [20], 151-189 (2008).
[3] G. Benettin, A. Ponno: Understanding the FPU state in FPU-like models, 3, doi: 10.3934/mine. 2021025, 1-22 (2020).
[4] L. Berchialla, L. Galgani, A. Giorgilli: Localization of energy in FPU chains, Discr. \& Cont. Dyn. Syst., 11, 855-866 (2004).
[5] C. Caracciolo, U. Locatelli: Computer-assisted estimates for Birkhoff normal forms, J. of Comput. Dyn., 7, 425-460. (2020).
[6] C. Caracciolo, U. Locatelli, M. Sansottera, M. Volpi: "Librational KAM tori in the secular dynamics of the $v$-Andromedæ planetary system", preprint.
[7] L. Casetti, M. Cerruti-Sola, M. Pettini, E.G.D. Cohen: The Fermi-Pasta-Ulam problem revisited: Stochasticity thresholds in nonlinear Hamiltonian systems, Phys. Rev. E, 55, 6566-6574 (1997).
[8] R.C. Calleja, A. Celletti, R. de la Llave: Local Behavior Near Quasi-Periodic Solutions of Conformally Symplectic Systems, J. of Dyn. \& Diff. Eqs., 25, 821-841 (2013).
[9] A. Celletti, A. Giorgilli, U. Locatelli: Improved estimates on the existence of invariant tori for Hamiltonian systems, Nonlinearity, 13, 397-412 (2000).
[10] H. Christodoulidi, C. Efthymiopoulos, T. Bountis: Energy localization on q-tori, long-term stability, and the interpretation of Fermi-Pasta-Ulam recurrences, Phys. Rev. E, 81, 016210/1-16 (2010).
[11] H. Christodoulidi, C. Efthymiopoulos: Low-dimensional q-Tori in FPU lattices: Dynamics and localization properties, Physica D, 261, 92-113 (2013).
[12] J. Couetdic, J. Laskar, A.C.M. Correia, M. Mayor, S. Udry: Dynamical stability analysis of the HD 202206 system and constraints to the planetary orbits, Astron. \& Astroph., 519, A10 (2010).
[13] A. Deprit: Canonical transformations depending on a small parameter, Celestial Mechanics, 1, 12-30 (1969).
[14] L.H. Eliasson: Perturbations of stable invariant tori for Hamiltonian systems, Ann. Scuola Norm. Sup. Pisa, Cl. Sci., IV Serie, 15, 115-147 (1988).
[15] E. Fermi, J. Pasta, S. Ulam: Studies of Nonlinear Problems, Los Alamos Report LA1940 (1955); reprinted in both [20] and "Collected Papers of Enrico Fermi", edited by E. Segré, University of Chicago, Vol. 2, p. 978, Chicago (1965).
[16] J.-L. Figueras., A. Haro, A. Luque: Rigorous computer-assisted application of KAM theory: A modern approach, Found. Comput. Math., 17, n. 5, 1123-1193 (2017).
[17] S. Flach, A. Ponno: The Fermi-Pasta-Ulam problem: Periodic orbits, normal forms and resonance overlap criteria, 237, 908-917 (2008).
[18] H. Flaschka: The Toda lattice. II. Existence of integrals, Phys. Rev., 9, 1924-1925 (1974).
[19] F. Gabern, A. Jorba, U. Locatelli: On the construction of the Kolmogorov normal form for the Trojan asteroids, Nonlinearity, 18, n.4, 1705-1734 (2005).
[20] G. Gallavotti, man. editor: The Fermi-Pasta-Ulam problem. A status report, Lecture Notes in Physics, 728, Springer, Berlin (2008).
[21] A. Giorgilli: Quantitative methods in classical perturbation theory, proceedings of the Nato ASI school "From Newton to chaos: modern techniques for understanding and coping with chaos in N-body dynamical systems", A.E. Roy e B.D. Steves eds., Plenum Press, New York (1995).
[22] A. Giorgilli: Notes on exponential stability of Hamiltonian systems, in Dynamical Systems, Part I. Pubbl. Cent. Ric. Mat. Ennio De Giorgi, Sc. Norm. Sup. Pisa, 87-198 (2003).
[23] A. Giorgilli, A. Delshams, E. Fontich, L. Galgani, C. Simó: Effective stability for a Hamiltonian system near an elliptic equilibrium point, with an application to the restricted three body problem, J. Diff. Equations, 77, 167-198 (1989).
[24] A. Giorgilli, U. Locatelli: Kolmogorov theorem and classical perturbation theory, J. of App. Math. and Phys. (ZAMP), 48, 220-261 (1997).
[25] A. Giorgilli, U. Locatelli, M. Sansottera: Kolmogorov and Nekhoroshev theory for the problem of three bodies, Cel. Mech. \& Dyn. Astr., 104, 159-173 (2009).
[26] A. Giorgilli, U. Locatelli, M. Sansottera: On the convergence of an algorithm constructing the normal form for lower dimensional elliptic tori in planetary systems, Cel. Mech. \& Dyn. Astr., 119, 397-424 (2014).
[27] A. Giorgilli, U. Locatelli, M. Sansottera: Secular dynamics of a planar model of the Sun-Jupiter-Saturn-Uranus system; effective stability in the light of Kolmogorov and Nekhoroshev theories, Regular and Chaotic Dynamics, 22, 54-77 (2017).
[28] A. Giorgilli, M. Sansottera: Methods of algebraic manipulation in perturbation theory, in "Chaos, Diffusion and Non-integrability in Hamiltonian Systems - Applications to Astronomy", Proceedings of the Third La Plata International School on Astronomy and Geophysics, P.M. Cincotta, C.M. Giordano and C. Efthymiopoulos eds., Universidad Nacional de La Plata and Asociación Argentina de Astronomía Publishers, La Plata (2012).
[29] W. Gröbner: Die Lie-Reihen und Ihre Anwendungen, Springer Verlag, Berlin (1960); Italian transl.: Le serie di Lie e le loro applicazioni, Cremonese, Roma (1973).
[30] J.D. Hadjedemetriou: Periodic orbits in gravitational systems, 43-79, in "Chaotic Worlds: from Order to Disorder in Gravitational N-Body Dynamical Systems", B.A. Steves, A.J. Maciejewski, M. Hendry (eds)., Nato Science Series II, 227, Springer Netherlands (2006).
[31] J. Henrard: The algorithm of the inverse for Lie transform, in "Recent advances in Dynamical Astronomy", B.D: Tapley and V. Szebehely, V. (eds.), Reidel, Dortrecht (1973).
[32] G. Hori: Theory of general perturbations with unspecified canonical variables, Publ. Astron. Soc. Japan, 18, 287-296 (1966).
[33] A.N. Kolmogorov: Preservation of conditionally periodic movements with small change in the Hamilton function, Dokl. Akad. Nauk SSSR, 98, 527 (1954). Engl. transl. in: Los Alamos Scientific Laboratory translation LA-TR-71-67; reprinted in: Lecture Notes in Physics, 93.
[34] V. Lanchares, A.I. Pascual, M. Iñarrea, J.P. Salas, J.F. Palacián, P. Yanguas: Reeb's Theorem and Periodic Orbits for a Rotating Hénon-Heiles Potential, J. Dyn. Diff. Eqs., https://doi.org/10.1007/s10884-019-09814-6 (2019).
[35] J. Laskar: Frequency Map analysis and quasi periodic decompositions, in Benest, D., Froeschlé, C. and Lega E. (eds.): Hamiltonian systems and Fourier analysis, Taylor and Francis (2003).
[36] J. Laskar, P. Robutel: High order symplectic integrators for perturbed Hamiltonian systems, Cel. Mech. \& Dyn. Astr., 80, 39-62 (2001).
[37] A. Luque, J. Villanueva: A KAM theorem without action-angle variables for elliptic lower-dimensional tori, Nonlinearity, 24, 1033-1080 (2011).
[38] V.K. Melnikov: On some cases of conservation of almost periodic motions with a small change of the Hamiltonian function, Dokl. Akad. Nauk SSSR, 165, 1245-1248 (1965).
[39] A. Morbidelli, A. Giorgilli: Superexponential stability of KAM tori, J. Stat. Phys., 78, 1607-1617 (1995).
[40] N.N. Nekhoroshev: Exponential estimates of the stability time of near-integrable Hamiltonian systems. Russ. Math. Surveys, 32, 1 (1977).
[41] N.N. Nekhoroshev: Exponential estimates of the stability time of near-integrable Hamiltonian systems, 2. Trudy Sem. Petrovs., 5, 5 (1979).
[42] B. Noyelles: Expression of Cassini's third law for Callisto, and theory of its rotation, Icarus, 202, 225-239 (2009).
[43] B. Noyelles, N. Delsate, T. Carletti: Equilibrium search algorithm of a perturbed quasi-integrable system, preprint (2012); https://arxiv.org/pdf/1101.2138.pdf
[44] T. Penati, S. Flach: Tail resonances of Fermi-Pasta-Ulam-breathers and their impact on the pathway to equipartition, Chaos, 17, 023102 (2007).
[45] J. Pöschel: On elliptic lower dimensional tori in Hamiltonian sytems, Math. Z., 202, 559-608 (1989).
[46] G. Reeb: Sur certaines propriétés topologiques des trajectoires des systémes dynamiques, Acad. Roy. Belgique. Cl. Sci. Mém. Coll. in $8^{\circ} .27$ article 9 (1952).
[47] B. Rink: An Integrable Approximation for the Fermi-Pasta-Ulam Lattice, in G. Gallavotti, man. ed., see [20], 283-301 (2008).
[48] M. Sansottera, A.-S. Libert:"Resonant Laplace-Lagrange theory for extrasolar systems in mean-motion resonance", Cel. Mech. \& Dyn. Astr., 131, 38 (2019).
[49] M. Sansottera, U. Locatelli, A. Giorgilli: A semi-analytic algorithm for constructing lower dimensional elliptic tori in planetary systems, Cel. Mech. \& Dyn. Astr., 111, 337-361 (2011).
[50] C. Skokos, T. Bountis, C. Antonopoulos: Geometrical properties of local dynamics in Hamiltonian systems: The Generalized Alignment Index (GALI) method, Physica D, 231, 30-54 (2007).
[51] M. Volpi, U. Locatelli, M. Sansottera: A reverse KAM method to estimate unknown mutual inclinations in exoplanetary systems, Cel. Mech. \& Dyn. Astr., 130, 36 (2018).


[^0]:    * 2020 Mathematics Subject Classification. Primary: 37J40; Secondary: 37N05, 65P40, 70H08, 70 H 15. Key words and phrases: FPU problem, lower dimensional invariant tori, KAM theory, normal form methods, perturbation theory for Hamiltonian systems.

[^1]:    ${ }^{1}$ Actually, when the whole procedure constructing the invariant elliptic torus is successful, on that same torus the motion law of the action vector $\boldsymbol{I}$ will be such that $\left\|\boldsymbol{I}(t)-\boldsymbol{I}^{\star}\right\|=o\left(\left\|\boldsymbol{I}^{\star}\right\|\right)$. Reasonable ways to practically choose the initial translation $\boldsymbol{I}^{\star}$ will be described in section 5

[^2]:    ${ }^{2}$ Setting $K=2$ is quite natural for Hamiltonian systems close to stable equilibria, see, e.g., 27. The reason of such a choice can be understood by referring to the discussion in Section 2 The size of the perturbing terms making part of the Hamiltonian term $\mathcal{H}^{*}$ in (5) is $\mathcal{O}\left(\left\|\boldsymbol{I}^{\star}\right\|^{s / 2}\right)$, $s$ being the degree in the square root of the actions $\boldsymbol{I}$. Each monomial term appearing in the perturbation will have a maximal trigonometric degree in $\varphi$ that is not greater than $s$. The subsequent introduction of the actions $\boldsymbol{p}$ modifies the dependency on the actions, but does not change the relation between the order of magnitude in the (small) parameter $\left\|\boldsymbol{I}^{\star}\right\|$ and the degree in the angles, the latter being at most twice bigger with respect to the former.

[^3]:    ${ }^{3}$ Because of the so called "exchange theorem" (see [29]), the new Hamiltonian $H^{(r)}$ is obtained from the old one, by applying the Lie series to $H^{(r-1)}$ in reverse order with respect to what is written in (13). This is consistent with the order of the discussion in the following subsections: the first stage of the $r$-th normalization step deals with the canonical transformation generated by $\chi_{0}^{(r)}$, the second one with $\chi_{1}^{(r)}$ and the last one with both $\chi_{2}^{(r)}$ and $\mathfrak{D}^{(r)}$. A rather self-consistent introduction to the Lie series formalism in the Hamiltonian framework can be found, e.g., in [22], where the method is described in an explicitly algorithmic way, going back to the seminal articles written by Hori, Deprit and Henrard (see [32], [13] and [31, resp.).

[^4]:    ${ }^{4}$ In practical implementations, such a change of coordinates $\mathfrak{D}^{(r)}$ can be conveniently defined by composing a subsequence of Lie series, each of them being related to a quadratic generating function $\mathcal{D}_{2}^{(r ; m)}(\boldsymbol{\xi}, \boldsymbol{\eta})$. By the way, this allows to automatically determine $\mathfrak{D}^{(r)}$ in such a way it is a near to the identity canonical transformation. Moreover, as a further alternative method, when one is dealing with the estimates needed to prove the convergence of the algorithm, in [26] the use of the Lie transforms (that are equivalent to the composition of infinite sequences of Lie series) has been found to be very suitable.

[^5]:    ${ }^{5}$ A complete proof is included in C. Caracciolo: On the stability in the neighborhood of invariant elliptic tori, Ph.D. thesis, Univ. of Rome "Tor Vergata" (2021), which is available on request to the author.

[^6]:    ${ }^{6}$ To fix the ideas, in our numerical experiments such an upper bound has been set so as to be equal to 20 .

[^7]:    ${ }^{7}$ Since the pioneering work [15], it is usual to set $N$ equal to an integer power of 2 . This allows to compute very efficiently the passage from the original coordinates $(\boldsymbol{y}, \boldsymbol{x})$ to those related to the normal modes, that are ( $\boldsymbol{Y}, \boldsymbol{X}$ ), and vice versa (see formula (2)).

[^8]:    ${ }^{8}$ Let us recall that in section 3 we decided to set $K=2$.

[^9]:    ${ }^{9}$ Let us consider the dashed line approaching $\theta_{ \pm 3}=0$ from below. Looking at such a plot, it is evident that the ratio $\theta_{j} /(2 \pi)=\Omega_{j} / \omega_{1} \rightarrow m$, for some integer numbers $j$ and $m$. By following such a curve backward for energies smaller and smaller, we can see that it is going to converge to the value $2 \pi \nu_{2} / \nu_{1}-4 \pi \simeq-1$, where $\nu_{1}$ and $\nu_{2}$ are given in (3). They are the angular velocities of the first and second normal mode in the case of small oscillations, so they are equal to the limit values of $\omega_{1}$ and $\Omega_{1}$ for $E_{S} \rightarrow 0$, respectively. Therefore, we can conclude that $j=1$ and $m=2$.

[^10]:    ${ }^{10}$ From a conceptual point of view, it would be more elegant to fix the initial conditions in advance, in such a way to look for the Birkhoff normal form minimizing the remainder. We stress that such a procedure would require to handle many huge expansions to be conveniently stored at the same time. Instead, the computational algorithm described in the text above (that prescribes to consider a Birkhoff normal form at once) allows us to enormously reduce the requirements about the memory storage.

