Ergodicity, KAM, FPUT

Giovanni Gallavotti
Dipartimento di Fisica,
Università di Roma-La Sapienza and I.N.F.N.-Roma1

Abstract

Boltzmann introduced the microcanonical ensemble in 1868, [1], and immediately attempted to give an example of a system whose stationary states would be described by the emsemble (as suggested also by his ergodic hypothesis). The example, [2], has been recently shown to be incorrect, if taken literally: the point was to suppose that constants of motion, if any besides the energy, would necessarily be smooth functions; and soon later he warned on the dangers implicit in a similar assumption. Fifty years later Fermi wrote a paper attempting to prove that in general a nonlinear system should be ergodic, [3]: but his proof relied again on Boltzmann's assumption. Thirty-four more years elapsed, and Fermi returned on the problem collaborating with Pasta, Ulam, Tsingou: the surprise was that the considered non linear chain was apparently not following the ergodic hypothesis. In the same year Kolmogorov had proved the conservation of many quasi periodic motions in nonlinear perturbations of integrable systems, [4]: his theorem was considered, already a few years later, a possible explanation of the FPUT work, [5]. This was only the beginning of intense research: here a brief sketch is presented to illustrate the above themes and the connection with the multiscale aspects of the problems, and the "Renormalization group method" intended as a map $\mathcal{R}$ whose iterations can be interpreted as successive magnifications, zooming on ever smaller regions of phase space in which motions develop closer and closer to the searched quasi periodic motion of given spectrum.

1 Ergodicity

The Hamiltonian of the systems of N oscillators $r_x = (p_x, r_x), x = 1, ..., N$, considered by FPUT was, $(N = 64, r_0 = r_N = 0)$:

$$\sum_{x=1}^{N-1} \frac{p_x^2}{2m} + \sum_{x=1}^{N-1} \frac{1}{2}(r_{x+1} - r_x)^2 + \frac{\alpha}{3}(r_{x+1} - r_x)^3 + \frac{\beta}{4}(r_{x+1} - r_x)^4 \tag{1.1}$$

representing a linear chain of oscillators, which can be generalized to a d-dimensional lattice with r, x as d-dimensional vectors, for instance:

$$\sum_{\mathbf{x}\in Z_N} \frac{1}{2}\mathbf{p}_{\mathbf{x}}^2 + \sum_{\mathbf{x}\in Z_N} \sum_{i=1}^{d} \frac{1}{2}(\mathbf{r}(\mathbf{x} + \mathbf{e}_i) - \mathbf{r}_{\mathbf{x}})^2 + \frac{\beta}{2}(\mathbf{r}_(\mathbf{x} + \mathbf{e}_i) - \mathbf{r}_{\mathbf{x}})^4 \tag{1.2}$$

where $\mathbf{e}_i, i = 1, \ldots, d$ is an ortonormal basis and $Z_N = (0, \ldots, N-1)^d$ is the cube with side N in the lattice Z^d, and $\mathbf{r}_{\mathbf{x}} = 0$ if any coorinate of $\mathbf{x}$ is 0 or N.

Boltzmann's ergodic hypothesis, [6], formally stated an hypothesis, which had been perhaps earlier conceived (for instance by Descartes, [7, p.4]) and which had been accepted by Maxwell, [8], supposes that the point representing the system istantaneous state wonders visiting all points that can possibly be visited (*i.e.* compatible with the constraints).

The hypothesis, leaving aside the obvious exceptions, and its implication that all points covered in time a closed orbit allowed Boltzmann to discover a mechanical interpretation of the Second Law as a manfestation of the Maupertuis' *least action* principle via the following extraordinary argument.

(i) An equilibrium state of a system is idetified with a periodic motion $t \to x(t)$ developing under the action of internal forces with potential energy $V(x(t))$ and of external forces with potential $V_{ext}(x(t))$; its kinetic energy is $K(x(t))$, and the *action* of x is defined, if its period is i (adopting here the original notation for the period), by

$$\mathcal{A}(x) = \int_0^i (\frac{m}{2}\dot{x}(t)^2 - V(x(t)))\, dt \tag{1.3}$$

(ii) Properties of a small periodic variations $\delta x(t)$ of a state, *i.e.* properties of an equilibrium state close to x, represented as:

$$\delta x(t) = x'(\frac{i'}{i}t) - x(t) \stackrel{def}{=} x'(i'\varphi) - x(i\varphi) \tag{1.4}$$

where $\varphi \in [0,1]$ is the *phase*, as introduced by Clausius, [9]. The role of φ is simply to establish a correspondence between points on the initial trajectory x and on the varied one x': it is manifestly an arbitrary correspondence (which could be defined differently without affecting the final result) convenient to follow the algebraic steps.[1]

(iii) Following Clausius version, [9], the heat theorem is deduced from a few identities about the average values $\overline{F} = i^{-1}\int_0^i F(x(t))dt$ for generic observables F and $\overline{K} = i^{-1}\int_0^i \frac{1}{2}\dot{x}(t)^2 dt$ for the average of the kinetic energy.

Let $x(t) = \xi(\frac{t}{i})$ with $\xi(\varphi)$ 1-periodic; it is $\overline{F} \equiv \int_0^1 F(\xi(\varphi))d\varphi$, and $\overline{K} = \frac{1}{2}\int_0^1 \frac{\xi'(\varphi)^2}{i^2}d\varphi$. Changing x by δx means to change ξ by $\delta\xi$ and i by δi. So variation of an average is the sum of the variation due to $\delta\xi$ and that due to δi: namely $\delta\overline{F} = \delta_\xi\overline{F} + \delta_i\overline{F}$. Therefore

$$\delta\overline{V} \equiv \delta_\xi\overline{V}, \qquad \delta\overline{V}_{ext} = \delta_\xi\overline{V}_{ext}, \qquad \delta\overline{K} = -2\frac{\delta i}{i}\overline{K} + \delta_\xi\overline{K} \tag{1.5}$$

The variation $\delta(\overline{K}+\overline{V}+\overline{V}_{ext}) \stackrel{def}{=} dU + dL$, with $U \stackrel{def}{=} \overline{K}+\overline{V}$ and $dL \stackrel{def}{=} d\overline{V}_{ext}$, has a natural interpretation of heat δQ in/out in the process $x \to x'$.[2]

Then using the above Eq. (1.5): $\delta_\xi\overline{K} = \delta\overline{K} + 2\frac{\delta i}{i}\overline{K}$

$$\begin{aligned}\delta Q \equiv \delta(\overline{K}+\overline{V}+\overline{V}_{ext}) &= -2\frac{\delta i}{i}\overline{K} + \delta_\xi(\overline{K}+\overline{V}+\overline{V}_{ext})\\ = -2\frac{\delta i}{i}\overline{K} + 2\delta_\xi\overline{K} &+ \boldsymbol{\delta}_\xi(-\overline{\mathbf{K}}+\overline{\mathbf{V}}+\overline{\mathbf{V}}_{\mathbf{ext}}) \equiv +2\frac{\delta i}{i}\overline{K} + 2\delta\overline{K} + \mathbf{0}\end{aligned} \tag{1.6}$$

where the last $\mathbf{0}$ is due to *Maupertuis' principle*, as the motion follows the equations of motion (at fixed i). Therefore, (*Heat theorem*):

$$\frac{\delta Q}{\overline{K}} = +2\frac{\delta i}{i} + 2\frac{\delta\overline{K}}{\overline{K}} = 2\delta\log(i\overline{K}) \stackrel{def}{=} \delta\mathbf{S} \tag{1.7}$$

is an exact differential. In reality it is somewhat strange that both Boltzmann and Clausius call the last equation a "generalization of the action principle": the latter principle uniquely determines a motion, *i.e.* it determines

[1] It should be noted that Boltzmann does not care to introduce the phase and this makes his computations difficult to follow (in the sense that once realized which the final result is, it may be easier to reproduce it rather than to follow his calculations).

[2] Because $\delta(\overline{K}+\overline{V}) \stackrel{def}{=} dU$ and $\delta\overline{V}_{ext} \stackrel{def}{=} \delta L$ are the variation of the internal energy U and dL is the work that the system performs: *i.e.* $\delta Q = \delta U + \delta L$.

its equations. The equation instead does not determine a motion but it only establishes a relation between the variation of average kinetic and potential energies of close periodic motions under the assumption that they satisfy the equations of motion; and it does not establish a variational property, see [10, Sec.I.4].

The above analysis was successive to the ergodic hypothesis and a complement to it: after achieving the construction of the microcanonical ensemble, [1], as recognized by Maxwell in one of his last works,[8], and later by Gibbs, [11] (who refers to a later publication, [12], curiously quoting it under the title of the first section) and soon Boltazmann wanted to provide simple exapmles of ergodic systems.

He examined the system of one particle, in dimension 2, subject to a central force of potential $-\frac{\alpha}{|\mathbf{x}|}$, to a potential $+\frac{\beta}{|\mathbf{x}|^2}$ and to a flat repelling line. For $\beta = 0$ the motion is integrable and Boltzmann tried to prove that as soon as $\beta > 0$ the system would be ergodic. This is not right (as conjectured in [13, 14] and proved in [15]) the example fails as a consequence of the KAM theorem, discovered in the early '950s by Kolmogorov, [4]. Nevertheless Boltzmann's paper was probably known to Fermi, who in his early youth attempted to provide a proof of the ergodic hypothesis,[16] (overlooking a similar difficulty: *i.e.* the possibility of the existence of non smooth constants of motion); hence when (relatively) large computers became available the FPUT project was attempted.

Existence of non smooth constants of motion rests, ultimately, upon realizing that the phase space may have a large volume covered by smooth invariant surfaces of low dimension: however proving their existence requires a multi scale analysis: by examining the motions with ever increasing precision, invariant curves can be separated from the nearby motions (which are likely to undergo chaotic motions).

2 KAM

In spite of its interest the first attempt to establish a rigorous connection between the problem of existence of non smooth constant of motions and the ergodicity of the FPUT chain had to wait 1971, [17], treating "only" N particles chains with N either prime or a power of 2 (hence including at least the cases studied in FPUT work $N = 16, 32, 64!$) and a complete proof, including the "other N-s" (N not a power of 2 or not a prime), dates

2006, [18]; both, although conceptually very different, relied on the KAM theorem, which is discussed below as an application of the renormalization group method.

Consider the Hamiltonian:

$$H_0(\mathbf{A},\boldsymbol{\alpha})=\frac{1}{2}(\mathbf{A}\cdot J_0\mathbf{A})+\boldsymbol{\omega}_0\cdot\mathbf{A}+f_0(\mathbf{A},\boldsymbol{\alpha})\equiv h_0+f_0 \tag{2.1}$$

real analytic for $(\mathbf{A},\boldsymbol{\alpha})\in(\mathcal{D}_\rho\times\mathcal{T}^N)$, $N\ge 2$, with: $\mathcal{D}_\rho=\{\mathbf{A}\in R^N,|A_j|<\rho\}$, $\mathcal{T}^N$ = N-dimensional torus $[0,2\pi]^N$, identified with unit polydisk, $\{\mathbf{z}|\,z_j=e^{i\alpha_j},j=1,\ldots,N\}$, $\boldsymbol{\omega}_0\in R^N$ and J_0 is a $N\times N$ matrix, positive for simplicity, with eigenvalues in $[J_{0,-},J_{0,+}]$.

The Hamiltonian is supposed holomorphic in the complex region $\mathcal{C}_{\rho_0,\kappa_0}=\{(\mathbf{A},\mathbf{z})||A_j|\le\rho_0,e^{-\kappa_0}\le|z_j|\le e^{\kappa_0}\}$; the label 0 is attached, since the beginning, because H_n,f_n,ρ_n,κ_n will arise later, with $n=1,2,\ldots$; the perturbation f_0 size will be measured by ε_0 as defined here:

$$\begin{aligned}
&\mathcal{C}_{\rho_0,\kappa_0}\stackrel{def}{=}\{(\mathbf{A},\mathbf{z})||A_j|\le\rho_0,\ e^{-\kappa_0}\le|e^{i\alpha_j}|\le e^{\kappa_0},j=1,\ldots,N\}\subset\mathcal{C}^{2N}\\
&\quad\varepsilon_0=||\boldsymbol{\partial}_{\mathbf{A}}f_0||_{\rho_0,\kappa_0}+\frac{1}{\rho_0}||\boldsymbol{\partial}_{\boldsymbol{\alpha}}f_0||_{\rho_0,\kappa_0}\qquad\text{with}:\\
&\quad||f||_{\rho_0,\kappa_0}\stackrel{def}{=}\max_{\mathcal{C}_{\rho_0,\kappa_0}}|f(\mathbf{A},\mathbf{z})|,\qquad\forall f\text{ holomorphic in }\mathcal{C}_{\rho_0,\kappa_0}
\end{aligned}\tag{2.2}$$

with $\rho_0>0, 1\ge\kappa_0>0$, $z_j\equiv e^{i\alpha_j}$; generally $\mathcal{C}_{\rho,\kappa}(\overline{\mathbf{A}})$ will denote a 'polydisk' centered at $\overline{\mathbf{A}}$, *i.e.* defined as in Eq.(2.2) with $|A_j-\overline{A}_j|\le\rho$ replacing $|A_j|\le\rho$ and κ replacing κ_0; polydisks centered at the "origin" $\overline{\mathbf{A}}=0$ will be simply denoted $\mathcal{C}_{\rho,\kappa}$ and called "centered polydisks", while $C_\rho(\overline{\mathbf{A}})$ is the $\{\mathbf{A}||A_i-\overline{A}_i|<\rho\}$. Set $|\mathbf{A}|=\max|A_j|,|\mathbf{z}|=\max|z_j|,\ \forall\mathbf{A},\mathbf{z}\in\mathcal{C}^N$.

For simplicity functions of $\boldsymbol{\alpha}$ will be implicitly regarded as functions of $z_j=e^{i\alpha_j}$ and their arguments will be written as $\mathbf{z}$ or $\boldsymbol{\alpha}$, as convenient.

The idea is to focus attention on the center of $\mathcal{C}_{\rho_0,\kappa_0}$ where, if $\varepsilon_0=0$, a motion ("free motion": $(\mathbf{A},\boldsymbol{\alpha})\to(\mathbf{A},\boldsymbol{\alpha}+\boldsymbol{\omega}_0t)$) takes place which is quasi periodic "with spectrum" $\boldsymbol{\omega}_0$. This is done by studying motions in a small polydisk $\mathcal{C}_{\widetilde{\rho},\widetilde{\kappa}}(\mathbf{a})\subset\mathcal{C}_{\rho_0,\kappa_0}$, eccentric if $\mathbf{a}\ne\mathbf{0}$.

Thus motions starting in $\mathcal{C}(\widetilde{\rho},\widetilde{\kappa})(\mathbf{a})\subset\mathcal{C}(\rho_0,\kappa_0)$ near the center of $\mathcal{C}(\rho_0,\kappa_0)$ can be studied as "through a microscope": in the good cases (*i.e.* under suitable assumption on the initial parameters $J_0,\boldsymbol{\omega}_0$ and f_0) the Hamiltonian will turn out to be expressible (after a further coordinate change, to turn

$\mathcal{C}(\widetilde\rho,\widetilde\kappa)(\mathbf{a})$ into $\mathcal{C}(\rho_1,\kappa_1)$, centered at 0) in a form substantially closer to that of a quasi periodic oscillator (described by its "normal" Hamiltonian $\boldsymbol{\omega}_0\cdot\mathbf{A}$, or $\boldsymbol{\omega}_0\cdot\mathbf{A}+\frac12\mathbf{A}\cdot J_0\mathbf{A}$, in the variables $\mathbf{A},\boldsymbol{\alpha}$).

Iterating the process the Hamiltonian changes but, *remaining analytic in a polydisk* C_{ρ_n,κ_n}, converges to that of a harmonic oscillator: the interpretation will be that, looking very carefully in the vicinity of the "unperturbed" torus $\mathcal{T}_{\boldsymbol{\omega}_0}=\{\mathbf{A}=\mathbf{0},\boldsymbol{\alpha}\in[0,2\pi]^N\}$, the perturbed Hamiltonian exhibits with increasing precision a harmonic motion with spectrum $\boldsymbol{\omega}_0$.

The result is the KAM theorem in a form not only reminiscent of the methods called "renormalization group", RG, in quantum field theory but just a realization of them, [19]. Existence of the invariant torus can be seen as the existence of a trivial fixed point (a harmonic oscillator) of a single canonical coordinate change together with a time rescaling.

The "spectrum" $\boldsymbol{\omega}_0$ will be supposed "diophantine", *i.e.* for some $C>0$ it is, denoting $\mathcal{Z}^N$ the lattice of the integers, for all $\mathbf{0}\ne\boldsymbol{\nu}\in\mathcal{Z}^N$:

$$|\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}|\;\ge\;C\,|\boldsymbol{\nu}|^{-N},\qquad |\boldsymbol{\nu}|\equiv\sum_{i=1}^N|\nu_i|>0 \tag{2.3}$$

The inequality Eq.2.3 will be repeatedly used to define canonical transformations $(\mathbf{A},\boldsymbol{\alpha})\longleftrightarrow(\mathbf{A}',\boldsymbol{\alpha}')$ with generating functions $\Phi(\mathbf{A}',\boldsymbol{\alpha})+(\mathbf{A}'+\mathbf{a})\cdot\boldsymbol{\alpha}$:

$$\mathbf{A}=\mathbf{A}'+\mathbf{a}+\boldsymbol{\partial}_{\boldsymbol{\alpha}}\Phi(\mathbf{A}',\boldsymbol{\alpha}),\quad \boldsymbol{\alpha}'=\boldsymbol{\alpha}+\boldsymbol{\partial}_{\mathbf{A}'}\Phi(\mathbf{A}',\boldsymbol{\alpha}) \tag{2.4}$$

with Φ chosen so that in the new coordinates $(\mathbf{A}',\boldsymbol{\alpha}')$ the perturbation is '*weaker*'; at the price that the new coordinates will cover a (much) smaller domain, inside $\mathcal{C}(\rho_0,\kappa_0)$.

The $\mathbf{a}$ is introduced because the quasi periodic motion with spectrum $\boldsymbol{\omega}_0$ which, in the first approximation with $f_0=0$, is located at $\mathbf{A}=0$, at a better approximation is shifted, as expected at least if the $\boldsymbol{\alpha}$-average $\overline{f}_0(\mathbf{A})$ of $f_0(\mathbf{A},\boldsymbol{\alpha})$ is not 0:[3] the constant $\mathbf{a}$ provides flexibility to identify the average position, in $\mathbf{A}'$-space, of the new approximations.

Suppose that a holomorphic $\Phi(\mathbf{A}',\boldsymbol{\alpha})$ defines, via Eq.(2.4), in a domain $(\mathbf{A}',\boldsymbol{\alpha}')\in\mathcal{C}_{\widetilde\rho,\widetilde\kappa}$, a symplectic map $(\mathbf{A}',\boldsymbol{\alpha}')\to(\mathbf{A},\boldsymbol{\alpha})$ with image $\subset\mathcal{C}_{\rho_0,\kappa_0}$.

Choose $\mathbf{a}$ as solution of the implicit equation $\mathbf{a}=-J_0^{-1}\boldsymbol{\partial}_{\mathbf{a}}\overline{f}_0(\mathbf{a})$ and insert Eq.(2.4) into H_0.

[3]As shown already by the cases with f_0 not depending on $\boldsymbol{\alpha}$.

After easy&patient algebra, the Hamiltonian becomes, [4] up to a constant, the sum of the four lines (that will be referred as $F_j, j = 0, \ldots, 4$):

$$\begin{aligned}
(0):&\quad H'(\mathbf{A}',\boldsymbol{\alpha}') = \frac{1}{2}(\mathbf{A}'\cdot J_0\mathbf{A}') + \boldsymbol{\omega}_0\cdot\mathbf{A}'\\
(1):&\quad +\overline{f}_0(\mathbf{A}'+\mathbf{a}) - \overline{f}_0(\mathbf{a}) - \boldsymbol{\partial}_{\mathbf{a}}\overline{f}_0(\mathbf{a})\cdot\mathbf{A}'\\
(2):&\quad +\Big(f_0(\mathbf{A}'+\mathbf{a}+\boldsymbol{\partial}_{\boldsymbol{\alpha}}\Phi,\boldsymbol{\alpha}) - f_0(\mathbf{A}'+\mathbf{a},\boldsymbol{\alpha})\Big) + \frac{1}{2}\boldsymbol{\partial}_{\boldsymbol{\alpha}}\Phi\cdot J_0\boldsymbol{\partial}_{\boldsymbol{\alpha}}\Phi\\
(3):&\quad +(\boldsymbol{\omega}_0 + J_0(\mathbf{A}'+\mathbf{a}))\cdot\boldsymbol{\partial}_{\boldsymbol{\alpha}}\Phi + f_0(\mathbf{A}'+\mathbf{a},\boldsymbol{\alpha}) - \overline{f}_{\mathbf{0}}(\mathbf{A}'+\mathbf{a})
\end{aligned}\tag{2.5}$$

where a few terms have been added or subtracted (including free addition or subtraction of constants, like $\boldsymbol{\omega}_0\cdot\mathbf{a}$, $\frac{1}{2}J_o\mathbf{a}\cdot\mathbf{a}$ and $\overline{f}_0(\mathbf{a})$): so that each line $F_j, j > 0$, can *formally* be estimated to have size $\sim O(\varepsilon_0^2)$ in a smaller polydisk, while the line (0) is the new unperturbed Hamiltonian (in the $(\mathbf{A}',\boldsymbol{\alpha}')$ variables [4]). [5]

As familiar in pertubation theory this may suggest to define:

$$\begin{aligned}
&\mathbf{a} = -J_0^{-1}\boldsymbol{\partial}_{\mathbf{a}}\overline{f}_0(\mathbf{a})\\
&\Phi(\mathbf{A}',\boldsymbol{\alpha}) = -\sum_{\mathbf{0}\ne\boldsymbol{\nu}\in\mathcal{Z}^N}\frac{f_{0,\boldsymbol{\nu}}(\mathbf{A}'+\mathbf{a})}{i(\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}+((\mathbf{A}'+\mathbf{a})\cdot J_0\boldsymbol{\nu}))}e^{i\boldsymbol{\nu}\cdot\boldsymbol{\alpha}},
\end{aligned}\tag{2.6}$$

Existence of $\mathbf{a}$ inside C_{ρ_0} can be easily established via *dimensional bounds*.

- In general an example of dimensional bound is a bound of a holomorphic function $f(\zeta)$ and of its derivatives at a point ζ in terms of its maximum $|f|$ in the definition domain and of the distance $d(\zeta)$ of ζ to the boundary: the n^{th}-derivative at ζ is bounded by $n!(\max|f|)\,d(\zeta)^{-n}$. Likewise a function $g(\zeta)$ holomorphic in an annulus bounded by the circles $\zeta = e^{i\alpha\pm\kappa}, \alpha\in[0,2\pi]$, will have Fourier's coefficients bounded by its maximum ε in the annulus as $|g_{\boldsymbol{\nu}}|\le\varepsilon\exp(-\kappa\,|\boldsymbol{\nu}|)$. The mentioned bounds will be called dimensional bounds.

[4] $\boldsymbol{\alpha}$ has to be imagined as a function of $(\mathbf{A}',\boldsymbol{\alpha}')$, solution of the second implicit equation in Eq.2.4.

[5] Heuristically: line (1) is a sum of quantities of order ε_0 but it is of second order in $\mathbf{A}'$ and could be made of $O(\varepsilon_0^2)$ by suitably restricting size of $\mathbf{A}'$: remark that $\overline{f}_0(\mathbf{a})$ is a constant added freely but the last term is there because of the choice $\mathbf{a} = -J_0^{-1}\boldsymbol{\partial}_{\mathbf{a}}\overline{f}_0(\mathbf{a})$. Line (2) can be made of $O(\varepsilon_0^2)$ by choosing Φ of $O(\varepsilon_0)$. Line (3) could even be set $= 0$ and invites taking it a *definition* of Φ (if possible).

- In conclusion, if $\varepsilon = \max_{\mathcal{C}(\rho,\kappa)} |\boldsymbol{\partial}_{\mathbf{A}} F(\mathbf{A},\boldsymbol{\alpha})| + \max_{\mathcal{C}(\rho,\kappa)} \frac{1}{\rho}|\boldsymbol{\partial}_{\boldsymbol{\alpha}} F(\mathbf{A},\boldsymbol{\alpha})|$ and F is holomorphic in $\mathcal{C}(\rho,\kappa)$, a dimensional bound on $F(\mathbf{A},\boldsymbol{\alpha})$ for $(\mathbf{A},\boldsymbol{\alpha}) \in \mathcal{C}((1-\lambda)\rho,\kappa)$, $\lambda < 1$, is, for $\boldsymbol{\nu} \neq 0$:

$$|\boldsymbol{\partial}^{\mathbf{n}}_{\mathbf{A}} \boldsymbol{\partial}_{\mathbf{A}} F(\mathbf{A},\boldsymbol{\alpha})| < \mathbf{n}!\,(\lambda\rho)^{-|\mathbf{n}|}\,\varepsilon, \qquad |\boldsymbol{\nu}||F_{\boldsymbol{\nu}}(\mathbf{A},\boldsymbol{\alpha})| < e^{-\kappa|\boldsymbol{\nu}|}\,\rho\,\varepsilon \tag{2.7}$$

  where $\mathbf{n} \in \mathcal{Z}^N_+$, $|\mathbf{n}| = \sum n_i$, $\mathbf{n}! = \prod n_i!$, $\boldsymbol{\nu} \in \mathcal{Z}^N$, $|\boldsymbol{\nu}| = \sum_i |\nu_i|$. [6]

- In the following, dimensional bounds will be used repeatedly: when a bound on a function holomorphic in $\mathcal{C}(\rho,\kappa)$ is needed, it will be obtained dimensionally but restricted to a smaller polydisk $\mathcal{C}(\lambda\rho,\kappa-\delta)$, for $\lambda < 1, \delta < \kappa$ (conveniently fixed each time, *without* attention to optimize the choice to obtain better bounds).

- As a first example, a condition for the existence of the shift $\mathbf{a}$, in the first of Eq.(2.6), is ε_0 small: *if* $J^{-1}_{0,-}\varepsilon_0\rho_0^{-1} < \frac{1}{\gamma_0}$ *and* γ_0 *is large enough, the equation* $\mathbf{a} = -J_0^{-1}\boldsymbol{\partial}_{\mathbf{a}}\overline{f}(\mathbf{a})$ *has a solution* $\mathbf{a}$ *with* $|\mathbf{a}| < J^{-1}_{0,-}\varepsilon_0 < \frac{1}{\gamma_0}\rho_0$: see Appendix A where the choice $\gamma_0 = 3N$ is proposed.

However the Φ in Eq.(2.6) does not make sense (in general); *instead* it will "be approximated", and *properly defined*, by $\Phi_0(\mathbf{A}',\boldsymbol{\alpha})$ as:

$$-\sum_{\mathbf{0}\neq\boldsymbol{\nu}\in\mathcal{Z}^N} \frac{f_{0,\boldsymbol{\nu}}(\mathbf{A}'+\mathbf{a})e^{i\boldsymbol{\alpha}\cdot\boldsymbol{\nu}}}{i\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}}\Big(1 - \frac{J_0(\mathbf{A}'+\mathbf{a})\cdot\boldsymbol{\nu}}{\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}} + \Big(\frac{J_0(\mathbf{A}'+\mathbf{a})\cdot\boldsymbol{\nu}}{\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}}\Big)^2\Big) \tag{2.8}$$

From $\varepsilon_0 = \|\partial_{\mathbf{A}'} f_0\|_{\rho_0,\kappa_0} + \frac{1}{\rho_0}\|\partial_{\boldsymbol{\alpha}} f_0\|_{\rho_0,\kappa_0}$ the dimensional inequalities (Eq.2.7):

$$|\boldsymbol{\partial}_{\mathbf{A}'} f_{0,\boldsymbol{\nu}}(\mathbf{A}'+\mathbf{a})| \le \varepsilon_0 e^{-\kappa_0|\boldsymbol{\nu}|}, \qquad \rho_0|\boldsymbol{\nu}||f_{0,\boldsymbol{\nu}}(\mathbf{A}'+\mathbf{a})| < \varepsilon_0 e^{-\kappa_0\cdot|\boldsymbol{\nu}|} \tag{2.9}$$

hold if $\mathbf{A}'+\mathbf{a} \in C(\rho_0)$; and allow to estimate Φ_0 and to define properly the canonical map in Eq.2.4 as a map from $\mathcal{C}(\frac{1}{4}\rho_0,\kappa_0-2\delta_0)$ to $\mathcal{C}(\rho_0-\frac{1}{2}\rho_0,\kappa_0-\delta_0)$ for some analyticity loss parameter $0 < \delta_0 < \frac{1}{4}\kappa_0$.

This is done via a few dimensional bounds obtained by successive reduction of analiticity domains $\mathcal{C}(\rho,\kappa)$ to $\mathcal{C}(\rho',\kappa')$ conveniently chosen. [7]

---

[6] Remark that to estimate $\mathbf{A}$-derivatives the ρ has to be reduced by a factor $1-\lambda$ while to estimate $\boldsymbol{\alpha}$-derivatives κ_0 can be left unchanged, as such m-th $\boldsymbol{\alpha}$-derivatives modify the second of Eq.2.7 by a factor $|\boldsymbol{\nu}|^m$.

[7] Warning: every time a new estimate is proposed, the dimensionless constants involved in it will typically be γ_j, c_j with a new label j even when equal to constants with label $j' < j$.

If $\gamma_0 = 3N$, as above, and C is the Diophantine constant, Eq.(2.3), the

$$J_{0,-}^{-1}\varepsilon_0 < \frac{1}{\gamma_0}\rho_0, \qquad J_{0,+}\rho_0 C^{-1} < 1 \tag{2.10}$$

and the second of Eq.(2.9) imply a simple bound on $\|\Phi_0\|_{\frac{2}{3}\rho_0,\kappa_0}$ [8]. Therefore

$$\|\Phi_0\|_{\frac{2}{3}\rho_0,\kappa_0} \le \varepsilon_0\rho_0 \sum_{\boldsymbol{\nu}\neq\mathbf{0}} \frac{|\boldsymbol{\nu}|e^{-\kappa_0|\boldsymbol{\nu}|}}{|\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}|}\Big(1+\frac{|J_0\boldsymbol{\nu}|\rho_0}{|\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}|}+\Big(\frac{|J_0\boldsymbol{\nu}|\rho_0}{|\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}|}\Big)^2\Big) \le \gamma_1\frac{\varepsilon_0}{C}\,\kappa_0^{-c_1} \tag{2.11}$$

which yields

$$\begin{aligned} &||\boldsymbol{\partial}_{\mathbf{A}'}\Phi_0||_{\frac{1}{2}\rho_0,\kappa_0} + \frac{1}{\rho_0}\|\boldsymbol{\partial}_{\boldsymbol{\alpha}}\Phi_0\|_{\frac{2}{3}\rho_0,\kappa_0} \\ &\le \widetilde{\gamma}\varepsilon_0 \sum_{\boldsymbol{\nu}\neq\mathbf{0}} \frac{|\boldsymbol{\nu}|^2 e^{-\kappa_0|\boldsymbol{\nu}|}}{|\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}|}\Big(1+\frac{|J_0\boldsymbol{\nu}|\rho_0}{|\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}|}+\Big(\frac{|J_0\boldsymbol{\nu}|\rho_0}{|\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}|}\Big)^2\Big) \le \widehat{\gamma}\frac{\varepsilon_0}{C}\,\kappa_0^{-\widehat{c}} \end{aligned} \tag{2.12}$$

with $\widetilde{\gamma},\widehat{\gamma},\widehat{c}$ dimensionless constants. [9]

The bounds Eq.(2.12) lead easily, via dimensional estimates, to a precise definition of the canonical map in Eq.2.4 as a map $(\mathbf{A}',\boldsymbol{\alpha}') \in \mathcal{C}(\widetilde{\rho},\widetilde{\kappa}) \to (\mathbf{A},\boldsymbol{\alpha}) \in \mathcal{C}(\rho_0,\kappa_0)$, where $\widetilde{\rho} < \rho_0, \widetilde{\kappa} < \kappa_0$ will be suitably defined. And, once the canonical map is defined, the Hamiltonian will take the form $H_1(\mathbf{A}',\boldsymbol{\alpha}') = \boldsymbol{\omega}_0\cdot\mathbf{A}' + \frac{1}{2}\mathbf{A}'\cdot J_0\mathbf{A}' + f_1(\mathbf{A}',\boldsymbol{\alpha}')$ for $(\mathbf{A}',\boldsymbol{\alpha}')$ in the definition domain $\mathcal{C}(\widetilde{\rho},\widetilde{\kappa})$.

The 'new' perturbation will be, by Eq.(2.5), $f_1(\mathbf{A}',\boldsymbol{\alpha}') = \sum_{j=1}^3 F_j(\mathbf{A}',\boldsymbol{\alpha})$ if expressed in the *mixed variables* $(\mathbf{A}',\boldsymbol{\alpha})$. Hence it will eventually be necessary to express $\boldsymbol{\alpha}$ via $(\mathbf{A}',\boldsymbol{\alpha}')$ by inverting the second in Eq.2.4 as $\boldsymbol{\alpha} = \boldsymbol{\alpha}' + \Delta(\mathbf{A}',\boldsymbol{\alpha}')$: so that the perturbation acting in H_1 will be, properly, $f_1(\mathbf{A}',\boldsymbol{\alpha}') = \sum_{j=1}^3 F_j(\mathbf{A}',\boldsymbol{\alpha}'+\Delta(\mathbf{A}',\boldsymbol{\alpha}'))$.

To solve for the implicit function $\Delta(\mathbf{A}',\boldsymbol{\alpha}')$ in the second of Eq.2.4, an argument can be used similar to the one for the solution of $\mathbf{a} = -J_0^{-1}\boldsymbol{\partial}_{\mathbf{a}}\overline{f}_0(\mathbf{a})$ in Appendix A. In Appendix B the similar analysis is adopted to construct

[8] As $\mathbf{A}' \in C(\frac{2}{3}\rho_0)$, implies $\mathbf{A}'+\mathbf{a} \in C(\frac{5}{6}\rho_0)$: remark $|\mathbf{a}| < \frac{1}{3N}\rho_0 < \frac{1}{6}\rho_0$ for $N \ge 2$.

[9] E.g. $\widehat{c} = 4N+2$, $\widehat{g} = \widetilde{\gamma}\sum_{\boldsymbol{\nu}} |\boldsymbol{\nu}|^{4N+2}e^{-\kappa_0\,|\boldsymbol{\nu}|}$, $\widetilde{\gamma} = 7$. Remark that the bound on the second term in the first line of Eq.2.12 does not involve a change in κ_0, unlike the first term (where ρ is reduced).

$\Delta(\mathbf{A}',\boldsymbol{\alpha}')$ and a proof that it can be defined on $\mathcal{C}(\frac12\rho_0,\kappa_0-2\delta_0)$ subject to the bound:

$$\|\Delta\|_{\frac12\rho_0,\kappa_0-2\delta_0}<\overline{\gamma}\frac{\varepsilon_0}{C}\,\kappa_0^{-\overline{c}} \tag{2.13}$$

for dimensionless $\overline{\gamma},\overline{c}$, and for $\delta_0>0$ to be determined and depending on ε_0,κ_0.

However it is convenient to postpone the study of Δ and to obtain first bounds on $f_1'(\mathbf{A}',\boldsymbol{\alpha})=\sum_{j=1}^3 F_j(\mathbf{A}',\boldsymbol{\alpha})$ in a domain $\mathcal{C}(\widetilde{\rho},\kappa_0)$.

The conditions for the definitions of $F_j(\mathbf{A}',\boldsymbol{\alpha})$, *i.e.* for the definition of the generating function Φ_0 to be defined in the mixed variables $(\mathbf{A}',\boldsymbol{\alpha})$, are summarized in Eq.2.10

Under such conditions the F_j composing $f_1'(\mathbf{A}',\boldsymbol{\alpha})$, in the lines labeled $(j), j>0$ in Eq.(2.5), can be easily bounded dimensionally as follows.

- (a) F_0 is $\frac12(J_0\mathbf{A}'\cdot\mathbf{A}')+\boldsymbol{\omega}_0\cdot\mathbf{A}'$ for $\mathbf{A}'\in C(\frac12\rho_0)$.

- (b) $F_1(\mathbf{A}')=$ second order remainder of the $\mathbf{A}'$-expansion of $\overline{f}_0(\mathbf{A}'+\mathbf{a})$:

$$\|F_1\|_{\widetilde{\rho},\kappa_0}\le\widetilde{\gamma}_1\varepsilon_0\frac{\widetilde{\rho}^2}{\rho_0}\qquad \mathbf{A}'\in\mathcal{C}(\widetilde{\rho},\kappa_0),\quad \widetilde{\rho}\le\frac12\rho_0 \tag{2.14}$$

  with $\widetilde{\gamma}_1$ dimensionless. Hence $\|\partial_{\mathbf{A}'}F_1\|_{\widetilde{\rho}}\le\gamma_1\varepsilon_0\frac{\widetilde{\rho}}{\rho_0}$, $\widetilde{\rho}<\frac14\rho_0,\gamma_1=4\widetilde{\gamma}_1$.

- (c) For $\widetilde{\rho}\le\frac14\rho_0$ by Eq. 2.5 2.12 it is $|F_2|_{\frac12\rho_0,\kappa_0}<\varepsilon_0 4\rho_0\gamma_2\frac{\varepsilon_0}{C}\kappa_0^{-c_2}$ in $\mathcal{C}(\frac12\rho_0,\kappa_0)$. Hence, for γ_3,c_3 dimensionless:

$$|\partial_{\mathbf{A}'}F_2|_{\widetilde{\rho},\kappa_0}+\frac{1}{\widetilde{\rho}}|\partial_{\boldsymbol{\alpha}}F_2|_{\widetilde{\rho},\kappa_0}\le\gamma_3\varepsilon_0\frac{\rho_0}{\widetilde{\rho}}\frac{\varepsilon_0}{C}\kappa_0^{-c_3} \tag{2.15}$$

  on the two contributions to F_2 (remark that $\rho_0-\frac14\rho_0>\frac14\rho_0$ and that only the first is responsible of the factor $\frac{\rho_0}{\widetilde{\rho}}$). [10]

- (d) Finally F_3, with Φ_0 defined in Eq.(2.8), does not vanish, but *an occurring algebraic cancellation* simplifies it, under the conditions in

[10]Remark: the domain in $|\mathbf{A}'|$ has been reduced to $<\frac14\rho_0$ to permit the dimensional estimate of the derivative of the first of the two terms in F_2, as is done below also to bound F_3.

Eq.2.10 and, using Eq. 2.9 , as:

$$\begin{aligned}F_3(\mathbf{A}',\mathbf{z})&\stackrel{def}{=}(\boldsymbol{\omega}_0+J_0(\mathbf{A}'+\mathbf{a})\cdot\boldsymbol{\partial}_{\boldsymbol{\alpha}}\Phi_0+f_0(\mathbf{A}'+\mathbf{a},\boldsymbol{\alpha})-\overline{f}_0(\mathbf{A}'+\mathbf{a}))\\&=-\sum_{\mathbf{0}\neq\boldsymbol{\nu}}f_{0,\boldsymbol{\nu}}(\mathbf{A}'+\mathbf{a})\frac{(J_0(\mathbf{A}'+\mathbf{a})\cdot\boldsymbol{\nu})^3}{(\boldsymbol{\omega}_0\cdot\boldsymbol{\nu})^3}e^{i\boldsymbol{\alpha}\cdot\boldsymbol{\nu}}\le\widetilde{\gamma}\frac{\varepsilon_0}{\rho_0}\widetilde{\rho}^3\kappa_0^{-c}\end{aligned}\tag{2.16}$$

and can be bounded in $\mathcal{C}_{\widetilde{\rho},\kappa_0}$ by $\widetilde{\gamma}\frac{\varepsilon_0}{\rho_0}\widetilde{\rho}^3\kappa_0^{-c}$ for suitable dimendionless $\widetilde{\gamma},c$, and in the variables $(\mathbf{A}',\boldsymbol{\alpha})\in\mathcal{C}_{\widetilde{\rho},\kappa_0}$ as in c) above:

$$|\partial_{\mathbf{A}'}F_3|_{\widetilde{\rho},\kappa_0}+\frac{1}{\widetilde{\rho}}|\partial_{\boldsymbol{\alpha}}F_3|_{\widetilde{\rho},\kappa_0}\le\gamma_4\varepsilon_0\Big(\frac{\widetilde{\rho}}{\rho_0}\Big)^2\kappa_0^{-c_4}\tag{2.17}$$

where γ_4,c_4 are dimensionless constants and $\widetilde{\rho}\le\frac14\rho_0$.

Therefore if $\widetilde{\rho}\le\frac14\rho_0,\widetilde{k}\le\kappa_0$ and if Eq.2.10 holds:

$$\begin{aligned}&J_1=J_0,\qquad|\boldsymbol{\partial}_{\mathbf{A}'}F_1(\mathbf{A}')|\le\overline{\gamma}\varepsilon_0\Big(\frac{\widetilde{\rho}}{\rho_0}\Big),\\&|\boldsymbol{\partial}_{\mathbf{A}'}F_2(\mathbf{A}',\boldsymbol{\alpha})|+\frac{1}{\widetilde{\rho}}|\boldsymbol{\partial}_{\boldsymbol{\alpha}}F_2(\mathbf{A}',\boldsymbol{\alpha})|\le\overline{\gamma}\varepsilon_0\frac{\varepsilon_0}{C}\frac{\rho_0}{\widetilde{\rho}}\kappa_0^{-\overline{c}},\\&|\boldsymbol{\partial}_{\mathbf{A}'}F_3(\mathbf{A}',\boldsymbol{\alpha})|+\frac{1}{\widetilde{\rho}}|\boldsymbol{\partial}_{\boldsymbol{\alpha}}F_3(\mathbf{A}',\boldsymbol{\alpha})|\le\overline{\gamma}\varepsilon_0\Big(\frac{\widetilde{\rho}}{\rho_0}\Big)^2\kappa_0^{-\overline{c}}\end{aligned}\tag{2.18}$$

for $\overline{\gamma},\overline{c}$ constants (dimensionless), $(\mathbf{A}',\boldsymbol{\alpha})\in\mathcal{C}(\widetilde{\rho},\kappa_0)$.

Set $F'_j(\mathbf{A}',\boldsymbol{\alpha}')=F_j(\mathbf{A}',\boldsymbol{\alpha}'+\Delta(\mathbf{A}',\boldsymbol{\alpha}'))$; the size ε_1 of $f_1(\mathbf{A}',\boldsymbol{\alpha}')$, can be easibly bounded, in $\mathcal{C}(\frac14\rho_0,\kappa_0-2\delta_0)$ with $\delta_0\le\overline{\gamma}(\frac{\varepsilon_0}{C})(\kappa_0)^{-\overline{c}}<\frac12\kappa_0$, by composing the $\boldsymbol{\alpha}$-derivatives of the $F_j(\mathbf{A}',\boldsymbol{\alpha})$ with the $\mathbf{A}'$ and $\boldsymbol{\alpha}'$-derivatives of $\Delta_j(\mathbf{A}',\boldsymbol{\alpha}')$. Via Eq.(2.13) it is seen that the bounds on $\partial_{\mathbf{A}'}F'_j(\mathbf{A}',\boldsymbol{\alpha}')$ and $\partial_{\boldsymbol{\alpha}'}F'_j(\mathbf{A}',\boldsymbol{\alpha}')$ *remain the same* as those on $\partial_{\mathbf{A}'}F_j(\mathbf{A}',\boldsymbol{\alpha})$ and $\partial_{\boldsymbol{\alpha}}F_j(\mathbf{A}',\boldsymbol{\alpha})$ provided the constants γ_j,c_j, introduced above, are increased to a suitably larger pair $\overline{\gamma},\overline{c}$. See Appendix B . [11]

Thus the corresponding $F'_j(\mathbf{A}',\boldsymbol{\alpha}')\stackrel{def}{=}F_j(\mathbf{A}',\boldsymbol{\alpha})$ admit the same bounds as Eq.2.18, but in the smaller domain $\mathcal{C}(\frac12\rho_0,\kappa_0-2\delta_0)$ with $\delta_0\le\overline{\gamma}(\frac{\varepsilon_0}{C})\kappa_0^{-\overline{c}}<\frac12\kappa_0$, by simply suitably increasing the constants $\overline{\gamma},\overline{c}$, as discussed in AppendixB .

[11] Remark that $J_{0,\pm},C$ are given and, without loss, ρ_0,κ_0 can be reduced at will, so that Eq.(2.10) is really a condition on ε_0.

Hence the Eq.2.18 considered as estimates of the derivatives of $F'_j(\mathbf{A}',\boldsymbol{\alpha}')$ (with the larger $\overline{\gamma},\overline{c}$ and $\delta_0=\overline{\gamma}(\frac{\varepsilon_0}{C})\kappa_0^{-\overline{c}}<\frac12\kappa_0$) lead immediately to an estimate of the size $\varepsilon_1=|\partial_{\mathbf{A}'}f_1|_{\widetilde{\rho},\widetilde{\kappa}}+\frac{1}{\widetilde{\rho}}|\partial_{\boldsymbol{\alpha}'}f_1|_{\widetilde{\rho},\widetilde{\kappa}}$ of the perturbation $f_1(\mathbf{A}',\boldsymbol{\alpha}')=\sum_{j=1}^3 F_j(\mathbf{A}',\boldsymbol{\alpha}'+\Delta(\mathbf{A}',\boldsymbol{\alpha}'))$: namely $\varepsilon_1\le C\gamma(\frac{\varepsilon_0}{C})^{\frac32}\kappa_0^{-c}$.[12] See Appendix B for details.

Given $\rho_0,\kappa_0,n\ge1,\lambda=\frac12$ define, for $\overline{\gamma},\overline{c}$, as in Eq.2.18:

$$\rho_n=\Big(\frac{\varepsilon_{n-1}}{C}\Big)^{\lambda}\rho_{n-1},\quad \kappa_n=\kappa_{n-1}-2\delta_{n-1},\quad \delta_n=\overline{\gamma}\frac{\varepsilon_n}{C}\kappa_n^{-\overline{c}}\tag{2.19}$$

Iterate the renormalization transformation $(\mathbf{A}_{n-1},\boldsymbol{\alpha}_{n-1})\to(\mathbf{A}'=\mathbf{A}_n,\boldsymbol{\alpha}'=\boldsymbol{\alpha}_n)$ transforming the Hamiltonian $H_{n-1}(\mathbf{A}_{n-1},\boldsymbol{\alpha}_{n-1})$, on $\mathcal{C}(\rho_{n-1},\kappa_{n-1})$ with pertubation f_{n-1}, into $H_n(\mathbf{A}_n,\boldsymbol{\alpha}_n)$, on $\mathcal{C}(\rho_n,\kappa_n)$ with pertubation f_n. Suppose inductively, for all $n\ge0$:

$$\gamma_0\varepsilon_n J_{0,-}^{-1}<\rho_n,\qquad C^{-1}\rho_n\ J_{0,+}<1,\qquad \kappa_n>\frac12\kappa_0\tag{2.20}$$

Then, for $(\mathbf{A}',\boldsymbol{\alpha}')\in\mathcal{C}(\rho_n,\kappa_n)$, the $\varepsilon_n=\|\boldsymbol{\partial}_{\mathbf{A}'}f_n\|_{\rho_n,\kappa_n}+\frac{1}{\rho_n}\|\boldsymbol{\partial}_{\boldsymbol{\alpha}}f_n\|_{\rho_n,\kappa_n}$ are bounded, *if* $\kappa_n>\frac12\kappa_0$, by:

$$\frac{\varepsilon_n}{C}=\gamma\Big(\frac{\varepsilon_{n-1}}{C}\Big)^{\xi}(\frac{\kappa_0}{2})^{-c},\qquad \kappa_n=\kappa_{n-1}-\delta_{n-1}\tag{2.21}$$

with γ,c dimensionless, $\xi=\frac32$.

The recursion, Eq. 2.19 , for $\frac{\varepsilon_n}{C}$ can be iterated for $n\ge1$, *as long as the conditions Eq.2.10 remain valid*, finding $\frac{\varepsilon_n}{C}\le\Lambda(\frac{\varepsilon_{n-1}}{C})^{\xi}$ for $\Lambda=\overline{\gamma}(\frac{\kappa_0}{2})^{-\overline{c}}$ and $\gamma,c,\xi=\frac32$ dimensionless constants. If $\eta_n\stackrel{def}{=}\sigma\frac{\varepsilon_n}{C}$ with $\sigma=\Lambda^2$ this becomes simply $\eta_n\le\eta_0^{\xi^{n-1}}$, and consequently $\rho_n\le\Lambda^{-\frac{n(n+1)}{2}}\eta_0^{\xi^n-1}\rho_0=O(\eta_0^{\xi^n})\rho_0=O((\Lambda^2\frac{\varepsilon_0}{C})^{\xi^n})\rho_0$.

Then Eq.2.19 imply that $\kappa_n=\kappa_0-2\sum_{k=1}^n\overline{\gamma}\frac{\varepsilon_{n-k}}{C}(\frac12\kappa_0)^{-\overline{c}}$ is bounded by $\kappa_n=\kappa_0-2\sum_{k=0}^{n-1}\Lambda^{-1}\eta_k$: therefore the conditions in Eq.2.20 can be reproduced if $\eta_0=\sigma\frac{\varepsilon_0}{C}=\Lambda^2\frac{\varepsilon_0}{C}$ is small enough.

The result is that in the coordinates $(\mathbf{A}',\boldsymbol{\alpha}')\stackrel{def}{=}(\mathbf{A}_n,\boldsymbol{\alpha}_n)$, up to a time scale $\sim\boldsymbol{\omega}_0^{-1}O(\Lambda^2\Big(\frac{\varepsilon_0}{C}\Big)^{-\xi^n})$ and in region of phase space of order $\rho_n\sim\rho_0\Big(\Lambda^2\frac{\varepsilon_0}{C}\Big)^{\xi^n}$,

[12] Remark that here it is necessary to regard the F'_j as defined in $\mathcal{C}(\widetilde{\rho},\widetilde{\kappa})$ with $\widetilde{\kappa}=\kappa_0-2\delta_0$ to make use of dimensional bounds on $\Delta(\mathbf{A}',\boldsymbol{\alpha}')$ implied by Eq.2.13.

the motion is quasi periodic motion with spectrum $\boldsymbol{\omega}_0$ and Hamiltonian H_n with perturbation size $\sim \varepsilon_n \sim (\Lambda^2\frac{\varepsilon_0}{C})^{\xi^n}$.

The conclusion is that analyticity in the $\mathbf{A}$ variables is lost: in the limit $n\to\infty$ remains a quasi periodic motion confined for time $\in(0,+\infty)$ on the real (and analytic) torus to which the approximations $\mathbf{A}'_n, \boldsymbol{\alpha}'_n$ lead.

- It is also possible to define a sequence of maps $\widetilde{\mathcal{K}}_n$ defined in the *fixed domain* $\mathcal{C}_{\frac14\rho_0,\frac14\kappa_0}$ by rescaling the action coordinates of the polydisks by a factor $(\frac{\varepsilon_{n-1}}{C})^{-\frac12} = \rho_{n-1}/\rho_n$, $n\ge 1$ so that they are all turned into $\mathcal{C}_{\frac14\rho_0,\frac14\kappa_0}$: calling $\widetilde{\eta}_n \stackrel{def}{=} \frac{\varepsilon_{n-1}}{C}$ the rescaling transformation will rescale time by $\widetilde{\eta}_n^{-\frac12}$ and will change $\mathbf{A}_n$ into $\mathbf{A}'_n = \widetilde{\eta}_n^{-\frac12}\mathbf{A}_n$ and change the Hamiltonian into:

$$\widetilde{H}_n = \boldsymbol{\omega}_0\cdot\mathbf{A}'_n + \widetilde{\eta}_n^{\frac12}\frac12(\mathbf{A}'_n\cdot J_0\mathbf{A}'_n) + \widetilde{\eta}_n^{-\frac12} f_n(\widetilde{\eta}_n^{\frac12}\mathbf{A}'_n, \boldsymbol{\alpha}'_n) \xrightarrow[n\to\infty]{} \boldsymbol{\omega}_0\cdot\mathbf{A}'_\infty \quad (2.22)$$

  and in the rescaled variables the sizes of the anharmonic terms tend to 0 superexponentially, taking into account the recursion defined in Eq.(2.20) which implies that the size of f_n is of order $\frac{\varepsilon_n}{C} = (\widetilde{\eta}_n)$.

- Hence the perturbation f_0 and the twist J_0 are, after renormalization, "irrelevant" operators (in Eq.(2.17) both tend to 0 as $n\to\infty$), while the harmonic oscillator $\boldsymbol{\omega}_0\cdot\mathbf{A}$ is a "fixed point": in some sense the transformation has the harmonic oscillator as an *attractive fixed point.* This completes a proof of the KAM theorem, *interpreted in the Renormalization Group frame* [4, 20, 21, 22]: it can be classified as a "*super-renormalizable*" problem, as it *requires only a second order perturbation analysis around the trivial fixed point*, Eq.(2.7).

Remarks:
(1) The analysis of this section follows closely [23] correcting a few dimensional typos.

It is possible to redefine $J_1 \equiv J_0 + \frac12\partial^2_{\mathbf{a}}\overline{f}(\mathbf{a})$ and correspondingly add to F_1, in Eq.(2.5), the second order term in $\mathbf{A}'$ namely $-\mathbf{A}'\cdot\frac12\partial^2_{\mathbf{a}}\overline{f}(\mathbf{a})\mathbf{A}'$. This modifies F_0 in the first line of Eq. 2.5 replacing J_0 with J_1 and the improves the bound of F'_1 to: $|F'_1(\mathbf{A}')| < \gamma_1\varepsilon_0(\frac{\widetilde{\rho}}{\rho_0})^2$. Then defining $\rho_n = (\frac{\varepsilon_0}{C})^\lambda\rho_{n-1}$ with $\lambda=\frac13$ it follows, as above, $\frac{\varepsilon_n}{C} < \Lambda(\frac{\varepsilon_0}{C})^{\frac43}$ (as, in this case, the best choice of λ is $\lambda=\frac13$ so that $\xi=\frac43>\frac32$).

In the $\mathbf{A}_n, \boldsymbol{\alpha}_n$ the motion is controlled by a quadratic Hamiltonian $\boldsymbol{\omega}_0\mathbf{A}_n + \frac{1}{2}\mathbf{A}_n J_n \mathbf{A}_n$ in a portion of action space $O((\frac{\varepsilon_n}{C})^{\frac{1}{3}}\rho_0)$ "much larger" compared to the previous case, where the size was $O(\frac{\varepsilon_0}{C})^{\frac{1}{2}}\rho_0$, but $J_n = J_0$ was simpler.

The implication is that up a time scale of order $\omega_0^{-1}(\frac{\varepsilon_0}{C})^{-\frac{4}{3}n}$ the motion is quasi periodic in a domain of size $\rho_0(\frac{\varepsilon_0}{C})^{(\frac{1}{3})^n}$ controlled by Hamiltonian with quadratic part $J_n \neq J_0$. The previous choice gave control up to a time shorter $\omega_0^{-1}(\frac{\varepsilon_0}{C})^{-(\frac{3}{2})^n}$ controlled by the simpler Hamiltonian with quadratic part $J_n = J_0$ but in a domain of much smaller size $\rho_0(\frac{\varepsilon_0}{C})^{(\frac{1}{2})^n}$.
(2) The role of the constant C is essential, of course, but it should be remarked that it appears in the smallness condition of $\frac{\varepsilon_0}{C}$ and in $J_{0,+}\rho_0 C^{-1} < 1$: both conditions can be fulfilled if $C = O(\varepsilon_0^{\xi})$ with $\xi < 1$. Question: could this be of interest (aside the simple application in the remark in the following Sectiom, see 3.Sec.3).
(3) The analysis of the first iteration is extremely simplified if f_0 depends only on $\boldsymbol{\alpha}$: this is a *Thirring model* and in Eq.2.5 it is $\mathbf{a} = 0$, $F_1 = 0$, F_2 contains only the quadratic term but F_3 still has to be defined via Eq.2.8 so that at the next iteration dependence on $\mathbf{A}'$ is introduced in f_1. Nevertheless the above proof remains valid and yields existence of the invariant torus with spectrum $\boldsymbol{\omega}_0$ *without* any condition on the lower bound $J_{0,-}$ as long as it is > 0. The invariant tori in Thirring models are called "twistless" as no upper bound on $J_{0,-}$ is necessary: for such models a direct proof based on the explicit computation of a series, "Birkhoff series", converging to the parametric equations of the invariant torus, can be derived, [24] , and covers also some cases in which $J_{0,-} = 0$. [13]

3 FPUT

The FPUT digital work (reprinted in [25]), as soon as the contemporary KAM theorem became widely known, was qualitatively "explained", at least by several physicists, [5], by the remark that the observed motions were

[13] Which do not fit immediately in the above proof: the latter can be adapted to the case $H = \boldsymbol{\omega}_0 \cdot \mathbf{A} + \frac{1}{2}\mathbf{A}^{\perp} \cdot J_0 \mathbf{A}^{\perp} + f_0(\mathbf{A}^{\perp}, \boldsymbol{\alpha})$ where $\mathbf{A}^{\perp} = (A_1, A_2, \ldots, A_p, 0, \ldots 0)$ if $J_{0,j} > 0$ for $j \leq p$ and $(J_0)_j = 0$ for $j > p$, which generalizes the Thirring models and is even an easy extension of the twistless Hamiltonians theory (as the perturbation may now depend also on $\mathbf{A}^{\perp}$): the extension is also implied by the analysis in [24] .

generated by a "small" perturbation of the simple harmonic chain and could *probably* be treated via the results of the KAM theory: which, at the same time, was leading to a deeper understanding of the three body problem. The latter problem, as well as the FPUT chain, were in fact both perturbations of integrable systems, although affected by "resonances".

However it took a long time to develop a complete proof that low energy motions observed in some FPUT experiments should mostly be quasi periodic, [26].

The models originally chosen in FPUT had Hamiltonian:

$$H = \sum_{j=1}^{N-1} \frac{1}{2} p_j^2 + W_{\alpha,\beta}(q_j - q_{j+1}), \qquad W_{\alpha,\beta}(x) = \frac{1}{2}x^2 + \alpha x^3 + \beta x^4 \tag{3.1}$$

with *fixed endpoints* $q_0 = 0, q_N = 0$, [5]. If $\beta = 0$ or $\alpha = 0$ the models are respectively called fput-α-chain or fput-β-chain.

The α-chain, as a dynamical system, presents the difficulty that $W_{\alpha,0}(x)$, is unbounded below, if $\alpha > 0$, or above if $\alpha < 0$: hence it is not even clear whether all or, at least, some of the motions remain in a bounded phase space region (in a finite or infinite time). The problem does not arise in the β-chains if $\varepsilon > 0$ as it is always supposed here.

In the original FPUT work nine choices of α and β are considered and the cases numbered $4, 5, 6, 7, 9$ are apparently β models ($i.e.$ $\alpha = 0, \beta > 0$) while those numbered $1, 2, 3, 8$ are α models ($\alpha > 0, \beta = 0$). In all simulations the motions kept a total energy constant, "within 1% or so".

The moving particles coordinates, $N - 1$ in number, can be expressed as $q_k = \frac{\sqrt{2}}{\sqrt{N}} \sum_{h=1}^{N-1} \sin(\frac{\pi h k}{N}) Q_h$ and $p_k = \frac{\sqrt{2}}{\sqrt{N}} \sum_{h=1}^{N-1} \sin(\frac{\pi h k}{N}) P_h$ for $k = 0, \dots, N$.

The well known theory of the vibrating string with elastic potential energy $\sum_{k=0}^{N-1} \frac{1}{2}(q_k - q_{k+1})^2 = \frac{1}{2} \sum_{h=1}^{N-1} \omega_h^2 Q_h^2$ induces to study its oscillations via Forurier's transforms coordinates $(\mathbf{P}, \mathbf{Q})$ expressed as functions of $(\mathbf{p}, \mathbf{q})$ to which are canonically conjugated. If $\omega_h = 2 \sin \frac{\pi h}{N}$, therefore:

$$H_0(\mathbf{p}, \mathbf{q}) = H_0(\mathbf{P}, \mathbf{Q}) = \sum_{k=1}^{N-1} \frac{1}{2}(P_k^2 + \omega_k^2 Q_k)^2 \stackrel{def}{=} \sum_{k=1}^{N-1} \omega_k a_k \tag{3.2}$$

where $a_k = \frac{(P_k^2 + \omega_k^2 Q_k^2)}{2\omega_k}$. The latter is a natural coordinate as the pairs $(a_k, \varphi_k) = (a_k, \arcsin(\frac{\omega_k Q_k}{\sqrt{2\omega_k a_k}}))$ are *action-angle* variables and evolve, in the elastic chain, as $(a_k,\varphi_k)\to(a_k,\varphi_k+\omega_k t)$; and $(\mathbf{p},\mathbf{q})\longleftrightarrow(\mathbf{P},\mathbf{Q})\longleftrightarrow(\mathbf{a},\boldsymbol{\varphi})$ are symplectic maps.

Therefore $Q_k=(\frac{2a_k}{\omega_k})^{\frac12}\sin\varphi_k$ and, as an example, the fput-β-chain Hamiltonian can be written, if $s_h=\sin\frac{\pi h}{N}, c_h=\cos\frac{\pi h}{N}$:

$$\begin{aligned}
&H_\varepsilon(\mathbf{p},\mathbf{q})=\sum_{k=1}^{N-1}\omega_k\alpha_k+\varepsilon\sum_{\mathbf{h}}T_{\mathbf{h}}\prod_{i=1}^{4}Q_{h_i}=\boldsymbol{\omega}_k\cdot\mathbf{a}+\varepsilon f_0(\mathbf{a},\boldsymbol{\varphi})\\
&T_{\mathbf{h}}=(\frac{2}{N})^2\sum_{k=1}^{N-1}\prod_{j=1}^{4}(s_{h_jk}(1-c_{h_j})-c_{h_jk}s_{h_j}),\\
&f_0(\mathbf{a},\boldsymbol{\varphi})=\frac{1}{2^4}\sum_{\mathbf{h}\in Z^4,\boldsymbol{\sigma}}T_{\mathbf{h}}\prod_{j=1}^{4}\sigma_j e^{i\sigma_j\varphi_{h_j}}(\frac{2a_{h_j}}{\omega_j})^{\frac12},
\end{aligned}\tag{3.3}$$

where $\boldsymbol{\sigma}=(\sigma_j)_{j=1}^4,\sigma_j=\pm1$ and the variables Q_k have been expressed in terms of the $\mathbf{a},\boldsymbol{\varphi}$.

The first mathematical results, [17], dealt with the β-chain with fixed extremes $q_0=q_N=0$, considered as a perturbation of order ε of the elastic chain with $N-1$ particles (*i.e.* of the chain with $\alpha=0,\beta=0$).

Perturbation theory can be applied to H_ε: just look for a canonical transformation generated by a function $\varepsilon\Phi(\mathbf{a}',\boldsymbol{\varphi})$ as $\mathbf{a}=\mathbf{a}'+\varepsilon\boldsymbol{\partial}_{\boldsymbol{\varphi}}\Phi(\mathbf{a}',\boldsymbol{\varphi}),\boldsymbol{\varphi}'=\boldsymbol{\varphi}+\varepsilon\boldsymbol{\partial}_{\mathbf{a}'}\Phi(\mathbf{a}',\boldsymbol{\varphi})$ which, if possible, transforms $H_\varepsilon(\mathbf{A},\boldsymbol{\varphi})$ into $H_\varepsilon(\mathbf{a}',\boldsymbol{\varphi})$ (where $\boldsymbol{\varphi}$ is imagimed expressed as function of $(\mathbf{a}',\boldsymbol{\varphi}')$).

$$H_\varepsilon(\mathbf{a}',\boldsymbol{\varphi})=\boldsymbol{\omega}\cdot\mathbf{a}'+\varepsilon\boldsymbol{\omega}\cdot\boldsymbol{\partial}_{\boldsymbol{\varphi}}\Phi(\mathbf{a}',\boldsymbol{\varphi}),\boldsymbol{\varphi})+\varepsilon f_0(\mathbf{a}'+\varepsilon\boldsymbol{\partial}_{\boldsymbol{\varphi}}\Phi,\boldsymbol{\varphi})\tag{3.4}$$

Naturally Φ would be chosen so that $H_\varepsilon(\mathbf{a}',\boldsymbol{\varphi})$ is $\boldsymbol{\omega}\cdot\mathbf{a}'$. Calculating up to $O(\varepsilon^2)$ this means that Φ should be such that $\varepsilon\boldsymbol{\omega}\cdot\boldsymbol{\partial}_{\boldsymbol{\varphi}}\Phi(\mathbf{a}',\boldsymbol{\varphi})+\varepsilon f_0(\mathbf{a}',\boldsymbol{\varphi})$ vanishes. However even this might be impossible if there are $\boldsymbol{\nu}\in Z^{N-1}$ such that $\boldsymbol{\omega}\cdot\boldsymbol{\nu}=0$.

Therefore the best that can be done is to subtract from the above f_0 the terms which, expressed in the Fourier's series, have "resonant" harmonics (orthogonal to $\boldsymbol{\omega}$ but $\neq0$) and define Φ to cancel, at lowest order, the part of f_0 left. So Φ can be chosen so that the new Hamiltonian (once laboriously expressed in the variables $(\mathbf{a}',\boldsymbol{\varphi}')$) will contain, up to order ε^2, only 'resonant harmonics' (including $\boldsymbol{\nu}=0$).

The new Hamiltonian will have the form $H_\varepsilon(\mathbf{a}',\boldsymbol{\varphi}')=\boldsymbol{\omega}\cdot\mathbf{a}'+\varepsilon\overline{H}_1(\mathbf{a}',\boldsymbol{\varphi}')+\varepsilon^2H_2(\mathbf{a}',\boldsymbol{\varphi}')$ where $\overline{H}_1$ is entirely resonant.

Repeating the argument s-times the Hamiltonian is represented by $H_\varepsilon = H_2 + \sum_{k=1}^{s} \varepsilon^s \overline{H}_s + \varepsilon^{s+1} H_{\varepsilon,s+1}$ where the $\overline{H}_s$ contain only resonant terms, which remain the same if s is increased. The series obtained in this way is the "Birkhoff series", which however is generically divergent as shown by Poincaré, [27] .

However truncating the sum to an order $s \ge 1$ and neglecting the $\varepsilon^{s+1} H_{\varepsilon,s+1}$ might turn out to be a simpler system. This happens in the present case if N is prime or is not a power of 2, as pointed out in [17], simply because there are no $\boldsymbol{\nu} \neq 0$ with $\boldsymbol{\omega}_0 \cdot \boldsymbol{\nu} = 0$ in the Fourier's expansion of the potential (*i.e.* no resonances $\boldsymbol{\nu}$ with $|\boldsymbol{\nu}| \le 4$). The terms of order $\le \varepsilon^2$ in $\overline{H}_1, \overline{H}_2$ not only *do not depend* at all on the angles $\boldsymbol{\varphi}'$, but $\overline{H}_1 = 0$ and the terms in $\overline{H}_2$ can be collected in the form $\varepsilon^2 \mathbf{a}' \cdot J \mathbf{a}'$, where J is explicitly computed in terms of the spectrum $\boldsymbol{\omega}_0$ of the elastic chain and furthermore it is det $J \neq 0$.

Hence the Hamiltonian $\overline{H}_\varepsilon = \boldsymbol{\omega} \cdot \mathbf{a}' + \varepsilon^2 \overline{H}_2(\mathbf{a}')$ is integrable and, as computed in [17], is found to be:

$$\overline{H}_\varepsilon(\mathbf{a}') = \sum_{k=1}^{N-1} \omega_k a'_k + \frac{\varepsilon^2}{2N} \Big(\sum_{h,k}^{1,N-1} \frac{\omega_h \omega_k}{8} \; a'_h a'_k - \sum_k \frac{\omega_k^2}{32} {a'_k}^2 \Big) \tag{3.5}$$

The surprising simplicity of this normal form, and the $\omega_h \neq 0, h = 1, \ldots, N-1$, immediately allows to check that $\boldsymbol{\omega} \cdot \mathbf{a} + \frac{\varepsilon}{2} \mathbf{a} \cdot J \mathbf{a}$ has a Jacobian matrix J constant and det $J \neq 0$.

The *few* values of N, *i.e.* other than a prime or a power of 2, which were still missing has been solved in [26] where a careful analysis shewed that no restriction should be set on the N of fput-β-chains with fixed extremes, and identified the basic reason for the existence of an integrable Birkhoff transform permitting to avoid the issue of resonances affecting non trivially the Birkhoff transform, see also [28] .

The difficulty remaing after [17] was that, given N, it was not excluded that resonant harmonics were absent in the Fourier expansion of H_2: the only way out seemed to compute the corresponding Fourier's coefficients. A task that could be performed (as it was done as time elapsed) also for single choices of N, but not covering all values of N.

The breakthrough idea, [18], is to *forget* (temporarily) the computation of the normal form and to check that the resonant harmonics, that could contribute resonances to the normal form, could not be present because they break symmetries of the Hamiltonian (which are preserved in the normal form construction). Thus the normal form can be constructed correctly as if the assumption in [17] could be valid (although it was known to be false in few interesting cases: for examples (with $N = 15, 21$) see [28, Sec.8.5] . The symmetries are briefly presented in AppendixC.

The spectra of the motions of $\overline{H}_\varepsilon(\mathbf{a}')$ will have the form $\boldsymbol{\omega} = \boldsymbol{\omega}_0 + \varepsilon J\mathbf{a}'$ and KAM can be applied to perturbations of the considered integrable model. Let Ω be an open region in phase space, in which the $\overline{H}_\varepsilon(\mathbf{a}')$ is defined: then Ω is covered by the tori corresponding to the actions $\mathbf{a}$: *e.g.* $0 < H_\varepsilon(\mathbf{a}') < E$.

The det $J \neq 0$ implies that the set in Ω of invariant tori for whose spectra verify the inequality $|\boldsymbol{\omega} \cdot \boldsymbol{\nu}| > C|\boldsymbol{\nu}|$ fails, for some $0 \neq \boldsymbol{\nu} \in Z^N$, has volume $< C\varepsilon^{-2}\kappa|\Omega|$, for a constant $\kappa > 0$: hence the for almost points in Ω the spectrum $\boldsymbol{\omega}$ verifies a Diophantine inequality.

Hence under a perturbation of $\overline{H}_\varepsilon$ of size η a large part of phase space will remain covered by invariant tori quasi periodically run with $C_\varepsilon = C\varepsilon^2$-Diophantine frequencies if C_ε is large enough compared to η. The β-chain being approximated to order $\eta = \varepsilon^2$ by the above integrable system $\overline{H}_\varepsilon(\mathbf{a}')$ will therefore have a large part of a set Q in phase space covered by $C_\varepsilon = C\varepsilon^2$-Diophantine tori. In particular this applies to the set Q of energy close to the ground state of the linear chain where $H(\mathbf{p}, \mathbf{q}) < E$ (if $E > 0$ and ε is small enough), and the fraction of Q covered by invariant tori will approach 1 as $\varepsilon \to 0$ at fixed $E > 0$.

Remark: (1) The statement that KAM implies that at small energy $H_\varepsilon(\mathbf{p}, \mathbf{q}) < E$ "most phase space points are on invariant tori" requires some care. Describing the system in the coordinates $(\mathbf{a}', \boldsymbol{\varphi}')$ the set $\overline{H}_\varepsilon(\mathbf{a}')$ depends on ε: but the application of KAM to $\overline{H}_\varepsilon(\mathbf{a}')$ imposes a condition on the same ε. The remark at the end of Sec.2 is useful here, as it stresses that if the KAM theory is applied to cases in which the Diophantine constant has the form $C\varepsilon^{-1}$ the KAM theorem holds provided the perturbation size $\eta\varepsilon^{-1}$ with h small enough.

(2) About the α-model it is possible that, applying perturbation theory to the fput-α-chain, an equivalent Hamiltonian is obtained which is an $O(\varepsilon^2)$ perturbation of an integrable Hamiltonian with J_ε still of $O(\varepsilon)$: but this is not (yet?) proved, [28]. See also the heuristic analysis in Section4.

(3) The approach to the β-model at fixed extremes can be applied to consider the Toda lattice chains with periodic or fixed boundary conditions as approximations to the fput-β-chains models. The above analysis, [26], suggests to try to use integrability of the Toda lattice to play the pertubation

theory role in the search of an integrable Birkhoff approximation.

The suggestion is to consider the fput-Hamiltonians as perturbations of the Toda-lattice and consequences of integrability of several forms of its Hamiltonian:

$$
\begin{aligned}
&H(\mathbf{p},\mathbf{x})=\sum_{j=1}^{N-1}\frac{1}{2}p_j^2+\sum_{j=0}^{N-1}W_T(q_j-q_{j+1})\\
&W_T(x)=\frac{1}{\varepsilon^2}e^{\varepsilon x}=\frac{1}{\varepsilon^2}(1+\varepsilon x)+\frac{1}{2!}x^2+\frac{\varepsilon}{3!}x^3+\frac{\varepsilon^2}{4!}x^4+\frac{\varepsilon^3}{5!}x^5+\ldots
\end{aligned}
\tag{3.6}
$$

integrable for instance with boundary conditions $q_0=q_N$, *periodic* lattice, [29, 30, 31], or $q_0=-\infty, q_N=+\infty$, lattice with *endpoints fixed at* $\pm\infty$, [32], or with fixed endpoints $q_0=q_N=0$.

In a sense the idea is that the Toda chains might play the role of an integrable Birkhoff normal form, of which fput-chains can be considered a perturbation. In the next section the basic results, [31], on the normal form of the periodic Toda-chains are summarized and heuristically applied to the fput-chains.

4 Toda and FPUT: heuristics

The last suggestion in Sec.3 is to try to take advantage of the very detailed studies on the Toda chain described below. For this purpose a periodic Toda Hamiltonian will be written, [31]:

$$
H(\mathbf{p},\mathbf{q})=\sum_{k=1}^{N}\Big(\frac{1}{2}p_k^2+\alpha^2e^{q_k-q_{k+1}}\Big)=\sum_{k=1}^{N}(\frac{1}{2}p_k^2+\alpha^2V_T(q_k-q_{k+1}))
\tag{4.1}
$$

with $q_1=q_{N+1}$. This is obtained by a canonical transformation on Eq.3.2: $p=\varepsilon p'$, $q=\frac{1}{\varepsilon}q'$ followed by a rescaling of time $t'=\varepsilon^2t$ leading to Eq.3.4, where α is a new coupling equal to ε^{-2}.

The *periodic* Toda lattice model is studied in [31] where $2N$ global canonical action-angle variables $(\mathbf{a},\boldsymbol{\varphi})\in(R_+^N,[0,2\pi]^N)$ are constructed and evolve at time t into $(\mathbf{a},\boldsymbol{\varphi}+\boldsymbol{\omega}(\mathbf{a})t)$, following a Hamiltonian $H_\alpha(\mathbf{a})$. Near $\mathbf{a}=0$ it is $\omega_i(\mathbf{a})=2\alpha\omega_{0,i}+\frac{1}{4N}a_i+O(|\mathbf{a}|^2)$, if $\boldsymbol{\omega}_0$ is the spectrum of the linear oscillations $\omega_{0,i}=2\sin\frac{\pi i}{N}$.

The $H_\alpha(\mathbf{a})$, [31, th.1.4], is *globally defined* for $\mathbf{a}\in R_+^N$ and the Jacobian $J_\alpha(\mathbf{a})=\det\frac{\partial\boldsymbol{\omega}(\mathbf{a})}{\partial\mathbf{a}}$, 3, is analytic in $\alpha>0$; for $\alpha>0$ it is strictly convex in $\mathbf{a} \in R_+^N$ and $J_\alpha(\mathbf{a})_{ij} = \frac{1}{4N}\delta_{ij} + O(|\mathbf{a}|) \neq 0$ and the spectrum is $\omega_i = 2\alpha\sin(\frac{i\pi}{N}) + \frac{1}{4N}a_i + O(|\mathbf{a}|^2)$.

The general KAM theory of Sec.2 can be applied, as in Sec.3, to perturbations of the Toda-model because the Jacobian J_α strict convexity for α in any interval bounded away from 0 implies that the spectrum has the Dipohantine property: so pertubations of size η will leave most of the invariant tori only slightly deformed if η is small.

However application to the fput-chains requires more analysis mainly because the use of the change of variables leads to the Hamiltonian in Eq.4.1 with α small and compare it with Eq.3.6 ε small: aside from the time rescaling, this means to compare motions with α in Eq.4.1 to motions with ε in Eq.3.6 setting $\alpha = \varepsilon^{-2}$; so when α is very large.

The analysis in [31] is global and the coordinates $(\mathbf{a}, \boldsymbol{\varphi})$ exist for all $\alpha > 0$ and transform analytically the Hamiltonian Eq.4.1 into a function $\widetilde{H}_\alpha(\mathbf{a})$ with convex Jacobian $J_{ij} = \frac{\partial^2_{a_i} H(\mathbf{a})}{\partial^2 a_j}$, see th.(1.1),th.(1.2) in [31]. Hence in the region of couplings relevant for the fput-chains, *i.e.* ε close to 0, $\alpha = \varepsilon^{-2}$ large, the Toda lattice Hamiltonian admits action-angles coordinates.

Proceeding *heuristically*: the Toda lattice can be considered a perturbation of order ε^3 of the fput-lattice $H_{3,4}$, see Eq.3.6:

$$H_3(\mathbf{p}, \mathbf{q}) = \sum_{j=0}^{N-1} \Big(\frac{1}{2}(p_k^2 + (q_k - q_{k+1})^2) + \frac{\varepsilon}{3!}(q_k - q_{k+1})^3 + \frac{\varepsilon^2}{4!}(q_k - q_{k+1})^4\Big) \tag{4.2}$$

Hence the periodic elastic chain perturbed by $h_{3,4} = \frac{\varepsilon}{3!}\sum_{j=0}^{N-1}(q_k - q_{k+1})^3 + \frac{\varepsilon^2}{4!}\sum_{j=0}^{N-1}(q_k - q_{k+1})^4$ and the periodic Toda lattice truncated to order ε^2 are the same: thus their first two orders Birkhoff's normal forms, to remove the resonant terms present in $H_{3,4}$, coincide up to the next order above ε^2 (the first steps in the construction of Birkhoff normal form only involves the common part $h_{3,4}$). The integrability of the Toda lattice should then imply that the normal form for the above truncation is integrable implying that the normal form of the above fput-chain is integrable to order ε^2.

Therefore the argument of [26] can be applied (heuristically here, and avoiding any explict calculation of the Birkhoff normal form) to reach the same conclusion and the periodic fput-chain model in Eq.4.2 shows quasi periodic motions at small ε, filling large regions of phase space in any region Ω where $e < H < E$, up to a set of volume approching 0 as $\varepsilon \to 0$.

The chain considered in Eq.4.2, $H_{3,4}$, has a Birkhoff's normal form which to order ε^2 is the sum of the (resonant) contributions to the α-chain (to second order) and those to the β chain (to first order): making use of the arbitrariness of ε it follows that both α and β chains have an integrable normal form, to order 2 included, and the KAM theory can be applied to both, if the Jacobian J 3 is non singular (as it is in the Toda lattice and in the α-model, [31, 26]).

Appendices

A Shift equation

Existence of the shift $\mathbf{a} = -J_0^{-1}\partial_{\mathbf{a}}\overline{f}_0(\mathbf{a})$.
Remark the bound on the matrix elements $\sigma_{ij} = J_0^{-1}\boldsymbol{\partial}^2_{\mathbf{a}}\overline{f}_0(\mathbf{a})$ in $C_{\lambda\rho_0}$ by $\sigma = ((1-\lambda)\rho_0)^{-1}J_{0,-}^{-1}\varepsilon_0$: hence the map $S\mathbf{a} = \mathbf{a} + J_0^{-1}\boldsymbol{\partial}_{\mathbf{a}}\overline{f}(\mathbf{a})$ has non singular Jacobian if $N\sigma < \frac{1}{3}$ (for instance) [14] and the map is locally invertible; furthermore any pair $\mathbf{a}', \mathbf{a}'' \in C((1-\lambda)\rho_0)$ is mapped into a pair $S\mathbf{a}', S\mathbf{a}''$ at distance $|S\mathbf{a}' - S\mathbf{a}''| > (1-\sigma N)|\mathbf{a}' - \mathbf{a}''|$, showing invertibility of S in $C_{(1-\lambda)\rho_0}$ with image in $C((1+N\sigma)(1-\lambda)\rho_0) \subset C(\rho_0)$ if λ is fixed $\lambda = \frac{2}{3}$ and $N\sigma < \frac{1}{3}$.

As $\mathbf{a}$ varies on the boundary of $C_{(1-\lambda)\rho_0}$ the image $S\mathbf{a}$ describes a set surrounding the polydisk of radius $(1-N\sigma)(1-\lambda)\rho_0 = \frac{2}{9}\rho_0$: therefore the images $S\mathbf{a}$ cover a set containing 0 and imply existence of $\mathbf{a}$ with $S\mathbf{a} = 0$ and therefore, $|\mathbf{a}| = |J_0^{-1}\partial_{\mathbf{a}}\overline{f}_0(\mathbf{a})| < J_{0,-}^{-1}\varepsilon_0$. In particular if $\lambda = \frac{2}{3}, \gamma_0 = 3N$:

$$3NJ_{0,-}^{-1}\varepsilon_0\rho_0^{-1} < 1 \Rightarrow |\mathbf{a}| < J_{0,-}^{-1}\varepsilon_0 < \frac{1}{\gamma_0}\rho_0, \qquad \mathbf{a} \in C(\frac{1}{3N}\rho_0) \tag{a.1}$$

in general any other $\lambda < 1$ would be acceptable provided $\gamma_0 = N(1-\lambda)^{-1}$; the choices of λ, γ_0 are just convenient for later use and far from optimal.

B Renormalized H: dimensional estimates

Remark that in $\mathcal{C}(\frac{1}{2}\rho_0, \kappa_0)$ the Jacobian of the map $\boldsymbol{\alpha}' = \boldsymbol{\alpha} + \partial_{\mathbf{A}'}\Phi_0$ is $(I + \boldsymbol{\partial_\alpha}\boldsymbol{\partial}_{\mathbf{A}'}\Phi_0)$ and, from the bound on Φ_0, Eq.2.12 , $|\boldsymbol{\partial_\alpha}\boldsymbol{\partial}_{\mathbf{A}'}\Phi_0| < 6\widehat{g}\frac{\varepsilon_0}{C}\kappa_0^{-\widehat{c}}$. [15]

---

[14] Let $|v| = \sum_i |v_i|$ for $v \in C^N$. Then the matrix $\delta_{ij} + \sigma_{ij}$ applied to $v \in C^N$ results in $w_i = v_i + \sum_j \sigma_{ij}v_j$: *i.e.* $|w| = \sum_i |w_i| \geq |v|(1-N\sigma) > 0$ if, say, $N\sigma < \frac{1}{3}$.

[15] Recall that Φ_0 is bounded in $\mathcal{C}(\frac{2}{3}\rho_0, \kappa_0)$ and $|\mathbf{a}| < \frac{1}{6}\rho_0$; then use $\frac{2}{3} - \frac{1}{2} = \frac{1}{6}$.

By Eq.2.12, $|\partial_{\mathbf{A}'}\Phi_0(\mathbf{A}',\boldsymbol{\alpha})| < \widehat{\gamma}\frac{\varepsilon_0}{C}\kappa_0^{-\widehat{c}}$ and the Jacobian determinant will be $\ge 1-\delta_0 > 0$ if $\delta_0 \le N!6\widehat{g}\frac{\varepsilon_0}{C}\,\kappa_0^{-\widehat{c}}$ is smaller than 1: for definitness suppose

$$\delta_0 = N!6\widehat{g}\frac{\varepsilon_0}{C}\,\kappa_0^{-\widehat{c}} < \frac{1}{2}\kappa_0 < 1. \tag{b.1}$$

Hence, for such δ_0, the $\boldsymbol{\alpha}' = S\boldsymbol{\alpha} = \boldsymbol{\alpha}+\boldsymbol{\partial}_{\mathbf{A}'}\Phi_0(\mathbf{A}',\boldsymbol{\alpha})$, maps (by Eq. 2.12) , $\mathcal{C}(\frac{1}{2}\rho_0,\kappa_0-\delta_0)$ into a set $\subset \mathcal{C}(\frac{1}{2}\rho_0,\kappa_0)$ *covering* the polydisk $\mathcal{C}(\frac{1}{2}\rho_0,\kappa_0-2\delta_0)$ and is also locally invertible (as its Jacobian is nonsingular) and globally one-to-one, hence holomorphic. [16] S^{-1} will be denoted $\boldsymbol{\alpha} = \boldsymbol{\alpha}'+\Delta(\mathbf{A}',\boldsymbol{\alpha}')$ and Δ inherits the bound Eq.2.13 $\|\Delta\|_{\frac{1}{2}\rho_0,\kappa_0-2\delta_0} < \overline{\gamma}\frac{\varepsilon_0}{C}\,\kappa_0^{-\overline{c}}$ provided the conditions in Eq.(2.10) hold together with $\delta_0 < \frac{1}{2}\kappa_0$, and the constants $\overline{\gamma},\overline{c}$ are chosen large enough. : its inverse in $\mathcal{C}(\frac{1}{2}\rho_0,\kappa_0-2\delta_0)$ will be written $\boldsymbol{\alpha} = \boldsymbol{\alpha}'+\Delta(\mathbf{A}',\boldsymbol{\alpha}')$ and fullfills the bound 2.13 for suitable $\overline{\gamma},\overline{c},\delta_0$.

For the purpose of dimensional bounds it is convenient to consider the second map S in Eq.(2.4) in the domain $\mathcal{C}(\frac{1}{2}\rho_0,\kappa_0-2\delta_0)$ with δ_0 in Eq.b.1 and values in $\mathcal{C}(\frac{1}{2}\rho_0,\kappa_0-\delta_0)$.

The conditions for the bounds are summarized as, Eq. 2.10 b.1:

$$\gamma_0\varepsilon_0 J_{0,-}^{-1} < \rho_0, \qquad C^{-1}\rho_0\; J_{0,+} < 1, \qquad \delta_0 = \widetilde{\gamma}\frac{\varepsilon_0}{C}\kappa_0^{-\widetilde{c}} < \frac{1}{2}\kappa_0 \tag{b.2}$$

for $\gamma_0,\widetilde{\gamma},\widetilde{c}$ suitable dimensionless constants. [17]

Given a generic function $F(\mathbf{A}',\boldsymbol{\alpha})$ define $F'(\mathbf{A}',\boldsymbol{\alpha}') \stackrel{def}{=} F(\mathbf{A}',\boldsymbol{\alpha}'+\Delta(\mathbf{A}',\boldsymbol{\alpha}'))$. Then, by Eq.2.20, to estimate the derivatives $\partial_{\mathbf{A}'}$ and $\partial_{\boldsymbol{\alpha}'}$ of F_j consider first the $\partial_{\mathbf{A}'}$ and $\partial_{\boldsymbol{\alpha}}$. Then make use of, see also Eq. 2.10 2.13 :

---

[16] Remark that, for the purpose of solving the implicit functions problems in Eq.2.4 , an argument can be used similar to the one for the solution of $\mathbf{a} = -J_0^{-1}\boldsymbol{\partial}_{\mathbf{a}}\overline{f}_0(\mathbf{a})$ in Appendix A . Any pair $\boldsymbol{\alpha}',\boldsymbol{\alpha}'' \in C(\frac{1}{2}\kappa_0)$ is mapped, if $|v| = \sum_i |v_i|, v \in C^N$, into a pair $S\boldsymbol{\alpha}', S\boldsymbol{\alpha}''$ at distance $|S\boldsymbol{\alpha}''-S\boldsymbol{\alpha}'| > (1 - N\|\partial_{\boldsymbol{\alpha}}\partial_{\mathbf{A}'}\Phi_0\|_{\frac{1}{2}\rho_0,\kappa_0})\,|\boldsymbol{\alpha}'-\boldsymbol{\alpha}''|$; hence the map S is one-to-one globally and analyticity of S^{-1} will also follow from the positivity of the Jacobian. For a more general and systematic study of the inversion of the maps in Eq.2.4 see Appendicx G in [33].

[17] Likewise the $\mathbf{A} = \mathbf{A}'+\mathbf{a}+\partial_{\boldsymbol{\alpha}}\Phi_0(\mathbf{A}',\boldsymbol{\alpha})$ would lead to $\mathbf{A} = \mathbf{A}'-\mathbf{a}+\Xi(\mathbf{A}',\boldsymbol{\alpha}')$ with $\Xi(\mathbf{A}',\boldsymbol{\alpha}') = -\partial_{\boldsymbol{\alpha}}\Phi_0(\mathbf{A}',\boldsymbol{\alpha}'+\Delta(\mathbf{A}',\boldsymbol{\alpha}'))$ under similar conditions; but Ξ is not needed here.

$$
\begin{aligned}
\partial_{\mathbf{A}'}F'(\mathbf{A}',\boldsymbol{\alpha}') =&\partial_{\mathbf{A}'}F(\mathbf{A}',\boldsymbol{\alpha}) + \partial_{\boldsymbol{\alpha}}F(\mathbf{A}',\boldsymbol{\alpha})\cdot\partial_{\mathbf{A}'}\Delta(\mathbf{A}',\boldsymbol{\alpha}')\\
\partial_{\boldsymbol{\alpha}'}F'(\mathbf{A}',\boldsymbol{\alpha}') =&\partial_{\boldsymbol{\alpha}}F(\mathbf{A}',\boldsymbol{\alpha})\cdot(I+\partial_{\boldsymbol{\alpha}'}\Delta(\mathbf{A}',\boldsymbol{\alpha}'))\\
\|\Delta\|_{\frac12\rho_0,\kappa_0-2\delta_0} <&\overline{\gamma}\frac{\varepsilon_0}{C}\kappa_0^{-\overline{c}}\\
|\partial_{\mathbf{A}'}\Delta|_{\frac14\rho_0,\kappa_0-2\delta_0}+\frac{4}{\rho_0}|\partial_{\boldsymbol{\alpha}'}\Delta|_{\frac14\rho_0,\kappa_0-2\delta_0} <& \overline{\gamma}\frac{\varepsilon_0}{C}\kappa_0^{-\overline{c}}
\end{aligned}
\tag{b.3}
$$

Hence $f_1(\mathbf{A}',\boldsymbol{\alpha}') = f'_1(\mathbf{A}',\boldsymbol{\alpha}) \equiv \sum_{j=1}^{3} F'_j(\mathbf{A}',\boldsymbol{\alpha})$ and Eq.2.19:

$$
\begin{aligned}
&|\partial_{\mathbf{A}'}f_1(\mathbf{A}',\boldsymbol{\alpha}')| = |\partial_{\mathbf{A}'}f'_0(\mathbf{A}',\boldsymbol{\alpha}) + \partial_{\boldsymbol{\alpha}}f'_0(\mathbf{A}',\boldsymbol{\alpha})\partial_{\mathbf{A}'}\Delta(\mathbf{A}',\boldsymbol{\alpha}')|\\
&\le \overline{\gamma}\varepsilon_0(\frac{\varepsilon_0}{C})^{\frac12}\kappa_0^{-\overline{c}} + \widetilde{\gamma}\rho_0\varepsilon_0(\frac{\varepsilon_0}{C})^{\frac12}\kappa_0^{-\overline{c}}\frac{1}{\rho_0}\overline{\gamma}\frac{\varepsilon_0}{C}\kappa_0^{-\overline{c}}\\
&|\frac{1}{\widetilde{\rho}}\partial_{\boldsymbol{\alpha}'}f'_1(\mathbf{A}',\boldsymbol{\alpha}')| = |\frac{1}{\widetilde{\rho}}\partial_{\boldsymbol{\alpha}}f'_0(\mathbf{A}',\boldsymbol{\alpha})(I+\partial_{\boldsymbol{\alpha}'}\Delta(\mathbf{A}',\boldsymbol{\alpha}'))|\\
&\le |\frac{1}{\widetilde{\rho}}\partial_{\boldsymbol{\alpha}}f'_0(\mathbf{A}',\boldsymbol{\alpha})|(1+N\overline{\gamma}\frac{\varepsilon_0}{C}\kappa_0^{-\overline{c}})
\end{aligned}
\tag{b.4}
$$

Hence the ε_1 can be bounded as the ε'_1 in Eq.2.20 in the domain $\mathcal{C}(\frac14\rho_0,\kappa_0-4\delta_0)$ (smaller than $\mathcal{C}(\frac12\rho_0,\kappa_0-2\delta_0)$ to allow using dimensional bounds in the bounds of the derivatives of Δ) simply by replacing the constants γ_j, c_j by a single pair $\overline{\gamma},\overline{c}$.

C FPUT symmetry

The Hamiltonian $H_0(\mathbf{p},\mathbf{q}) = \sum_{k=1}^{N}\frac12(p_k^2+(q_k-q_{k+1})^2)$ for $N = 2(n+1)$ particles can be solved by changing variables $\mathbf{q}\longleftrightarrow\widehat{\mathbf{q}}, \mathbf{p}\longleftrightarrow\widehat{\mathbf{p}}$ with:

$$
q_h = \frac{1}{\sqrt{N}}\sum_{k=1}^{N} e^{\frac{2\pi i}{N}hk}\widehat{q}_k,\quad p_h = \frac{1}{\sqrt{N}}\sum_{k=1}^{N} e^{\frac{2\pi i}{N}hk}\widehat{p}_k \tag{c.1}
$$

with $\widehat{q}_h = \overline{\widehat{q}}_{N-h}$ and $\widehat{p}_h = \overline{\widehat{p}}_{N-h}$ and periodic boundary $\widehat{q}_N = \widehat{q}_0, \widehat{p}_0 = \widehat{p}_N$.

It becomes $H_0(\widehat{\mathbf{p}},\widehat{\mathbf{q}}) = \sum_{k=1}^{N}\omega_k a_k$ where $\omega_k = 2\sin\frac{\pi k}{N}$ and $a_k = \frac12(|\widehat{p}_k|^2 + \omega_k^2|\widehat{q}_k^2|)$. The coordinates $\widehat{q}_N, \widehat{p}_N$ describe the free motion of the center of mass; discarding $\widehat{q}_N, \widehat{p}_N$ the Hamiltonian is reduced to $2n = N-2$ degrees of freedom.

It is convenient to separate real and imaginary part of the $\widehat{\mathbf{q}}$ and $\widehat{\mathbf{p}}$ calling $Q_h = -\widehat{q}_{h,im}$ for $h \le \frac{N}{2}$, and $Q_{\frac{N}{2}+h} = \widehat{q}_{h,re}$. Then

$$q_k = \frac{\sqrt{2}}{\sqrt{N}} \sum_{h=1}^{N/2-1} (Q_h \sin\frac{2\pi kh}{N} + Q_{\frac{1}{2}N+h} \cos\frac{2\pi kh}{N}) \tag{c.2}$$

and define similarly P_h in terms of the $\widehat{\mathbf{p}}$: in this way the phase space coordinates for H_0 will be denoted $\mathbf{Q} = (Q_1, \ldots, Q_{\frac{1}{2}N}, Q_{\frac{1}{2}N+1} \ldots, Q_N) = (\mathbf{Q}_+, \mathbf{Q}_-)$ and, likewise $\mathbf{P} = (\mathbf{P}_+, \mathbf{P}_-)$.

Define the canonical maps S, R:

$$S\begin{pmatrix} \mathbf{Q}_+, \mathbf{Q}_- \\ \mathbf{P}_+, \mathbf{P}_- \end{pmatrix} = \begin{pmatrix} \mathbf{Q}_+, -\mathbf{Q}_- \\ \mathbf{P}_+, -\mathbf{P}_- \end{pmatrix}, \qquad R\begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{q}^t \\ \mathbf{p}^t \end{pmatrix} \tag{c.3}$$

where $\mathbf{q}^t = (q_2, \ldots, q_{N-1}, q_1)$ if $\mathbf{q} = (q_1, \ldots, q_{N-1})$ and likewise is defined $\mathbf{p}^t$; the map S is written for the variables $\mathbf{Q}$ to make clear that $S\mathbf{Q} = \mathbf{Q}$ means, by Eq(c.2), that the manifold F in phase space of the fixed points of the linear symplectic map S consists of the $\mathbf{Q}_- = 0, \mathbf{P}_- = 0$ (*i.e.* of the $\mathbf{p}, \mathbf{q}$ with purely imaginary Fourier's tranforms $\widehat{\mathbf{p}}, \widehat{\mathbf{q}}$). And from Eqc.2 it is seen that on F it is $q_N = q_0 = 0$: furthermore the maps R, S commute with H_ε therefore evolution generated by H_ε leaves invariant the manifold E.

Hence the N-degrees of freedom periodic b.c. Hamiltonian H_ε describes, if restricted to F, a free chain with fixed endpoints and n degrees of freedom if $N = 2(n+1)$. Furthermore its Hamiltonian can be written:

$$H_0 = \sum_{k=1}^{n} \omega_k a_k, \quad a_k = \frac{1}{2}\frac{(P_k^2 + \omega_k^2 Q_k^2)}{\omega_k}, \quad \sin\varphi_k = \frac{\omega_k Q_k}{\sqrt{2\omega_k a_k}} \tag{c.4}$$

with $(a_k, \varphi_k)_{k=1}^n$ pairs of conjugate symplectic coordinates $(\mathbf{a}, \boldsymbol{\varphi})$.

The relation between the periodic chain, with N particles, and the fixed extremes chain, with $n = \frac{N}{2} - 1$ particles, leads to determine classes of integers N for which N-particles chains Birkhoff's series of some low order are integrable at small ε, [26], including various cases beyond those with N=power of 2 or N-prime.

Soon, [18], an original approach was developed to cover all chains with N arbitrary integer (which by the time had still remained to relatively few integers): and the result was achieved without the need to compute the normal form, but using the symmetries, to infer in general, that fput-β-chains normal forms could not be affected by possible resonances that, altough possible, were conjectured not to be involved in the normal form construction.

It is interesting to remark that the the potential of fput-β-chain with fixed extremes, setting $\boldsymbol{\sigma} = (\sigma_i)_{i=1}^4, \sigma_i = \pm 1$ and $T_{\mathbf{h}} = (\frac{2}{N})^2 \sum_{k=1}^{\frac{N}{2}} \prod_{j=1}^4((1-c_{h_i})s_{h_ik} - s_{h_i}c_{h_ik})$ for $\mathbf{h} = (h_i)_{i=1}^4$, where $c_h = \cos\varphi_h, s_h = \sin\varphi_h$. can be written:

$$V_4(\mathbf{Q}) = \frac{1}{4}\sum_{\mathbf{h},\boldsymbol{\sigma}} T_{\mathbf{h}} \prod_{j=1}^4 \sigma_j e^{i\sigma_j\varphi_{h_j}} (\frac{a_{h_j}}{\omega_{h_j}})^{\frac{1}{2}} \tag{c.5}$$

and easily identifies the terms which need to be absent to obtain a normal form which is integrable: if $n_{\mathbf{h},\boldsymbol{\sigma}}(k) = \sum_{h_j=k}\sigma_j$ and $\mathbf{n} = (n_{\mathbf{h},\boldsymbol{\sigma}}(k))_{k=1}^{N=1}$ the uneliminable resonances $\boldsymbol{\nu}$ are the ones for which $\xi = \mathbf{n}\cdot\boldsymbol{\omega} = 0$. Hence to prove the result it is necessary to show that the sum of the terms labeled by the above $\mathbf{n}$ for which $\xi = 0$ either do not arise (but they do in interesting cases) or that they add up to 0 as a consequence of the symmetry or because $\mathbf{n} = 0$, as found in [18, 28].

STATEMENT: This paper has no supplementary data

References

[1] L. Boltzmann. Studien über das gleichgewicht der lebendigen kraft zwischen bewegten materiellen punkten. *Wiener Berichte*, 58, (W.A.,#5):517–560, (49–96), 1868.

[2] L. Boltzmann. Lösung eines mechanischen problems. *Wiener Berichte*, 58, (W.A.,#6):1035–1044, (97–105), 1868.

[3] E. Fermi. Dimostrazione che in generale un sistema meccanico è quasi ergodico. *Nuovo Cimento*, 25:267–269, 1923.

[4] A.N. Kolmogorov. On the preservation of conditionally periodic motions. *In Lecture Notes in Physics, Stochastic behavior in classical and quantum Hamiltonians, ed. G. Casati, J. Ford, Vol. 93, 1979*, 93, 1979.

[5] J. Ford. The fermi-pasta-ulam problem: paradox turns discovery. *Physics Reports*, 213:271–310, 1992.

[6] L. Boltzmann. Über die mechanische Bedeutung des zweiten Hauptsatzes der Wärmetheorie. *Wiener Berichte*, 53, (W.A.,#2):195–220, (9–33), 1866.

[7] O. Darrigol. *Physics & Necessity*. Oxford University Press, Oxford, 2014.

[8] J. C. Maxwell. On Boltzmann's theorem on the average distribution of energy in a system of material points. *Transactions of the Cambridge Philosophical Society*, 12:547–575, 1879.

[9] R. Clausius. Ueber die Zurückführung des zweites Hauptsatzes der mechanischen Wärmetheorie und allgemeine mechanische Prinzipien. *Annalen der Physik*, 142:433–461, 1871.

[10] G. Gallavotti. *Nonequilibrium and irreversibility (II edition)*, volume LNP 1040 of *Lecture Notes in Physics*. Springer-Verlag, 2025.

[11] J. Gibbs. *Elementary principles in statistical mechanics*. Schribner, Cambridge, 1902.

[12] L. Boltzmann. Einige allgemeine sätze über Wärmegleichgewicht. *Wiener Berichte*, 63, (W.A.,#19):679–711, (259–287), 1871.

[13] G. Gallavotti. Ergodicity: a historical perspective. equilibrium and nonequilibrium. *European Physics Journal H*, 41,:181–259, 2016.

[14] G. Gallavotti and I. Jauslin. A Theorem on Ellipses, an Integrable System and a Theorem of Boltzmann. *arXiv*, 2008.01955:1–9, 2020.

[15] G. Felder. Poncelet Property and Quasi-periodicity of the Integrable Boltzmann System. *Letters in Mathematical Physics*, 111:1–19, 2021.

[16] G. Gallavotti. *Classical mechanics and the quantum revolution in Fermi's early works*, volume In *Enrico Fermi: his work and legacy*, editors C. Bernardini and L. Bonolis. Springer, 2004.

[17] T. Nishida. A note on an existence of conditionally periodic oscillation in a one-dimensional anharmonic lattice. *Mem. Fac. Engrg. Kyoto Univ.*, 33:27–34, 1971.

[18] B. Rink. Proof of Nishida's conjecture on anharmonic lattices. *Communications Mathematical Physics*, 261:613–627, 2006.

[19] G. Gallavotti. Quasi integrable mechanical systems. *Phenomènes Critiques, Systèmes aleatories, Théories de jauge, Proceedings, Les Houches, XLIII (1984), North Holland, Amsterdam*, II:539–624, 1986.

[20] V. Arnold. Small denominators and problems of stability of motion in classical and celestial mechanics. *Russian Mathematical Surveys*, 18:85–191, 1963.

[21] J. Moser. On invariant curves of an area preserving mapping of the annulus. *Nachrichten Akademie Wissenshaften Göttingen, II*, 11:1–20, 1962.

[22] G. Benettin, L. Galgani, A. Giorgilli, and J. Strelcyn. A proof of Kolmogorov's theorem on invariant tori using canonical transformations defined by the Lie method. *Nuovo Cimento B*, 79:201–223, 1984.

[23] G. Gallavotti. Quasi periodic hamiltonian motions, scale invariance, harmonic oscillators. *Journal of Mathematical Physics*, 60:062901, 2019.

[24] G. Gallavotti. Twistless KAM tori. *Communications in Mathematical Physics*, 164:145–156, 1994.

[25] G. Gallavotti. *The Fermi-Pasta-Ulam problem. A status report.* Lecture Notes in Physics, ed. G. Gallavotti. Springer, 2008.

[26] B. Rink. Symmetry and resonance in periodic fpu chains. *Communications Mathematical Physics*, 218:665–685, 2001.

[27] G. Gallavotti. Some rigorous results about 3D Navier Stokes. *Les Houches 1992 NATO-ASI meeting on Turbulence in spatially extended systems, ed. by R. Benzi, C. Basdevant, S. Ciliberto, Nova Science Publishers, NY*, pages 45–81, 1993.

[28] B. Rink. *An Integrable Approximation for the Fermi–Pasta–Ulam Lattice*, volume 728 of *Lecture Notes in Physics*. Springer, 2008.

[29] H. Flaschka. The toda lattice. ii. existence of integrals. *Physical Review B*, 9:1924–1925, 1974.

[30] M. Henon. Integrals of the toda lattice. *Physical Review B*, 9:1921–1923, 1974.

[31] A. Henrici and T. Kappeler. Birkhoff normal form for the periodic toda lattice. *arXiv preprint nlin/0609045*, 2006.

[32] J. Moser. Finitely many mass points on the line under the influence of an exponential potential—an integrable system. *Dynamical Systems, Theory and Applications: Battelle Seattle 1974 Rencontres*, 38:467–497, 1975.

[33] G. Gallavotti. *The Elements of Mechanics (II edition)*. Gallavotti website [I edition was Springer 1984], Roma, 2008.