Journal of Statistical Physics, Vol. 71, No. 3/4, 1993

Exponential Stability of States Close to Resonance in Infinite-Dimensional Hamiltonian Systems

Dario Bambusi[1] **and Antonio Giorgilli**[1]

Received August 4; 1992; final November 30, 1992

We develop canonical perturbation theory for a physically interesting class of infinite-dimensional systems. We prove stability up to exponentially large times for dynamical situations characterized by a finite number of frequencies. An application to two model problems is also made. For an arbitrarily large FPU-like system with alternate light and heavy masses we prove that the exchange of energy between the optical and the acoustical modes is frozen up to exponentially large times, provided the total energy is small enough. For an infinite chain of weakly coupled rotators we prove exponential stability for two kinds of initial data: (a) states with a finite number of excited rotators, and (b) states with the left part of the chain uniformly excited and the right part at rest.

KEY WORDS: Hamiltonian systems; infinite dimensional systems; Nekhoroshev theory; perturbation theory.

1. INTRODUCTION

The problem of stability in near to integrable Hamiltonian systems has been recently investigated by several authors in the light of both the KAM and Nekhoroshev theorems. The theory is now quite well developed for systems with a finite number of degrees of freedom. For infinite systems only few results have been obtained, mainly in connection with KAM theory[(1–6)] (see also ref. 7), while the extension of Nekhoroshev-like results has been less considered[(8,9)] (see also ref. 10). The aim of the present paper is to extend Nekhoroshev-type results to infinite systems, the unperturbed motion of which has only a finite number of frequencies. The possibility of a complete extension of the Nekhoroshev theorem to infinite systems is still an open problem.

---

[1] Dipartimento di Matematica dell'Università, 20133 Milano, Italy.

0022-4715/93/0500-0569$07.00/0

We prove a general theorem that we apply to two well-known models which have been extensively investigated and have some physical interest. The first one is a modification of the standard Fermi–Pasta–Ulam system; precisely, we consider, as in the FPU system, a one-dimensional chain of point masses interacting through nonlinear springs, but with alternate light and heavy masses. The second model is a system of weakly coupled rotators.(11, 12, 2, 13, 14, 1)

The modified FPU model has been numerically investigated in ref. 15. The spectrum of the system splits into two well-separated branches, usually called the acoustical one (low frequencies) and the optical one (high frequencies). The separation between the two branches increases with the ratio m_2/m_1 of the heavy to the light masses; correspondingly, the whole system can be considered as composed of two separate subsystems with a small coupling provided by the nonlinearity of the springs. The relevant phenomenon observed in the paper quoted above is the following: if one starts with some energy concentrated in the acoustic branch, then only a small amouont of energy flows to the optical modes, up to a time exponentially increasing with the ratio of the frequencies. For a finite number n of degrees of freedom, such a result could be expected on the basis of a theorem proved in ref. 16, but there remains the problem that the analytical estimates of the constants of that theorem do not guarantee that the phenomenon persists when $n \to \infty$. On the other hand, the numerical computations suggest that the phenomenon might be independent of n when the *specific* energy is fixed. In the present paper we prove that the phenomenon persists in the limit $n \to \infty$, but with the restriction that the *total* (and not the specific) energy be fixed. In our opinion, the gap between the numerical indication (fixed specific energy) and the present theorem (fixed total energy) can hardly be filled on a purely dynamical basis, but requires perhaps the use of a statistical argument: we shall come back to this point later. We point out that our approach differs from Nekhoroshev's in that we do not try to prove that the single action variables of the system are frozen: such a result would be impossible, due to the fact that the optical modes form an almost completely resonant system. Instead, we only bound the exchange of energy between the two subsystems of the optical and of the acoustical modes.

Concerning the system of rotators, we consider a chain with a long-range interaction. We prove that the theorem of Nekhoroshev, ensuring the freezing of the action of each rotator, can be extended to an infinite chain provided one considers the particular class of initial conditions which correspond to the so-called localized states. Precisely, we consider initial conditions close to a state in which only a finite number of rotators have a nonvanishing action, and prove that such states are stable up to times which increase exponentially with a power of the inverse of the size of the perturbation.

As a matter of fact, we deduce all these results from a general formulation of a Nekhoroshev-like theory for infinite systems, which, by the way, is completely coordinate independent. Let us explain in a few words the key points of our theory. In extending a Nekhoroshev-type theory to infinite systems one has usually to tackle two main difficulties: on the one hand, one has to build up algebraic and analytical tools, like function spaces, norms, and so on, which are needed to produce estimates; on the other hand, one is confronted with the difficulty due to the presence of an infinite number of frequencies.

We solve the first problem by introducing an abstract scheme which allows us to develop a normal form theory for Hamiltonians defined on a generic Banach space, obtaining in particular estimates independent of the dimension of the space.

The difficulty connected with the infinite number of frequencies is instead much harder, because the usual diophantine estimates contain a very strong dependence on the dimension of the space. In ref. 1 this difficulty was overcome by assuming a strong decay of the interactions, which allowed a weaker formulation of the diophantine conditions. Here, instead, we restrict our attention to particular states of the system, roughly characterized by the fact that only a finite number of frequencies is actually relevant. More precisely, the general abstract theorem we prove is concerned only with the neighborhood of a periodic orbit, so that we do not need diophantine-like conditions at all. This constitutes the so-called analytic part of Nekhoroshev's theorem and is actually enough for the application to the FPU-like system. The case of rotators requires also the so-called geometric part of the Nekhoroshev theorem. Here, we exploit the idea of Lochak[17] that, in finite dimensions, all the frequencies are close, in some sense, to a complete resonance. In the case of an infinite system this is no longer true in a global sense, but holds for any state in which all rotators but a finite number are initially at rest; thus, Nekhoroshev's result can be extended to a suitable neighborhood of such localized states.

The paper is organized as follows. In Section 2 we recall the FPU-type model and formulate our result for such a system. In Section 3 we do the same for the model of rotators. In Section 4 we state our general results in abstract form. Section 5 contains the proof of the abstract theorems. Section 6 contains the proof of the result on the FPU-type model, deducing it from the abstract theory. Section 7 contains the proof of the analytic part of the theorem for the system of rotators. The geometric part for the same model is isolated in Section 8, because it is given in a general form, essentially independent of the specific model considered here.

2. FREEZING OF HIGH FREQUENCIES IN AN FPU-TYPE MODEL

We consider a one-dimensional chain of an even number of particles interacting through anharmonic springs; the ends of the chain are fixed. The Hamiltonian of the system is

$$H(y, x) = \sum_{j=1}^{n} \frac{y_j^2}{2m_j} + \sum_{j=1}^{n+1} \frac{1}{2} k(x_j - x_{j-1})^2 + \sum_{j=1}^{n+1} \varphi(x_j - x_{j-1}) \tag{2.1}$$

where y_j, x_j are conjugate variables representing, respectively, the momentum and the position of the particles, $\varphi\colon \mathbb{R} \to \mathbb{R}$ is an analytic function with $\varphi(0) = \varphi'(0) = \varphi''(0) = 0$ (primes denote differentiation), k is a positve constant, and m_j are the masses of the particles. The condition of fixed ends is $x_0 = x_{n+1} = 0$. As a variant to the original FPU model, we consider the case in which the masses of the particles are alternate and very different:

$$m_j = m_{j+2}, \qquad \frac{m_1}{m_2} \ll 1$$

It is well known that this leads to a frequency spectrum (of the linear part of the system) with two disjoint branches. In fact, the spectrum is given by

$$(\omega_l^{\pm})^2 := k\,\frac{m_1 + m_2 \pm (m_1^2 + m_2^2 + 2m_1 m_2 \cos 2\beta_l)^{1/2}}{m_1 m_2} \tag{2.2}$$

where

$$\beta_l := \frac{\pi l}{n+1}, \qquad 1 \leqslant l \leqslant \frac{n}{2} \tag{2.3}$$

Here and below, we use superscripts minus and plus to denote quantities referring to the acoustical and optical branches, respectively. Thus, we have a set of low frequencies $\omega^- = (\omega^{-1},..., \omega^-_{n/2})$ ranging from zero to $\omega^-_{\max}$, given by

$$(\omega^-_{\max})^2 = 2\,\frac{k}{m_2}$$

and a set of high frequencies $\omega^+ = (\omega_1^+,..., \omega^+_{n/2})$ ranging from $\omega^+_{\min}$ to $\omega^+_{\max}$ given by

$$(\omega^+_{\min})^2 = 2\,\frac{k}{m_1}, \qquad (\omega^+_{\max})^2 = 2k\left(\frac{1}{m_1} + \frac{1}{m_2}\right)$$

The relevant fact here is that the ratio $\omega_{\min}^{+}/\omega_{\max}^{-}$ of the lowest optical frequency to the highest acoustic one increases with the ratio m_2/m_1, so that it turns out to be very large in the case $m_1 \ll m_2$; at the same time, the ratio $(\omega_{\max}^{+} - \omega_{\min}^{+})/\omega_{\min}^{+}$ goes to zero, so that the optical branch tends to be completely resonant.

The first step consists in introducing normal modes for the quadratic part of the Hamiltonian. The coordinate transformation to normal modes is given by

$$x_j = \sum_{\substack{l=1\\ \pm}}^{n/2} \frac{2}{(n+1)^{1/2}} \left(\frac{2k}{|2k - m_j \omega_l^{\pm 2}|}\right)^{1/2} |\cos \beta_l| \times \frac{\sin(j\beta_l)}{[m_1^2 + m_2^2 + 2m_1 m_2 \cos(2\beta_l)]^{1/4}} q_l^{\pm} \tag{2.4}$$

Here, the label $\pm$ under the sum means that the sum has to be extended to both the acoustical variables q^- and the optical ones q^+. The coordinate transformation has to be completed to a canonical transformation $(p, q) \mapsto (y, x)$ in the obvious way. In the new variables H can be written in the form

$$H(p, q) = h^-(p^-, q^-) + h^+(p^+, q^+) + H_{\text{pert}}(q)$$

where

$$h^{\pm}(p^{\pm}, q^{\pm}) := \sum_{l=1}^{n/2} \tfrac{1}{2}(p_l^{\pm 2} + \omega_l^{\pm 2} q_l^{\pm 2}) \tag{2.5}$$

are the acoustic (h^-) and the optical (h^+) harmonic energies, while $H_{\text{pert}}(q)$ is a complicated function of q providing the interaction between the two subsystems; the form of this function is not essential here.

We also introduce the mean (quadratic) value of the frequencies of the optical branch, defined by

$$\omega^2 := \frac{(\omega_{\max}^{+})^2 + (\omega_{\min}^{+})^2}{2} = k\left(\frac{2}{m_1} + \frac{1}{m_2}\right) \tag{2.6}$$

We point out that, although we have written a Hamiltonian function for a system with a finite number n of degrees of freedom, all the operations performed so far are meaningful also in the case $n = \infty$; more details on this limit will be given in the technical Section 6.

As explained in the Introduction, our aim is to bound the exchange of energy between the optic and acoustic branches. The corresponding result is stated in the following theorem.

Theorem 2.1. Consider the Cauchy problem for the Hamiltonian system (2.1) with initial data (p^0, q^0). Assume that the function φ can be extended to a complex analytic function on a complex sphere of radius ρ_* around the origin, for some positive ρ_*, and that there exists a positive constant Υ such that

$$|\varphi(x)| \leqslant \tfrac{1}{3}\,\Upsilon |x|^3, \qquad |\varphi'(x)| \leqslant \Upsilon |x|^2, \qquad \forall x \in \mathbf{C} \quad \text{with } |x| < \rho_* \tag{2.7}$$

Define the total harmonic energy $\mathscr{E}_0$ of the initial datum as

$$\mathscr{E}_0 := \sum_{\substack{l=1\\ \pm}}^{n/2} \tfrac{1}{2}\left[(p_l^{0\pm})^2 + (\omega_l^{\pm})^2 (q_l^{0\pm})^2\right]$$

and the (dimensionless) parameter μ as

$$\mu := 3^3 2^7 \frac{\Upsilon}{k^{3/2}} \sqrt{\mathscr{E}_0} + 2^9 \frac{\omega_{\max}^-}{\omega} \tag{2.8}$$

where ω is given by (2.6). If $\mu < 1$, and if

$$\left(\frac{2\mathscr{E}_0}{k}\right)^{1/2} < \rho_* \tag{2.9}$$

then one has the estimate

$$\frac{|h^+(t) - h^+(0)|}{\mathscr{E}_0} \leqslant \mu$$

with h^+ given by (2.5), for all times t with

$$|t| \leqslant \frac{1}{\omega\mu} \exp\left(\frac{1}{\mu}\right)$$

Let us add a few comments. In plain words, the result is the following: if the *total* energy of the initial datum is small *and* if the ratio $\omega/\omega_{\max}^-$ is large, then the exchange of energy between the optical and acoustical branches is small compared to their total energy, up to a time growing exponentially with the inverse of the perturbation parameter μ. Now, as it stands, the result looks meaningless in the thermodynamic limit, where one should be able to consider a fixed *specific* energy, instead of the total one; furthermore, we point out that we cannot control the energy of each mode, either acoustic or optical, but only the exchange of energy between the two subsystems. Concerning the second point, we remark that both subsystems become strongly resonant in the limit $n \to \infty$; moreover, the optical frequencies tend to become equal for $\mu \to 0$. Thus, one cannot expect that the harmonic energies of the single modes are conserved, and in particular the internal sharing of the energy of the optical subsystem cannot be bounded: this was put into evidence by numerical simulations in ref. 15. Concerning the first point, namely whether one has a similar freezing for fixed specific energy, the problem is the following: the control of the sharing of energy depends on the maximal energy that can be concentrated on a single particle of the system; thus, by dynamical considerations one cannot exclude *a priori* that *all* the available energy can concentrate on a single particle for a long time, so that our result could hardly be improved. We believe that such a situation could possibly be excluded on the basis of statistical considerations.

3. STABILITY OF LOCALIZED STATES IN A SYSTEM OF WEAKLY COUPLED ROTATORS

We consider an infinite chain of weakly coupled identical rotators, and study the stability of states in which only a finite number of rotators is excited.

As a general model we consider the case of long-range interaction among the rotators, as described by the Hamiltonian

$$H(J, \phi) := \sum_{l \in \mathbf{Z}} \frac{J_l^2}{2I} + \varepsilon \sum_{\substack{l, j \in \mathbf{Z} \\ l \neq j}} \frac{1}{|l-j|^\alpha} [1 - \cos(\phi_l - \phi_j)], \qquad \alpha > 1 \tag{3.1}$$

where J, ϕ are canonically conjugate variables, I is the momentum of inertia of each rotator, and ε is a small parameter with the dimension of an energy.

A simpler model, considered by many authors, is that of rotators with a finite-range interaction, in particular limited only to first neighbors. The Hamiltonian in such a case is

$$h^R(J, \phi) := \sum_{l \in \mathbf{Z}} \frac{J_l^2}{2I} + \varepsilon \sum_{l \in \mathbf{Z}} [1 - \cos(\phi_l - \phi_{l-1})] \tag{3.2}$$

As said in the Introduction, our aim is to prove that, if we consider initial data with finite total energy, and with the additional hypothesis that only a finite number of rotators have angular momentum significantly different from zero, then such a situation changes only a little, up to times growing exponentially with the inverse of the size of the perturbation.

To prove our result, we need a suitable Hamiltonian framework for our infinite system. In fact, the two different models require a different framework with regard to the characterization of the phase space. In particular, this makes all the technical estimates quite complicated for the long-range case with respect to the finite-range one. In view of this fact, and to avoid unnecessary technicalities, we give a complete discussion of the finite-range case. At the end of the section we also include a brief discussion of the long-range model. Proving the results for the latter case is just a matter of applying again the same scheme as for the finite range: this just requires more time to work out all the estimates.

We give now the definitions of the phase space and of the symplectic form for the finite-range case.

The phase space $\mathscr{P}$ is defined as

$$\mathscr{P} := l^2 \times \Delta^2 \ni (J, \phi) \tag{3.3}$$

where l^2 is the space of square summable sequences, and Δ^2 is the space of the sequences $\phi = \{\phi_l\}$ such that

$$\sum_{l \in \mathbf{Z}} (\phi_l - \phi_{l-1})^2 < \infty$$

The space $\mathscr{P}$ coincides in fact with the set of states with finite energy. With this definition, the Hamiltonian (3.2) turns out to be analytic on $\mathscr{P}$ (see Section 7). If $(J, \phi) \in \mathscr{P}$ is a point of the phase space, we define its norm by

$$\|(J, \phi)\|_{\mathscr{P}}^2 := \sum_{l \in \mathbf{Z}} \frac{|J_l|^2}{2I} + \varepsilon \sum_{l \in \mathbf{Z}} \frac{1}{2} |\phi_l - \phi_{l-1}|^2 + 2\varepsilon |\phi_0|^2 \tag{3.4}$$

On the phase space we define the symplectic form Ω in the natural way, namely by

$$\Omega((J, \phi), (J^1, \phi^1)) := \sum_{l \in \mathbf{Z}} J_l \phi_l^1 - J_l^1 \phi_l \tag{3.5}$$

We remark that the symplectic form is not defined on the whole $\mathscr{P}$, but only on a subspace of it. Actually, this is enough for our purposes because we can extend the usual methods of Hamiltonian mechanics to such a situation: the details can be found in Section 4.

To state our theorem we also need to introduce a quantity which measures, in a suitable sense, the distance between two states. The problem is that the distance induced by the norm above is too strong, because it does not take into account the fact that the ϕ's are angles. For instances, $\|(J, \phi)\| = C$ means that the differences between nearby angles are bounded; this is clearly incompatible with the dynamics, since at least some (a finite number) of the angles should be allowed to rotate freely with respect to their neighbors. Indeed, this happens already in the case $\varepsilon=0$. In order to consider also these cases, we take as a natural substitute of the norm the energy of the system, the function $h^R(J, \phi)$ defined by (3.2). Remark indeed that the condition $h^R(J, \phi) \leqslant C$ with $C > 2\varepsilon$ allows the difference between a finite number of nearby angles to be arbitrarily large. As a norm induces a distance, our substitute of the norm induces a substitute of the distance, namely the function

$$\begin{aligned} &h^R((J, \phi)-(J^1, \phi^1)) \\ &\quad = \sum_{l \in \mathbf{Z}} \frac{(J_1 - J_l^1)^2}{2I} + \varepsilon \sum_{l \in \mathbf{Z}} \{1 - \cos[(\phi_l - \phi_l^1) - (\phi_{l-1} - \phi_{l-1}^1)]\} \end{aligned}$$

Besides the remark above, we stress that although this function does not satisfy the formal properties of a distance, it gives a good characterization of close states; indeed, if the value of the function is zero for two given states (J, ϕ) and (J^1, ϕ^1), then all the corresponding actions coincide.

We give now a formal statement of our result.

Theorem 3.1. Consider the Cauchy problem for the Hamiltonian system (3.2) in the phase space (3.3). Considering a finite set of indices S with cardinality n, take any sequence $\bar{J} = \{\bar{J}_l\}_{l \in \mathbf{Z}}$ with support S, and denote by $\tilde{\nu}$ the maximal frequency corresponding to $\bar{J}$, namely

$$\tilde{\nu} := \max_{l \in S} \left\{ \frac{\bar{J}_l}{I} \right\}$$

Define also the dimensionless parameter $\tilde{\mu}$ by

$$\tilde{\mu} := \left(\frac{\varepsilon}{I \tilde{\nu}^2} \right)^{1/2}$$

Then there exist positive constants μ_* and C_1, depending on n but independent of all the parameters of h^R and of all properties of $\bar{J}$, such that the following holds true: if

$$\tilde{\mu} < \mu_*$$

then, for all initial data $(J^0, \phi^0) \in \mathscr{P}$ close enough to $\bar{J}$, precisely satisfying

$$h^R((J^0, \phi^0) - (\bar{J}, 0)) \leqslant \tfrac{1}{4} n^2 \varepsilon \tag{3.6}$$

one has that the corresponding solution $(J(t), \phi(t))$ exists and remains close of $\bar{J}$, and precisely satisfies

$$h^R((J(t), \phi(t)) - (\bar{J}, 0)) \leqslant C_1 I \tilde{\nu}^2 \tilde{\mu}^{2/n}$$

for all times t with

$$|t| \leqslant \frac{1}{48e\tilde{\nu}\sqrt{n}} \exp\left(\frac{\mu_*}{\tilde{\mu}}\right)^{1/n}$$

Possible values of μ_* and C_1 are

$$\mu_* = \frac{1}{2^{5n}3^{3n}27n^{(n-1)/2}}, \qquad C_1 = 313n$$

We remark that the class of initial data allowed by (3.6) contains states with angles asymptotically well aligned.

As outlined in the Introduction, we prove this theorem by using the rational approximation technique introduced by Lochak. According to this technique, we prove a stability theorem, in Nekhoroshev's sense, for a neighborhood of a periodic orbit, and then use a number-theoretic argument to extend the result to the whole phase space in the case of a finite-dimensional system, and to suitable neighborhoods of localized states in the case on an infinite system. The local result concerning a neighborhood of a periodic orbit of the unperturbed system is given by the following theorem; its proof is in fact the main step we need in order to prove Theorem 3.1.

Theorem 3.2. Consider the Cauchy problem for the dynamical system with Hamiltonian

$$H(J, \phi) := \sum_{l \in \mathbf{Z}} \omega_l J_l + \sum_{l \in \mathbf{Z}} \frac{J_l^2}{2I} + \varepsilon \sum_{l \in \mathbf{Z}} [1 - \cos(\phi_l - \phi_{l-1})] \tag{3.7}$$

Assume that $\omega = \{\omega_l\}_{l \in \mathbf{Z}}$ is completely resonant, i.e., there exist $\nu \in \mathbb{R}$ and $\{k_l\}_{l \in \mathbf{Z}} \in \mathbf{Z}^{\mathbf{Z}} \setminus \{0\}$ such that

$$\omega_l = k_l \nu$$

Denote by S^* the set

$$S^* := \{i \in \mathbf{Z} \colon \omega_i \neq \omega_{i-1}\}$$

and assume that its cardinality is finite and equal to n^*. For any initial datum $\zeta_0 = (J^0, \phi^0) \in \mathscr{P}$, define

$$\beta := \max\left\{2\left[\frac{h^R(J^0, \phi^{(0)})}{\varepsilon}\right]^{1/2}, 12n^*\right\} \tag{3.8}$$

where h^R is again the function defined in (3.2); define also

$$\mu := 2^5 3^3(\beta + n^*)\left(\frac{\varepsilon}{Iv^2}\right)^{1/2} \tag{3.9}$$

If

$$\mu < 1 \tag{3.10}$$

then the solution $(J(t), \phi(t))$ of the Cauchy problem exists, and one has

$$h^R(J(t), \phi(t)) \leqslant \beta^2 \frac{\varepsilon}{2} \tag{3.11}$$

for all times t satisfying

$$|t| \leqslant \frac{1}{12e} \frac{1}{\|\delta\omega\|} \exp\left(\frac{1}{\mu}\right)$$

where

$$\|\delta\omega\| := \left(4|\omega_0|^2 + \sum_{j \in S^*} |\omega_j - \omega_{j-1}|^2\right)^{1/2}$$

The Hamiltonian (3.7) is nothing but (3.2) expanded around the point $J = I\omega$. We point out that the theorem covers also cases of infinite initial energy; for instance, it covers the case of a left chain with initial data $\omega_l = v$ for $l < 0$ interfaced with a right chain with initial data $\omega_l = 0$ for $l \geqslant 0$, with nonvanishing specific energy. The reduction to the case of finite energy considered in Theorem 3.1 is due to the application of the number-theoretic result referred to above.

Before closing this section, we indicate how to modify the definition of the phase space in order to deal with the long-range case. The whole game is based on the introduction of a suitable norm, in such a way that the phase space is as large as possible, and the Hamiltonian (3.1) is analytic. To start with, we consider only the angle variables, and introduce the family of functions $\langle|\phi|\rangle_{k,\alpha}$, depending on a positive integer parameter k,

$$\langle|\phi|\rangle_{k,\alpha} := \left(\sum_{j,l} \frac{1}{|l-j|^\alpha} |\phi_j - \phi_l|^k\right)^{1/k}$$

With this, we introduce the further function

$$\langle|\phi|\rangle_{\infty,\alpha} := \operatorname*{Sup}_{k \geqslant 2} \langle|\phi|\rangle_{k,\alpha}$$

We shall denote by $\Delta^{\infty,\alpha}$ the set of infinite sequences $\phi = \{\phi_j\}_{j\in\mathbf{Z}}$ such that $\langle|\phi|\rangle_{\infty,\alpha} < \infty$. Now we define the phase space $\mathscr{P}$ as the set of the sequences

$$(J, \phi) \in l^2 \times \Delta^{\infty,\alpha}$$

equipped with the norm

$$\|(J, \phi)\|_{\mathscr{P}}^2 := \sum_{l\in\mathbf{Z}} \frac{J_j^2}{2I} + \frac{1}{2}\varepsilon\langle|\phi|\rangle_{\infty,\alpha}^2 + 2\varepsilon|\phi_0|^2$$

With these definitions, Theorem 3.1 can be proved, with slightly different numerical constants. Similarly, a result of the kind of Theorem 3.2, with a restriction to states of finite energy, can be proved. In particular, we are still unable to extend the result to cases of initial data with infinite energy. The proof is essentially the same as in the case of finite range; one has just to add the proof that the space l^2 is continuously embedded in $\Delta^{\infty,\alpha}$.

4. ABSTRACT PERTURBATION THEORY

In this section we adapt the usual symplectic formalism for infinite-dimensional spaces, in order to obtain a framework suitable for perturbation theory. Then we develop a perturbation scheme leading to a general normal form theorem. Precisely, we shall concentrate on three points: (i) the extension of the usual methods of Hamiltonian mechanics in order to be able to handle our model problems; (ii) the introduction of suitable domains, which are needed for quantitative perturbation theory; and (iii) the characterization of the general class of Hamiltonians we can deal with.

Concerning the symplectic space, we make a construction which differs a little from the usual ones (as common references see, for example, refs. 18 and 19). In fact, we follow here an approach that we consider more suitable for perturbation theory, almost in the same spirit of the work of Kuksin[7]; actually, our scheme is more general: besides our system of rotators, it also allows one to deal, e.g., with the wave equation in unbounded domains in $\mathbb{R}^n$.

Consider a Banach space $\mathscr{P}$, and assume we are given a bilinear skew-symmetric form Ω on a domain $D(\Omega) \subset \mathscr{P} \times \mathscr{P}$. We assume moreover that there exists a linear subspace $\mathscr{C}_0 \subset \mathscr{P}$ with the following properties:

(α) $D(\Omega)$ contains $\mathscr{C}_0 \times \mathscr{P}$, and $\Omega(x, \cdot)$ is a continuous linear functional on $\mathscr{P}$ for all $x \in \mathscr{C}_0$.

(β) Ω is nondegenerate on $\mathscr{P} \times \mathscr{C}_0$, namely, if $y \in \mathscr{P}$ is such that $\Omega(x, y) = 0$ $\forall x \in \mathscr{C}_0$, then one has $y = 0$.

We shall call $(\mathscr{P}, \mathscr{C}_0, \Omega)$ a symplectic space. For the functional $\Omega(x, \cdot)$ we shall also use the standard notation of interior product: $x \,\lrcorner\, \Omega := \Omega(x, \cdot)$.

We consider the following norm on $\mathscr{C}_0$:

$$\|x\|_{\mathscr{F}} := \operatorname*{Sup}_{\|y\|_{\mathscr{P}}=1} |\Omega(x, y)|$$

which is well defined by axiom (α), and is a norm by virtue of the axiom (β); then, we complete $\mathscr{C}_0$ with respect to this norm, thus obtaining a Banach space that we shall denote by $\mathscr{F}$. Notice that, by axiom (α), Ω can be extended to a continuous bilinear functional on $\mathscr{P} \times \mathscr{F}$.

The usual definition of canonical transformation can be generalized to the present case: a differentiable map $\mathscr{T}\colon \mathscr{P} \to \mathscr{P}$ is said to be canonical in case (i) the differential $\mathscr{T}'(\zeta)$ satisfies $\mathscr{T}'(\zeta)(\mathscr{P} \cap \mathscr{F}) = \mathscr{P} \cap \mathscr{F}$ for all $\zeta \in \mathscr{P}$, and (ii) $\Omega(\mathscr{T}'(\zeta)x, \mathscr{T}'(\zeta)y) = \Omega(x, y)$ for all $x \in \mathscr{P}$ and all $y \in \mathscr{P} \cap \mathscr{F}$.

We introduce now the definition of symplectic gradient (or Hamiltonian vector field). We shall consider two different cases, which cover all the situations we shall encounter. As a first case we consider a function f: $\mathscr{P} \to \mathbb{R}$ which is of class C^1; in this case we define its symplectic gradient $\nabla^{\Omega} f\colon \mathscr{P} \to \mathscr{F}$ by

$$\Omega(\nabla^{\Omega} f(\zeta), x) = df(\zeta)x, \qquad \forall x \in \mathscr{P}$$

provided it exists. As a second case we consider a linear function $g\colon \mathscr{P} \supset D(g) \to \mathbb{R}$, and assume that its domain $D(g)$ contains $\mathscr{C}_0$; then we define $\nabla^{\Omega} g \in \mathscr{P} \cup \mathscr{F}$ as the constant function satisfying

$$\Omega(\nabla^{\Omega} g, x) = g(x), \qquad \forall x \in \mathscr{C}_0$$

provided it exists. The symplectic gradient of a function f at x will also be denoted by $X_f(x)$ $(\equiv \nabla^{\Omega} f(x))$.

Finally, we introduce the Poisson bracket. We consider a function $f \in C^1(\mathscr{P})$ and a function g with the property that $\nabla^{\Omega} g$ exists and satisfies $\nabla^{\Omega} g\colon \mathscr{P} \to \mathscr{P}$; the *Poisson bracket* $\{f, g\}(\zeta)$ is defined by

$$\{f, g\}(\zeta) := f'(\zeta)\, \nabla^{\Omega} g(\zeta) \tag{4.1}$$

where $f(x)$ is the differential of f at x (we shall use the symbol d only for the operator of exterior differentiation acting on forms). We remark that the definition can be applied also to the case $f \in C^1(\mathscr{P}, \mathscr{S})$ with $\mathscr{S}$ a generic Banach space, because the Poisson bracket is in fact the Lie derivative of f with respect to the Hamiltonian vector field $\nabla^{\Omega} g$. We shall make use of this extension.

In order to better illustrate our scheme, let us now add a few examples.

(i) Wave equation on a bounded domain $K \subset \mathbb{R}^n$. It is well known that this is a Hamiltonian system with Hamiltonian function

$$\frac{1}{2}\int_K [p(x)^2 - u(x)\,\Delta u(x)]\,d^n x$$

In this case we put $\mathscr{P} = L^2(K) \times H^1(K) \ni (p, u)$ (where H^1 is the usual Sobolev space), and $\mathscr{C}_0 = \mathscr{P}$; the symplectic form is, as usual,

$$\Omega((p_1, u_1), (p, u)) := \int_K [p_1(x)\,u(x) - p(x)\,u_1(x)]\,d^n x$$

It is now an easy matter to check that one has $\mathscr{F} = H^{-1} \times L^2$.

(ii) Wave equation on unbounded domains of $\mathbb{R}^n$. The Hamiltonian and the symplectic form are the same as in the previous example. The major change concerns the definition of the spaces $\mathscr{P}$ and $\mathscr{C}_0$. The natural choice for $\mathscr{P}$ is the space of states with finite energy. However, one has to remark that, on the one hand, the symplectic form is defined on $L^2 \times L^2$, but, on the other hand, $\mathscr{P}$ is not contained in $L^2 \times L^2$. So, the symplectic form turns out to be only densely defined on $\mathscr{P}$. One can take for $\mathscr{C}_0$ the space C_c^∞ of infinitely differentiable functions with compact support. For more details on a situation of this kind see ref. 9.

(iii) The system of rotators of Section 3. The space $\mathscr{P}$ is given by (3.3), and the symplectic form by (3.5). The space $\mathscr{C}_0$ can be defined as $\mathscr{A} \times l^2$, where $\mathscr{A}$ is the set of the sequences with finite support. Notice that in this case $\mathscr{C}_0$ is not dense in $\mathscr{P}$.

We come now to the definition of the domains. With reference to the symplectic space $(\mathscr{P}, \mathscr{C}_0, \Omega)$, consider first the complexifications $\mathscr{P}^{\mathbb{C}}$ and $\mathscr{F}^{\mathbb{C}}$ of $\mathscr{P}$ and $\mathscr{F}$, respectively. We consider a domain $\mathscr{G} \subset \mathscr{P}$, and define its extension $\mathscr{G}_{R,d} \subset \mathscr{P}^{\mathbb{C}}$ with parameters $R > 0$ and $d < 1$ as the union of open complex balls $B(\zeta, R(1-d)) \subset \mathscr{P}^{\mathbb{C}}$ of radius $R(1-d)$ centered at any point ζ of $\mathscr{G}$; formally,

$$\mathscr{G}_{R,d} := \bigcup_{\zeta \in \mathscr{G}} B(\zeta, R(1-d))$$

Finally, we characterize the general Hamiltonian systems to which we shall apply our perturbation scheme. We assume that the Hamiltonian can be given the form

$$H(\zeta) = h_\omega(\zeta) + \hat{h}(\zeta) + f(\zeta) \tag{4.2}$$

where h_ω will be considered as the unperturbed part of the Hamiltonian, while $\hat{h}$ and f will be considered as perturbations.

The most important property that we assume for the unperturbed Hamiltonian h_ω is that it gives rise to a periodic flow. This looks like a very strong hypothesis, in consideration of the fact that one usually considers an integrable Hamiltonian with a quasiperiodic flow. However, as shown by Lochak, a quasiperiodic Hamiltonian can be approximated by a periodic one in virtue of Dirichlet's theorem on the approximation of irrational numbers; this makes the periodic case sufficient for our purposes. The extension to the quasiperiodic case with a finite number of frequencies is just a technical matter. We also assume some technical hypotheses on h_ω, and in particular on the flow generated by it.

A first hypothesis is the following one:

1. The symplectic gradient $\nabla^\Omega h_\omega$ exists, and generates a continuous flow Φ_t on $\mathscr{P}^{\mathrm{C}}$.

A second group of hypotheses concerns the properties of the smoothness of the flow Φ_t:

2. (i) The map Φ_t is of class $C^1(\mathscr{P}^{\mathrm{C}}, \mathscr{P}^{\mathrm{C}})$ for every fixed t; moreover, denoting by $\Phi_t'(\zeta)$ its differential (with t fixed) at ζ, we assume (ii) that $\Phi_t'(\zeta)(\mathscr{P} \cap \mathscr{F})^{\mathrm{C}} \subset (\mathscr{P} \cap \mathscr{F})^{\mathrm{C}}$ and (iii) that the map $\zeta \mapsto \Phi_t'(\zeta)$ is analytic as a map from $\mathscr{P}^{\mathrm{C}}$ to $C^1(\mathscr{P}, \mathscr{P})$. Finally, (iv) for every $u \in (\mathscr{P} \cap \mathscr{F})^{\mathrm{C}}$, every $x \in \mathscr{P}^{\mathrm{C}}$, and all $t \in \mathbb{R}$ one has

$$\Omega(\Phi_t'(\zeta)u, \Phi_t'(\zeta)x) = \Omega(u, x)$$

3. The time derivative $d\Phi_t(\zeta)/dt$ exists for every $\zeta \in D$, where D is a dense subset of $\mathscr{P}^{\mathrm{C}}$ which is invariant for Φ_t.

We remark that all these hypotheses are trivially satisfied in the finite-dimensional case, due to the smoothness and to the periodicity of the Hamiltonian h_ω. In the infinite-dimensional case they could also be deduced from suitable smoothness hypotheses on h_ω and its symplectic gradient. However, we did not try to isolate the best hypotheses needed for this, since the properties above are actually enough for our purposes.

The fourth hypothesis is nothing but a natural request on the choice of the domain $\mathscr{G}$:

4. For every $d < 1$ the complex domain $\mathscr{G}_{R,d}$ is invariant for Φ_t.

Finally, we add a last technical hypothesis on the norm of Φ_t':

5. For every $t \in \mathbb{R}$ we have $\|\Phi_t'(\zeta)\|_{\mathscr{P},\mathscr{P}} = 1$.

The latter hypothesis could be replaced by $\|\Phi_t'(\zeta)\|_{\mathscr{P},\mathscr{P}} < C$ for some positive C.

We come now to the perturbation. We assume that both $\hat{h}$ and f are functions of class C^∞, and that their symplectic gradients are well defined and analytic. Furthermore, we assume that $\hat{h}$ commutes with h_ω, namely that $\{h_\omega, \hat{h}\} = 0$. In practice, the relevant property is that $\hat{h}$ is already in normal form with respect to h_ω; this is the case, for instance, when h_ω and $\hat{h}$ represent two separate subsystems interacting through the coupling term f.

According to the usual perturbation schemes, we first give a theorem stating that the Hamiltonian can be given a normal form up to an exponentially small remainder, with explicit quantitative estimates.

Theorem 4.1. On the domain $\mathscr{G}$ as above, consider a real Hamiltonian function $H: \mathscr{G} \to \mathbb{R}$ which can be decomposed into the sum of three functions h_ω, $\hat{h}$, f as in (4.2), with $\hat{h}$ and f of class $C^\infty(\mathscr{P})$, and such that their symplectic gradient is defined for all $x \in \mathscr{G}$. Assume that the flow Φ_t generated by h_ω is periodic with period $T := 2\pi/\omega$, namely

$$\Phi_{t+T} = \Phi_t, \qquad \forall t \in \mathbb{R}$$

and that h_ω satisfies hypotheses 1–5 above. Concerning $\hat{h}$, assume that $\{h_\omega, \hat{h}\} = 0$. Fix a positive parameter R and assume also that $\nabla^\Omega \hat{h}$ and $\nabla^\Omega f$ can be extended to complex analytic functions from $\mathscr{G}_{R,0}$ to $\mathscr{P}^{\mathbb{C}}$, and that the inequalities

$$\frac{1}{R} \operatorname*{Sup}_{\zeta \in \mathscr{G}_{R,0}} \|\nabla^\Omega \hat{h}(\zeta)\|_{\mathscr{P}} \leqslant \omega_0, \qquad \frac{1}{R} \operatorname*{Sup}_{\zeta \in \mathscr{G}_{R,0}} \|\nabla^\Omega f(\zeta)\|_{\mathscr{P}} \leqslant \omega_f$$

hold for some constants ω_0 and ω_f. Let $0 < d \leqslant 1/4$ be a positive parameter and define the pure number

$$\mu := \frac{24}{d} e\pi \left(\frac{\omega_0 + \omega_f}{\omega} \right) \tag{4.3}$$

Then the following statement holds true: if $\mu < 1$, then there exists an analytic canonical transformation $\mathscr{T}: \mathscr{G}_{R,2d} \to \mathscr{G}_{R,d}$, with $\mathscr{T}(\mathscr{G}_{R,2d}) \supset \mathscr{G}_{R,3d}$, such that $H \circ \mathscr{T}$ has the form

$$H(\mathscr{T}(\zeta)) = h_\omega(\zeta) + \hat{h}(\zeta) + Z(\zeta) + \mathscr{R}(\zeta) \tag{4.4}$$

where:

R1. $\{h_\omega, Z\} = 0$

R2. $\displaystyle \frac{1}{R} \operatorname*{Sup}_{\zeta \in \mathscr{G}_{R,2d}} \|\nabla^\Omega \mathscr{R}(\zeta)\|_{\mathscr{P}} \leqslant 3(e+1)\, \mu \omega_f \exp\left(-\frac{1}{\mu}\right)$

R3. For any analytic $g: \mathscr{G}_{R,0} \to S$, with $\mathscr{S}$ a generic Banach space, we have

$$\operatorname*{Sup}_{\zeta \in \mathscr{G}_{R,2d}} \|g(\zeta) - g(\mathscr{T}(\zeta))\|_{\mathscr{S}} \leqslant \frac{2\pi(e+1)}{d} \frac{\omega_f}{\omega} \operatorname*{Sup}_{\zeta \in \mathscr{G}_{R,d}} \|g(\zeta)\|_{\mathscr{S}}$$

We use now the normal form Theorem 4.1 in order to bound the variation of h_ω up to exponentially large times. This is given by the following result.

Corollary 4.2. Under the same hypotheses of Theorem 4.1, and with the additional assumption that h_ω can be extended to a complex analytic function on $\mathscr{G}_{R,d}$, the following holds true: along the solutions of the Cauchy problem of system (4.2), one has

$$|h_\omega(t) - h_\omega(0)| \leqslant \frac{6\pi(e+1)}{d} \frac{\omega_f}{\omega} \operatorname*{Sup}_{\zeta \in \mathscr{G}_{R,d}} |h_\omega(\zeta)| \tag{4.5}$$

for all $|t| \leqslant \min(T_0, T_*)$, where T_0 is the escape time of the solution from the domain $\mathscr{G}_{R,2d}$, and

$$T_* = \frac{2\pi}{3} \frac{1}{\omega\mu} \exp\left(\frac{1}{\mu}\right)$$

A stronger form of Theorem 4.1 can be given in case some further hypotheses on $\hat{h}$ and f are assumed. In this case, an improved and qualitatively different formulation of Corollary 4.2 can be obtained. Apart from aesthetic considerations, this is relevant because we can remove the analyticity hypothesis on h_ω, just requiring analyticity for the symplectic gradient $\nabla^\Omega h_\omega$. This improvement, which could seem minor, allows one to deal with the interesting case of states of infinite energy in the model of rotators.

Theorem 4.3. Consider the Hamiltonian system (4.2) satisfying all the hypotheses of Theorem 4.1. Assume also that both $\hat{h}$ and f can be extended to analytic functions on the domain $\mathscr{G}_{R,0}$, and that there exist constants E_0 and E such that the inequalities

$$\operatorname*{Sup}_{\zeta \in \mathscr{G}_{R,0}} |\hat{h}(\zeta)| \leqslant E_0, \qquad \operatorname*{Sup}_{\zeta \in \mathscr{G}_{R,0}} |f(\zeta)| \leqslant E$$

hold. Define the constant

$$E^* := \max\left\{E, \frac{1}{12} \frac{\omega_f}{\omega_f + \omega_0} E + e\pi \frac{E_0}{d} \frac{\omega_f}{\omega}\right\}$$

Then, denoting by $\mathscr{T}$ and $\mathscr{R}$, respectively, the canonical transformation and the remainder as in Theorem 4.1, one has the following estimates:

R4. $\sup_{\zeta \in \mathscr{G}_{R,2d}} |h_\omega(\zeta) - h_\omega(\mathscr{T}(\zeta))| \leqslant 2E^*$

R5. $\sup_{\zeta \in \mathscr{G}_{R,2d}} |\mathscr{R}(\zeta)| \leqslant 4eE^* \exp\left(-\frac{1}{\mu}\right)$

The corresponding form of Corollary 4.2 is the following.

Corollary 4.4. Under the hypotheses of Theorem 4.3, with the additional assumption that $0 < d \leqslant 1/6$ and that that $\nabla^\Omega h_\omega$ can be extended to a complex analytic function on $\mathscr{G}_{R,3d}$, and that h_ω is densely defined in $\mathscr{G}$. Then along the solutions of the Cauchy problem of system (4.2), one has

$$|\hat{h}(t) + f(t) - \hat{h}(0) - f(0)| \leqslant 6E^* \tag{4.6}$$

for all $|t| \leqslant \min(T_0, T_*)$, where T_0 is the escape time of the solution from the domain $\mathscr{G}_{R,3d}$, and

$$T_* = \frac{d}{2e}\left[\frac{1}{R} \sup_{\zeta \in \mathscr{G}_{R,3d}} \|\nabla^\Omega h_\omega(\zeta)\|\right]^{-1} \exp\left(\frac{1}{\mu}\right)$$

We omit the detailed proof, giving just the essential point. First, one considers initial data belonging to the domain of h_ω, which is dense in $\mathscr{P}$; this allows one to prove $|h_\omega(t) - h_\omega(0) \leqslant 6E^*$ over the same time interval, for these initial data. Then, using conservation of energy, one proves (4.6) for the same set of initial data. Finally, using the density of the domain of h_ω and the continuity of $\hat{h} + f$ on $\mathscr{P}$, one concludes the proof.

5. ABSTRACT TECHNICALITIES

The scheme of the proof of Theorem 4.1 is essentially the same used, in a finite-dimensional context, in ref. 16 (see also ref. 9). In fact, the scheme used there is based on a recursive algorithm, equivalent to the Lie transform, to perform canonical transformations and to give the Hamiltonian a suitable normal form. The algorithm can be efficiently used for quantitative estimates. One has to introduce norms for functions which dominate the sup norm, give estimates, in terms of the norms above, for Poisson brackets and for the solutions of the homological equation (5.12). The extension to the case of an infinite system requires a clever choice of the phase space, and so also of the norms, and a careful implementation of the averaging method in order to give the estimate for the solution of the homological equation. With these elements, the proof of Theorem 4.1 can be achieved by a straightforward application of the scheme of ref. 16. In view of these remarks, we report here only the original elements of the proof, namely the choice of the norms, the estimate of the Poisson bracket, and the estimate of the solution of the homological equation. Concerning instead the proof of Theorem 4.3, some additional estimates are necessary, which are reported below in explicit form.

We begin by briefly outlining, at a purely formal level, the precedure used in order to put in normal form the Hamiltonian (see ref. 20 for more details). We first recall the method we use in order to perform a canonical transformation.

We consider a sequence $\{\chi_s\}_{s\geqslant 1}$ of functions on the phase space $\mathscr{P}$, which will be called a "generating sequence," and define a corresponding linear operator T_χ acting on functions by

$$T_\chi g := \sum_{r\geqslant 0} g_r \tag{5.1}$$

where

$$g_0 := g, \qquad g_r := \sum_{l=1}^{r} \frac{l}{r}\{\chi_l, g_{r-l}\}, \qquad r\geqslant 1 \tag{5.2}$$

Letting this operator act on the identical function on the phase space, we obtain a transformation synthetically written as

$$\zeta = T_\chi \zeta' := (T_\chi \mathbb{1})(\zeta')$$

This transformation turns out to be canonical, and moreover the following relevant identity holds[20]:

$$(T_\chi f)(\zeta') = f(T_\chi \zeta')$$

We look now for a finite generating sequence $\chi = \{\chi_s\}_{s=1}^{r}$ such that the transformed Hamiltonian $T_\chi H$ is in normal form up to a small remainder $\mathscr{R}^{(r)}$, of order r in some small parameter to be determined; precisely, we ask the transformed Hamiltonian to be of the form

$$H(T_\chi \zeta) = (T_\chi H)(\zeta) = h_\omega(\zeta) + \hat{h}(\zeta) + Z(\zeta) + \mathscr{R}^{(r)}(\zeta) \tag{5.3}$$

where Z is in normal form with respect to h_ω, i.e., we have $\{h_\omega, Z\} \equiv 0$.

In order to find the equations for χ, we denote $T_\chi h_\omega = \sum_{s\geqslant 0} h_s$, $T_\chi \hat{h} = \sum_{s\geqslant 0} \hat{h}_s$, $T_\chi f = \sum_{s\geqslant 0} f_s$, and $Z = \sum_{s=1}^{r} Z_s$ with Z_s of order s, and consider h_s, $\hat{h}_s$, and f_s of order s, $s+1$, and $s+1$, respectively. So, equating terms of the same order in (5.3), we obtain for χ_s and Z_s the following equations:

$$\{h_\omega, \chi_s\} + Z_s = \Psi_s, \qquad 1 \leqslant s \leqslant r \tag{5.4}$$

with

$$\begin{aligned} \Psi_1 &= f \\ \Psi_s &= \hat{h}_{s-1} + \sum_{l=1}^{s-1} \frac{1}{s} \{\chi_l, h_{s-l}\} + f_{s-1}, \qquad 2 \leqslant s \leqslant r \end{aligned} \tag{5.5}$$

These equations can be recursively solved, thus giving the generating sequence χ.

Before coming to quantitative estimates, we give a proposition on Poisson brackets.

Proposition 5.1. Let $f \in C^\infty(\mathscr{P})$ and $g \in C^\infty(\mathscr{P})$ be such that their symplectic gradients $\nabla^\Omega f$, $\nabla^\Omega g$ exist and belong to $C^\infty(\mathscr{P}, \mathscr{P})$. Then $\nabla^\Omega\{f, g\}$ exists, belongs to $C^\infty(\mathscr{P}, \mathscr{P})$, and satisfies

$$\nabla^\Omega\{f, g\} = [\nabla^\Omega f, \nabla^\Omega g] \tag{5.6}$$

Proof. First, we prove that if $f \in C^\infty(\mathscr{P})$ is such that its symplectic gradient is defined for all $x \in \mathscr{P}$, then one has $\nabla^\Omega f \in C^\infty(\mathscr{P}, \mathscr{F})$. To this end, consider the map $\mathscr{F} \ni x \mapsto x \,\lrcorner\, \Omega \in \mathscr{P}^*$; due to the definition of the norm of $\mathscr{F}$, it is an isometry, and, moreover, by axiom (β) it is invertible on its range. So, its inverse map is a linear continuous map. The differentiability of $\nabla^\Omega f$ as a map from $\mathscr{P}$ to $\mathscr{F}$ is an immediate consequence of this fact. So, one has

$$d(\mathscr{L}_{X_f} g) = \mathscr{L}_{X_f}(dg) \tag{5.7}$$

where we denoted by $\mathscr{L}_{X_f}$ the operator of Lie derivative with respect to the vector field $X_f \equiv \nabla^\Omega f$; the form above coincides with

$$\mathscr{L}_{X_f}(X_g \,\lrcorner\, \Omega) = \mathscr{L}_{X_f} X_g \,\lrcorner\, \Omega + X_g \,\lrcorner\, \mathscr{L}_{X_f} \Omega$$

(for this kind of calculation see ref. 21). Notice that this expression is meaningful since $X_f \in C^1(\mathscr{P}, \mathscr{P}) \cap C^1(\mathscr{P}, \mathscr{F})$. Now we show that, if $\xi_1 \in \mathscr{P} \cap \mathscr{F}$ and $\xi_2 \in \mathscr{P}$, then $\mathscr{L}_{X_f} \Omega(\xi_1, \xi_2) = 0$. Notice first that, using the definition of derivative, one has

$$\langle (X \,\lrcorner\, \Omega)'(x)\eta, \xi \rangle = \langle \Omega; X'(x)\eta, \xi \rangle$$

$\forall X \in C^1(\mathscr{P}, \mathscr{P}) \cap C^1(\mathscr{P}, \mathscr{F})$, $\eta \in \mathscr{P}$, and $\xi \in \mathscr{P} \cap \mathscr{F}$ or $\eta \in \mathscr{P} \cap \mathscr{F}$, and $\xi \in \mathscr{P}$; here we also denoted by $\langle \Omega; \xi, \eta \rangle$ the value of the form Ω on the vectors ξ and η, and by a prime the derivative of the map. So, we have[21]

$$\begin{aligned}\langle(\mathscr{L}_{X_f}\Omega)(x);\xi_1,\xi_2\rangle &= \langle\Omega; X_f'(x)\xi_1,\xi_2\rangle + \langle\Omega;\xi_1, X_f'(x)\xi_2\rangle \\ &= \langle\Omega; X_f'(x)\xi_1,\xi_2\rangle - \langle\Omega; X_f'(x)\xi_2,\xi_1\rangle \\ &= \langle(X_f \,\lrcorner\, \Omega)'(x)\xi_1;\xi_2\rangle - \langle(X_f \,\lrcorner\, \Omega)'(x)\xi_2;\xi_1\rangle \\ &= \langle d(X_f \,\lrcorner\, \Omega)(x);\xi_1,\xi_2\rangle = \langle ddf(x);\xi_1,\xi_2\rangle = 0\end{aligned}$$

So, since for all $x\in\mathscr{P}$ one has $X_g(x)\in\mathscr{P}\cap\mathscr{F}$, it follows that

$$d\{f,g\} = [X_f, X_g] \,\lrcorner\, \Omega$$

From this one immediately gets (5.6). The smoothness statement on $\nabla^{\Omega}\{f,g\}$ is an easy consequence of this formula. ∎

Now we introduce the norms needed in order to complete the scheme above with quantitative estimates. Given an analytic function $g\colon \mathscr{G}_{R,d}\to\mathscr{S}$ ($\mathscr{S}$ being a Banach space, and $\mathscr{G}_{R,d}$ the domain defined in Section 4), we define its norm $N_d(g)$ by

$$N_d(g) := \operatorname*{Sup}_{\zeta\in\mathscr{G}_{R,d}} \|g(\zeta)\|_{\mathscr{S}} \tag{5.8}$$

If $\mathscr{S}=\mathbf{C}$ it is useful to define also the norm

$$N_d^{\nabla}(g) := \frac{1}{R}\operatorname*{Sup}_{\zeta\in\mathscr{G}_{R,d}} \|\nabla^{\Omega}g(\zeta)\|_{\mathscr{P}} \tag{5.9}$$

which measures the size of the Hamiltonian vector field generated by g. The Poisson bracket between two functions is estimated by the following

Lemma 5.2. Let $g\colon\mathscr{G}_{R,0}\to\mathscr{S}$ be an analytic function on $\mathscr{G}_{R,d}$ and let $g_1\colon\mathscr{G}_{R,0}\to\mathbf{C}$ be another analytic function such that $\nabla^{\Omega}g_1$ exists, is analytic as a function from $\mathscr{G}_{R,d}$ to $\mathscr{P}$, and satisfies $N_d^{\nabla}(g_1)<\infty$; then $\{g,g_1\}$ is a analytic on $\mathscr{G}_{R,d}$ and the following inequality holds:

$$N_d(\{g,g_1\}) \leqslant \frac{1}{d} N_0(g)\, N_d^{\nabla}(g_1) \tag{5.10}$$

Moreover, if g is such that $\nabla^{\Omega}g$ exists, is analytic from $\mathscr{G}_{R,d}$ to $\mathscr{P}$, and satisfies $N_d^{\nabla}(g)<\infty$, then $\nabla^{\Omega}\{g,g_1\}$ exists, $\forall d'>0$ it is analytic as a function from $\mathscr{G}_{R,d+d'}$ to $\mathscr{P}$, and it satisfies

$$N_{d+d'}^{\nabla}(\{g,g_1\}) \leqslant \frac{2}{d'} N_d^{\nabla}(g)\, N_d^{\nabla}(g_1) \tag{5.11}$$

Proof. Using the definition (4.1) of Poisson brackets, we immediately have

$$\|\{g, g_1\}(\zeta)\|_{\mathscr{S}} \leqslant \|g'(\zeta)\| \cdot \|\nabla^{\Omega} g_1(\zeta)\|_{\mathscr{P}}$$

But, from Cauchy's inequality (see, e.g., ref. 22) we also have

$$\operatorname*{Sup}_{\zeta \in \mathscr{G}_{R,d}} \|g'(\zeta)\| \leqslant \frac{1}{Rd} \operatorname*{Sup}_{\zeta \in \mathscr{G}_{R,d}} \|g(\zeta)\|$$

and (5.10) immediately follows. In order to obtain (5.11), notice that, by (5.6), we have

$$\begin{aligned}\nabla^{\Omega}\{g, g_1\}(x) &= [X_g, X_{g_1}](x)\\ &= X'_g(x)\, X_{g_1}(x) - X'_{g_1}(x)\, X_g(x)\\ &= \{\nabla^{\Omega} g, g_1\}(x) + \{g, \nabla^{\Omega} g_1\}(x)\end{aligned}$$

so that the lhs of (5.11) is less than

$$\begin{aligned}&N_{d+d'}(\{\nabla^{\Omega} g, g_1\}) + N_{d+d'}(\{g, \nabla^{\Omega} g_1\})\\ &\quad\leqslant \frac{1}{Rd'} N_d(\nabla^{\Omega} g)\, N^{\nabla}_{d+d'}(g_1) + \frac{1}{Rd'} N^{\nabla}_{d+d'}(g)\, N_d(\nabla^{\Omega} g_1)\end{aligned}$$

From this (5.11) immediately follows. All the analyticity properties are straightforward. ∎

We come now to the estimate of the solution of the homological equation. Let us bring into evidence a technical element of this estimate. The idea is to perform an average over a periodic orbit of the unperturbed system as in refs. 23 and 24. The relevant fact is that such an average can be performed in an intrinsic manner, without any reference to a coordinate system. This is particularly useful when studying the neighborhood of an elliptic equilibrium, for instance, as in the FPU model, where the unperturbed action–angle variables introduce an annoying singularity.

Lemma 5.3. Let Ψ: $\mathscr{G}_{R,d} \to \mathbf{C}$ be a C^1 function such that its symplectic gradient exists and is analytic over $\mathscr{G}_{R,d}$. Then, the homological equation

$$\{h_\omega, \chi\}(\zeta) + Z(\zeta) = \Psi(\zeta) \tag{5.12}$$

has a solution given by

$$\begin{aligned}Z(\zeta) &:= \frac{1}{T}\int_0^T \Psi(\Phi_t(\zeta))\, dt\\ \chi(\zeta) &:= \frac{1}{T}\int_0^T t[\Psi(\Phi_t(\zeta)) - Z(\Phi_t(\zeta))]\, dt\end{aligned} \tag{5.13}$$

where T is the period of Φ in time. The symplectic gradients of χ and Z exist and satisfy

$$\begin{aligned} N_d^{\nabla}(Z) &\leqslant N_d^{\nabla}(\Psi) \\ N_d^{\nabla}(\chi) &\leqslant T N_d^{\nabla}(\Psi) \end{aligned} \tag{5.14}$$

Proof. Denote $g(\zeta) := \Psi(\zeta) - Z(\zeta)$, with Z given by (5.13). Referring to the domain D of differentiability of Φ with respect to time, take $\zeta \in D$, then we have

$$\begin{aligned} \{h_\omega, \chi\}(\zeta) &= \frac{d}{dt}\bigg|_{t=0} \chi(\Phi_t(\zeta)) \\ &= \frac{d}{dt}\bigg|_{t=0} \frac{1}{T}\int_0^T g(\Phi_{t+s}(\zeta))s\, ds \\ &= \frac{1}{T}\int_0^T s\frac{d}{ds} g(\Phi_{t+s}(\zeta))\, ds\bigg|_{t=0} \\ &= \frac{1}{T} g(\Phi_s(\zeta))s\big|_0^T - \frac{1}{T}\int_0^T g(\Phi_s(\zeta))\, ds \\ &= g(\zeta) \end{aligned}$$

Therefore (5.13) solves (5.12) on D. But it is easy to check that, for any $g \in C^1(\mathscr{P}, \mathbf{C})$ which admits a symplectic gradient, we have

$$\nabla^\Omega(g \circ \Phi_t)(\zeta) = [\Phi_t'(\zeta)]^{-1}(\nabla^\Omega g)(\Phi_t(\zeta))$$

(just use the smoothness hypothesis 2 of Section 4), so that

$$\nabla^\Omega \chi(\zeta) = \frac{1}{T}\int_0^T [\Phi_t'(\zeta)]^{-1}(\nabla^\Omega g)(\Phi_t(\zeta))t\, dt$$

which is continuous in the argument ζ. It follows that $\{h_\omega, \chi\}$ is continuous on $\mathscr{G}_{R,d}$. Since the two sides of (5.12) are continuous and coincide on the dense set D, they coincide everywhere. To obtain the estimate (5.14), remark that

$$\|\nabla^\Omega Z(\zeta)\| \leqslant \frac{1}{T}\int_0^T \|[\Phi_t'(\zeta)]^{-1}\| \cdot \|(\nabla^\Omega \Psi)(\Phi_t(\zeta))\|\, dt \leqslant \operatorname*{Sup}_{\mathscr{G}_{R,d}} \|\nabla^\Omega \Psi(\zeta)\|$$

$$\|\nabla^\Omega \chi(\zeta)\| \leqslant \frac{1}{T}\int_0^T t\|[\Phi_t'(\zeta)]^{-1}\| \cdot \|(\nabla^\Omega g)(\Phi_t(\zeta))\|\, dt$$

from which the conclusions immediately follow. ∎

We give now the following.

Scheme of the Proof of Theorem 4.1. With the choice above of the norms, and also using the two lemmas above, one can repeat almost literally the arguments of ref. 16, Sections 6–8 (see also ref. 9). In particular, one gets an explicit estimate for the generating sequence (see Theorem 6.1 of the above reference) in the form

$$N_d^{\nabla}(\chi_l) \leqslant \frac{\beta^{l-1}}{l}\,\Phi$$

with constants β and Φ given by

$$\Phi = 2\pi\,\frac{\omega_f}{\omega}, \qquad \beta = \frac{12\pi(r-1)}{d}\,\frac{\omega_f + \omega_0}{\omega}$$

Here, r is the order up to which normalization of the Hamiltonian is carried on. This gives the complete proof of Theorem 4.1. ▮

We come now to the proof of Theorem 4.3. We need two more technical estimates. The first one is a straightforward generalization of a preliminary result already proved, in a slightly different context, in ref. 16. We give here the statement, which can be easily proved along the lines of the proof of Lemma 10.3 of the paper above.

Lemma 5.4. Let $\{\chi_l\}_{l \geqslant 1}$ be a generating sequence with

$$N_d^{\nabla}(\chi_l) \leqslant \frac{\beta^{l-1}}{l}\,\Phi$$

and let g be any analytic function; then for the s-th term ($s \geqslant 1$) of the sequence $T_\chi g$ [see (5.1)] one has

$$N_d(g_s) \leqslant \Phi\left(\frac{e\Phi}{d} + \beta\right)^{s-1} \frac{N_0(g)}{d} \tag{5.15}$$

The second technical estimate is given here with a detailed proof.

Lemma 5.5. Define E^* as in the statement of Theorem 4.3. Then we have

$$N_{2d}(\Psi_s) \leqslant E^*\left(r\,\frac{\mu}{e}\right)^{s-1} \tag{5.16}$$

where μ is defined by (4.3), and r is the order up to which normalization of the Hamiltonian has been performed.

Proof. We start from the definition (5.5) of Ψ_s, which, following ref. 16 (see p. 593), can be given the form

$$\Psi_1 = f$$
$$\Psi_s = \sum_{l=1}^{s-1} \frac{l}{s} \{\chi_l, Z_{s-l}\} + \frac{s-1}{s} \{\chi_{s-1}, \hat{h}\} + \frac{1}{s} f_{s-1} + \frac{1}{s} \hat{h}_{s-1} \tag{5.17}$$

Then, we look for a sequence η_s such that

$$N_{d_s}(\Psi_s) \leqslant E\eta_s, \qquad s \geqslant 1$$

where $d_s := d + (s-1)d/(r-1)$. We shall also use the following notation:

$$\mu_1 := \frac{e\Phi}{d} + \beta$$

So, using Eqs. (5.10) and (5.15), we obtain for this sequence the inequality

$$E\eta_s \leqslant \sum_{l=1}^{s-1} \frac{l(r-1)}{sd} E\Phi\eta_{s-1}\beta^{l-1} + \frac{s-1}{sd} \Phi\beta^{s-2}E_0 + \frac{1}{s} \Phi\mu_1^{s-2} \frac{E}{d} + \frac{1}{s} \Phi\mu_1^{s-2} \frac{E_0}{d}$$
$$< \frac{\Phi}{d} \left[\sum_{l=1}^{s-1} (r-1) E\beta^{l-1}\eta_{s-l} + \mu_1^{s-2}(E+E_0) \right]$$

Thus, we can define the sequence η_s by

$$\eta_1 = 1$$
$$E\eta_s = \frac{\Phi}{d} \left[\sum_{l=1}^{s-1} (r-1) E\beta^{l-1}\eta_{s-l} + \mu_1^{s-2}(E+E_0) \right], \qquad s \geqslant 2 \tag{5.18}$$

The rhs of the last equation can be rewritten, isolating in the sum the term with $l = 1$, as

$$\frac{\Phi}{d} \left[\sum_{l=2}^{s-1} (r-1) E\beta^{l-1}\eta_{s-l} + (r-1) E\eta_{s-1} + \mu_1^{s-2}(E+E_0) \right]$$
$$= \frac{\Phi}{d} \left[\beta \sum_{l=1}^{s-2} (r-1) E\beta^{l-1}\eta_{s-1-l} + \mu_1\mu_1^{s-1-2}(E+E_0) + (r-1) E\eta_{s-1} \right]$$
$$\leqslant \mu_1 E\eta_{s-1} + \frac{\Phi}{d}(r-1) E\eta_{s-1} \leqslant r\frac{\mu}{e} E\eta_{s-1}$$

which holds for $s \geqslant 3$. Here we used the obvious inequality

$$\frac{e\Phi}{d} + \beta + r\frac{e\Phi}{d} \leqslant r\frac{\mu}{e}$$

which follows from the definitions above of μ, μ_1, β, and Φ. So we can define a majorant $\tilde{\eta}_s$ for the sequence η_s by putting

$$\tilde{\eta}_1 = 1, \qquad E\tilde{\eta}_2 = \frac{\Phi}{d}(rE + E_0), \qquad \tilde{\eta}_s = \frac{r\mu}{e}\tilde{\eta}_{s-1}, \qquad s \geqslant 3$$

Then it is easy to check that this sequence is bounded by the rhs of (5.16). ∎

Scheme of the Proof of Theorem 4.3. By following the lines of the proof of Lemma 11.3 in ref. 16 one immediately gets the estimate

$$N_{2d}(h_s) \leqslant E\left(r\frac{\mu}{e}\right)^{s-1}$$

Using this estimate, the proof of Theorem 4.3 can be easily completed. ∎

6. PROOF OF THE THEOREM ON THE FPU-TYPE MODEL

The proof of Theorem 2.1 depends on the following steps. The first step is the definition of the phase space. The second step is the choice of the domain $\mathscr{G}$ in which perturbation theory will be developed; the form of the domain is strictly related to the form of the Hamiltonian, namely, the energy of the system: in fact, the domain will be such that, by conservation of energy, the solution of the equations of motion turns out to be trapped in it. The third step is the approximation of h^+ with a function h_ω giving rise to periodic orbits, and the explicit computation of the constants ω_0, ω_f, μ. The fourth and final step is the application of Corollary 4.2.

Concerning the phase space, our goal is to characterize $\mathscr{P}$ as the space of states with finite energy and corresponding to configurations with fixed ends. To this end, we start by defining

$$\mathscr{P} := l^2 \times \Delta_0^2 \ni (y, x)$$

where in the case of finite n, l^2 and Δ_0^2 coincide with $\mathbb{R}^n$, while in the case of infinite n, l^2 is the space of square-summable sequences, and Δ_0^2 is the closure of sequences with $x_0 = 0$ and finite support contained in $\mathbf{N}^+$ in the norm

$$\|x\|_{\Delta_0^2}^2 := \sum_{j \geqslant 1} |x_j - x_{j-1}|^2$$

To complete the definition of $\mathscr{P}$, we need to introduce a norm. The natural choice would be to define the norm of a point $\zeta \in \mathscr{P}$ as the square root of

$$\mathscr{E}(\zeta) := \sum_{j=1}^{n} \frac{|y_j|^2}{2m_j} + \sum_{j=1}^{n+1} \frac{1}{2} k \cdot |x_j - x_{j-1}|^2 \tag{6.1}$$

However, it is a better procedure to make use of the linear transformation (2.4) to normal modes, with new variables p, q. By the way, in the case of an infinite system this transformation takes the form

$$x_j = \sum_{\pm} \frac{2}{\pi} \int_0^{\pi} \left(\frac{2k}{|2k - m_j \omega_\beta^{\pm 2}|} \right)^{1/2} |\cos \beta| \frac{\sin(j\beta)}{[m_1^2 + m_2^2 + 2m_1 m_2 \cos(2\beta)]^{1/4}} q_\beta^\pm \, d\beta$$

In terms of these variables we define the norm of a point ζ as

$$\|\zeta\|_{\mathscr{P}}^2 = \sum_{l=1}^{n/2} \tfrac{1}{2}(|p_l^-|^2 + \omega_l^{-2} |q_l^-|^2) + \sum_{l=1}^{n/2} \tfrac{1}{2}(|p_l^+|^2 + \omega^2 |q_l^+|^2) \qquad (6.2)$$

where ω is given by (2.6). We remark that this norm is equivalent to the norm (6.1) given by the harmonic energy, with equivalence constants independent of the number n of degrees of freedom; in fact in terms of the variables p, q the harmonic energy is nothing but

$$\mathscr{E}(\zeta) = \sum_{\substack{\pm \\ l=1}}^{n/2} \tfrac{1}{2}(|p_l^\pm|^2 + (\omega_l^\pm)^2 |q_l^\pm|^2)$$

and the relation with the norm (6.2) is given by

$$\left(2 \frac{m_1 + m_2}{2m_2 + m_1} \right)^{-1} \mathscr{E}(\zeta) \leqslant \|\zeta\|^2 \leqslant \left(1 + \frac{m_1}{m_2} \right) \mathscr{E}(\zeta) \qquad (6.3)$$

In particular, taking into account (2.8) and the hypothesis $\mu < 1$, one has

$$(1 + 2^{-14})^{-1} \mathscr{E}(\zeta) \leqslant \|\zeta\|^2 \leqslant (1 + 2^{-15}) \mathscr{E}(\zeta) \qquad (6.4)$$

Concerning $\mathscr{C}_0$, we simply put $\mathscr{C}_0 = \mathscr{P}$, so that we also have $\mathscr{F} = \mathscr{P}$.

The definition of the domain is almost trivial: the domain $\mathscr{G}$ is the origin, namely $\mathscr{G} = \{0\}$; the complex extended domain turns out to be

$$\mathscr{G}_{R,0} := B(0, R)$$

namely the ball (in $\mathscr{P}^{\mathrm{C}}$) centered in 0 and having radius R. We also choose $d = 1/4$.

Choosing R small enough will ensure that all the orbits with initial harmonic energy sufficiently small do not escape from $\mathscr{G}_{R,2d} = B(0, R/2)$ for all times.

In order to apply the theory of Section 4, we split the Hamiltonian (2.1) as

$$h_\omega(\zeta) := \sum_{l=1}^{n/2} \tfrac{1}{2}(p_l^{+2} + \omega^2 q_l^{+2})$$

$$\hat{h}(\zeta) := \sum_{l=1}^{n/2} \tfrac{1}{2}(p_l^{-2} + \omega_l^{-2} q_l^{-2})$$

$$f(\zeta) := f_1(\zeta) + f_2(\zeta)$$

where

$$f_1(\zeta) = \sum_{l=1}^{n/2} \tfrac{1}{2}(\delta\omega_l)^2 q_l^{+2}$$

$$f_2(\zeta) := \sum_{j=1}^{n+1} \varphi(x_j - x_{j-1})$$

$$(\delta\omega_l)^2 := \omega_l^{+2} - \omega^2$$

Here we used the variables (y, x) or (p, q) according to convenience. Due to our choice of the norm, it is clear that assumption 5 of Theorem 4.1 holds (the differential of the unperturbed flow has norm equal to 1).

The estimate of the constants ω_f and μ relies on a quite elementary estimate of the symplectic gradient of f_2. This estimate is the key step in order to prove Theorem 2.1, and is reported below in the proof of the following result.

Lemma 6.1. Using the above notation, we have

$$N_0^\nabla(f_1) + N_0^\nabla(f_2) \leqslant \omega_f, \qquad N_0^\nabla(\hat{h}) \leqslant \omega_{\max}^-$$

with

$$\omega_f = (1 + 2^{-11}) \frac{2\Upsilon}{k\sqrt{m_1}} + \frac{k/m_2}{\omega}$$

where Υ is as in Theorem 2.1; moreover one has

$$\mu < 192\pi e \left[(1 + 2^{-11}) \frac{\Upsilon R}{k^{3/2}} + \frac{\omega_{\max}^-}{\omega} \right] \tag{6.5}$$

Proof. We begin by evaluating the norm of the symplectic gradient of f_2. Notice that the y_j component of $\nabla^\Omega f_2(\zeta)$ is

$$-\frac{\partial f_2(x)}{\partial x_j} = \varphi'(x_{j+1} - x_j) - \varphi'(x_j - x_{j-1})$$

while its x_j components vanish; therefore, using the definition of norm (6.2), and the equivalence relation (6.4), we have

$$\|\nabla^{\Omega} f_2(\zeta)\|^2 \leqslant (1+2^{-15})\, \mathscr{E}(\nabla^{\Omega} f_2(\zeta))$$
$$= (1+2^{-15}) \sum_{j=1}^{n+1} \frac{1}{2m_j} |\varphi'(x_j - x_{j-1}) - \varphi'(x_{j+1} - x_j)|^2$$

Using $|\varphi'(x)| \leqslant \Upsilon |x|^2$, we have

$$\|\nabla^{\Omega} f_2(\zeta)\|^2 \leqslant (1+2^{-15})\, 4\Upsilon^2 \sum_{j=1}^{n} \frac{|x_j - x_{j-1}|^4}{2m_j}$$
$$\leqslant (1+2^{-15}) \frac{4\Upsilon^2}{k^2 m_1} \left(\frac{k}{2} \sum_{j=1}^{n+1} |x_j - x_{j-1}|^2 \right)^2$$
$$\leqslant (1+2^{-11}) \frac{4\Upsilon^2}{k^2 m_1} \|\zeta\|_{\mathscr{P}}^4$$
$$\leqslant (1+2^{-11}) \frac{4\Upsilon^2}{k^2 m_1} R^4$$

Concerning f_1, we have

$$\|\nabla^{\Omega} f_1(\zeta)\|^2 = \frac{1}{2} \sum_{l=1}^{n/2} |\delta\omega_l^2 q_l^+|^2 \leqslant \max_l \frac{\delta\omega_l^4}{\omega^2} \|\zeta\|^2 \leqslant \frac{(k/m_2)^2}{\omega^2} R^2$$

from which the value of ω_f is easily deduced. The estimate of ω_0 and the calculation of μ are then trivial: just notice that, by (2.9) and by $\mu < 1$, we have $\omega_{\max}^- /(2\omega) \leqslant 1/(48e)$. ∎

With the settings and the lemma above, we can apply Corollary 4.2, and conclude that $|h_\omega(t) - h_\omega(0)|$ is bounded up to the exponential time T_*, provided the orbit does not escape from the domain $\mathscr{G}_{R,2d}$. We show that R can be determined in such a way that the escape time is actually infinite, thanks to energy conservation. Indeed, remark that the condition $\zeta \in B(0, R)$ can be expressed in terms of harmonic energy, because of the equivalence of the norm $\|\zeta\|$ and the harmonic energy $\mathscr{E}(\zeta)$ stated above. By conservation of energy, on the one hand one has

$$|\mathscr{E}(t)| \leqslant |\mathscr{E}(0)| + |f_2(0)| + |f_2(t)| \tag{6.6}$$

and on the other hand, one also has $|f_2(x(t))| \leqslant \Upsilon R^3/(3k^{3/2})$, so that, by (6.4), we get

$$\|\zeta(t)\| \leqslant (1+2^{-15})\, \mathscr{E}(t) \leqslant (1+2^{-15}) \mathscr{E}_0 + (1+2^{-15}) \frac{2\Upsilon R^3}{3k^{3/2}}$$

If the rhs of this expression is less than $R^2/4$, then we certainly have $\zeta(t) \in \mathscr{G}_{R,1/2}$. So, we obtain the inequality

$$(1+2^{-15})\frac{2\Upsilon R^3}{3k^{3/2}} - \frac{R^2}{4} + (1+2^{-15})\mathscr{E}_0 \leqslant 0 \tag{6.7}$$

which has solutions only if

$$\mathscr{E}_0 \leqslant \frac{3}{2^7}\frac{k^3}{\Upsilon^2}\frac{1}{1+2^{-15}} \tag{6.8}$$

This is satisfied in view of (2.8). We satisfy also inequality (6.7) by taking

$$R = [\tfrac{8}{3}\mathscr{E}_0(1+2^{-15})]^{1/2}$$

Inserting this value in (6.5) and using it to estimate the rhs of (4.5), one concludes the proof of the theorem.

7. PROOF OF THE ANALYTIC THEOREM ON WEAKLY COUPLED ROTATORS

First remark that our choice of $(\mathscr{P}, \mathscr{C}_0, \Omega)$ (see Sections 3 and 4) satisfies axioms (α) and (β) on the phase space: the proof is straighforward.

Now we prove Theorem 3.2; therefore we fix our attention on the Hamiltonian (3.7). Our aim is to choose a domain, to give the Hamiltonian a form suited for the application of Corollary 4.4, and to compute the corresponding constants. After application of the corollary, we shall estimate the escape time from the domain.

We start with the definition of the domain. Denote

$$G = \beta(\varepsilon/2)^{1/2} \tag{7.1}$$

where β is defined by (3.8); we put

$$\tilde{\mathscr{G}} := \{(J, \phi) \in \mathscr{P} : h^R(\zeta) \leqslant G^2\}, \qquad \mathscr{G} = \bigcup_{t \in [0,T]} \Phi_t(\tilde{\mathscr{G}})$$

with h^R defined by (3.2); notice that, if $G^2 \geqslant 2\varepsilon$, this is an unbounded domain, because no bound is imposed on the norm of the angles (see Section 3). The domain $\mathscr{G}$ is extended in the complex as explained in Section 4, via another positive parameter R, which will be fixed later (see (7.2)).

Now we split the perturbation into a part which is already in normal form with respect to the unperturbed system and a remaining part which will be the actual perturbation. Precisely, we define

$$h_{\omega}(J, \phi) := \sum_{l \in \mathbf{Z}} \omega_l J_l$$

$$\hat{h}(J, \phi) := \sum_{l \in \mathbf{Z}} \frac{J_l^2}{2I} + \varepsilon \sum_{l \notin S^*} [1 - \cos(\phi_l - \phi_{l-1})]$$

$$f(\phi) := \varepsilon \sum_{l \in S^*} [1 - \cos(\phi_l - \phi_{l-1})]$$

where S^* is again the set

$$S^* := \{i \in \mathbf{Z}: \omega_i \neq \omega_{i-1}\}$$

According to our abstract scheme, h_{ω} is the unperturbed system, and $\hat{h}$ commutes with it, while f plays the role of the perturbation. This form of the splitting, in particular for the ϕ-dependent terms, is crucial in order to get a good estimate of the deformation induced by the canonical transformation. In turn, this will play a fundamental role in controlling the escape time.

We come now to the computation of the constants. For clarity, we give the result by stating a few lemmas.

Lemma 7.1. We can choose

$$\omega_f = \frac{n^* \sqrt{2} \exp[R(2/\varepsilon)^{1/2}] \varepsilon}{R\sqrt{I}}$$

Proof. The J_l component of $\nabla^{\Omega} f$ is given by

$$\varepsilon \sum_{i \in S^*} \sin(\phi_i - \phi_{i-1})(\delta_l^i - \delta_l^{i-1})$$

where δ_l^i is the Kroneker symbol; so we have

$$\begin{aligned}
\|\nabla^{\Omega} f(\zeta)\|^2 &= \sum_{l \in \mathbb{Z}} \frac{\varepsilon^2}{2I} \left[\sum_{i \in S^*} \sin(\phi_i - \phi_{i-1}) (\delta_l^i - \delta_l^{i-1}) \right]^2 \\
&\leqslant \frac{\varepsilon^2}{2I} \exp[2R(2/\varepsilon)^{1/2}] \sum_{l \in \mathbf{Z}} \left(\sum_{i \in S^*} |\delta_l^i - \delta_l^{i-1}| \right)^2 \\
&\leqslant \frac{\varepsilon^2}{2I} \exp[2R(2/\varepsilon)^{1/2}] \sum_{l \in \mathbf{Z}} n^* \sum_{i \in S^*} (\delta_l^i - \delta_l^{i-1})^2 \\
&\leqslant \frac{\varepsilon^2}{2I} \exp[2R(2/\varepsilon)^{1/2}]\, 2n^* \sum_{l \in \mathbf{Z}} \sum_{i \in S^*} [(\delta_l^i)^2 + (\delta_l^{i-1})^2]
\end{aligned}$$

where we used Schwartz's inequality. From the above estimate the thesis easily follows. ▮

The first use of this lemma is in determining the parameter R controlling the size of the complex extension of the domain $\mathscr{G}$. Since ω_f is proportional to $\exp[R(2/\varepsilon)^{1/2}]$, is is natural to choose R proportional to $\varepsilon^{1/2}$. For simplicity, we take

$$R=\left(\frac{\varepsilon}{2}\right)^{1/2} \tag{7.2}$$

so that we get

$$\omega_f = 2n^* e\left(\frac{\varepsilon}{I}\right)^{1/2}$$

We come now to the calculation of ω_0. First of all, remark that, since $\hat{h}$ commutes with h_ω, one has $\hat{h}(\Phi_t(\zeta)) = \hat{h}(\zeta)$; from this one easily deduces

$$\operatorname*{Sup}_{\zeta\in\mathscr{G}_{R,d}} \|\nabla^{\Omega}\hat{h}(\zeta)\| = \operatorname*{Sup}_{\zeta\in\tilde{\mathscr{G}}_{R,d}} \|\nabla^{\Omega}\hat{h}(\zeta)\|$$

We state now two simple lemmas.

Lemma 7.2. We have

$$\operatorname*{Sup}_{\|\phi\|\leqslant R} \varepsilon \sum_{l\in\mathbf{Z}} |1-\cos(\phi_l-\phi_{l-1})| \leqslant \varepsilon \tag{7.3}$$

Proof. Using the Taylor expansion of the cosine, we have

$$\begin{aligned}
&\varepsilon \sum_{l\in\mathbf{Z}} |1-\cos(\phi_l-\phi_{l-1})|\\
&\quad\leqslant \varepsilon \sum_{j\geqslant 1} \frac{1}{(2j)!} \sum_{l\in\mathbf{Z}} |\phi_l-\phi_{l-1}|^{2j}\\
&\quad\leqslant \varepsilon \sum_{j\geqslant 1} \frac{1}{(2j)!}\left(\frac{2}{\varepsilon}\right)^j \left[\sum_{l\in\mathbf{Z}} \frac{\varepsilon}{2} |\phi_l-\phi_{l-1}|^2\right]^j\\
&\quad\leqslant \varepsilon \sum_{j\geqslant 1} \frac{1}{(2j)!}\left(\frac{\sqrt{2}R}{\sqrt{\varepsilon}}\right)^{2j}
\end{aligned}$$

which, due to (7.2), is less than the rhs of (7.3). ▮

We remark that, as a byproduct of the proof of this lemma, one also concludes that Hamiltonian (3.2) is analytic.

Lemma 7.3. We have

$$\operatorname*{Sup}_{\phi\in\mathscr{G}_{R,0}} \varepsilon \sum_{l\in\mathbf{Z}} |1-\cos(\phi_l-\phi_{l-1})| \leqslant \varepsilon\left(\frac{\beta^2}{2}+2\beta+1\right)$$

Proof. Denote for simplicity $\psi_k := \phi_k - \phi_{k-1}$ with $\phi\in\mathscr{G}$ and $\alpha_k := \alpha^1_k - \alpha^1_{k-1}$, with $\alpha^1\in\mathscr{P}$, $\|\alpha^1\|\leqslant R$. Then compute

$$\begin{aligned}
&\varepsilon \sum_{l\in\mathbf{Z}} |1-\cos(\psi_l+\alpha_l)| \\
&\quad= \varepsilon \sum_{l\in\mathbf{Z}} |1-\cos\psi_l\cos\alpha_l+\sin\psi_l\sin\alpha_l| \\
&\quad\leqslant \varepsilon \sum_{l\in\mathbf{Z}} |1-\cos\psi_l(\cos\ \alpha_l-1+1)| + \varepsilon\sum_{l\in\mathbf{Z}} |\sin\psi_l\sin\alpha_l| \\
&\quad\leqslant \varepsilon \sum_{l\in\mathbf{Z}} |1-\cos\psi_l| + \varepsilon\sum_{l\in\mathbf{Z}} |\cos\psi_l|\cdot|1-\cos\alpha_l| \\
&\qquad + \varepsilon\left(\sum_{l\in\mathbf{Z}} |\sin\psi_l|^2\right)^{1/2}\left(\sum_{l\in\mathbf{Z}} |\sin\alpha_l|^2\right)^{1/2} \\
&\quad\leqslant G^2+\varepsilon+\varepsilon\left(\sum_{l\in\mathbf{Z}} |1-\cos^2\psi_l|\right)^{1/2}\left(\sum_{l\in\mathbf{Z}} |1-\cos^2\alpha_l|\right)^{1/2} \\
&\quad\leqslant G^2+\varepsilon+\left\{1+\mathrm{Ch}\left[R\left(\frac{2}{\varepsilon}\right)^{1/2}\right]\right\}^{1/2}(2\varepsilon)^{1/2}G
\end{aligned}$$

From this, Using (7.1), the thesis follows. ∎

We are now able to calculate ω_0:

Lemma 7.4. We can take

$$\omega_0 := 8\beta\left(\frac{\varepsilon}{I}\right)^{1/2} \tag{7.4}$$

Proof. Denote again $\psi_l := \phi_l - \phi_{l-1}$, now with $\phi\in\mathscr{G}_{R,0}$. We have

$$\begin{aligned}
\|\nabla^\Omega\hat{h}(\zeta)\|^2 &\leqslant \sum_{l\in\mathbf{Z}} \frac{\varepsilon^2}{2I}4|\sin^2\psi_l| + \sum_{l\in\mathbf{Z}} \frac{\varepsilon(p_l-p_{l-1})^2}{2I^2} + 2\varepsilon\operatorname*{Sup}_l\left\{\frac{|p_l|^2}{I^2}\right\} \\
&\leqslant \frac{2\varepsilon^2}{I}\sum_{l\in\mathbf{Z}} |(1-\cos\psi_l)(1+\cos\psi_l)| + \frac{8\varepsilon}{I}\sum_{l\in\mathbf{Z}} \frac{|p_l|^2}{2I}
\end{aligned}$$

but the first term of this sum is less than

$$\begin{aligned}
&\frac{2\varepsilon}{I}\operatorname{Sup}_{l}\{|1+\cos\psi_l|\}\,\varepsilon\sum_{l\in\mathbf{Z}}|1-\cos\psi_l| \\
&\quad\leqslant\frac{2\varepsilon}{I}\left\{1+\operatorname{Ch}\left[R\left(\frac{2}{\varepsilon}\right)^{1/2}\right]\right\}\left[\varepsilon\left(\frac{\beta^2}{2}+2\beta+1\right)\right] \\
&\quad<\frac{2\varepsilon}{I}(1+e)\frac{7}{2}\varepsilon\beta^2
\end{aligned}$$

where we used $\beta>1$. From this, using Eq. (7.2), the thesis easily follows. ∎

The last lemma gives the values of E, E_0, and E^*.

Lemma 7.5. We can take

$$E=8n^*\varepsilon, \qquad E_0=4\varepsilon\beta^2, \qquad E^*=\left(8n^*+\frac{n^*\beta}{6}\right)\varepsilon$$

Proof. Using the result of Lemma 7.3, one immediately gets the values of E and E_0. The value of E^* is obtained by using the trivial inequality

$$2^6e^2\beta\left(\frac{\varepsilon}{Iv^2}\right)^{1/2}<\frac{1}{6}\mu<\frac{1}{6}$$

which follows from the definition (3.9) of μ. ∎

We give now the following result.

Proof of Theorem 3.2. Recall the definitions (7.2) of R and (7.1) of β. The definition (3.9) of μ together with the hypothesis $\mu<1$ of the theorem ensure that the constants ω_f, ω_0 as given above and the constant v in the statement of the theorem satisfy the hypotheses of Corollary 4.4. The straightforward application of the corollary gives the estimate

$$|h_\omega(t)-h_\omega(0)|\leqslant 6\varepsilon\left(8n^*+\frac{n^*\beta}{6}\right) \tag{7.5}$$

for $|t|\leqslant\min(T_0, T_*)$. We show now that the choice (3.8) for the parameter β implies $T_0>T_*$. By the definition of β one has $h^R(\zeta(0))\leqslant G^2/2$; but by the conservation of energy we have

$$h^R(\zeta(t))\leqslant h^R(\zeta(0))+|h_\omega(t)-h_\omega(0)|$$

and so also

$$h^R(\zeta(t)) \leqslant \frac{G^2}{2} + |h_\omega(t) - h_\omega(0)|$$

Using (7.5), we can bound the rhs of this relation by

$$\frac{G^2}{2} + 6\varepsilon\left(8n^* + \frac{n^*\beta}{6}\right)$$

If this is less than G^2, then we surely have $\zeta(t) \in \tilde{\mathscr{G}} \subset \mathscr{G}$. So, using (7.2) and (7.1), we obtain that this is true, provided that β satisfies the inequality

$$\frac{1}{2}\beta^2 \geqslant 12\left(8n^* + \frac{n^*\beta}{6}\right)$$

which therefore ensures $T_0 > T_*$: we replace it by the stronger condition

$$\beta \geqslant 12n^* \tag{7.6}$$

The choice (3.8) for β satisfies the latter condition. This concludes the proof of the theorem. ∎

8. PROOF OF THE GEOMETRIC THEOREM ON WEAKLY COUPLED ROTATORS

First, we denote by

$$(\omega_1,..., \omega_n) := \left(\frac{\bar{J}_1}{I},..., \frac{\bar{J}_n}{I}\right)$$

the frequencies corresponding to the vector $\bar{J}$. Then we use the Dirichlet theorem in order to approximate this set of frequencies with a completely resonant vector: let $\tilde{\nu}$ be the maximum among $\omega_1,..., \omega_n$, and let $Q \geqslant 1$ be an arbitrary real number; according to Dirichlet theorem there exist $k_1,..., k_n \in \mathbf{N}$ and $q \in \mathbf{N}$ with $q \leqslant Q$, such that

$$\left|\omega_j - \frac{\tilde{\nu}k_j}{q}\right| \leqslant \frac{\tilde{\nu}}{qQ^{1/(n-1)}}$$

So, in terms of points of the phase space, we can say that corresponding to the point $\bar{J}$ there exists a resonant point $(J^{(r)}, 0) \in \mathscr{P}$ such that (i) the corresponding unperturbed orbit is periodic, with frequency

$$\nu := \frac{\tilde{\nu}}{q} \tag{8.1}$$

and (ii) it is near $(\bar{J}, 0)$:

$$\|(J^{(r)}, 0) - (\bar{J}, 0)\|_{\mathscr{P}} \leqslant (nIv^2)^{1/2} \frac{1}{Q^{1/(n-1)}}$$

So, if we denote by ζ_1 the difference between the initial datum (J^0, ϕ^0) and $(\bar{J}, 0)$ by the triangle inequality we have

$$\{h^R[(J^0, \phi^0) - (J^{(r)}, 0)]\}^{1/2} \leqslant [h^R(\zeta_1)]^{1/2} + (nIv^2) \frac{1}{Q^{1/(n-1)}}$$

It follows that we can apply Theorem 3.2 with v given by (8.1), $n^* = 2n$, and β given by

$$\beta := \max \left\{ 2 \left[\frac{h^R(\zeta_1)}{\varepsilon} \right]^{1/2} + 2\sqrt{n} \left(\frac{Iv^2}{\varepsilon} \right)^{1/2} \frac{1}{Q^{1/(n-1)}}, 24n \right\}$$

Here, we still have the free parameter Q; to fix it, we shall use the explicit expression (3.9) of μ as a function of v and β, and optimize the choice of Q in order to make μ as small as possible. Thus, we calculate

$$\begin{aligned} \mu &< 2^5 3^3 \left[\frac{\sqrt{n}}{Q^{1/(n-1)}} + (26n + \beta') \left(\frac{\varepsilon}{Iv^2} \right)^{1/2} \right] \\ &\leqslant 2^5 3^3 \sqrt{n} \left[\frac{1}{Q^{1/(n-1)}} + \left(\frac{\beta'}{\sqrt{n}} + 26\sqrt{n} \right) \tilde{\mu} Q \right] \end{aligned} \tag{8.2}$$

where

$$\tilde{\mu} := \left(\frac{\varepsilon}{I\tilde{v}^2} \right)^{1/2}, \qquad \beta' := 2 \left[\frac{h^R(\zeta_1)}{\varepsilon} \right]^{1/2}$$

Minimizing with respect to Q, we get

$$\mu < 2^5 3^3 \sqrt{n} \left[\left(\frac{\beta'}{\sqrt{n}} + 26\sqrt{n} \right) \tilde{\mu} \right]^{1/n}$$

$$Q = \left[(n-1) \left(\frac{\beta'}{\sqrt{n}} + 26\sqrt{n} \right) \tilde{\mu} \right]^{-(n-1)/n}$$

Furthermore, using (3.6), we can estimate β by

$$\beta < 25\sqrt{n}\tilde{\mu}^{(1-n)/n}$$

We also have

$$\|\delta\omega\| \leqslant \left[4\tilde{v}^2 + \sum_{j \in S^*} (2\omega_j)^2 \right]^{1/2} < 4\tilde{v}(n^*)^{1/2}$$

Replacing the expressions above for μ, β, and $\|\delta\omega\|$ in the statement of Theorem 3.2, we find that the result of Theorem 3.1 easily follows. ∎

ACKNOWLEDGMENTS

During the preparation of this paper we had several discussions with G. Benettin and F. Fassò. Their comments and suggestions were illuminating. We also thank L. Galgani for his constant encouragement and his stimulating remarks.

REFERENCES

1. J. Fröhlich, T. Spencer, and C. E. Wayne, Localization in disordered, nonlinear dynamical systems, *J. Stat. Phys.* **42**:247–274 (1986).
2. M. Vittot and J. Bellissard, Invariant tori for an infinite lattice of coupled classical rotators, Preprint (1985).
3. J. Pöschel, Small divisors with spatial structure in infinite dimensional Hamiltonian systems, *Commun. Math. Phys.* **127**:351–393 (1990).
4. S. B. Kuksin, Hamiltonian perturbations of infinite-dimensional linear systems with imaginary spectrum, *Funkt. Anal. Prilozhen.* **21**(3):22–37 (1987) [*Funct. Anal. Appl.* **21** (1987)].
5. C. E. Wayne, Periodic and quasi-periodic solutions of nonlinear wave equation via KAM theory, *Commun. Math. Phys.* **127**:479–528 (1990).
6. L. Chierchia and P. Perfetti, Maximal almost periodic solutions for Lagrangian equations on infinite dimensional tori, Preprint (1992).
7. S. B. Kuksin, Perturbation theory for quasiperiodic solutions of infinite dimensional Hamiltonian systems. 1. Symplectic structures and Hamiltonian scales of Hilbert spaces, Max Planck Institut für Mathematik, preprint MPI/90-99; Perturbation theory for quasiperiodic solutions of infinite dimensional Hamiltonian systems. 2. Statement of the main theorem and its consequences, Max Planck Institut für Mathematik, preprint MPI/90-100.
8. G. Benettin, J. Fröhlich, and A. Giorgilli, A Nekhoroshev-type theorem for Hamiltonian systems with infinitely many degrees of freedom, *Commun. Math. Phys.* **119**:95–108 (1988).
9. D. Bambusi, A Nekhoroshev-type theorem for the Pauli–Fierz model of classical electrodynamics, Department of Mathematics, University of Milan, preprint 43/1991.
10. S. B. Kuksin, An averaging theorem for distributed conservative systems and its application to Von Karman's equation, *PMM USSR* **53**:150–157 (1989).
11. G. Benettin, L. Galgani, and A. Giorgilli, Classical perturbation theory for systems of weakly coupled rotators, *Nuovo Cimento B* **89**:89–102 (1985).
12. G. Benettin, L. Galgani, and A. Giorgilli, Numerical investigations on a chain of weakly coupled rotators in the light of classical perturbation theory, *Nuovo Cimento B* **89**:103–119 (1985).

13. C. E. Wayne, The KAM theory of systems with short range interaction, I, *Commun. Math. Phys.* **96**:311–329 (1984); The KAM theory of systems with short range interaction, II, *Commun. Math. Phys.* **96**:331–344 (1984).
14. C. E. Wayne, On the elimination of non-resonant harmonics, *Commun. Math. Phys.* **103**:351–386 (1986); C. E. Wayne, Bounds on the trajectories of a system of weakly coupled rotators, *Commun. Math. Phys.* **104**:21–36 (1986).
15. L. Galgani, A. Giorgilli, A. Martinoli, and S. Vanzini, On the problem of energy equipartition for large systems of the Fermi–Pasta–Ulam type: Analytical and numerical estimates, *Physica D* **59**:334–348 (1992).
16. G. Benettin, L. Galgani, and A. Giorgilli, Realization of holonomic constraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory, Part II, *Commun. Math. Phys.* **121**:557–601 (1989).
17. P. Lochak, Canonical perturbation theory via simultaneous approximation, *Uspekhi Math. Nauk*, to appear.
18. M. P. Chernoff and J. E. Marsden, *Properties of Infinite Dimensional Hamiltonian Systems* (Springer-Verlag, Berlin, 1974).
19. Y. Choquet-Bruhat, C. De Witt-Morette, and M. Dillard-Bleick, *Analysis, Manifolds and Physics* (North-Holland, Amsterdam, 1978).
20. A. Giorgilli and L. Galgani, Formal integrals for an autonomous Hamiltonian system near an equilibrium point, *Cel. Mech.* **17**:267–280 (1978).
21. S. Lang, *Differential Manifolds* (Springer-Verlag, New York, 1985).
22. J. Mujica, *Complex Analysis in Banach Spaces* (North-Holland, Amsterdam, 1986).
23. J. Moser, Regularization of Kepler's problem and the averaging method on a manifold, *Commun. Pure Appl. Math.* **23**:609–636 (1970).
24. R. Cushman, Normal form of Hamiltonian vectorfields with periodic flows, in *Differential Geometric Methods in Mathematical Physics*, S. Sterneberg, ed. (Reidel, Dordrecht, 1984).