M P E J

Mathematical Physics Electronic Journal

ISSN 1086-6655
Volume 9, 2003

Paper 1
Received: Feb 28, 2003, Accepted: Apr 14, 2003
Editor: C. Liverani

A KAM theorem for infinite–dimensional discrete systems

Paolo Perfetti

Dipartimento di matematica, II Università degli Studi di Roma
Via della Ricerca Scientifica 00133 Roma, Italy E–mail: perfetti@mat.uniroma2.it

Abstract Infinite–dimensional, discrete hamiltonian systems of the type kinetic energy + potential energy over $\mathbb{R}^{\mathbb{Z}} \times \mathbb{T}^{\mathbb{Z}}$ are studied. The existence of many quasi–periodic motions with a maximal set of nonzero frequencies is shown.

Supported by Istituto Nazionale di Alta Matematica "Francesco Severi" and Ministero dell'Università e della Ricerca Scientifica e Tecnologica, research program "Metodi variazionali ed equazioni differenziali nonlineari"

Introduction

We study a family of hamiltonian dynamical systems of the form kinetic energy plus potential energy (possibly formal) having an infinite number of degrees of freedom (hereafter d.o.f.). The potential energy is a multi–periodic function of infinite variables whose period is 2π and the phace–space is $\mathbb{R}^{\mathbb{Z}} \times \mathbb{T}^{\mathbb{Z}}$ ($\mathbb{T}$ is the 2π–periodic torus). We prove that under suitable conditions of growth on the kinetic part and quite general conditions on the potential functions, the systems admit quasi–periodic motions with frequencies $\{\omega_i\}_{i\in\mathbb{Z}}$ such that $|\omega_i| \to +\infty$ as $|i| \to +\infty$. The equations of motions are

$$\dot{A}_i = f_i(\varphi) \qquad \dot{\varphi}_i = \frac{\partial h(A_i)}{\partial A_i} \qquad i \in \mathbb{Z}, \quad A_i \in \mathbb{R}, \ \varphi_i \in \mathbb{T}, \quad \mathbb{T} = \mathbb{R}/2\pi\mathbb{Z} \tag{1.1}$$

and if we think of f_i as minus the gradient of a function $V(\varphi)$ respect to φ_i, they are the canonical equations of the (formal) hamiltonian $H(A,\varphi) = \sum_{i\in\mathbb{Z}} h(A_i) + V(\varphi)$; H is formal because the growth of $h(A_i)$ makes the sum unbounded and because the function $V(\varphi)$ may not exist as a function over $\mathbb{T}^{\mathbb{Z}}$. What must be well defined are the functions $(f_1(\varphi), f_2(\varphi), \ldots)$. On h, which is the same for each identical particle, we impose the condition $0 < a' \le |h_{_{AA}}| \le a < \infty$ for $A \in \mathbb{R}$. If $f(\varphi) \doteq (f_1, f_2, \ldots) \equiv 0$, all the motions are linear in time over $\mathbb{R}^{\mathbb{Z}} \times \mathbb{T}^{\mathbb{Z}}$ that is $(A,\varphi) = \{(A_i(t), \varphi_i(t)\} = \{(A_i^o, \varphi_i^o + \omega_i t)\}$, $i \in \mathbb{Z}$. The infinite–dimensional vector $\{\omega_i\}_{i\in\mathbb{Z}}$ is the vector–frequency. If not all the functions f_i are identically zero instead, almost all (in the sense that the excluded subset of motions have a small relative measure) linear motions survive and are continued into $\{(A_i^o + \alpha_i(t), \varphi_i^o + \omega_i t + \beta_i(t))\}$ with $|\alpha_i|$ and $|\beta_i|$ small provided that: 1) subsets of ω like $\omega^{(l)} \doteq (\omega_1, \omega_2, \ldots \omega_l)$ satisfy the Diophantine condition $|\omega^{(l)} \cdot \nu|^{-1} \le \gamma e^{|\nu|^\tau/\tau}$ for all $\nu \in \mathbb{Z}^l \backslash \{0\}$, $0 < \tau < 1$ (actually we will take $0 < \tau < \frac{1}{2}$; γ is a positive number), 2) $|\omega_i|$ is sufficiently large.
By our techniques we cannot relax the condition $|h_{AA}| \le a$ allowing $|\frac{h_A}{A}|$ to be unbounded for $|A| \to +\infty$. The same model has been studied in [P] but a much greater lower bound on $|\omega_i|$ is necessary because we impose a different Diophantine condition on the frequencies; namely $|\omega^{(l)} \cdot \nu|^{-1} \le \gamma |\nu|^l \ \forall\ \nu \in \mathbb{Z}^l \backslash \{0\}$.
For a model represented by (1.1) one can think of an array of rotators: point particles constrained to move circularly and centered on the sites of $\mathbb{N}$. In this case the kinetic energy of a single particle is quadratic in the action–variable $h(A) = \frac{1}{2}A^2$.
The problem considered here falls within the family of the problems concerning the search of almost–periodic solutions in systems with infinitely many d.o.f. (originating in statistical mechanics for instance). Viewing (1.1) as the equations of infinite particles, the existence of the solutions described in this paper would imply that energy remains confined in suitable subsets of phace–space during all the time–evolution. The drawback (up to now) is that the energy–dependence of each d.o.f. is too large (faster than any power but slower than any exponential). We obtain $e^{(|j| \ln^2 |j|)^\alpha}$ $0 < \alpha < 1$ (see the upper bound on $\gamma\Omega_{l_k}$ before Theorem 4.7)). For more realistic rates it might be appropriate to consider a particular (but quite general) interaction $\{f_i\}$. An example is $f_i(\varphi_{i-1}, \varphi_i, \varphi_{i+1}) = a_i \cos(\varphi_i - \varphi_{i-1}) + a_{i+1} \cos(\varphi_{i+1} - \varphi_i)$ (or something like it) and then to study carefully the geometry of the small denominators (see for instance [BFG] where a Nekhoroshev–type theorem is obtained for infinitely many harmonic oscillator coupled by a suitable potential or [Wa1]–[Wa3] where good estimates on the perturbative parameter are obtained for chains of a very large *but finite* number of rotator coupled through anelastic forces [1]. Infinite–dimensional extensions

[1] Nevertheless the results of [Wa1] can be hardly extended to cover the infinite–dimensional theory because the potential is needed to be analytic over a strip around the real axis whose size is proportional to NlnN where N is the number of particles (see (1.14) at p.314)

of KAM theorem have been already studied but with some differences respect to this paper. [FSW], [VB] and [Pö] consider perturbations (the $\{f_i\}'s$) depending on the action–variables (say $(I_1, I_2, I_3, \ldots)$) which are extremely localized in phace–space ($|I_j(0)| \sim e^{-|j|^{1+\alpha}}$ $\alpha > 0$ ($e^{-|j|^{d+\alpha}}$ $\alpha > 0$ for a d–dimensional) problem in [FSW] and [VB] while at least $e^{-\ln^\gamma |j|}$ with $\gamma > 1$ in [Pö]). Moreover according to our definitions (see (2.3) and (2.4)), their models are all of short–range type [2]. A first improvement has been obtained in [CP] where the short–range condition has been relaxed allowing each particle to interact directly with each other particle "of the world." Also our models provide perturbation functions which are independent on the actions–variables. Although closely related, the result in [P] is slightly more general than [CP] because the kinetic part of the hamiltonian, like here, can have a quadratic part plus a *correction* in such a way that the leading term is at most quadratric for $|A| \to +\infty$. In [P] as well as in [CP], the size of the frequencies is of order $e^{p|j|\ln|j|}$ where $p \sim 10$ ($e^{pd|j|\ln|j|}$ for a d–dimensional problem). Here the size has been lowered to $e^{(|j|\ln^2|j|)^\alpha}$, $0 < \alpha < 1$ using a different Diophantine condition described at point 1) before. To reach a power law ($|j|^p$, $p \geq 1$) would be a very interesting result at least for two applications: 1) the thermodynamic limit in the context of statistical mechanics 2) the study of almost–periodic solutions of some discretized PDE's. Another difference respect to [CP] and [P] rests in the method of proof. In [CP] as well as in [P], a "configurational" KAM theorem has been proved (technically speaking one solves a system of partial differential equations whose variables are the functions defining the torus in configuration–space) while here we perform a proof based on countable changed of action–angle coordinates in phace–space (a "classic proof"). May be analogous results could be obtained through a configurational proof. We have used a "classic" proof because we think it allows the treatment of: 1) perturbations of infinitely many harmonic oscillators (where configurational proofs can be hardly adapted, if any) 2) Nekhoroshev type problems where a background made of action–angles variables and canonical transformations seems to be necessary.

Acknowledgments I thank C.Liverani for some useful discussions and the referees for their valuable comments. Moreover I thank the Editors in Chief for pointing out some typographical errors

2. Setup

Denote by $\mathcal{T}$ the Cartesian product of infinitely many copies of the one dimensional (flat) torus $\mathbb{T}$: $\mathcal{T} \equiv \bigotimes_{i\in\mathbb{Z}} \mathbb{T}_i \equiv \mathbb{T}^{\mathbb{Z}}$, $\mathbb{T}_i \equiv \mathbb{T} \equiv \mathbb{R}/2\pi\mathbb{Z}$ (d being a positive integer) and endow $\mathcal{T}$ with the standard weak topology (see, *e.g.* , [Ke] and [Je]). Such topology is also induced by *metrics*: To any summable positive sequence $w : \mathbb{Z} \to (0, \infty)$ s.t. $\sum_{i\in\mathbb{Z}} w_i < \infty$ ($w_i > 0$ $\forall i$) which we shall call a *weight*, we can associate a metric ρ_w by setting, $\forall\, \varphi, \varphi' \in \mathcal{T}$ ($\varphi = \{\varphi_i\}_{i\in\mathbb{Z}}$, $\varphi' = \{\varphi'_i\}_{i\in\mathbb{Z}}$, $\varphi_i, \varphi'_i \in \mathbb{T}$):

$$\rho_w(\varphi, \varphi') \equiv \sum_{i\in\mathbb{Z}} \rho(\varphi_i, \varphi'_i)\; w_i \tag{2.1}$$

where ρ is the standard (flat) metric on $\mathbb{T} \equiv \mathbb{T}_i$: $\rho([a],[b]) \equiv \inf_{n\in\mathbb{Z}} |a - b + 2\pi n|$ here $a, b \in \mathbb{R}$ and $[\cdot]$ denote equivalence (mod. 2π) class. We shall denote by $\mathcal{T}_w$ the pair $(\mathcal{T}_w, \rho_w)$ and by $\mathcal{B}_w$ the Banach space formed by the sequences $a \in \mathbb{R}^{\mathbb{Z}}$ having finite norm $\|a\|_w \equiv \sum_{i\in\mathbb{Z}} |a_i|\; w_i < \infty$, $\rho'_w(a, b) \doteq \|a - b\|$. We define $\mathcal{R} \equiv \bigotimes_{i\in\mathbb{Z}} \mathbb{R} \equiv \mathbb{R}^{\mathbb{Z}}$ and endow this space with the topology induced by the following metric: $\forall\, A, A' \in \mathcal{R}$ [(1)] $\lambda_w(A, A') \equiv \sum_{i\in\mathbb{Z}} w_i \arctan|A_i - A'_i|$. We shall denote by $\mathcal{R}_w$ the pair $(\mathcal{R}_w, \lambda_w)$; $\mathcal{R}_w \times \mathcal{T}_w$

[2] our *short–range* models correspond in [Pö] to *finite–range couplings*, see p.357 while what we call *long–range* models correspond to his *short–range couplings*

[(1)] we don't consider here the physical dimensions is a complete metric space with metric $\rho_w(\varphi,\varphi')+\lambda_w(A,A')$. The metric over $\mathcal{T}$ (as well as over $\mathcal{R}$) defines the compact topology with its open sets and induces a well defined notion of convergence. It turns out that the sequence $\{\varphi^{(n)}\}\in\mathcal{T}$, $n=1,2,\dots$ converges respect to the metric (2.1) if and only if converges respect to the following notion: $\varphi^{(n)}\xrightarrow[n\to+\infty]{}\varphi$ if $\forall\, i$ $\varphi_i^{(n)}\xrightarrow[n\to+\infty]{}\varphi_i$ (no uniformity in the components). The same equivalence of convergences holds in the space $\mathcal{R}$. We shall also work with real analytic functions $f\colon U\times\mathbb{T}^n\to\mathbb{R}$, $U=\overset{\circ}{U}\subset\mathbb{R}^n$ (n will be the d.o.f. of the intermediate hamiltonians we will consider; see before Lemma 4.2). $f(y,x)$ can be extented to an holomorphic function, which we call $f(y,x)$ too, on the set $\cup_{y\in U}D(y;\rho_y)\times\Delta_\xi$, $D(y;\rho_y)=\{z\in\mathbb{C}^n\colon |y-z|<\rho_y\}$, $\Delta_\xi=\{z\in\mathbb{C}^n\colon\ \mathrm{Re}\, z_i\in\mathbb{T},\ |\mathrm{Im}\, z_i|<\xi\}$, $0<\xi<1$ and continuous on the closure $\overline{\cup_{y\in U}D(y;\rho_y)}\times\overline{\Delta}_\xi$. $C^\omega(A\times B;\mathbb{C})$ is the space of complex functions holomorphic over $A\times B\subset\mathbb{C}^n\times\mathbb{C}^n$ and $C(A\times B;\mathbb{C})$ is the space of the functions continuous over the same domain. The norm of functions, vector–functions and matrix–valued functions are given in Appendix 1.

We now give a precise meaning to second order ODE's on $\mathcal{R}_w\times\mathbb{T}_w$: Given a continuous map (h_A,f) : $\mathcal{R}_w\times\mathcal{T}_w\to\mathcal{R}_w\times\mathcal{B}_w$, ($h_{A_i}\doteq\frac{\partial h}{\partial A_i}$) consider the system (1.1) where $(h_{A_i},f_i)\equiv(\pi_i^{(A_i)}\circ h_A,\ \pi_i^{(\varphi_i)}\circ f)$ ($\pi_i^{(\varphi_i)}:\mathcal{B}_w\to\mathbb{R}_i$, $\pi_i^{(A_i)}:\mathcal{R}_w\to\mathbb{R}_i$ being the standard projections). A *solution of (1.1)* is just a continuous map $t\in\mathbb{R}\to(A(t),\varphi(t))\in\mathcal{R}_w\times\mathcal{T}_w$, with $(A_i,\varphi_i)\in C^1(\mathbb{R},\mathbb{R})\times C^1(\mathbb{R},\mathbb{T})$, $\forall i$, and satisfying the system (1.1).

Remarks 2.1 (i) In some sense the notion of solution we have just introduced is a "weak notion" (ii) Let $f\equiv 0$ so that $(A(t),\varphi(t))=(A^o,[\omega t])\equiv(\{A_i^o\},\{[\omega_i t]\})_{i\in\mathbb{Z}}$ and $\omega\doteq h_{A^o}$; $(A(t),\varphi(t))$ is a solution for any $\omega\in\mathbb{R}^{\mathbb{Z}}$ (not necessarily in $\mathcal{B}_w$). Note in particular that $t\in\mathbb{R}\to(A^o,[\omega t])\in\mathcal{R}_w\times\mathcal{T}_w$ is continuous for any $\omega\in\mathbb{R}^{\mathbb{Z}}$. These facts are no longer true if one considers stronger topologies; for example, if $|\omega_i|\to\infty$ as $|i|\to\infty$, $[\omega t]$ is not continuous with respect to the uniform topology [$\rho_{\text{uniform}}(\varphi,\varphi')\equiv\sup_{i\in\mathbb{Z}}\rho(\varphi_i,\varphi_i')$] . Thus, our interest in solutions $(A(t),\varphi(t))$ with $(A_i(t),\varphi_i(t))$ "close" to $(A_i^o,\omega_i t)$ with $|\omega_i|\to\infty$ as $|i|\to\infty$ explains the choice of the topology induced by $\rho_w+\lambda_w$.

Global existence and uniqueness for the Cauchy problem associated to (1.1) with (h_A,f) Lipschitz, are elementary applications of standard contraction techniques. We just stress that the "initial velocities" can be taken to be completely arbitrary (and not necessarily in $\mathcal{B}_w$):

Proposition 2.1 *Let $(h_A,f)\colon\mathcal{R}_w\times\mathcal{T}_w\to\mathcal{R}_w\times\mathcal{B}_w$ be a Lipschitz map (i.e. $\exists\ L>0$ s.t. $\rho_w'(f(\varphi),f(\varphi')+\lambda_w(h_A,h_{A'})\le L(\rho_w(\varphi,\varphi')+\lambda_w(A,A'))$, $\forall\varphi,\varphi'\in\mathcal{T}_w\ \ A,A'\in\mathcal{R}_w$). Given any $\varphi^o\in\mathcal{T}_w$ and any $A^o\in\mathbb{R}^{\mathbb{Z}}$, there exists a unique solution, global in time, of the Cauchy problem*

$$\begin{cases}\dot\varphi_i(t)=h_{A_i}(A(t)) & \varphi_i(0)=\varphi_i^o\qquad i\in\mathbb{Z}\\ \dot A_i(t)=f_i(\varphi(t)) & A_i(0)=A_i^o.\end{cases}\tag{2.2}$$

The proof is an application of standard techniques in ODE's and we omit it. The existence for all times of the solutions follows by: 1) the infinite–dimensional torus $\mathcal{T}_w$ is a compact space 2) the space $\mathcal{R}_w$, though not compact, is complete, bounded and very similar to a compact space for what concerns the behavior of the metric $\lambda_w(A,A')$ respect to the indices $i\in\mathbb{Z}$ definitively large.

It is appopriate here to stress one point. As is well known, in the proof we must check the Lipschitz condition on $\lambda_w(h_A(A),h_A(B))$ and we are led to study $\sum_{i\in\mathbb{Z}}w_i\max(1,\sup_{A_i\in\mathbb{R}}|h_{A_iA_i}|)\arctan|A_i-B_i|$. We know that our KAM theorem needs $\{|A_i|\}_{i\in\mathbb{Z}}$ unbounded and then we are forced to take $\sup_{A_i\in\mathbb{R}}|h_{A_iA_i}|<\infty$. This problem would be absent if the d.o.f. were finite but its presence here forbids kinetic energies like $h(A_i)=A_i^{2+\delta}$ for any $\delta>0$ while allows e.g. $h(A_i)=A_i^2+o(A_i^2)$.

As for f we consider two examples of Lipschitz maps (the force). The first one is so called *short range*; fix $L \geq 1$ and consider, for $j \in \mathbb{Z}$, a collection of functions g_j depending on sites of the lattice within (Euclidean) distance L from j; in formulae: $g_j \equiv g_j(\varphi^{(L)})$, $\varphi^{(L)} \equiv \{\varphi_k\}_{k \in B_j(L)}$, $B_j(L) \equiv \{k : \|k - j\| \leq L\}$. We assume that the "localized potentials" g_j are *real–analytic* functions from $\mathbb{T}^{|B_j(L)|} \to \mathbb{R}$ and that, for some positive M: $\sup_{j,\varphi^{(L)} \in \mathbb{T}^{|B_j(L)|}} |g_j(\varphi^{(L)})| \leq M$. Then we set

$$f_i \equiv \sum_{\|j-i\| \leq L} \partial_{\varphi_i} g_j \tag{2.3}$$

The system (1.1) with such f_i is called a finite range system of infinitely many coupled variables (see also [Wa1] and [VB]). A particular case, often considered, is given in $d = 1$ by $L = 1$, $g_j = \cos(\varphi_j - \varphi_{j-1}) - \cos(\varphi_{j+1} - \varphi_j)$. It can be noted that each variable is coupled with just a finite number of different variables and this explains the words *short range*

The second example is so called *long range* as each variable is coupled with any other variable. In $d = 1$ it is given by

$$f_i \equiv \cos\varphi_i \sum_{j \in \mathbb{Z}} a_j \prod_{k \neq 0} (1 + a_{j+k} \sin\varphi_{i+k}) \ , \quad \sum_{j \in \mathbb{Z}} |a_j| < \infty \tag{2.4}$$

The property of being Lipschitz depends of course from the metric. In fact, consider two cases: (1) $a_j \equiv b^{-|j|}$, $b > 1$; (2) $a_j = (1 + |j|^p)^{-1}$, $p > 1$. Then, in the first case (1), f is Lipschitz if we take $w_j \equiv c^{-|j|}$ with any $1 < c < b$, while in case (2) we can take $w_i \equiv a_j$. We don't need the existence of a function $V: \mathbb{T}^{\mathbb{Z}} \to \mathbb{R}$ such that $-\frac{\partial V}{\partial \varphi_i} = f_i(\varphi)$. For instance we could take $V(\varphi) = -\sum_{i=1}^{\infty} mg(1 - \cos\varphi_i) + \sum_{i=1}^{\infty} \kappa(1 - \cos(\varphi_{i+1} - \varphi_i))$ for the short range case and $V(\varphi) = \sum_{i,j=1}^{\infty} e^{-|i-j|}(1 - \cos(\varphi_i - \varphi_j))$ for the long range. We need well defined the functions $g^{[I]}: \mathbb{T}^{|I|} \to \mathbb{R}$, $g^{[I]} = \int g(\varphi) d\mu_J$, $I \subset \mathbb{Z}$, $J = \mathbb{Z} \backslash I$, $d\mu_J = \bigotimes_{i \in J} d\mu_i$ for a measurable function $g: \mathcal{T}_w \to \mathbb{R}$,. In $\mathcal{T}_w$ there exists a unique *probability measure* (see [Ha], Section 38) defined over the σ–algebra, $\mathcal{R}$, generated by the cylinders $\mathcal{R}_I = \bigotimes_{i \in I} U_i \bigotimes_{j \notin I} \mathbb{T}_i$ where U_i is an open subset of $\mathbb{T}_i$ and $I \subset \mathbb{Z}$ is a finite subset of $\mathbb{Z}$: $|I| < \infty$ ($|\cdot|$ denoting here cardinality). The measure μ on $\mathcal{R}$ is such that $\mu(\mathcal{R}_I) = \prod_{i \in I} \mu_i(U_i)$ where μ_i is the normalized "Lebesgue measure" on $\mathbb{T}_i$. If $|I| < \infty$, $g^{[I]}$ is a measurable function on $\mathbb{T}^{|I|}$ and $g^{[I]} \to g$ *a.e.* on $\mathcal{T}$ as $|I| \to \mathbb{Z}$. For the examples in (2.3) and (2.4) the convergence is uniform (see [Je])

Definition 2.2 *A continuous function* $f: \mathcal{T}_w \to \mathcal{B}_w$ *is a* g-gradient *if for any finite* $I \subset \mathbb{Z}$ *there exists a* $C^1(\mathbb{T}^{|I|}; \mathbb{R})$ *function,* $V^{(I)}(\varphi)$*, so that* $f_i^{[I]}(\varphi) = -\partial_{\varphi_i} V^{(I)}(\varphi)$, $\forall\, i \in I$, $\forall\, \varphi \in \mathbb{T}^{|I|}$. It is easy to check that the examples (2.3), (2.4) in Section 2 are g–gradients.

We shall speak of Hamiltonian equations on $\mathcal{T}_w$ *whenever* f *in (1.1) is a g–gradient.*

Let us conclude this section by introducing the (strong) *regularity class* we shall work with.

Definition 2.3 *A g-gradient f is called* uniformly weakly real-analytic *if there exists a real number* $\sigma > 0$ *such that for any finite set* $I \subset \mathbb{Z}$, $V^{(I)}(\varphi)$ *is real-analytic on* $\mathbb{T}^{|I|}$ *and can be analytically continued to the set* $\{z \in \mathbb{C}^{|I|} : \ \mathrm{Re}\, z_i \in \mathbb{T} \ \ |\mathrm{Im}\, z_i| < \sigma\}$ *and continuous on the closure .*

Remarks (i) In fact we could deal with more general classes of vector fields allowing the width of analyticity of $V^{(I)}$ to tend to zero as $|I| \to \infty$ (the allowed rate of decay would then be dictated by the quantitative analysis carried out below). However well known models like (2.3) with $d = 1$, $L = 1$, $g_j = \cos(\varphi_j - \varphi_{j-1}) - \cos(\varphi_{j+1} - \varphi_j)$ have the same analyticity width for all the variables (ii) Example (2.4) in Section 1 is uniformly weakly analytic and as parameter σ one could take any positive number; example (2.3) is uniformly weakly analytic for some (small enough) $\sigma > 0$.

3. Almost–Periodic Solutions and Diophantine Sequences

Definition 3.1 *A sequence $\omega \in \mathbb{R}^{\mathbb{Z}}$ is said to be* rationally independent *if for any finite subset I of $\mathbb{Z}$ and for any $n_i \in \mathbb{Z}$ $\sum_{i\in I} \omega_i n_i \neq 0 \quad \text{unless} \quad n_i = 0 \; \forall \; i \in I$*

In other words, ω is rationally independent if any finite vector $\omega^{(I)} \equiv \{\omega_i\}_{i\in I} \in \mathbb{R}^{|I|}$ is rationally independent.

Definition 3.2 *A continuous real function $q(t)$ is said to be* almost–periodic over $\mathcal{T}_w$ *(with frequency ω) if there exist a rationally independent sequence $\omega \in \mathbb{R}^{\mathbb{Z}}$ and a continuous function $Q : \mathcal{T}_w \to \mathbb{R}$ such that $q(t) = Q([\omega t])$. A solution $(A(t), \varphi(t))$ of (1.1) is called* maximal almost–periodic *if $(A_i(t), \varphi_i(t) - [\omega_i t])$ is, for all i, almost–periodic over $\mathcal{T}_w$ with frequency ω.*

As usual in problems involving small denominators, we have to impose on the frequencies a quantitative irrationality condition. Our aim is to get the lower bound of $|\omega_i|$ as small as possible and a good condition is the following

Definition 3.3 *A rationally independent sequence $\omega \in \mathbb{R}^{\mathbb{Z}}$ is called (γ, τ)–*Diophantine *if for any finite set $I \subset \mathbb{Z}$, there exist constants $\gamma > 0$ and $\tau \in (0, 1)$ such that for any choice of $n_i \in \mathbb{Z}$ with $\sum_{i\in I} |n_i| > 0$ it is:*

$$|\sum_{i\in I} \omega_i n_i| \geq \gamma^{-1} \exp(-|n|^\tau/\tau) \tag{3.1}$$

Of course (3.1) implies that a sequence $\omega \in \mathbb{R}^{\mathbb{Z}}$ is rationally independent: for any finite subset I of $\mathbb{Z}$ and for any $n_i \in \mathbb{Z}$, $\sum_{i\in I} \omega_i n_i \neq 0 \quad \text{unless} \quad n_i = 0 \; \forall \; i \in I$. Diophantine vectors in the sense of (3.1) form a set of full Lebesgue measure in $\mathbb{R}$. The proof is standard and we omit. It can be performed along the same lines of the proof present in [CP] p.194 although the irrationality condition considered there is $|\sum_{i\in I} \omega_i n_i| \geq \gamma^{-1}|n|^{-|I|}$. Also in infinite dimensions they are rather abundant as the following Lemma shows. For any vector $v \in \mathbb{R}^n$ let be $\|v\| = (\sum_{i=1}^n v_i^2)^{1/2}$ while $|v| = \sum_{i=1}^n |v_i|$; $E(t) \doteq e^{t^\tau/\tau}$

Lemma 3.4 *Let $\omega \in \mathbb{R}^{l-1}$ be (γ, τ)–Diophantine and $\Omega \geq \frac{1}{\gamma}$ be a positive number. Let's define the subset of $[\Omega, +\infty)$:*

$$\mathcal{A}_l \doteq \mathcal{A}_l(\omega, \Omega) = \{\alpha \geq \Omega \colon \left|(\omega, \alpha) \cdot \nu\right|^{-1} \leq \Omega^{-1} E(|\nu|) \quad \forall \, \nu \in \mathbb{Z}^l \quad \nu_l \in \mathbb{Z}\backslash\{0\}\} \tag{3.2}$$

Then the following estimate of the Lebesgue measure of the complement $\mathcal{A}_l$ holds $Leb([\Omega, +\infty)\backslash\mathcal{A}_l) \leq L(\tau, l-1)\|\omega\|$, where $L(\tau, l-1) = C_3 2^{c\frac{l}{\tau}} \frac{\Gamma(\frac{l-1}{\tau})}{\Gamma(\frac{l-1}{2})}$; c and C_3 are constants l, τ and ω independent

Proof

$$|\omega \cdot \nu|^{-1} \leq \gamma E(|\nu|), \qquad \forall \; \nu \in \mathbb{Z}^{l-1}\backslash\{0\}$$

We need the following relation for $\alpha \geq \Omega$

$$\left|(\omega_o, \alpha) \cdot \nu\right|^{-1} \leq \Omega^{-1} E(|\nu|) \qquad \forall \nu \in \mathbb{Z}^l\backslash\{0\} \quad \nu_l \neq 0 \tag{3.3}$$

If $\nu = (\hat{\nu}, \nu_l) = (\hat{\nu}, 0)$ we have $\left|\omega \cdot \hat{\nu}\right|^{-1} \leq \Omega^{-1} E(|\hat{\nu}|) \leq \gamma E(|\hat{\nu}|)$ if we impose $\Omega \geq \frac{1}{\gamma}$
In $\hat{\nu} = \hat{0}$ (3.3) becomes $|\nu_l \alpha|^{-1} \leq \Omega^{-1} E(|\nu_l|)$ which is due to $\alpha \geq \Omega$ and $\frac{1}{|\nu_l|} \leq E(|\nu_l|)$.
$\mathcal{A}_l^c$ is the set of α for which (3.3) is false

$$\mathcal{A}_l^c = \{\alpha \geq \Omega : \exists \, \nu \in \mathbb{Z}^l \quad \nu_l \neq 0, \quad \left|(\omega, \alpha) \cdot \nu\right|^{-1} > \Omega^{-1} E(|\nu|)\} \tag{3.4}$$

and its Lebesgue measure is bounded by

$$Leb(\mathcal{A}_l^c) \le \sum_{\substack{\hat\nu\in\mathbb{Z}^{l-1}\\ \hat\nu\ne 0}} \sum_{\substack{\nu_l\in\mathbb{Z}\\ \nu_l\ne 0}} Leb\{\alpha \ge \Omega \,:\, \left|(\alpha + \frac{\omega\cdot\hat\nu}{\nu_l})\right| < \frac{\Omega}{|\nu_l|}\frac{1}{E(|\nu|)}\} \tag{3.5}$$

To get (3.5) let's consider a centered interval of lenght $\frac{\Omega}{|\nu_l|}\frac{1}{E(|\nu|)}$ around each value $\frac{\omega\cdot\hat\nu}{\nu_l}$ $(\nu_l \ne 0)$. For a given vector $\nu \in \mathbb{Z}^l$, $Leb\{\alpha \ge \Omega \,:\, \left|(\alpha + \frac{\omega\cdot\hat\nu}{\nu_l})\right| < \frac{\Omega}{|\nu_l|}\frac{1}{E(|\nu|)}\}$ is an upper bound on the measure of the set of the $\alpha' s$ such that the vector (ω,α), multiplied by $|\nu|$, has projection over ν bounded by $\frac{\Omega}{E(|\nu|)}$ (i.e. (ω,α) may be resonant with ν or non–resonant up to a size given by $\frac{\Omega}{E(|\nu|)}$). If for for some ν we had $\left|(\omega,\alpha)\cdot\nu\right|^{-1} \le \Omega^{-1}E(|\nu|)\}$, then summing over all $\nu \in \mathbb{Z}^{l-1}$, $\nu_l \ne 0$, we are considering more vectors than those satisfiyng (3.4) so that the sum in (3.5) is an upper bound of $Leb(\mathcal{A}_l^c)$

$$\begin{aligned} Leb(\mathcal{A}_l^c) &\le \sum_{\substack{\hat\nu\in\mathbb{Z}^{l-1}\\ \hat\nu\ne 0}} \sum_{\substack{\nu_l\in\mathbb{Z}\setminus\{0\}\\ \frac{\omega\cdot\hat\nu}{\nu_l}\ge\frac{\Omega}{2}}} \frac{2\Omega}{|\nu_l|}e^{-|\nu|^\tau/\tau} \le \sum_{\substack{\hat\nu\in\mathbb{Z}^{l-1}\\ \hat\nu\ne 0}} \sum_{\substack{\nu_l\in\mathbb{Z}\setminus\{0\}\\ \frac{\omega\cdot\hat\nu}{\nu_l}\ge\frac{\Omega}{2}}} \frac{2\Omega}{|\nu_l|}e^{-\|\nu\|^\tau/\tau} \le \\ &\le \sum_{\substack{\hat\nu\in\mathbb{Z}^{l-1}\\ \hat\nu\ne \hat 0}} \sum_{\substack{\nu_l\ge 1\\ \frac{|\nu_1|}{\nu_l}\ge\frac{a}{2}}} \frac{4}{\nu_l}\Omega e^{-\|\nu\|^\tau/\tau} \le \sum_{\substack{\hat\nu\in\mathbb{Z}^{l-1}\\ \hat\nu\ne \hat 0}} \sum_{\substack{\nu_l\ge 1\\ |\nu_1|\ge\frac{a}{2}}} \frac{4}{\nu_l}\Omega e^{-\|\nu\|^\tau/\tau} \end{aligned} \tag{3.6}$$

The condition $\frac{\omega\cdot\hat\nu}{\nu_l} \ge \frac{\Omega}{2}$ comes from $\alpha \ge \Omega$. Imposing in fact $-\frac{\omega\cdot\hat\nu}{\nu_l} + \frac{\Omega}{|\nu_l|E(|\hat\nu|)} \ge \Omega$ we find a bigger value of the measure. $-\frac{\omega\cdot\hat\nu}{\nu_l} \ge -\frac{\Omega}{|\nu_l|E(|\hat\nu|)} + \Omega$ is implied by $-\frac{\omega\cdot\hat\nu}{\nu_l} \ge \Omega(1-\frac{1}{E(1)}) \ge \frac{\Omega}{2}$. Moreover rotating in such a way that $R^T e = e^{(1)} = (1,0,\dots,0) \in \mathbb{R}^{l-1}$, $e = \frac{\omega_o}{\|\omega_o\|}$ and calling $a \doteq \frac{\Omega}{\|\omega_o\|}$, we obtain the last bound in (3.6). Being $(\sum_{i=1}^l k_i^2)^\tau \ge |k_1|^\tau + C(\sum_{i=2}^l k_i^2)^\tau$ for all $k_1 \ge 0$, $\sum_{i=2}^l |k_i| \ge 1$ $i = 2,\dots,l$, and for a suitable C, we have $([\frac{a}{2}] = A)$

$$Leb(\mathcal{A}_l^c) \le \sum_{\substack{\hat\nu\in\mathbb{Z}^{l-1}\\ \hat\nu\ne \hat 0}} \sum_{\substack{\nu_l\ge 1\\ |\nu_1|\ge A}} 4\Omega e^{-(|\nu_1|^\tau + C(\sum_{i=2}^l |\nu_i|^2)^\tau)/\tau} \tag{3.7}$$

$$\begin{aligned} &\sum_{|\nu_1|\ge\frac{a}{2}} e^{-|\nu_1|^\tau/\tau} \le 2\int_A^\infty dx e^{-x^\tau/\tau} = 2\tau^{\frac{1-\tau}{\tau}} e^{-A^\tau/\tau}\int_0^\infty dx(x+\frac{1}{\tau}A^\tau)^{\frac{1-\tau}{\tau}}e^{-x} \le \\ &\le 2\tau^{\frac{1-\tau}{\tau}}e^{-A^\tau/\tau}\Big\{\int_0^{A^\tau/\tau} dx(\frac{2}{\tau}A^\tau)^{\frac{1-\tau}{\tau}}e^{-x} + \int_{A^\tau/\tau}^\infty (2x)^{\frac{1-\tau}{\tau}}e^{-x}\Big\} \le \overline{C}A^{1-\tau}e^{-A^\tau/\tau} = \\ &= C_1[\frac{a}{2}]^{1-\tau}e^{-\frac{1}{\tau}[\frac{a}{2}]^\tau} \end{aligned}$$

the rest of (3.7) is bounded by $\sum_{k\in\mathbb{Z}^{l-1}} e^{-C(\sum_{i=1}^{l-1}|k_i|^2)^\tau/\tau} \le C_2 2^{c\frac{l}{\tau}}\frac{\Gamma(\frac{l-1}{\tau})}{\Gamma(\frac{l-1}{2})}$. Reuniting all the previous bounds and estimates we get

$$Leb(\mathcal{A}_l^c) \le 4\Omega C_1[\frac{a}{2}]^{1-\tau}e^{-\frac{1}{\tau}[\frac{a}{2}]^\tau}C_2 2^{c\frac{l}{\tau}}\frac{\Gamma(\frac{l-1}{\tau})}{\Gamma(\frac{l-1}{2})} \le C_3\|\omega_o\|2^{c\frac{l}{\tau}}\frac{\Gamma(\frac{l-1}{\tau})}{\Gamma(\frac{l-1}{2})}$$

c and C_3 are constants l and τ independent.

4. Main Results

Theorem 4.1 *Let $h_A = \{h_{A_i}(A_i)\}_{i\in\mathbb{Z}}: \mathcal{R}_w \to \mathcal{R}_w$ satisfy $0 < a' \le \inf_{A_i\in\mathbb{R}} |h_{A_iA_i}| \le \sup_{A_i\in\mathbb{R}} |h_{AA}| \le a$ and let $f : \mathcal{T}_w \to \mathcal{B}_w$ be a uniformly weakly analytic g–gradient (see Definition 2.3). Then there exist uncountably many maximal almost–periodic solutions of (1.1) with Diophantine frequencies.*

The proof, achieved in some intermediate steps, is a corollary of theorem 4.7 and is written after it. Before writing that intermediate steps we try to explain our strategy. First of all we define a bijective map of $\mathbb{Z}$ onto $\mathbf{N}$ in such a way to have a system of variables indexed by $j = 1, 2, 3, \ldots$.
Let's define $l_k = l - 1 + k$, $k = 0, 1, 2, \ldots$ and the function $V_{l_k}(\varphi^{(l_k)})$, $\varphi^{(l_k)} = (\varphi_1, \ldots, \varphi_{l-1+k})$ where $-\partial_{\varphi_i} V_{l_k}(\varphi^{(l_k)}) = f_i^{[I_k]}(\varphi^{(l_k)})$, $I_k = 1, 2, \ldots, l-1+k$ and $f_i^{[I_k]}(\varphi^{(l_k)}) = \int f_i d\mu_{J_k}$ $J_k = \mathbb{Z}\backslash I_k$. Let's define a sequence of hamiltonians $H_{l_k} \doteq \sum_{i=1}^{l-1+k} h(A_i) + V_{l_k}(\varphi^{(l_k)})$, $\quad V_{l_k}(\varphi^{(l_k)}) = \int d\varphi_{l+k} V_{l_{k+1}}(\varphi^{(l_{k+1})})$. For $k = 0$ we take a vector–frequency $\frac{\omega_o}{\mu}$ such that $\frac{\partial h}{\partial A_i}|_{A^o} = \frac{(\omega_o)_i}{\mu}$ $1 \le i \le l-1$. For $|\mu|$ small enough we apply to H_{l_0} n_{l_0} (an integer number) of canonical transformations (whose composition we call $\mathcal{R}^{(n_{l_0})}$) so that the hamiltonian becomes $H_{l_0}^{(n_{l_0})} \doteq h^{(n_{l_0})}(A^{(n_{l_0})}) + V_{l_0}^{(n_{l_0})}(A^{(n_{l_0})}, \varphi^{(n_{l_0})})$ and where $V_{l_0}^{(n_{l_0})}$ is of order $\varepsilon^{2^{n_{l_0}}}$ (actually a number slightly less than 2 will be present but at the moment we leave 2). We stress that we *we don't take the limit* $n_{l_o} \to +\infty$. Then we write the hamiltonian with one more d.o.f. $H_{l_1} \doteq \sum_{i=1}^{l} h(A_i) + V_{l_1}(\varphi^{(l_1)})$ as $H_{l_1} \doteq \sum_{i=1}^{l} h(A_i) + (V_{l_1}(\varphi^{(l_1)}) - V_{l_0}(\varphi^{(l_0)})) + V_{l_0}(\varphi^{(l_0)})$, and act on H_{l_1} with $\mathcal{R}^{(n_{l_0})}$ in such a way to obtain the new hamiltonian $(h^{(n_{l_0})}(A^{(n_{l_0})}) + V_{l_0}^{(n_{l_0})}) + h(A_l) + P(\varphi^{(l_1)})$ where P by definition is the function $(V_{l_1}(\varphi^{(l_1)}) - V_{l_0}(\varphi^{(l_0)}))$ after the action of $\mathcal{R}^{(n_{l_0})}$. Now using $\int d\varphi_l(V_{l_1}(\varphi^{(l_1)}) - V_{l_0}(\varphi^{(l_0)})) = 0$ and by means of a second canonical change of coordinate, if $\omega_l = \frac{\partial h}{\partial A_l}|_{A^o}$ is Diophantine in the sense of (3.2) and Ω big enough, we can reduce the perturbation P to order Ω^{-1}. *Taking n_{l_0} such that $\varepsilon^{2^{n_{l_0}}} = O(\Omega^{-1})$*, both the terms P and $V_{l_0}^{(n_{l_0})}$ are reduced to order Ω^{-1} and the hamiltonian $(h^{(n_{l_0})}(A^{(n_{l_0})}) + V_{l_0}^{(n_{l_0})}) + h(A_l) + P(\varphi_{l_0})$ is transformed into $(h^{(n_{l_0})}(A^{(n_{l_0})}) + h(A_l)) + \tilde{V}_{l_0}^{(n_{l_0})} + \tilde{P}(\varphi^{(l_0)})$ where $\tilde{V}_{l_0}^{(n_{l_0})} + \tilde{P}$ are of order Ω^{-1} and Ω is a (fast) increasing function of the number l of d.o.f.. Then a further canonical transformation reduces the perturbation to order $\Omega^{-2^{n_{l_1}}}$. Of course the proceedure works if Ω is sufficiently large. At this point we can continue the procedure "adding" more and more d.o.f.. In the limit $k \to +\infty$ we get the infinite–dimensional torus whose vector frequency (Diophantine) is $(\frac{\omega_o}{\mu}, \omega_l, \omega_{l_1}, \ldots)$; $|\omega_l| \ge \Omega_l$ (Ω_l is the previous Ω), $|\omega_{l_1}| \ge \Omega_{l_1}$, $|\omega_{l_2}| \ge \Omega_{l_2}$ and so on. The vector $(\frac{\omega_o}{\mu}, \Omega_l, \Omega_{l_1}, \ldots)$ *may not be* a Diophantine vector because it may not satisfy the irrationality conditions (see Definition 3.3) even if its components are sufficiently large. A priori nothing forbids the vector $(\frac{\omega_o}{\mu}, \Omega_l, \Omega_{l_1}, \ldots)$ from being Diophantine. What we know by Lemma 3.4 (applied for each value of $k = 0, 1, 2, \ldots$) is that outside the interval $[\Omega_{l_k}, \Omega_{l_k} + L(\tau, l_k)\|\omega^{(l_{k-1})}\|]$, certainly there is a value ω_{l_k} such that the vector $(\frac{\omega_o}{\mu}, \omega_l, \omega_{l_1}, \ldots)$ is Diophantine. Moreover $L(\tau, l_k)\|\omega^{(l_{k-1})}\|$ increases faster than Ω_{l_k} ($L(\tau, l_k)\|\omega^{(l_{k-1})}\|$. Precisely it increases as $\frac{\Gamma(\frac{l-1}{\tau})}{\Gamma(\frac{l-1}{2})}$ while Ω_{l_k} faster that any power of k but slower that an exponential). This is a marked difference respect to [P] and [CP]. There what increases very fast with k ($(k!)^a$ $a \sim 10$ for a $d = 1$–dimensional problem) is Ω_{l_k} and this forces the components of $(\frac{\omega_o}{\mu}, \omega_l, \omega_{l_1}, \ldots)$ to be very large. Here what increases faster is $L(\tau, l_k)\|\omega^{(l_{k-1}\|)}$ but this *does not forbid* the components ω_{l_k} from being close to Ω_{l_k} (possibly equal). Let's define $M(\delta) \doteq \sup_{t\ge1} E(t)e^{-t\delta} = \exp\{\frac{1-\tau}{\tau}\delta^{\frac{-\tau}{1-\tau}}\}$.

Lemma 4.2 *We consider the hamiltonian $H_{l_0}(A, \varphi) = \sum_{i=1}^{l-1} h(A_i) + V(\varphi): \mathbb{R}^{l-1} \times \mathbb{T}^{l-1} \to \mathbb{R}$ where: 1) $h: \mathbb{R} \to \mathbb{R}$ is real–analytic, $h \in C^\omega(\mathbb{C}; \mathbb{C})$, $\sup_{A_i\in\mathbb{C}} |h_{A_iA_i}| \le a$, 2) $\{h_{A_i^o}\}_{i=1}^{l-1} = \{\frac{1}{\mu}\omega_i^o\}_{i=1}^{l-1} \doteq \frac{\omega^o}{\mu}$,*

ω^o is a (γ,τ)–Diophantine vector 3) $V(\varphi)\in C^\omega(\Delta_\xi;\mathbb{C})\cap C(\overline{\Delta}_\xi;\mathbb{C})$, $\|V_\varphi\|_\xi=\mathcal{V}$. There exists the canonical transformation $(A,\varphi)=\mathcal{C}(A',\varphi')$, $A=A'+\Xi(A',\varphi')$, $\varphi=\varphi'+\Delta(A',\varphi')$, Ξ and Δ belong to $C^\omega(D(A^o;\frac{\rho}{4})\times\Delta_{\xi-3\delta};\mathbb{C}^{l-1})\cap C(\overline{D}(A^o;\frac{\rho}{4})\times\overline{\Delta}_{\xi-3\delta};\mathbb{C}^{l-1})$, $\xi-3\delta>0$, $\rho=\frac{1}{2|\mu|a\gamma NE(N)}$, $N=\frac{1}{2\delta}\ln[(\mu^2a\gamma^2\mathcal{V})^{-1}]$, such that

$$H_{l_0}(\mathcal{C}(A',\varphi'))\doteq H_{l_0}^{(0)}(A',\varphi')=\sum_{j=1}^{l-1}h(A'_j)+V'(A',\varphi')+<V> \tag{4.1}$$

$$\Big(\frac{16}{\rho}\|V'_{\varphi'}\|_{\frac{\rho}{16},\xi-4\delta}+\|V'_{A'}\|_{\frac{\rho}{16},\xi-4\delta}\Big)\le 2^6la(\delta\rho)^{-1}(\gamma\mathcal{V}M(\delta)|\mu|)^2=V_x \tag{4.2}$$

$$\gamma V_x\le 2^7\frac{l}{\delta}|\mu|^{3-\sigma}\exp\{B_2\delta^{\frac{-\tau}{1-\tau}}\}\frac{(\gamma^2\mathcal{V}a)^{2-\sigma}}{4\delta\sigma} \tag{4.3}$$

$B_2(\tau,\sigma)=\frac{1-\tau}{\tau}(2+\sigma^{\frac{-\tau}{1-\tau}})$, $(0<\sigma<1)$ *provided that* $(l-1)\frac{16}{\rho\delta}\gamma\,|\mu|\,\mathcal{V}M(\delta)\le 1$

$$\max\{\frac{4}{\rho}\|\Xi\|_{\frac{\rho}{4},\xi-3\delta}\,,\;\|\Delta\|_{\frac{\rho}{4},\xi-3\delta}\}\le 8(l-1)\frac{1}{\rho}\mathcal{V}M(\delta)\gamma|\mu|\;\doteq T_o \tag{4.4}$$

Remarks i) The lemma makes clearer (in an tractable way) that the problem is of perturbative nature. The bound V_x in (4.3) is small provided $|\mu|$ is small. ii) The kinetic part of the hamiltonian (4.1) is the same of H_{l_0}. This makes unnecessary in the proof the step where the frequency of the torus is "restored" iii) here δ is a fixed quantity but it will become smaller and smaller as the number of d.o.f. increases. Then the quantities equivalent to $\frac{1}{|\mu|}$ (the size of the components of the frequencies following $\frac{\omega_o}{\mu}$) are forced to become greater and greater.

Proof It is quite standard so we just sketch it. As usual one looks for the solution of the Hamilton–Jacobi equation

$$\Phi(A',\varphi)=-\sum_{\substack{\nu\in\mathbb{Z}^{l-1}\\0<|\nu|\le N}}\frac{e^{i\nu\cdot\varphi}V_\nu}{i\nu\cdot h_{A'}} \tag{4.5}$$

knowing that $h_A(A^o)=(\frac{\omega^o}{\mu},\alpha)$, $\mu\in[-1,1]\backslash\{0\}$, α has the dimensions of a frequency. $A'\cdot\varphi+\Phi(A',\varphi)$ is the generating function of the canonical transformation that reduces the order of the perturbation thanks to $|\frac{\omega_o}{\mu}\cdot\nu|^{-1}\le|\mu|\,\gamma E(|\nu|)$. ∎

Moreover we have $\gamma V_x\ge 2^7l\frac{M^2(\delta)}{\delta}(\gamma^2a\mathcal{V}^{(l)})^2|\mu|^3$. Now let's consider the hamiltonian (4.1) and recall $\frac{\rho}{16}\doteq\rho_o$, $\xi-4\delta=\xi_o$, $a=a_o$; $\|(h_{AA})^{-1}\|_{\rho_o,\xi_o}\doteq\eta_o\ge(a')^{-1}$ (see the introduction) $\big(\frac{1}{\rho_o}\|V'_{\varphi'}\|_{\rho_o,\xi_o}+\|V'_{A'}\|_{\rho_o,\xi_o}\big)=V_x$.

Lemma 4.3 *Let's consider the hamiltonian (4.1). Provided that $(0<\sigma<1)$*

$$\frac{2^8l}{\sigma\delta_o^2}|\mu|(\eta_oa_o)^2(\gamma V_x)^{1-\sigma}\exp\{B_1(\tau,\sigma)\delta_o^{\frac{-\tau}{1-\tau}}\}\le 1. \tag{4.6}$$

$(B_1(\tau,\sigma)=\frac{1-\tau}{\tau}(1+\sigma^{\frac{-\tau}{1-\tau}}))$, there exists the canonical transformation $(A,\varphi)=\mathcal{C}(A',\varphi')$, $A=A'+\Xi(A',\varphi')$ $\varphi=\varphi'+\Delta(A',\varphi')$, Ξ and Δ belong to $C^\omega(D(A^o;\frac{\rho'}{4})\times\Delta_{\xi_o-3\delta_o};\mathbb{C}^{l-1})\cap C(\overline{D}(A^o;\frac{\rho'}{4})\times\overline{\Delta}_{\xi_o-3\delta_o};\mathbb{C}^{l-1})$

$$\max\{\frac{4}{\rho'}\|\Xi\|_{\frac{\rho'}{4},\xi_o-3\delta_o}\,,\;\|\Delta\|_{\frac{\rho'}{4},\xi_o-3\delta_o}\}\le\frac{2^3}{\rho'}\gamma|\mu|M(\delta_o)\|V'_\varphi\|_{\rho_o,\xi_o}\doteq T_1$$

$\rho' = \frac{1}{2|\mu|\gamma a N_o E(N_o)}$, $N_o = \frac{1}{2\delta_o}\ln[(\gamma V_x)^{-1}]$, $\xi_o - 4\delta_o > 0$ *such that*

$$H^{(0)}_{l_0}(\mathcal{C}(A',\varphi')) = H^{(1)}_{l_0}(A',\varphi') = \sum_{j=1}^{l-1} h(A'_j) + < V'(A') > + V''(A',\varphi') \tag{4.7}$$

$(A',\varphi') \subset D(A^{(1)}_o; \frac{\rho'}{32}) \times \Delta_{\xi_o - 4\delta_o}$, $D(A^{(1)}_o; \frac{\rho'}{16}) \subset D(A^o; \frac{\rho'}{8})$

$$\gamma\big(\frac{32}{\rho'}\|V''_{\varphi'}\|_{\frac{\rho'}{32},\xi_o-4\delta_o} + \|V''_{A'}\|_{\frac{\rho'}{32},\xi_o-4\delta_o}\big) \le \gamma V^{(1)}_x \le \frac{2^4 l}{\sigma\delta_o^2}(\gamma V_x)^{2-\sigma}\exp(\{B_2(\tau,\sigma)\delta_o^{\frac{-\tau}{1-\tau}}\})$$

$$\|h_{A'A'} + < V'(A') >_{A'A'}\|_{\frac{\rho'}{32}} \le a_o(1+\frac{1}{8}), \qquad \partial_{A'}(h + < V'(A') >)|_{A^{(1)}_o} = \frac{\omega_o}{\mu}$$

$$\|(h_{A'A'} + < V'(A') >_{A'A'})^{-1}\|_{\frac{\rho'}{32}} \le \eta_o(1+\frac{1}{8})$$

The proof (omitted) is analogous to that of Lemma 4.2 except for one point. Here is necessary to restore the vector–frequency of the torus and the calculation is $\frac{\partial}{\partial A_i}(\sum_{j=1}^{l-1} h(A'_j) + < V'(A') >) = \frac{(\omega_o)_i}{\mu}$. It is an application of the analytic–implicit function theorem

Lemma 4.3 leads to the following Theorem (repeat enough times the Lemma)

Theorem 4.4 *Let's consider the hamiltonian (4.7). Let be* $0 < \sigma \le \sigma_o$, $\sigma_o < 2 - 2^{\frac{\tau}{1-\tau}}$

$$(\eta_o a_o)^{a_1} l^{a_2} (\xi_o - \xi_n)^{-a_3} \exp\{c_1(\xi_o-\xi_n)^{\frac{-\tau}{1-\tau}}\}(\gamma V_x) \le 1 \tag{4.8}$$

(a_1, a_2, a_3, c_1 are constants depending on τ,σ and independent on the number of d.o.f. $l-1$), we define a finite number of canonical tranformations $\mathcal{C}^{(j)}(A^{(j)},\varphi^{(j)})$ $j = 1,\dots,n$ ($n \doteq n_{l_0}$) whose domain is $D(A^{(j)}_o;\rho_j) \times \Delta_{\xi_j}$ ($\rho_j \ge \rho_{j+1} > 0$, $\xi_j > \xi_{j+1} > 0$) such that the hamiltonian (4.7) becomes

$$\begin{aligned} &H^{(n_{l_0})}_{l_0}(A^{(n)},\varphi^{(n)}) \doteq \\ &\doteq H^{(0)}(\mathcal{C}^{(1)} \circ \mathcal{C}^{(2)} \ldots \circ \mathcal{C}^{(n)})(A^{(n)},\varphi^{(n)}) = h^{(n)}(A^{(n)}) + V^{(n)}(A^{(n)},\varphi^{(n)}) \end{aligned} \tag{4.9}$$

$\|h^{(n)}_{A^{(n)}A^{(n)}}\|_{\rho_n} \le 2a_o$, $\quad \|(h^{(n)}_{A^{(n)}A^{(n)}})^{-1}\|_{\rho_n} \le 2\eta_o \qquad h^{(n)}_{A^{(n)}_o} = \frac{\omega_o}{\mu}$

$$\gamma\big(\frac{1}{\rho_n}\|V^{(n)}_{\varphi^{(n)}}\|_{\rho_n,\xi_n} + \|V^{(n)}_{A^{(n)}}\|_{\rho_n,\xi_n}\big) \le \gamma V^{(n)}_x \le (Q\gamma V_x)^{(2-\sigma)^n} \tag{4.10}$$

$Q = \Big(\big(\frac{l-1}{(\xi_o-\xi_n)\sigma}\big)^{\frac{2}{1-\sigma}} K_o e^{(\xi_o-\xi_n)^{\frac{-\tau}{1-\tau}}K'_o}\Big)$; *$K_o$ and K'_o are constants depending on σ and τ; $\xi_o - \xi_n > 0$. Moreover for each $j = 1,\dots,n-1$*

$$\begin{aligned} &(A^{(j-1)},\varphi^{(j-1)}) = \mathcal{C}^{(j)}(A^{(j)},\varphi^{(j)}) = \\ &= (A^{(j)} + \Xi^{(j)}(A^{(j)},\varphi^{(j)}),\ \varphi^{(j)} + \Delta^{(j)}(A^{(j)},\varphi^{(j)})) \end{aligned} \tag{4.11}$$

$\rho_j = (2a_j\gamma|\mu|N_jE(N_j))^{-1}$, $N_j = \frac{1}{2\delta_j}\ln(\gamma V^{(j)}_x)^{-1}$

$$\begin{aligned} &\max\{\frac{1}{2\rho_j}\|\Xi^{(j)}\|_{2\rho_j,\xi_{j-1}-3\delta_{j-1}} \quad , \quad \|\Delta^{(j)}\|_{2\rho_j,\xi_{j-1}-3\delta_{j-1}}\} \le \\ &\le \frac{16(l-1)}{\sigma\delta_{j-1}^2}(\gamma V^{(j-1)}_x)^{1-\sigma}\exp\{B_1(\tau,\sigma)\delta_{j-1}^{\frac{-\tau}{1-\tau}}\} \doteq T_j \end{aligned} \tag{4.12}$$

Proof It is a repeated application of Lemma 4.3 plus a standard induction argument

Remarks i) If we were to prove the existence of $l-1$ dimensional tori with high frequencies $\frac{\omega^o}{\mu} \in \mathbb{R}^{l-1}$ we should make the limit $n \to +\infty$. Actually our problem is infinite–dimensional and we will fix n according to how big will be the frequency (ω_l) (see the upper bound in the third formula of (4.19)) corresponding to the next d.o.f (l). Here is a key point in the proof. We do not take $\lim_{n\to+\infty}$ now because we are not interested in a finite–dimensional problem. The integer n will be high accordingly to ω_l. The unboundedness of the sequence $\{\omega_i\}$ (which is a consequence of the unboundednes of the d.o.f.) will force n to go to $+\infty$. The upper bound on μ can be rewritten as $l^{a_2}(\xi_o-\xi_n)^{-a_3}\exp\{c_1(\xi_o-\xi_n)^{\frac{-\tau}{1-\tau}}\}O(\mu^{2-\sigma}) \le 1$ and we stress that the bigger is $l-1$ (the d.o.f.), the smaller is μ. "Adding" degrees of freedom, the number l increses and we take $0<\tau<\frac{1}{2}$ to avoid the factorial behavior $\mu \sim (l!)^{-b}$ with b positive which otherwise would arise

Let's call $\mathcal{C}^{(0)}$ the canonical transformation of Lemma 4.2 (there is indicated with $\mathcal{C}$) and let's define $(\mathcal{C}^{(0)}\circ\mathcal{C}^{(1)}\circ\ldots\circ\mathcal{C}^{(n)}) = \mathcal{R}^{(n)}$ where $\mathcal{C}^{(1)}\circ\ldots\circ\mathcal{C}^{(n)}$ is the transformation written in (4.9).

$$\|\partial(\mathcal{C}^{(0)}\circ\mathcal{C}^{(1)}\circ\ldots\circ\mathcal{C}^{(n)})\|_{\rho_n,\xi_n} \le \Pi_{j=0}^{n}\big(1+(1+\frac{1}{\delta_j})T_j\big) \le C_o \qquad C_o>1 \tag{4.13}$$

(possibly adding 1 to the exponent a_3 of (4.8)). C_o is a constant depending on the number of d.o.f.

The next step is the important passage from $l-1$ to l degrees of freedom. Let's consider the hamiltonian $H_{l_1}(I,\psi) = \sum_{i=1}^{l} h(I_i) + V_{l_1}(\psi_1,\ldots,\psi_l)\colon \mathbb{T}^l\times\mathbb{R}^l \to \mathbb{R}$, $\|(V_{l_1})_\psi\|_\xi = \mathcal{V}^{(l)}$ and rewrite it as

$$\sum_{i=1}^{l-1} h(I_i) + V_{l_0}(\psi_1,\ldots,\psi_{l-1}) + h(I_l) + (V_{l_1}(\psi) - V_{l_0}(\psi_1,\ldots,\psi_{l-1})) \tag{4.14}$$

where $V_{l_0}(\psi_1,\ldots,\psi_{l-1}) = \int_0^{2\pi}\frac{d\psi_l}{2\pi}V_{l_1}(\psi)$. $V_{l_0}(\psi_1,\ldots,\psi_{l-1})$ is exactly the function $V_{l_1}(\psi)$, $\psi\in\mathbb{T}^{l-1}$, present in the hamiltonian of Lemma 4.2. To the set of variables $(I_1,\ldots,I_{l-1},\psi_1,\ldots,\psi_{l-1})$ we apply the transformation $\mathcal{R}^{(n)}$. (I_l,ψ_l) undergo an identity transformation. $H_{l_1}(I,\psi)$ becomes

$$h^{(n)}(A^{(n)}) + V_{l_0}^{(n)}(A^{(n)},\varphi^{(n)}) + h(I_l) + (V_{l_1}(\hat\psi(A^{(n)},\varphi^{(n)}),\psi_l) - V_{l_0}^{(n)}(A^{(n)},\varphi^{(n)})) \tag{4.15}$$

Now let's recall $(A^{(n)},I_l) \doteq (\hat A, A_l) = A$, $(\varphi^{(n)},\psi_l) \doteq (\hat\varphi,\varphi_l) = \varphi$, and rewrite the hamiltonian (4.15) as

$$\tilde H_{l_1}(A,\varphi) = \tilde h_1(\hat A) + h(A_l) + \tilde V_1(\hat A,\hat\varphi) + \tilde V_2(\hat A,\varphi) \doteq \tilde h_{l_1}(A) + \tilde V_{l_1}(\hat A,\varphi) \tag{4.16}$$

$$\tilde h_{l_1}(A) \doteq h^{(n)}(A^{(n)}) + h(I_l), \quad \tilde V_1(\hat A,\hat\varphi) \doteq V_{l_0}^{(n)}(A^{(n)},\varphi^{(n)}),$$
$$\tilde V_2(\hat A,\varphi) \doteq (V_{l_1}(\hat\psi(A^{(n)},\varphi^{(n)}),\psi_l) - V_{l_0}^{(n)}(A^{(n)},\varphi^{(n)}))$$

$(\tilde h_{l_1})_A(\hat A_o^{(n)},A_l^o) = (\frac{\omega_o}{\mu},\alpha)$, $\hat A\in D(\hat A_o^{(n)};\rho_n)$, $A_l\in D(A_l^o;\tilde r)$, $\hat\varphi\in\Delta_{\xi_n}$, $\varphi_l\in\Delta_\xi$,

Being $\rho_n<\tilde r$ and $\xi_n<\xi$, we restrict to $D((\hat A_o^{(n)},A_l^o);\rho_n)$ the domain of the action–variables $(A\in D((\hat A_o^{(n)},A_l^o);\rho_n))$ and to Δ_{ξ_n} the domain of the angular–variables $(\varphi\in\Delta_{\xi_n})$. Now let's define $(\hat A_o^{(n)},A_l^o) \doteq A^o$ $\|(\tilde h_{l_1})_{AA}\|_{D(A_o;\rho_n)} \le 2a_o$ $\|(\tilde h_{l_1})_{AA}^{-1}\|_{D(A_o;\rho_n)} \le 2\eta_o$

$$\gamma\Big(\frac{1}{\rho_n}\|(\tilde V_1)_{\hat\varphi}\|_{\rho_n,\xi_n} + \|(\tilde V_1)_{\hat A}\|_{\rho_n,\xi_n}\Big) \le (Q\gamma V_x)^{(2-\sigma)^n}$$

By (4.13) and for $\frac{\rho_n}{\gamma}(Q\gamma V_x)^{(2-\sigma)^n} \le C_o \mathcal{V}^{(l)}$ we have $\|(\tilde V_2)\varphi\|_{\rho_n,\xi_n} 2\mathcal{V}^{(l)}C_o$ and $\|(\tilde V_2)_{\hat A}\|_{\rho_n,\xi_n} \le \frac{2}{\rho_n}\mathcal{V}^{(l)}C_o$ δ_o is the "loss of analyticity" of Lemma 4.2 and T_o is defined in (4.4);
Let's recall $\rho_n = \tilde\rho$ and $\xi_n = \tilde\xi$

$$\frac{1}{\tilde\rho}\|(\tilde V_2)_\varphi\|_{\tilde\rho,\tilde\xi} + \|(\tilde V_2)_{\hat A}\|_{\tilde\rho,\tilde\xi} \le \frac{4}{\tilde\rho}\mathcal{V}^{(l)}C_o \qquad \frac{1}{\tilde\rho}\|(\tilde V_{l_1})_\varphi\|_{\tilde\rho,\tilde\xi} + \|(\tilde V_{l_1})_{\hat A}\|_{\tilde\rho,\tilde\xi} \le \frac{5}{\tilde\rho}\mathcal{V}^{(l)}C_o$$

We are now ready to state a Lemma analogous to Lemma 4.2 (let's recall $\tilde V_{l_1} \doteq \tilde V$)

Lemma 4.6 *We consider the hamiltonian $\tilde H(A,\varphi) = \tilde h_{l_1}(\hat A) + \tilde V_{l_1}(\hat A,\varphi)\colon \mathbb{T}^l \times \mathbb{R}^l \to \mathbb{R}$ (see (4.16)) where: 1) $\tilde h_{l_1} \in C^\omega(D(A^o;\tilde\rho);\mathbb{C})$, $\|(\tilde h_{l_1})_{AA}\|_{D(A^o;\tilde\rho)} \le 2a_o$, $\|(\tilde h_{l_1})^{-1}_{AA}\|_{D(A^o;\tilde\rho)} \le 2\eta_o$, 2) $(\tilde h_{l_1})_A(A^o) = (\frac{\omega^o}{\mu},\alpha)$ which is a Diophantine vector in the sense of (3.3), 3) $\tilde V(\hat A,\varphi) \in C^\omega(D(\hat A^o;\tilde\rho)\times \Delta_{\tilde\xi};\mathbb{C}) \cap C(\overline{D(\hat A^o;\tilde\rho)}\times\overline{\Delta_{\tilde\xi}};\mathbb{C})$, $\frac{1}{\tilde\rho}\|\tilde V_\varphi\|_{\tilde\rho,\tilde\xi} + \|\tilde V_{\hat A}\|_{\tilde\rho,\tilde\xi} \le \frac{5}{\tilde\rho}C_o\mathcal{V}^{(l)}$. There exists the canonical transformation $(A,\varphi) = \tilde{\mathcal{C}}(A',\varphi')$, defined in $D(A^o;\frac{\tilde\rho}{4})\times\Delta_{\tilde\xi-3\tilde\delta} \ni (A',\varphi')$, $\tilde\xi - 3\tilde\delta > 0$, $N = \frac{1}{2\tilde\delta}\ln\big[\frac{\Omega}{M(\tilde\delta)}\frac{\tilde\rho}{8C_o\mathcal{V}^{(l)}}\big]$ such that*

$$\tilde H_{l_1}(\tilde{\mathcal{C}}(A',\varphi')) \doteq H^{(0)}_{l_1}(A',\varphi') = \tilde h_{l_1}(A') + \tilde V'(A',\varphi') \tag{4.17}$$

$$\frac{16}{\tilde\rho}\|\tilde V'_{\varphi'}\|_{\frac{\tilde\rho}{16},\tilde\xi-4\tilde\delta} + \|\tilde V'_{A'}\|_{\frac{\tilde\rho}{16},\tilde\xi-4\tilde\delta} \le \frac{1}{\tilde\delta\tilde\rho}2^{11}(C_o\mathcal{V}^{(l)})^2\frac{M(\tilde\delta)}{\tilde\rho\Omega} \doteq \tilde V_x^{(o)} \tag{4.18}$$

provided that

$$\begin{aligned} &2^6\frac{M(\tilde\delta)}{\tilde\delta}\frac{l}{\tilde\rho\Omega}(4C_o\mathcal{V}^{(l)}) \le 1, \qquad \frac{a_o\tilde\rho}{\tilde\delta\sigma\Omega}\Big(\frac{\tilde\rho\Omega}{8M(\tilde\delta)C_o\mathcal{V}^{(l)}}\Big)^\sigma \exp\left\{\frac{2}{(\sigma\tilde\delta)^{\frac{\tau}{1-\tau}}}\frac{1-\tau}{\tau}\right\} \le 1,\\ &2^6\frac{M(\tilde\delta)}{\tilde\delta}\frac{l\gamma}{\tilde\rho}(C_o\mathcal{V}^{(l)}) \le \gamma\Omega \le 2^5(C_o\mathcal{V}^{(l)})^2 M(\tilde\delta)\frac{\gamma^2}{\tilde\rho^2}(Q\gamma V_x)^{-(2-\sigma)^n} \end{aligned} \tag{4.19}$$

Proof Following Lemma 4.2 we construct the canonical transformation whose generating function is $\tilde S(A',\varphi) = A'\cdot\varphi + \tilde\Phi(A',\varphi)$ and (suppress in this proof the index l_1 from $\tilde h$)

$$\Phi(A',\varphi) = -\sum_{\substack{\nu\in\mathbb{Z}^l\\ \nu_l\ne 0,\, 0<|\nu|\le N}} \frac{e^{i\nu\cdot\varphi}(\tilde V_2)_\nu(\hat A')}{i\nu\cdot\tilde h_{A'}} \tag{4.20}$$

because of the superexponential behavior of $\tilde V_1$, see (4.10) and the freedom in the choice of n, (see theorem 4.4), just $\tilde V_2$ is present in Φ.

$$\begin{aligned} &|\tilde h_{A'}\cdot\nu| = |\tilde h_{\hat A'}\cdot\hat\nu + \tilde h_{A'_l}\nu_l| =\\ &= |\frac{\omega_o}{\mu}\cdot\hat\nu + \alpha\nu_l + \hat\nu\cdot(\tilde h_{\hat A'} - \tilde h_{\hat A_o}) + \nu_l(\tilde h_{A'_l} - \tilde h_{A^o_l})| \ge \frac{1}{2}\frac{\Omega}{E(|\nu|)} \quad \forall\, 0<|\nu|\le N \end{aligned} \tag{4.21}$$

for $|A' - A^o| \le \frac{\Omega}{4a_oNE(N)}$. Being Ω as high as we need, later we will impose $\frac{\Omega}{4a_oNE(N)} \ge \tilde\rho$. Now continue exactly as in Lemma 4.2. It is essential that $\int_0^{2\pi} d\varphi_l \tilde V_2(\hat A,\hat\varphi,\varphi_l) = 0$ otherwise a term not small with $\frac{1}{\gamma\Omega}$ "would survive" the canonical transformation ∎

Remarks As the reader can note in the second of (4.19), $\Omega \to +\infty$ implies $n \to +\infty$ too and we find the invariant torus with vector–frequency $\frac{\omega_o}{\mu}$ which exists by Corollary 4.5. The Hamilton–Jacobi equation (4.20) does not involve $\tilde V_1$ because of its superexponential decay.

We are now in the same situation after Lemma 4.2 but with one more d.o.f. (l) and proceed as follows. For each $k \geq 0$ we have (the number of d.o.f. is $l-1+k$) : 1) a sequence of hamiltonians $\{H_{l_k}^{(n_{l_k})}\}$ whose domain is $D(A_o^{(n_{l_k})};\rho_{n_{l_k}}^{(l_k)}) \times \Delta_{\xi_{n_{l_k}}^{(l_k)}}$, $H_{l_k}^{(n_{l_k})}(A^{(n_{l_k})},\varphi^{(n_{l_k})}) = h^{(n_{l_k})}(A^{(n_{l_k})}) + V^{(n_{l_k})}(A^{(n_{l_k})},\varphi^{(n_{l_k})})$ and $\frac{\partial h^{(n_{l_k})}(A^{(n_{l_k})})}{\partial A_i^{(n_{l_k})}} = \frac{(\omega_o)_i}{\mu}$ if $1 \leq i \leq l-1$ and α_{l_j} if $j \geq l$; $V^{(n_{l_k})}(A^{(n_{l_k})},\varphi^{(n_{l_k})})$ if of order $O(\varepsilon_k^{(2-\sigma)^{n_{l_k}}})$ and $\varepsilon_k < 1$. 2) $\xi_{n_{l_{k-1}}}^{(l_{k-1})} - \xi_{n_{l_k}}^{(l_k)} = \frac{\mathcal{S}}{l_k \ln^2(l_k)} \doteq \mathcal{S}_k$ where $\mathcal{S} = (\xi - \xi_*)(\sum_{k=0}^{\infty}(l_k \ln^2(l_k))^{-1})$, 3) a vector–frequency $\omega^{(k)} = (\frac{\omega_o}{\mu},\alpha_{l_1},\ldots,\alpha_{l_k}) \in \mathbb{R}^{l-1+k}$; each α_{l_k} belongs to $\mathcal{A}_{l_k}$ where $\mathcal{A}_{l_k}$ is defined as in (3.2). Just write α_{l_k}, Ω_{l_k} and $l-1+k$ in place of α, Ω and l. Moreover $\gamma\Omega_{l_k} \leq C(l_k)^{-a_1-a_2}(\ln(l_k))^{-3}\exp\{-c_1(\frac{\mathcal{S}}{l_k \ln^2(l_k)})^{\frac{\tau}{1-\tau}}\}$ a quantity that goes to zero with $k \to +\infty$ faster than any power of l_k but slower that an exponential if $\frac{\tau}{1-\tau} < 1$ that is $0 < \tau < \frac{1}{2}$ 4) let's define $\mathcal{R}^{(n_{l_k})} \doteq \mathcal{C}_{l_k}^{(0)} \circ \mathcal{C}_{l_k}^{(1)} \circ \ldots \circ \mathcal{C}_{l_k}^{(n_{l_k})}$ the canonical transformation that "allows to pass" from $\tilde{H}_{l_k}$ to $H_{l_k}^{(n_{l_k})}$. $\tilde{H}_{l_k} \doteq H_{l_k} \circ \mathcal{R}^{(n_{l_{k-1}})}$ where $\mathcal{R}^{(n_{l_{k-1}})}$ is the transformation that changes $H_{l_{k-1}}$ into $H_{l_{k-1}}^{(n_{l_{k-1}})}$. To each $\mathcal{C}_{l_k}^{(r)}$ $(0 \leq r \leq n_{l_k})$ a quantity $T_r^{(l_k)}$ is associated (see (4.12)) and $\xi_{n_{l_{k-1}}}^{(l_{k-1})} - \xi_{n_{l_k}}^{(l_k)} = 3\tilde{\delta}^{(l_k)} - 4\sum_{j=1}^{n_{l_k}}\delta_j^{(l_k)}$ 5) $S_t^{(k)}$ is the time–evolution of the hamiltonian H_{l_k}. Now we can state the following

Theorem 4.7 *Assume 1)–5). The real part of the subset of* $\mathbb{C}^{l_k} \times \mathbb{C}^{l_k}$ *defined by*

$$\begin{aligned}
&\lim_{k\to+\infty} \cup_{z^{(n_{l_k})}\in\mathbb{T}^{l_k}} \big(\mathcal{R}^{(n_{l_0})} \circ \mathcal{R}^{(n_{l_1})} \circ \ldots \circ \mathcal{R}^{(n_{l_{k-1}})} \circ \mathcal{R}^{(n_{l_k})}\big)_j^{(A)}(A_o^{(n_{l_k})}, z^{(n_{l_k})}) \\
&\lim_{k\to+\infty} \cup_{z^{(n_{l_k})}\in\mathbb{T}^{l_k}} \big(\mathcal{R}^{(n_{l_0})} \circ \mathcal{R}^{(n_{l_1})} \circ \ldots \circ \mathcal{R}^{(n_{l_{k-1}})} \circ \mathcal{R}^{(n_{l_k})}\big)_j^{(\varphi)}(A_o^{(n_{l_k})}, z^{(n_{l_k})})
\end{aligned} \tag{4.22}$$

$j = 1, 2\ldots$ is invariant respect the flow (S_t) of (2.2) and is run quasi–periodically with vector–frequency ω

Proof This proof is standard as well (see for example [CC4] p.2082 or [G] pp.515–517). One has to use the canonicity of all the transformations constructed before and the density of $z^{(n_{l_k})} + \omega^{(k)}t$ over $\mathbb{T}^{l_k}$ for $t \in \mathbb{R}$. The convergence is component by component over j. ∎

The limits (4.22) define two maximal almost–periodic functions (see Definition 3.2), $P^{(\infty)}$ and $Q^{(\infty)}$, such that (4.22) is an invariant torus in $\mathcal{R}_w \times \mathcal{T}_w$. Let's call $A(t) \doteq A_o^{(\infty)} + \mathcal{P}^{(\infty)}(\theta)$ and $\varphi(t) \doteq \theta + \mathcal{Q}^{(\infty)}(\theta))$ $(\theta = \omega t = (\omega_1 t, \omega_2 t, \ldots))$

Proof of Theorem 4.1 we must show that for each i (no uniformity in the index i) the functions $(A_i(t), \varphi_i(t))$, are respectively $C^1(\mathbb{R},\mathbb{R})$ and $C^1(\mathbb{R},\mathbb{T})$ for each i. Let's consider the sequences $\{\mathcal{Q}_i^{(k)}\}$ $\{\mathcal{P}_i^{(k)}\}$ defined as in (4.22) before the limits and the union over $\mathbb{T}^{l_k}$. By definition $\dot{\mathcal{Q}}_j^{(k)}(A_o^{(n_{l_k})},\varphi^{(n_{l_k})}) = h_{A_j}(\mathcal{P}_j^{(k)}(A_o^{(n_{l_k})},\varphi^{(n_{l_k})}))$ (let's take values $l-1+k \geq j$) and the upper bound $|h_{AA}| \leq a$ allows the limit (use Lagrange formula) $\lim_{k\to+\infty} h_{A_j}(\mathcal{P}_j^{(k)}(A_o^{(n_{l_k})},\varphi^{(n_{l_k})})) = h_{A_j}((A_o^{(\infty)})_j + \mathcal{P}_j^{(\infty)}(\theta))$ which defines $\dot{\mathcal{Q}}^{(\infty)}$. As for the action–variables we have $\dot{\mathcal{P}}_j^{(k)}(A_o^{(n_{l_k})},\varphi^{(n_{l_k})}) = -(V_{l_k})_{\varphi_j^{(l_k)}}(\mathcal{Q}_1^{(k)},\ldots,\mathcal{Q}_{l_k}^{(k)})$. When k goes to infinity the sequence $\{-(V_{l_k})_{\varphi_j^{(l_k)}}(\mathcal{Q}_1^{(k)},\ldots,\mathcal{Q}_{l_k}^{(k)})\}$ converges uniformly on $\mathcal{T}_w$ to $f_i(\varphi)$ (see before Definition 2.2) and defines $\dot{\mathcal{P}}^{(\infty)}$. The continuity of $\dot{\varphi}_i(t)$ follows by the relation $|\varphi_i(t) - \varphi_i(t')| \leq a|t-t'|$ while that of $\dot{A}_i(t)$ by the relation $|\dot{A}_i(t) - \dot{A}_i(t')| \leq |f_i(\varphi(t)) - f_i(\varphi(t'))|$. By the Lipschitz property of the vector field $\{f\}$ we can write $|f_i(\varphi(t)) - f_i(\varphi(t'))| \leq L\rho_w(\varphi(t), \varphi(t'))$ and the continuity of $\varphi(t)$ as a function of t implies the continuity of the derivative respect to time of the actions–variables. ∎

Appendix 1 Cauchy estimates for holomorphic functions

Let be $f \in C^\omega(D(A^o;\rho) \times \Delta_\xi;\mathbb{C}) \cap C(\overline{D}(A^o;\rho) \times \overline{\Delta}_\xi;\mathbb{C})$; the space of the holomorphic functions over $D(A^o;\rho) \times \Delta_\xi = \{A \in \mathbb{C}^l\colon\ |A - A^o| < \rho,\ i = 1,\ldots,l\} \times \{\varphi \in \mathbb{C}^l\colon \mathrm{Re}\,\varphi_i \in \mathbb{T},\ |\mathrm{Im}\,\varphi_i| < \xi,\ i = 1,\ldots,l\}$ and continuous on the closure $\overline{D}(A^o;\rho) \times \overline{\Delta}_\xi$.

We can write its Fourier series

$$f(A,\varphi) = \sum_{\nu\in\mathbb{Z}^l} e^{i\nu\cdot\varphi} f_{\nu,k}(A) = \sum_{\nu\in\mathbb{Z}^l} e^{i\nu\cdot\varphi}\frac{1}{(2\pi)^l}\int_{\mathbb{T}^l} d\varphi e^{-i\nu\cdot\varphi} f(A,\varphi)$$

$\nu\cdot\varphi = \sum_{i=1}^l \nu_i\varphi_i$, $\quad \|f\|_{\rho,\xi} \doteq \sum_{\nu\in\mathbb{Z}^l} e^{|\nu|\xi}\sup_{A\in D(A^o;\rho)}|f_\nu(A)| \doteq \sum_{\nu\in\mathbb{Z}^l} e^{|\nu|\xi}\|f_\nu\|_\rho$, $|\nu| = \sum_{i=1}^{l-1}|\nu_i|$

$$\|f_\varphi\|_{\rho,\xi-\delta} \leq \frac{1}{e\delta}\|f\|_{\rho,\xi}, \quad \|f_{A_j}\|_{\rho-r,\xi} \leq \frac{1}{r}\|f\|_{\rho,\xi}, \quad \|f_A\|_{\rho-r,\xi} \leq \frac{l}{r}\|f\|_{\rho,\xi} \tag{4.23}$$

$$\|f_{A,\varphi}\|_{\rho-r,\xi-\delta} \leq \frac{1}{\delta}\frac{l}{r}\|f\|_{\rho,\xi} \qquad \|f\,g\|_{\rho,\xi} \leq \|f\|_{\rho,\xi}\|g\|_{\rho,\xi} \tag{4.24}$$

For a vector valued function whose component–functions are in $C^\omega(D(A^o;\rho) \times \Delta_\xi;\mathbb{C}) \cap C(\overline{D}(A^o;\rho) \times \overline{\Delta}_\xi;\mathbb{C})$; the norm is the sum of the norm of their components. For a matrix valued function $\{M(x,y)\}_{i,j=1}^k\colon C^\omega(D(A^o;\rho) \times \Delta_\xi;\mathbb{C}) \cap C(\overline{D}(A^o;\rho) \times \overline{\Delta}_\xi;\mathbb{C});\to \mathbb{C}^{2k}$ we set (only for the tensorial components)

$$\|M\| \doteq \sup_{v\in\mathbb{R}^k;|v|=1}|M(x,y)v| = \sup_{v\in\mathbb{R}^k;|v|=1}\sum_{i=1}^k\Big|\sum_{j=1}^k M_{ij}(x,y)v_j\Big| \tag{4.25}$$

Let's consider now the following function repeteadly used before because it represents the generating function of the canonical transformations (that is the solution of the Hamilton–Jacobi equation) involved in the theorems and propositions

$$\Phi(A,\varphi) = -\sum_{\substack{\nu\in\mathbb{Z}^{l-1}\\ 0<|\nu|\leq N}}\frac{e^{i\nu\cdot\varphi}V_\nu(A)}{ih_A\cdot\nu} \tag{4.26}$$

with $(A,\varphi) \in \overline{D}(A^o;\rho)\times\overline{\Delta}_\xi$; $|h_A\cdot\nu|^{-1} \leq \gamma E(|\nu|)\ \forall\ \nu \in \mathbb{Z}^{l-1}\backslash\{0\}$; γ is a real positive number and $M(\delta)$ is defined before Lemma 4.2. $\|\Phi\|_{\rho,\xi-\delta} \leq \sum_{\substack{\nu\in\mathbb{Z}^{l-1}\\ 0<|\nu|\leq N}} e^{|\nu|(\xi-\delta)}|\nu|\sup_{A\in D(A^o;\rho)}\frac{|V_\nu(A)|}{|h_A\cdot\nu|} \leq \gamma M(\delta)\|V_\varphi\|_{\rho,\xi}$. The proofs are standard calculations of complex analysis.

References

[BFG] Benettin G., Frölich, J., Giorgilli A.: *A Nekhoroshev–Type Theorem for Hamiltonian Systems with Infinitely Many Degrees of Freedom,* Commun. Math. Phys. **119** (1988), 95-108

[CC1] Celletti, A., Chierchia, L.: *Construction of Analytic KAM Surfaces and Effective Stability Bounds,* Commun. Math. Phys. **118** (1988) 119-161

[CC2] Celletti, A., Chierchia, L.: *A constructive theory of Lagrangian tori and computer- assisted applications,* Dynamics Reported 1995, 60–1995

[CC3] Celletti, A., Chierchia, L.: *On the stability of realistic three–body problems*, Commun.Math.Phys. **186** (1997) 413-449

[CC4] Celletti, A.,Chierchia, L.: *Rigorous estimates for a computer–assisted Kam theory* Jour. of Math. Phys. vol.28 (9) September 1987, 2078–2086

[CP] Chierchia L., Perfetti P.: *Second Order Hamiltonian Equations on* $\mathbb{T}^\infty$ *and Almost-Periodic Solutions*, Journal of Differential Equations Vol.**116** No.1 1995, 172-201

[FSW] Frölich, J., Spencer, T., Wayne, C.E.: *Localization in Disordered Nonlinear Dynamical Systems,* Jour. of Stat. Phys. **42** (1986), 247-274

[G] Gallavotti, G., *The Elements of Mechanics* Springer–Verlag

[Je] Jessen, B.: *The Theory of Integration in a space of an Infinite Number of Dimensions*, Acta mathematica **63** 1934, 250–323

[Ke] Kelley, J.: *General Topology,* Van Nostrand–Reinhold, Princeton, New Jersey, 1955

[P] Perfetti, P.: *Hamiltonian Equations on* $\mathbb{T}^\infty$ *and almost–periodic solutions*, proceedings of the "International Conference on Dynamical Systems and Differential Equations" May 18–21 2000 Kennesaw State University (GA) USA, An Added Volume To *Discrete and Continuous Dynamical Systems* 2001, 303–309

[Pö1] Pöschel, J.: *Small Divisors with Spatial Structure in Infinite Dimensional Hamiltonian Systems,* Commun. Math. Phys. **127** (1990) 351-393

[SZ] Salamon, D., Zehnder, E.: *KAM theory in configuration space,* Comment.Math. Helv.**64** (1989) 84-132

[Wa1] Wayne, C., E.: *The KAM Theory of Systems with Short Range Interactions, I and II,* Commun. Math. Phys. **96** (1984) 311-329 and 331-344

[Wa2] Wayne, C., E.: *On the Elimination of Non–Resonant Harmonics,* Commun. Math. Phys. **103** 451–386 (1986)

[Wa3] Wayne, C., E.: *Bound on the Trajectories of a System of Weakly Coupled Rotators* Commun. Math. Phys. **104** 21–36 (1986)

[VB] Vittot, M., Bellissard, J.: *Invariant tori for an infinite lattice of coupled rotators,* Report Centre de Physique Théorique CNRS-Luminy, Marseille (1985)