JOURNAL OF DIFFERENTIAL EQUATIONS 116, 172–201 (1995)

Second Order Hamiltonian Equations on $\mathbb{T}^\infty$ and Almost-Periodic Solutions

LUIGI CHIERCHIA

Dipartimento di Matematica, Terza Università di Roma, via C. Segre 2, 00146 Rome, Italy; and Centro "Vito Volterra," II Università di Roma "Tor Vergata," v. della Ricerca Scientifica, 00133 Rome, Italy

AND

PAOLO PERFETTI

Dipartimento di Matematica, II Università di Roma "Tor Vergata," v. della Ricerca Scientifica, 00133 Rome, Italy

Received October 28, 1992

Motivated by problems arising in nonlinear PDE's with a Hamiltonian structure and in high dimensional dynamical systems, we study a suitable generalization to infinite dimensions of second order Hamiltonian equations of the type $\ddot{x} = \partial_x V$, $[x \in \mathbb{T}^N,\ \partial_x \equiv (\partial_{x_1}, ..., \partial_{x_N})]$. Extending methods from quantitative perturbation theory (Kolmogorov–Arnold–Moser theory, Nash–Moser implicit function theorem, etc.) we construct uncountably many almost-periodic solutions for the infinite dimensional system $\ddot{x}_i = f_i(x)$, $i \in \mathbb{Z}^d$, $x \in \mathbb{T}^{\mathbb{Z}^d}$ (endowed with the compact topology); the Hamiltonian structure is reflected by f being a "generalized gradient." Such a result is derived under (suitable) analyticity assumptions on f_i but without requiring any "smallness conditions." © 1995 Academic Press, Inc.

CONTENTS

0022-0396/95 $6.00

1. Introduction

A natural approach to qualitative theory of nonlinear partial differential equations "with a Hamiltonian structure" is to regard such PDE's as infinite dimensional conservative dynamical systems and to try to extend, whenever possible, results and methods from the well developed finite dimensional theory.

One of the basic results in finite dimensions is the existence, under suitable assumptions, of maximal quasi-periodic solutions (see [A] and references therein). These are solutions, which, up to a change of coordinates, are described by a linear flow: $t \to \omega t$, with $\omega \in \mathbb{R}^N$, $N \equiv$ number of degrees of freedom $\equiv \frac{1}{2}$ dimension of the phase space (associated to the Hamiltonian system under consideration).

Some generalizations to infinite dimensions of the existence of quasi-periodic solutions have been studied by several authors; see, e.g., [FSW, VB, Wa2, Pö, Ku, CW]. All these papers make use of quite stringent "smallness assumptions" (a drawback already present in finite dimensions).

In this paper we study the following infinite dimensional "second order Hamiltonian" system of equations on $\mathbb{T}^\infty$ ($\mathbb{T} \equiv \mathbb{R}/2\pi\mathbb{Z}$),

$$\ddot{x}_i = f_i(x), \qquad i \in \mathbb{Z}^d, \qquad x \in \mathbb{T}^{\mathbb{Z}^d} \equiv \mathscr{T}, \tag{1.1}$$

where $\mathscr{T}$ is equipped with the weak topology and f is a (uniformly) Lipschitz map from $\mathscr{T}$ to a suitable Banach space; the Hamiltonian structure comes from f being a *generalized gradient* (see below).

For such systems we will construct uncountably many almost-periodic solutions under suitably analyticity assumptions on f but *without requiring any smallness condition*. Such solutions will have the form

$$t \to x(t) \equiv \{x_i(t)\}_{i \in \mathbb{Z}^d}, \qquad \text{with} \quad x_i(t) = [\omega_i t + u_i([\omega t])], \tag{1.2}$$

where $[\cdot]$ denotes the standard projection of $\mathbb{R}$ onto $\mathbb{T}$, u is a smooth function, and $\omega \in \mathbb{R}^{\mathbb{Z}^d}$ is a "Diophantine sequence" [i.e. $\exists \gamma > 0$ and a isomorphism, i_k, from $\mathbb{Z}_+$ onto $\mathbb{Z}^d$ such that, $\forall N \geqslant 1$, $|\sum_{k=1}^N \omega_{i_k} n_{i_k}|^{-1} \leqslant \gamma(\sum_{k=1}^N |n_{i_k}|)^N$, with $n_{i_k} \in \mathbb{Z}$, $\sum_{k=1}^N |n_{i_k}| > 0$]. Note that, as a consequence of a classical number theoretical theorem by Liouville, $|\omega_i| \to \infty$ as $|i| \to \infty$.

For the construction of quasi-periodic solutions for the finite dimensional situation, namely,

$$\ddot{x} = V_x(x), \qquad x \in \mathbb{T}^N \tag{1.3}$$

we refer the reader to [A] (and references therein), [M1, M2, SZ, CC1, CC2, and especially CZ].

Our interest in a qualitative analysis of (1.1) has been motivated mainly by: (i) regular motions for nearly integrable PDE's with a Hamiltonian structure; (ii) discrete approximations of, say, nonlinear wave equations in $\mathbb{R}^d$ (e.g., Sine–Gordon); (iii) many particle systems interacting via conservative forces.

(i) Important examples of nonlinear PDE's, such as the Korteweg–de Vries equation and the nonlinear Schrödinger equation, fall in the class of (infinite dimensional) *integrable Hamiltonian systems* (see, e.g., [AN] and references therein). An integrable aspect of these equations (with suitable boundary conditions) may be described, up to a change of coordinates, by a linear flow $t \to (\omega_1 t, ..., \omega_i t, ...) \in \mathbb{T}^\infty$, where the "frequencies" $\omega_i \nearrow \infty$ as $i \to \infty$ (typically $\omega_i \sim Ci^k$ with $k > 1$). Recently, [Ku], the existence, for perturbation of the above models, of special solutions described by a *finite number* of frequencies $t \to (\omega'_1 t, ..., \omega'_N t)$ has been established. The problem of the persistence (under small perturbations) of solutions with *infinitely many* frequencies is open.

The model we study is *not* directly related to the above models but it might be regarded as a *model problem* mimicking some of the basic features coming into play: Hamiltonian structure, "compactness" of the configuration space, regularity, "frequency growth." We point out, however, that in our model the frequencies associated to the constructed almost-periodic solutions grow *very* rapidly as $|i| \to \infty$: Such a fast growth is related to the Diophantine property and it is conceivable that relaxing this property one could establish, for (1.1), the existence of almost-periodic solutions having frequencies growing at "more interesting" rates (such as the polynomial rate mentioned above).

(ii) It has been well known since Lagrange that the d-dimensional wave equation can be derived as (a suitable) limit, as $\varepsilon \to 0$, of harmonic oscillators vibrating orthogonally to a lattice $\mathbb{Z}^d_\varepsilon \equiv \{n\varepsilon, n \in \mathbb{Z}^d\}$ (see, e.g., [G]). Such a limit is not affected by replacing the harmonic potential $(x_{i+1} - x_i)^2/2$, $i \in \mathbb{Z}^d_\varepsilon$, by $-\cos(x_{i+1} - x_i)$. For example, the so-called Sine–Gordon equation on (a domain of) $\mathbb{R}^d$,

$$w_{tt} = \Delta w + \sin w, \qquad w = w(\xi, t),\ \xi \in \mathbb{R}^d,\ t \in \mathbb{R} \tag{1.4}$$

is obtained (in the small amplitude regime) as limit as $\varepsilon \to 0$ of

$$\ddot{x}_i = \varepsilon^{-2} \sum_{\|j-i\|=1} \sin(x_j - x_i) + \sin x_i, \qquad i \in \mathbb{Z}^d \tag{1.5}$$

if one sets, for $\xi \in \mathbb{R}^d$, $\xi \equiv i\varepsilon$, $(i \in \mathbb{Z}^d)$, and $w(\xi, t) = x_i(t)$.

The finite approximation (1.5) is, for any fixed $\varepsilon > 0$, an example of system (1.1) treatable with our techniques.

We do not discuss here boundary conditions [and hence a proper formulation of a problem associated to (1.4)].

(iii) Models of many particles interacting via conservative forces provide natural concrete examples of systems (1.1). More precisely, one can regard (1.1) as describing a system of *infinitely many coupled rotators* (i.e., particles ideally constrained to move on "circles") centered on the site $i \in \mathbb{Z}^d$ and interacting ("coupled") via the "forces" f_i; "conservative" means that such forces are, in a suitable sense (see below), gradients of "potentials."

An example generalizing (1.5) is the following. Fix $L \geq 1$ and consider, for $j \in \mathbb{Z}^d$, a collection of functions g_j depending on sites of the lattice within (Euclidean) distance L from j; in formulae:

$$g_j \equiv g_j(x^{(L)}), \qquad x^{(L)} \equiv \{x_k\}_{k \in B_j(L)}, \qquad B_j(L) \equiv \{k: \|k-j\| \leq L\}. \tag{1.6}$$

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:

$$\sup_{j,\, x^{(L)} \in \mathbb{T}^{|B_j(L)|}} |g_j(x^{(L)})| \leq M. \tag{1.7}$$

Then we set

$$f_i \equiv \sum_{\|j-i\| \leq L} \partial_{x_i} g_j. \tag{1.8}$$

The system (1.1) with such f_i is called a finite range system of infinitely many coupled rotators (see also [Wa1] and [VB]). Example (1.5) is obtained by letting

$$g_j \equiv \varepsilon^{-2} \sum_{h=1}^{d} \cos(x_{j+e_h} - x_j) - \cos x_j, \tag{1.9}$$

where $e_1 \equiv (1, 0, ..., 0), ..., e_d \equiv (0, ..., 0, 1)$.

An example, with $d=1$, of "long range interaction" is given by (1.1) with:

$$f_i \equiv \cos x_i \sum_{j \in \mathbb{Z}} a_j \prod_{k \neq 0} (1 + a_{j+k} \sin x_{i+k}), \qquad \sum_{j \in \mathbb{Z}} |a_j| < \infty. \tag{1.10}$$

We remark that to the above examples one could associate the *formal Hamiltonians*

$$H_{\text{short}}(y, x) \equiv \sum_j \frac{y_j^2}{2} - \sum_j g_j \tag{1.11}$$

for (1.1) and (1.8), and

$$H_{\rm long}(y, x) \equiv \sum_j \frac{y_j^2}{2} - \sum_j \prod_k (1 + a_k \sin x_{j+k}) \tag{1.12}$$

for (1.1) and (1.10); in fact, we can immediately check that the *formal Hamilton equations* (associated to the *formal symplectic form* $\sum_{i \in \mathbb{Z}^d} dy_i \wedge dx_i$) yield (1.1). Rather than trying to give a precise meaning to such formal objects, we shall use the notion of *generalized gradients*, which shall allow us to treat directly the differential equations (1.1). Roughly speaking, a generalized gradient is a vector field (i.e., a continuous map from $\mathcal{T}$ to a suitable Banach space) such that when *averaged* (with respect to the natural probability measure associated to $\mathcal{T}$) over x_j with $j \notin I$ with I a finite subset of $\mathbb{Z}^d$, the function (of finite variables) thus obtained, is a gradient of a periodic function.

The rest of the paper is organized as follows: In Section 2 we give a precise notion of solutions of (1.1) for Lipschitz vector field (so that existence and uniqueness for all time of the Cauchy problem trivially follow); in Section 3 we introduce the Hamiltonian structure via generalized gradients; in Section 4 we define (maximal) almost-periodic solutions and in Section 5 we define Diophantine sequences and prove their abundance; Section 6 contains the statements of the main results; Section 7 is devoted to a finite dimensional "average theorem," which allows us to construct quasi-periodic solutions for a system over $\mathbb{T}^{N+1}$ starting from a quasi-periodic solution of a subsystem over $\mathbb{T}^N$ obtained by averaging the potential associated to the original $(N+1)$-dimensional system: iterating suitably one obtains the proof in infinite dimensions, which is spelled out in Section 8. The average theorem of Section 7 is, in turn, based on tools from perturbation theory (such as Kolmogorov–Arnold–Moser theory, Nash–Moser implicit function theorems, etc.): a short summary (with complete proofs) of KAM theory is given in Appendix 2, while Appendix 1 contains a classical results concerning the full (Lebesgue) measure of Diophantine vectors in $\mathbb{R}^N$.

2. Second Order ODE's on $\mathbb{T}^\infty$

Denote by $\mathcal{T}$ the Cartesian product of infinitely many copies of the one dimensional (flat) torus

$$\mathcal{T} \equiv \bigotimes_{i \in \mathbb{Z}^d} \equiv \mathbb{T}^{\mathbb{Z}^d}, \qquad \mathbb{T}_i \equiv \mathbb{T} \equiv \mathbb{R}/2\pi\mathbb{Z} \tag{2.1}$$

(d being a positive integer) and endow $\mathscr{T}$ with the standard weak topology (see, e.g., [Ke]). Such topology is also induced by *metrics*: To any summable positive sequence

$$w: \mathbb{Z}^d \to (0, \infty) \qquad \text{s.t.} \quad \sum_{i \in \mathbb{Z}^d} w_i < \infty \qquad (w_i > 0 \ \forall i), \tag{2.2}$$

which we shall call a *weight*, we can associate a metric ρ_w by setting, $\forall x, y \in \mathscr{T}$ $(x = \{x_i\}_{i \in \mathbb{Z}^d}, y = \{y_i\}_{i \in \mathbb{Z}^d}, x_i, y_i \in \mathbb{T})$,

$$\rho_w(x, y) \equiv \sum_{i \in \mathbb{Z}^d} \rho(x_i, y_i)\, w_i, \tag{2.3}$$

where ρ is the standard (flat) metric on $\mathbb{T} \equiv \mathbb{T}_i$:

$$\rho([a], [b]) \equiv \inf_{n \in \mathbb{Z}} |a - b + 2\pi n|. \tag{2.4}$$

Here $a, b \in \mathbb{R}$ and $[\cdot]$ denote equivalence (mod 2π) class.

We shall denote by $\mathscr{T}_w$ the pair $(\mathscr{T}_w, \rho_w)$ and by $\mathscr{B}_w$ the Banach space formed by the sequences $a \in \mathbb{R}^{\mathbb{Z}^d}$ having finite norm

$$\|a\|_w \equiv \sum_{i \in \mathbb{Z}^d} |a_i|\, w_i < \infty. \tag{2.5}$$

We can now give a precise meaning to second order ODE's on $\mathscr{T}_w$: Given a continuous map $f: \mathscr{T}_w \to \mathscr{B}_w$ consider the system

$$\ddot{x}_i = f_i(x), \qquad i \in \mathbb{Z}^d, \tag{2.6}$$

where $f_i \equiv \pi_i \circ f$ ($\pi_i: \mathscr{T}_w \to \mathbb{T}_i$ being the standard projection). A *solution of* (2.6) is just a continuous map $t \in \mathbb{R} \to x(t) \in \mathscr{T}_w$, with $x_i \in C^2(\mathbb{R})$, $\forall i$, and satisfying the system (2.6).

Remark 2.1. (i) In some sense the notion of solution we have just introduced is a "weak notion."

(ii) Let $f \equiv 0$ and $x(t) \equiv [\omega t] \equiv \{[\omega_i t]\}_{i \in \mathbb{Z}^d}$. Then $x(t)$ is a solution for any $\omega \in \mathbb{R}^{\mathbb{Z}^d}$ (not necessarily in $\mathscr{B}_w$). Note in particular that $t \in \mathbb{R} \to [\omega t] \in \mathscr{T}_w$ is continuous for any $\omega \in \mathbb{R}^{\mathbb{Z}^d}$. 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 $[\rho_{\text{uniform}}(x, y) \equiv \sup_{i \in \mathbb{Z}^d} \rho(x_i, y_i)]$. Thus, our interest in solutions $x(t)$ with x_i "close" to $\omega_i t$ with $|\omega_i| \to \infty$ as $|i| \to \infty$ explains the choice of the compact topology.

Global existence and uniqueness for the Cauchy problem associated to (2.6), with f Lipschitz, are an elementary application of standard contraction techniques (see [Pe]). We just stress that the "initial velocities" can be taken to be completely arbitrary (and not necessarily in $\mathscr{B}_w$):

PROPOSITION 2.2. *Let $f: \mathscr{T}_w \to \mathscr{B}_w$ be a Lipschitz map (i.e., $\exists C > 0$ s.t. $\|f(x) - f(y)\|_w \leqslant C\rho_w(x, y)$, $\forall x, y \in \mathscr{T}_w$). Given any $x^0 \in \mathscr{T}_w$ and any $y^0 \in \mathbb{R}^{\mathbb{Z}^d}$, there exists a unique solution, global in time, of the Cauchy problem*

$$\begin{cases} \ddot{x}_i(t) = f_i(x(t)), & i \in \mathbb{Z}^d \\ x_i(0) = x_i^0, & \dot{x}_i(0) = y_i^0. \end{cases} \tag{2.7}$$

The property of being Lipschitz depends of course on the metric, as shown by example (1.10) of Section 1. In fact, consider two cases: (1) $a_j \equiv b^{-|j|}$, $b > 1$; (2) $a_j = (1 + |j|^p)^{-1}$, $p > 1$. Then, in 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 $w_j \equiv a_j$.

3. Hamiltonian Structure and Averages

In finite dimensions, the Hamiltonian (or Lagrangian) structure in second order ODE's, $\ddot{x} = f(x)$, is expressed by the vector field f which is a *gradient*, $f = \partial_x V$; the Hamiltonian function is then $H \equiv \frac{1}{2}\dot{x}^2 - V(x)$ (and the Lagrangian is $L \equiv \frac{1}{2}\dot{x}^2 + V(x)$; see [A] for generalities).

In infinite dimensions, one could require as well that f be a gradient of a $C^1(\mathscr{T}_w; \mathbb{R})$-function. However, this attitude is much too restrictive as it imposes strong decay properties (as $|i| \to \infty$) to the components f_i of the field and even the simple examples (1.8), (1.10) of the Introduction would not fit in the picture. Indeed what one really needs is that f is a "local gradient" (where "local" refers to localized portions of $\mathbb{Z}^d$).

We shall therefore introduce the notion of *generalized gradient* (or "*g-gradient*"), bringing in the local structure with the help of finite dimensional projectors, which just *average out* the dependence upon variables x_i with i varying in the complementary of a finite subset of $\mathbb{Z}^d$.

To construct such operators, we introduce a *probability measure* on $\mathscr{T}_w$: Consider the σ-algebra, $\mathscr{R}$, generated by the "cylinders" (see, e.g., [Ha])

$$\mathscr{R}_I \equiv \bigotimes_{i \in I} A_i \bigotimes_{j \notin I} \mathbb{T}_j, \tag{3.1}$$

where A_i is an open subset of $\mathbb{T}_i$ and $I \subset \mathbb{Z}^d$ is a finite subset of $\mathbb{Z}^d$: $|I| < \infty$ ($|\cdot|$ denoting here cardinality). Then there exists (see [Ha, Section 38]) a unique (probability) measure μ on $\mathscr{R}$ such that

$$\mu(\mathscr{R}_I) = \prod_{i \in I} \mu_i(A_i), \tag{3.2}$$

where μ_i is the normalized "Lebesgue measure" on $\mathbb{T}_i$.

Now, given any (not necessarily finite) subset I of $\mathbb{Z}^d$, we can construct, as above, a product measure $\mu_I \equiv \bigotimes_{i\in I} d\mu_I$ on $\bigotimes_{i\in I} \mathbb{T}_i$ (endowed with the topology induced by the metric $\rho_I \equiv \sum_{i\in I} \rho_i$). Then if $J \equiv I^c \equiv \mathbb{Z}^d \backslash I$ the product measure $d\mu_I \otimes d\mu_J$ coincides with $d\mu$ and Fubini's theorem holds. Thus, to any bounded measurable function g on $\mathcal{T}$ we can naturally associate a measurable function, $g^{[I]}$, on $\bigotimes_{i\in I} \mathbb{T}_i$, by setting (for $x_I \in \bigotimes_{i\in I} \mathbb{T}_i$, $d\mu_i$-almost everywhere):

$$g^{[I]} \equiv \int g \, d\mu_J, \qquad J \equiv \mathbb{Z}^d \backslash I. \tag{3.3}$$

If $|I| < \infty$, $g^{[I]}$ is an honest measurable function on $\mathbb{T}^{|I|}$. We can now define g-gradients.

DEFINITION 3.1. A continuous function $f: \mathcal{T}_w \to \mathcal{B}_w$ is a g-gradient if for any finite $I \subset \mathbb{Z}^d$ there exists a $C^1(\mathbb{T}^{|I|}; \mathbb{R})$-function, $V^{(I)}(x)$, so that

$$f_i^{[I]}(x) = \partial_{x_i} V^{(I)}(x), \qquad \forall i \in I, \qquad \forall x \in \mathbb{T}^{|I|} \tag{3.4}$$

It is easy to check that the examples (1.8), (1.10) in Section 1 are g-gradients (see [Pe]).

We shall speak of second order Hamiltonian equations on $\mathcal{T}_w$ whenever f in (2.6) is a g-gradient.

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

DEFINITION 3.2. A g-gradient f is called uniformly weakly real-analytic if there exists a real number $\xi > 0$ such that for any finite set $I \subset \mathbb{Z}^d$, $V^{(I)}(x)$ is real-analytic on $\mathbb{T}^{|I|}$ and can be analytically continued to the set $\{z \in \mathbb{C}^{|I|} : |\mathrm{Im}\, z_i| \leqslant \xi\}$.

Remark 3.3. (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).

(ii) Example (1.10) in Section 1 is uniformly weakly analytic and as parameter ξ one could take any positive number; example (1.8) is uniformly weakly analytic for some (small enough) $\xi > 0$.

4. ALMOST-PERIODIC SOLUTIONS (DEFINITIONS)

We start with the definition of maximal almost-periodic functions with "rationally independent" frequency $\omega \in \mathbb{R}^{\mathbb{Z}^d}$.

DEFINITION 4.1. A sequence $\omega \in \mathbb{R}^{\mathbb{Z}^d}$ is said to be *rationally independent* if for any finite subset I of $\mathbb{Z}^d$ and for any $n_i \in \mathbb{Z}$

$$\sum_{i \in I} \omega_i n_i \neq 0 \qquad \text{unless} \quad n_i = 0 \; \forall i \in I. \tag{4.1}$$

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 4.2. A continuous real function $q(t)$ is called almost-periodic over $\mathcal{T}_w$ (with frequency ω) if there exist a rationally independent sequence $\omega \in \mathbb{R}^{\mathbb{Z}^d}$ and a continuous function $Q: \mathcal{T}_w \to \mathbb{R}$ such that $q(t) = Q([\omega t])$. A solution $x(t)$ of (2.6) is called *maximal almost-periodic* if $x_i(t) - [\omega_i t]$ is, for all i, almost-periodic over $\mathcal{T}_w$ with frequency ω.

Remark 4.3. (i) Recall [see Remark 2.1(i)] that $t \to [\omega t]$ is continuous $\forall \omega \in \mathbb{R}^{\mathbb{Z}^d}$.

(ii) A function q almost-periodic over $\mathcal{T}_w$ is almost-periodic in the sense of Bohr with frequency modulus given by

$$\sigma(q) = \left\{ r \in \mathbb{R} : r = \sum_{i \in I} \omega_i n_i \text{ for some } I \subset \mathbb{Z}^d,\ |I| < \infty,\ n_i \in \mathbb{Z} \right\} \tag{4.2}$$

(see [Ka]).

(iii) The word "maximal" in the above definition refers to the rationally independence of the frequency ω. Indeed, one can consider *quasi-periodic* solutions of (2.6): these are almost-periodic solutions with the associated frequency modulus being generated by a fixed vector $\omega^{(N)} \in \mathbb{R}^N$. The existence of such solutions has been established in a somewhat different context by [Wa2] and [Ku] (see also [Pö]).

5. DIOPHANTINE SEQUENCES

Actually the almost-periodic solutions constructed below will have frequencies ω verifying much stronger numerical properties than just being rationally independent: *They will be Diophantine* in the sense of the following definition.

DEFINITION 5.1. A rationally independent sequence $\omega \in \mathbb{R}^{\mathbb{Z}^d}$ is called *Diophantine* if for any finite set $I \subset \mathbb{Z}^d$, there exist constants $\gamma > 0$ and τ $(\geqslant |I|)$ such that for any choice of $n_i \in \mathbb{Z}$ with $\sum_{i \in I} |n_i| > 0$ we have

$$\left| \sum_{i \in I} \omega_i n_i \right| \geqslant \frac{1}{\gamma (\sum_{i \in I} |n_i|)^{\tau}}. \tag{5.1}$$

It is well known that Diophantine vectors in $\mathbb{R}^N$ form a set of full Lebesgue measure (for completeness we reproduce this elementary result in Appendix 1), but also in infinite dimensions Diophantine frequencies are rather abundant. One can in fact construct many Diophantine sequences with the help of the following Lemma (compare with Lemma 3 of [CZ]).

We recall that a vector $\omega \in \mathbb{R}^N$ is called (γ, τ)-*Diophantine* if

$$|\omega \cdot n| \geqslant \frac{1}{\gamma\, |n|^{\tau}}, \qquad \forall n \in \mathbb{Z}^N \setminus \{0\}, \tag{5.2}$$

where $\omega \cdot n$ denotes the standard scalar product in $\mathbb{R}^N$ and $|n| \equiv \sum_{i=1}^{N} |n_i|$.

LEMMA 5.2. *Let $\omega \in \mathbb{R}^N$ be (γ, N)-Diophantine; let Ω be a positive number satisfying*

$$\Omega \geqslant 4\sqrt{N}\, |\omega| \qquad \left(|\omega| \equiv \sum_{i=1}^{N} |\omega_i| \right) \tag{5.3}$$

and define the following subset of $[\Omega, \infty)$:

$$\begin{aligned} \mathscr{A}_N \equiv \mathscr{A}_N(\omega, \Omega) \equiv \{ \alpha \geqslant \Omega : |\omega \cdot n + \alpha h| \geqslant \Omega / [2(|n| + |h|)^{N+1}], \\ \forall n \in \mathbb{Z}^N, \forall 0 \neq h \in \mathbb{Z} \}. \end{aligned} \tag{5.4}$$

There exists a universal number $K > 1$ such that

$$l([\Omega, \infty) \setminus \mathscr{A}_N) < K\, |\omega|, \tag{5.5}$$

where l denotes Lebesgue measure.

Remark 5.3. (i) Since by definition $\Omega > 2/\gamma$, [as $|\omega| \geqslant |\omega_i| \geqslant \gamma^{-1}$ by (5.2)], (ω, α) is $(\gamma, N+1)$-Diophantine whenever $\alpha \in \mathscr{A}_N$.

(ii) From this, in particular, the above lemma tells us that given a (γ, N)-Diophantine vector ω in $\mathbb{R}^N$ (and almost all vectors in $\mathbb{R}^N$ are (γ, N)-Diophantine for some γ: see Appendix 1) we can pick α in $\mathscr{A}_N \subset [\Omega, \infty)$ [whose complementary measure in $[\Omega, \infty)$ is of $O(|\omega|)$] so that the vector (ω, α) is $(\gamma, N+1)$-Diophantine.

(iii) In the above statement and in its proof we could replace N with any $\tau > N-1$; however, to avoid introducing too many parameters we consider only the case $\tau = N$.

(iv) It is now easy to construct many Diophantine sequences. Fix $I_0 \in \mathbb{Z}^d$, $|I_0| = N < \infty$; pick a (γ_0, N)-Diophantine vector $\omega^{(N)} \in \mathbb{R}^N$ (with some $\gamma_0 > 0$); and fix a one-to-one map, j_h, from $\mathbb{Z}_+$ onto $\mathbb{Z}^d \setminus I_0$. Now set $\omega_i \equiv \omega_i^{(N)}$ $\forall i \in I_0$ and define ω_{j_h}, $h \geqslant 1$, inductively as follows. Let

$\Omega_1 \equiv 4\sqrt{N}\,|\omega^{(N)}|$ and choose ω_{j_1} in $\mathscr{A}_N(\omega^{(N)}, \Omega_1)$ so that, by the above Lemma, $\omega^{(N+1)} \equiv (\omega^{(N)}, \omega_{j_1})$ is $(\gamma_0, N+1)$-Diophantine. Analogously, given $\omega^{(N+h)} \equiv (\omega^{(N)}, ..., \omega_{j_h}) \in \mathbb{R}^{N+h}$, $(\gamma_0, N+h)$-Diophantine, we let $\Omega_h \equiv 4\sqrt{N+h}\,(|\omega^{(N)}| + \sum_{k=1}^{h} |\omega_{j_k}|)$ and pick $\omega_{j_{h+1}} \in \mathscr{A}_{N+h}(\omega^{(N+h)}, \Omega_h)$. It is clear that in this way one constructs many Diophantine sequences satisfying (5.1) with $\gamma \equiv \gamma_0$ and $\tau = N+k$ where $k \equiv 0$ if $I \subset I_0$, and otherwise $k \equiv \max\{h\colon j_h \in I\}$.

Proof of Lemma 5.2. For vectors in $\mathbb{R}^m$, we denote by $\|\cdot\|$ the Euclidean norm and by $|\cdot|$ the 1-norm (sum of absolute values of components). Set $a \equiv \Omega/(2\,\|\omega\|)$, $e \equiv \omega/\|\omega\|$, and let σ_N be the area of the unit-sphere $S^{N-1} \equiv \{x \in \mathbb{R}^N : \|x\| = 1\}$. Finally denote by n a generic vector in $\mathbb{Z}^N$ and by h a generic integer number and for a vector $y \in \mathbb{R}^N$, let $\bar{y} \equiv (y_2, ..., y_N)$. Now, first let $N \geqslant 2$; then

$$
\begin{aligned}
l([\Omega, \infty)\backslash \mathscr{A}_N) &= l\left\{\alpha \geqslant \Omega : |\omega\cdot n + \alpha h| < \frac{\Omega}{2(|n| + |h|)^{N+1}} \text{ for some } h \neq 0\right\} \\
&\leqslant \sum_{n\neq 0,\, h\neq 0} l\left\{\alpha \geqslant \Omega : \left|\frac{\omega\cdot n}{h} + \alpha\right| < \frac{\Omega}{2|h|\,(|n| + |h|)^{N+1}}\right\} \\
&\leqslant 2\Omega \sum_{h\geqslant 1} \frac{1}{h} \sum_{|\omega\cdot(n/h)| \geqslant \Omega/2} \frac{1}{(|n| + |h|)^{N+1}} \\
&= 2\Omega \sum_{h\geqslant 1} \frac{1}{h} \sum_{|e\cdot n| \geqslant ah} \frac{1}{(|n| + \mathrm{h})^{N+1}} \\
&\leqslant 2\Omega \sum_{h\geqslant 1} \frac{1}{h} \int_{\{x\in\mathbb{R}^N : |e\cdot x| \geqslant ah - \sqrt{N}\}} \frac{dx}{(|x| + h)^{N+1}} \\
&\leqslant 2\Omega \sum_{h\geqslant 1} \frac{1}{h} \int_{\{y\in\mathbb{R}^N : |y_1| \geqslant ah - \sqrt{N}\}} \frac{dy}{(\|y\| + h)^{N+1}} \\
&\leqslant 2\Omega \sum_{h\geqslant 1} \frac{1}{h} 2^{(N+1)/2} \int_{\{y\in\mathbb{R}^N : |y_1| \geqslant ah - \sqrt{N}\}} \frac{dy}{(|y_1| + \|\bar{y}\| + h)^{N+1}} \\
&= 2\Omega \sum_{h\geqslant 1} \frac{1}{h} \frac{2^{(N+3)/2}}{N} \int_{\mathbb{R}^{N-1}} \frac{d\bar{y}}{(\|\bar{y}\| + (a+1)h - \sqrt{N})^N} \\
&\leqslant 2\Omega \sum_{h\geqslant 1} \frac{1}{h} \frac{2^{(N+3)/2}}{N} \sigma_{N-1} \int_0^\infty \frac{r^{N-2}}{(r + (a/2)h)^N}\, dr \\
&\leqslant 2\Omega \sum_{h\geqslant 1} \frac{1}{h} \frac{2^{(N+3)/2}}{N} \sigma_{N-1} \int_0^\infty \frac{1}{(r + (a/2)h)^2}\, dr \\
&\leqslant K\,|\omega|,
\end{aligned}
$$

where in the sum after the first inequality we have $n \neq 0$ as $\Omega > 2/\gamma$ [see Remark 5.3(i)]; in the fourth inequality we set $By = x$ with B a unitary matrix sending the vector $(1, 0, ..., 0)$ to e; in the sixth inequality we used the assumption that $\Omega \geqslant 4 |\omega| \sqrt{N}$ and we have taken:

$$K \equiv 2^{9/2} \left(\sum_{h \geqslant 1} h^{-2} \right) \sup_N \left\{ \frac{2^{N/2} \sigma_{N-1}}{N} \right\}. \tag{5.6}$$

The case $N = 1$ is just shorter. ∎

6. Main Results

Theorem 6.1. *Let $f: \mathcal{T}_w \to \mathcal{B}_w$ be a uniformly weakly analytic g-gradient (Definitions 3.1, 3.2). Then there exist uncountably many maximal almost-periodic solutions of (2.6) (Definition 4.2) with Diophantine frequencies (Definition 5.1).*

This result is a simple corollary of the next more detailed theorem. Recall that a *non-degenerate (maximal) quasi-periodic solution* of

$$\ddot{x} = V_x(x), \qquad x \in \mathbb{T}^N \tag{6.1}$$

with frequency $\omega \in \mathbb{R}^N$ is a solution of the form

$$x(t) = [\omega t + u([\omega t])] \tag{6.2}$$

for a suitable function $u: \mathbb{T}^N \to \mathbb{R}^N$ satisfying

$$\det(\mathrm{Id}_N + u_\theta(\theta)) \neq 0, \qquad \forall \theta \in \mathbb{T}^N, \tag{6.3}$$

where Id_N is the identity $(N \times N)$-matrix and $(u_\theta)_{hk} \equiv \partial_{\theta_k} u_h$, $h, k = 1, ..., N$; the word maximal refers to the maximal dimension of the frequency ω.

Remark 6.2 (On quasi-periodic solutions). (i) To any quasi-periodic solution with ω rationally independent (i.e., $\omega \cdot n = 0$ for some $n \in \mathbb{Z}^N \Rightarrow n = 0$) one can associate an N-parameter family of solutions obtained by "phase-translation": for each $\theta \in \mathbb{T}^N$, $x(t; \theta) \equiv [\theta + \omega t + u([\theta + \omega t])]$ is still a solution (as the flow $t \to [\omega t]$ is dense in $\mathbb{T}^N$).

(ii) From now on we shall often omit the projection $[\cdot]$ [see (1.2), (2.4), and (ii) of Remark 2.1] in the notation.

(iii) Indeed, the above family corresponds to an *invariant N-torus*, embedded in the phase-space $\mathbb{R}^N \times \mathbb{T}^N$, given by

$$\mathcal{T}_\omega^N \equiv \{(y, x) \in \mathbb{R}^N \times \mathbb{T}^N : (y, x) = (\omega + D_\omega u(\theta), \theta + u(\theta)), \theta \in \mathbb{T}^N\}, \tag{6.4}$$

where $D_\omega \equiv \omega \cdot \partial_\theta \equiv \sum_{i=1}^N \omega_i \partial_{\theta_i}$; the y-component is just the velocity vector corresponding to the point x (as D_ω corresponds to d/dt along the linear flow $t \to \theta + \omega t$). The non-degeneracy (6.3) of the solution (6.2) allows us to see $\mathcal{T}_\omega^N$ as a regular embedded torus in the ambient space $\mathbb{R}^N \times \mathbb{T}^N$.

(iv) In view of the above observations it is clear that to find non-degenerate quasi-periodic solutions of (6.1) *is equivalent to finding solutions of the following non-linear* PDE *on* $\mathbb{T}^N$:

$$D_\omega^2 u(\theta) = V_x(\theta + u(\theta)), \qquad \min_{\theta \in \mathbb{T}^N} |\det(\mathrm{Id}_N + u_\theta)| > 0. \tag{6.5}$$

(Just substitute (6.2) in (6.1) and use the rational independence of ω to replace ωt with the generic point θ.) For investigations on (6.5) see [M2, SZ, CC1, CC2, CZ].

THEOREM 6.3. *Let f be as in Theorem* 6.1 *and assume that for some finite $I_0 \subset \mathbb{Z}^d$ the equation*

$$\ddot{x}^{(0)} = \partial_{x^{(0)}} V^{(0)}(x^{(0)}), \qquad x^{(0)} \in \mathbb{T}^{N_0}, \qquad N_0 \equiv |I_0| \tag{6.6}$$

admits a (maximal) non-degenerate quasi-periodic solution with a (γ, N_0)-Diophantine frequency vector $\omega^{(0)} \in \mathbb{R}^{N_0}$ [see (6.1)–(6.3) *and* (5.2)]. *Then for any $\varepsilon > 0$ there exist uncountably many maximal almost-periodic solutions, $x(t)$, of* (2.6) *with Diophantine frequencies $\omega \in \mathbb{R}^{\mathbb{Z}^d}$ such that $\omega_i = \omega_i^{(0)}$ for $i \in I_0$ ($|\omega_i| \to \infty$ as $|i| \to \infty$) and*:

$$\begin{aligned} &\sup_{\substack{t \in \mathbb{R} \\ p = 0,1}} \left| \frac{d^p}{dt^p} [x_i(t) - x_i^{(0)}(t)] \right| \leqslant \varepsilon \qquad \textit{for} \quad i \in I_0 \\ &\sup_{\substack{t \in \mathbb{R} \\ p = 0,1}} \left| \frac{d^p}{dt^p} [x_i(t) - \omega_i t] \right| \leqslant \varepsilon \qquad \textit{for} \quad i \notin I_0. \end{aligned} \tag{6.7}$$

The proof of the above theorems will be given in Section 8.

7. A (Finite-Dimensional) Average Theorem

In this section we discuss a finite-dimensional problem, which may be of some interest by itself, and whose solution will constitute the main step of the proof of Theorem 6.3.

Roughly speaking the question is how to construct quasi-periodic solutions for a (second order) Hamiltonian system on $\mathbb{T}^{N+1}$ if the existence of quasi-periodic solutions for a "subsystem" obtained from the original one by averaging out some variables is known.

More precisely, let $V: \mathbb{T}^{N+1} \to \mathbb{R}$, $(N \geqslant 1)$, be real-analytic and consider the "Hamiltonian" equations:

$$\ddot{x}' = V_{x'}(x'), \qquad x' \in \mathbb{T}^{N+1}. \tag{7.1}$$

Also let $\bar{V}(x)$, $x \in \mathbb{T}^N$, denote the "averaged potential"

$$\bar{V}(x) \equiv \frac{1}{2\pi} \int_0^{2\pi} V(x')\, dx_{N+1}, \qquad x' \equiv (x, x_{N+1}) \in \mathbb{T}^N \times \mathbb{T} \tag{7.2}$$

and consider the Hamiltonian equations in $\mathbb{T}^N$ associated to $\bar{V}$:

$$\ddot{x} = \bar{V}_x(x), \qquad x \in \mathbb{T}^N. \tag{7.3}$$

Now assume that (7.3) admits a non-degenerate (maximal) quasi-periodic solution

$$x = \omega t + u(\omega t), \qquad \omega \in \mathbb{R}^N, \qquad u: \mathbb{T}^N \to \mathbb{R}^N \tag{7.4}$$

with ω (γ, N)-Diophantine and u real-analytic. The question is: Can one find quasi-periodic solutions for (7.1) "close" (in some sense) to (7.4)?

The answer is positive *provided* the looked after quasi-periodic solutions have frequencies $\omega' \equiv (\omega, \alpha)$ with $\alpha \gg |\omega|$ and suitable.

To formulate a precise and *quantitative* result, we need some notation.

Given $N, M \geqslant 1$ and a function $g: \mathbb{T}^N \to \mathbb{R}^M$ real-analytic on

$$\Delta_\xi^N \equiv \{\theta \in \mathbb{C}^N : |\mathrm{Im}\, \theta_i| \leqslant \xi,\ i = 1, ..., N\} \tag{7.5}$$

[this means that the components g_h, for $h = 1, ..., M$, admit a holomorphic extension to some open domain containing Δ_ξ^N] we set:

$$\|g\|_{\Delta_\xi^N} \equiv \|g\|_\xi \equiv \sum_{h=1}^M \sup_{\Delta_\xi^N} |g_h|. \tag{7.6}$$

Now let $\mathscr{L}^p(\mathbb{C}^N)$, for $p \in \mathbb{N}$, be the space of linear maps from $\mathbb{C}^N$ into $\mathscr{L}^{p-1}(\mathbb{C}^N)$, $(\mathscr{L}^0(\mathbb{C}^N) \equiv \mathbb{C}^N$, $\mathscr{L}^1(\mathbb{C}^N) \equiv \mathscr{L}(\mathbb{C}^N) \equiv (N \times N)$-matrices, ...). If $T: \mathbb{T}^N \to \mathscr{L}^p(\mathbb{C}^N)$ is real-analytic on Δ_ξ^N [this means that $\forall c_1, ..., c_p \in \mathbb{C}^N$, the function $\theta \to (\cdots((Tc_1)c_2)\cdots c_N)$ is a real-analytic $\mathbb{C}^N$-*valued* function on Δ_ξ^N] then we set (inductively):

$$\|T\|_{\Delta_\xi^N} \equiv \|T\|_\xi \equiv \sup_{\substack{c \in \mathbb{C}^N \\ |c| = 1}} \|Tc\|_\xi, \qquad \left(|c| \equiv \sum_{h=1}^N |c_h| \quad \text{for} \quad c \in \mathbb{C}^N\right). \tag{7.7}$$

Finally, observe that, without loss of generality, *we can assume that u in* (7.4) *has vanishing mean value over* $\mathbb{T}^N$ (as we can replace $u(\theta)$ with $c+u(c+\theta)$ with $c=-\langle u\rangle$, $\langle\cdot\rangle$ denoting average over $\mathbb{T}^N$).

PROPOSITION 7.1. *Let* $V\colon \mathbb{T}^{N+1}\to\mathbb{R}$ *be a real-analytic function, let* $\bar V$ *be as in* (7.2) *and assume that* (7.3) *admits a (maximal) non-degenerate quasi-periodic solution* (7.4) *with a* (γ, N)*-Diophantine frequency vector* ω. *Fix* $0<r<\rho<\xi_0$ *so that*: (i) V *is real-analytic on* $\Delta_{\xi_0}^{N+1}$, (ii) u *is real-analytic on* Δ_r^N, (iii) $\{\theta+u(\theta):\theta\in\Delta_r^N\}\subset\Delta_\rho^N$, *and* (iv):

$$\|\mathrm{Id}_N+\partial_\theta u\|_r\equiv U,\qquad \|(\mathrm{Id}_N+\partial_\theta u)^{-1}\|_r\equiv\bar U<\infty. \tag{7.8}$$

Finally fix $0<r'<r$ *and let* Ω *be such that*

$$\begin{aligned}
&\Omega\geqslant\max\{4\sqrt{N}\,|\omega|,\gamma^{-1}\delta\}\\
&\delta^2\equiv C(N+1)!^6\,2^{60(N+1)}U^{10}\bar U^8\beta^3(r-r')^{-[11(N+1)+1]}(\xi_0-\rho)^{-1}\\
&\bar\delta\equiv\delta^2(\xi_0-\rho)\\
&\beta\equiv\max_{p=1,2,3}\{1,\gamma^2\,\|\partial_{x'}^p V\|_{\xi_0}\},
\end{aligned}\tag{7.9}$$

where $C>1$ *is a suitable universal constant. Then for any* $\alpha\in\mathscr{A}_N(\omega,\Omega)$ [*see* (5.4)] *there exists a (maximal) quasi-periodic solution of* (7.1), $x'=\omega't+u'(\omega't)$, *with* $\omega'\equiv(\omega,\alpha)$ *and* $u'\colon\mathbb{T}^{N+1}\to\mathbb{R}^{N+1}$ *real-analytic on* $\Delta_{r'}^{N+1}$. *Furthermore* $\langle u'\rangle=0$ *and*:

$$\begin{gathered}
\{\theta'+u'(\theta'):\theta'\in\Delta_{r'}^{N+1}\}\subset\Delta_{\rho'}^{N+1},\qquad \rho'\equiv\rho+\frac{\bar\delta}{(\gamma\Omega)^2}<\xi_0\\
\max_{p=0,1}\|\partial_{\theta'}^p[u'-(u,0)]\|_{r'}\leqslant\bar\delta(\gamma\Omega)^{-2}\\
\|D_{\omega'}[u'-(u,0)]\|_{r'}\leqslant\delta(\gamma\Omega)^{-1},
\end{gathered}\tag{7.10}$$

where $D_{\omega'}\equiv\omega'\cdot\partial_{\theta'}\equiv\sum_{h=1}^N\omega_h\,\partial_{\theta_h}+\alpha\,\partial_{\theta_{N+1}}$.

In the following we will not need the explicit (certainly not optimal) dependence upon N given in (7.9); nevertheless, we shall pay some attention to constants for the sake of concreteness and also because it may help the reader to keep track of the various estimates.

The proof is based on a result à la Nash–Moser from KAM theory which guarantees the existence of solutions of the equation (6.5) provided the frequency vector ω is Diophantine and provided one can find a "good enough *approximate solution*":

LEMMA 7.2. *Let*: $\xi < \xi_0' < \xi_0 < 1$, $V: \mathbb{T}^N \to \mathbb{R}$ *be a real-analytic function on* $\Delta_{\xi_0}^N$, $v: \mathbb{T}^N \to \mathbb{R}^N$ *be real-analytic on* Δ_ξ^N *and such that* $\{\theta + v(\theta): \in \Delta_\xi^N\} \subset \Delta_{\xi_0'}^N$, $\omega \in \mathbb{R}^N$ *be a* (γ, N)*-Diophantine vector; finally let also*

$$\begin{aligned} &\|\mathrm{Id}_N + v_\theta\|_\xi \equiv \eta, && \|(\mathrm{Id}_N + v_\theta)^{-1}\|_\xi \equiv \bar{\eta} < \infty \\ &\gamma^2 \|D^2 v - V_x(\theta + v)\|_\xi \equiv \varepsilon, && \max_{p=1,2,3} \{1, \gamma^2 \|\partial_x^p V\|_{\xi_0}\} \equiv \beta \end{aligned} \tag{7.11}$$

$(D \equiv D_\omega \equiv \omega \cdot \partial_\theta)$. *Fix* $0 < \xi' < \xi$. *There exists a universal constant* $B > 1$ *such that if*

$$BN!^4\, 2^{40N} \eta^{10} \bar{\eta}^8 \beta (\xi - \xi')^{-8N} (\xi_0 - \xi_0')^{-1}\, \varepsilon \leqslant 1 \tag{7.12}$$

then there exists a function $u: \mathbb{T}^N \to \mathbb{R}^N$, *real-analytic on* $\Delta_{\xi'}^N$, *which is solution of*

$$D^2 u = V_x(\theta + u), \qquad \langle u \rangle = \langle v \rangle. \tag{7.13}$$

Furthermore the estimates

$$\begin{aligned} &\|\mathrm{Id}_N + u_\theta\|_{\xi'} \leqslant 2\eta, \qquad \|(\mathrm{Id}_N + u_\theta)^{-1}\|_{\xi'} \leqslant 2\bar{\eta} \\ &\max_{\substack{p=0,1 \\ q=0,1,2}} \{\|\partial_\theta^p (u - v)\|_{\xi'}, \gamma^q \|D^q(u - v)\|_{\xi'}\} \leqslant A\varepsilon \end{aligned} \tag{7.14}$$

hold, where $A \equiv BN!^4\, 2^{40N} \eta^{10} \bar{\eta}^8 \beta (\xi - \xi')^{-8N}$.

One can actually show that the above solution u is "locally" unique (see [CC2]).

This lemma is a refinement of Lemma 6 of [CC1] (see also [M2, SZ, CC2]); the main difference is that we need here to leave the width of the domain of analyticity of the solution u as a *free parameter* (while in [CC1] it was fixed to be half of ξ). Rather than indicating the (tiny but dense) adjustments to the proof in [CC1] we present a complete (and short) proof in Appendix 2.

Proof of Proposition 7.1. We shall use Lemma 7.2 [with N, ω replaced by, respectively, $N+1, \omega' \equiv (\omega, \alpha)$]: We shall construct an approximate solution v so that the "error function" $e \equiv D_{\omega'}^2 v - V_{x'}(\theta' + v)$ has norm bounded by $O(1/\Omega^2)$ when $\omega' \equiv (\omega, \alpha)$, $\alpha \in \mathscr{A}_N(\omega, \Omega)$ with Ω chosen so large (7.9) that the condition (7.12) is satisfied.

We start by observing that the average over $\mathbb{T}^{N+1}$ of the vector-valued function

$$W(\theta') \equiv V_{x'}(\theta + u, \theta_{N+1}), \qquad \theta' \equiv (\theta, \theta_{N+1}) \in \mathbb{T}^N \times \mathbb{T} \tag{7.15}$$

is zero; in fact $\langle W_{N+1} \rangle = 0$ because

$$\int_0^{2\pi} \partial_{x_{N+1}} V(\theta + u, \theta_{N+1})\, d\theta_{N+1} = 0 \tag{7.16}$$

while the average of the first N component of W are given by ($x' \equiv (x, x_{N+1}) \in \mathbb{T}^N \times \mathbb{T}$ and $\theta' \equiv (\theta, \theta_{N+1}) \in \mathbb{T}^N \times \mathbb{T}$):

$$\int_{\mathbb{T}^{N+1}} \partial_x V(\theta + u, \theta_{N+1}) \frac{d\theta'}{(2\pi)^{N+1}} \equiv \int_{\mathbb{T}^N} \partial_x \bar{V}(\theta + u) \frac{d\theta}{(2\pi)^N}$$

$$= \int_{\mathbb{T}^N} D_\omega^2 u(\theta) \frac{d\theta}{(2\pi)^N} = 0. \tag{7.17}$$

Now, denoting, for a function G with $\langle G \rangle = 0$, $D^{-p}G$ the unique solution with zero average of $D^p g = G$, we set

$$v(\theta') \equiv D_{\omega'}^{-2} W(\theta') \equiv - \sum_{\substack{n' = (n, h) \in \mathbb{Z}^N \times \mathbb{Z} \\ n' \neq 0}} \frac{W_{n'}}{(\omega \cdot n + \alpha h)^2} e^{i n' \cdot \theta'}, \tag{7.18}$$

where W_n denote Fourier coefficients [and recall that $\omega' \equiv (\omega, \alpha)$ with a fixed $\alpha \in \mathscr{A}_N(\omega, \Omega)$]. Note that, because u satisfies $D_\omega^2 u = \bar{V}(\theta + u)$, v can be written in the form

$$v = - \sum_{n \in \mathbb{Z}^N, h \neq 0} \frac{W_{n'}}{(\omega \cdot n + \alpha h)^2} e^{i n' \cdot \theta'} + (u, 0) \equiv \tilde{v} + (u, 0) \tag{7.19}$$

and therefore

$$e \equiv D_{\omega'}^2 v - V_{x'}(\theta' + v) = V_{x'}(\theta + u, \theta_{N+1}) - V_{x'}((\theta + u, \theta_{N+1}) + \tilde{v}). \tag{7.20}$$

Next, we estimate $\tilde{v}$, v and e on $\Delta_{\bar{r}}^{N+1}$ for any $r' < \bar{r} < r$. The hypothesis (iii) of Proposition 7.1 and the analyticity assumptions yield:

$$\|W\|_r \leqslant \|V_{x'}\|_{\xi_0} \Rightarrow |W_{n'}| \leqslant \|V_{x'}\|_{\xi_0} e^{-|n'| r}. \tag{7.21}$$

Then, by (7.19), the definition of $\mathscr{A}_N$, and (A2.3) we obtain ($p = 0, 1, ...$):

$$\|\partial_{\theta'}^p \tilde{v}\|_{\bar{r}} \leqslant 4 \frac{\|V_{x'}\|_{\xi_0}}{\Omega^2} \sum_{n' \in \mathbb{Z}^{N+1}} |n'|^{2(N+1)+p} e^{-|n'|(r - \bar{r})}$$

$$\leqslant 4^{3N+4+p} \frac{\|V_{x'}\|_{\xi_0}}{\Omega^2} [2(N+1)+p]!\, (r - \bar{r})^{-(3N+3+p)}. \tag{7.22}$$

Next we have to evaluate the norms η, $\bar{\eta}$ in the text of Lemma 7.2. It is easy to check that:

$$\eta \leqslant U + \|\partial_{\theta'} \tilde{v}\|_{\bar{r}}, \qquad \bar{\eta} \leqslant \bar{U}(1 - \bar{U}\, \|\partial_{\theta'} \tilde{v}\|_{\bar{r}})^{-1}. \tag{7.23}$$

Note also that, since

$$\{\theta' + v : \theta' \in \Delta_{\bar{r}}^{N+1}\} \subset \Delta_{\rho + \|\tilde{v}\|_{\bar{r}}}^{N+1}, \tag{7.24}$$

we can take (in applying Lemma 7.2) the parameter $\xi'_0 = \rho + \|\tilde{v}\|_{\bar{r}}$. Now, choosing $\bar{r} = (r + r')/2$ we easily obtain, for $p = 0, 1$, the bounds [recall the definition of β in (7.9) and see (7.20)]

$$\|\partial_{\theta'}^{p} \tilde{v}\|_{\bar{r}} \leqslant \hat{\delta}(\gamma\Omega)^{-2}, \qquad \varepsilon \equiv \gamma^2\, \|e\|_{\bar{r}} \leqslant \beta\hat{\delta}(\gamma\Omega)^{-2} \tag{7.25}$$

with

$$\hat{\delta} \equiv 2^{9N+14}(2N+3)!\, \beta(r - r')^{-(3N+4)}. \tag{7.26}$$

From $\delta \leqslant (\gamma\Omega)$, [see (7.9)], it then follows that [recall that U, $\bar{U} \geqslant 1$ while $\xi_0 < 1$; and again $p = 0, 1$]

$$\|\partial_{\theta'}^{p} \tilde{v}\|_{\bar{r}} \leqslant \frac{\hat{\delta}}{(\gamma\Omega)^2} < \frac{\xi_0 - \rho}{2}, \qquad \eta \leqslant 2U, \qquad \bar{\eta} \leqslant 2\bar{U}. \tag{7.27}$$

We are now in a position to apply Lemma 7.2 with $\xi \equiv \bar{r} \equiv (r + r')/2$, $\xi'_0 \equiv \rho + \|\tilde{v}\|_{\bar{r}} \leqslant (\rho + \xi_0)/2$, $\xi' \equiv r'$, and N replaced by $(N+1)$ and u by u': using the above estimates [and bounding $(2N+3)!$ by a constant times $2^{3N}(N+1)!^2$] we see that the condition (7.12) in Lemma 7.2 is implied (for a suitable $C > B$) by $\delta^2(\gamma\Omega)^{-2} \leqslant 1$, which is verified because of our choice of Ω: In fact, by (7.14), (7.27), (7.25) it is

$$A\varepsilon \leqslant \frac{\delta^2}{2}(\gamma\Omega)^{-2}(\xi_0 - \rho) \equiv \frac{\bar{\delta}}{2}(\gamma\Omega)^{-2} \tag{7.28}$$

and condition (7.12) reads just $A\varepsilon(\xi_0 - \rho)^{-1} \leqslant 1$. Furthermore we see that

$$\{\theta' + u' : \theta' \in \Delta_{r'}^{N+1}\} \subset \Delta_{\rho + \|\tilde{v}\|_{\bar{r}} + A\varepsilon}^{N+1} \subset \Delta_{\rho + \bar{\delta}/(\gamma\Omega)^2}^{N+1} \equiv \Delta_{\rho'}^{N+1} \tag{7.29}$$

[recall (7.14), (7.24), (7.25), and the definition of $\bar{\delta}$ in (7.10)], where we used $\hat{\delta} < \delta^2(\xi_0 - \rho)/2 \equiv \bar{\delta}/2$; the fact that $\rho' < \xi_0$ [see (7.10)] follows from $\delta^2(\gamma\Omega)^{-2} \leqslant 1$ and from the first of (7.27). Now, for $p = 0, 1$, using (7.14), (7.28) and $\hat{\delta} < \bar{\delta}/2$, we get

$$\|\partial_{\theta'}^{p}(u' - (u, 0))\|_{r'} = \|\partial_{\theta'}^{p}(u' - v + \tilde{v})\|_{r'} \leqslant A\varepsilon + \frac{\hat{\delta}}{(\gamma\Omega)^2} \leqslant \frac{\bar{\delta}}{(\gamma\Omega)^2}. \tag{7.30}$$

Finally, mimicking the bounds (7.21)–(7.25), one obtains

$$\|D_{\omega'}\tilde{v}\|_r \leqslant \gamma^{-1}\frac{\hat{\delta}}{\gamma\Omega};\tag{7.31}$$

thus, by (7.14), (7.30), and [see (7.9)] using $\delta(\gamma\Omega)^{-1} \leqslant 1$, $\hat{\delta} \leqslant \delta/2$, the bound on $\|D_{\omega'}(u' - (u, 0))\|$ also follows easily. ∎

8. Proof of Existence of Almost-Periodic Solutions

Here we prove Theorem 6.3. Theorem 6.1 follows immediately from Theorem 6.3 and from Corollary A2.4 (see Appendix 2).

Proof of Theorem 6.3. The idea is to use iteratively the results of Section 7 to construct quasi-periodic solutions for larger and larger subsystems and to obtain, in the limit, almost periodic-solutions.

By hypothesis we are given a real-analytic quasi-periodic solution $y(t) \equiv \omega^{(0)}t + u^{(0)}(\omega^{(0)}t)$ of the subsystem

$$\ddot{y} = V_y^{(I_0)}(y), \qquad y \in \mathbb{T}^{N_0}, \qquad N_0 \equiv |I_0|,\tag{8.1}$$

the frequency vector $\omega^{(0)}$ being (γ, N_0)-Diophantine.

Thus calling ξ_0 the analyticity parameter associated to the field f (see Definition 3.2), we can assume that there exist $0 < r_0 < \rho_0 < \xi_0$ so that $u^{(0)}$ is real-analytic on $\Delta_{r_0}^{N_0}$ and such that

$$\begin{gathered}\{\theta^{(0)} + u^{(0)}(\theta^{(0)}) : \theta^{(0)} \in \Delta_{r_0}^{N_0}\} \subset \Delta_{\rho_0}^{N_0}\\ \|\mathrm{Id}_{N_0} + \partial_{\theta^{(0)}}u^{(0)}\|_{r_0} \equiv U_0, \qquad \|(\mathrm{Id}_{N_0} + \partial_{\theta^{(0)}}u^{(0)})^{-1}\|_{r_0} \equiv \bar{U}_0 < \infty.\end{gathered}\tag{8.2}$$

For concreteness we shall fix $\rho_0 \equiv \xi_0/2$ (and take r_0 small enough; note that ξ_0 is a fixed parameter which shall not change in the iteration).

In the following construction there is quite a bit of freedom (whence comes the uncountability of the solutions) as the extra frequencies ω_i, $i \notin I_0$ will be arbitrarily chosen in sets of positive measure. There is also some (less substantial) freedom in the *order* of "invading" $\mathbb{Z}^d$. More precisely, fix (arbitrarily) a one-to-one and onto map, $k \in \mathbb{Z}_+ \to j_k \in \mathbb{Z}^d \backslash I_0$, and set, recursively, for $k \geqslant 1$,

$$I_k \equiv I_{k-1} \cup \{j_k\}\tag{8.3}$$

so that

$$I_{k-1} \subset I_k, \qquad |I_k| = N_0 + k \equiv N_k, \qquad I_k \nearrow \mathbb{Z}^d.\tag{8.4}$$

We shall use the following notations. Denote by $(x_1, ..., x_{N_0})$ coordinates for $\otimes_{i \in I_0} \mathbb{T}_i$ and by x_{N_k} the coordinate associated to $\mathbb{T}_{j_k}$. Then set, for $k \geq 0$:

$$\begin{gathered} V^{(k)} \equiv V^{(I_k)}, \qquad x^{(k)} \equiv (x_1, ..., x_{N_k}) \in \mathbb{T}^{N_k} \\ \beta_k \equiv \max_{p=1,2,3} \{1, \gamma^2 \|\partial^p_{x^{(k)}} V^{(k)}\|_{\xi_0}\}. \end{gathered} \tag{8.5}$$

The $(k+1)$th step will consist in constructing non-degenerate quasi-periodic solutions with $u^{(k+1)}: \mathbb{T}^{N_{k+1}} \to \mathbb{R}^{N_{k+1}}$ and frequency $\omega^{(k+1)} \in \mathbb{R}^{N_{k+1}}$ of the form $(\omega^{(k)}, \alpha_k)$ with $\alpha_k \in \mathscr{A}_{N_k}(\omega^{(k)}, \Omega_k)$ for suitable $\Omega_k \gg \Omega_{k-1}$.

Note that the functions $V^{(k)}$ are defined up to an additive constant which we shall choose so that, for all $k \geq 0$:

$$V^{(k)} = \frac{1}{2\pi} \int_0^{2\pi} V^{(k+1)}\, dx_{j_{k+1}} \equiv \bar{V}^{(k+1)}: \mathbb{T}^{N_k} \to \mathbb{R}. \tag{8.6}$$

Next, we fix the sequence r_k measuring the analyticity domains of the $u^{(k)}$'s (recall that, in Proposition 7.1, r' is any number between zero and r): Also here the choice is rather arbitrary as the only requirement is that $r_k \searrow r_\infty > 0$. We shall take

$$r_k \equiv r_0 \left(1 - \mu \sum_{h=1}^{k} \frac{1}{h^\nu}\right), \qquad \mu \equiv \left(2 \sum_{h=1}^{\infty} \frac{1}{h^\nu}\right)^{-1}, \tag{8.7}$$

where $\nu > 1$ is a prefixed number (thus $r_\infty = r_0/2$).

Now, imagine that $u^{(k)}$, $(k \geq 0)$, is given together with $\omega^{(k)} \in \mathbb{R}^{N_k}$, that $u^{(k)}$ is real-analytic on $\Delta^{N_k}_{r_k}$, that $\omega^{(k)}$ is (γ, N_k)-Diophantine and denote N_k, U_k, $\bar{U}_k$, ρ_k, δ_k, $\bar{\delta}_k$, β_k the (obviously) corresponding objects (see Proposition 7.1; note that C, γ, ξ_0 remain fixed in the construction). Finally, let, for $0 < \sigma \leq 3\sqrt{5}/\pi^2 (\equiv [2 \sum_{k \geq 1} k^{-4}]^{-1/2})$,

$$\Omega_k \equiv \max\left\{4\sqrt{N_k}\, |\omega_k|, \frac{\delta_k (1+k)^2}{\gamma\sigma}\right\} \tag{8.8}$$

and choose

$$\alpha_k \in \hat{\mathscr{A}}_k \equiv [\Omega_k, \Omega_k + K\, |\omega_k|] \cap \mathscr{A}_{N_k}(\omega_k, \Omega_k) \tag{8.9}$$

(recall from (5.5) that $\hat{\mathscr{A}}_k$ has positive Lebesgue measure). From (8.8) it follows that

$$\sum_{k \geq 0} \frac{\bar{\delta}_k}{(\gamma\Omega_k)^2} \equiv \sum_{k \geq 0} \left(\frac{\delta_k}{\gamma\Omega_k}\right)^2 (\xi_0 - \rho_k) < \sum_{k \geq 0} \frac{\delta_k}{\gamma\Omega_k} \leq \sigma \frac{\pi^2}{6}. \tag{8.10}$$

Under the above positions, Proposition 7.1 guarantees the existence of a non-degenerate quasi-periodic solution with $u^{(k+1)}: \mathbb{R}^{N_{k+1}} \to \mathbb{T}^{N_{k+1}}$ (in Proposition 7.1 we set $u \equiv u^{(k)}$, $u' \equiv u^{(k+1)}$, $r \equiv r_k$, $r' \equiv r_{k+1}$, etc.) and $\omega_{k+1} \equiv (\omega_k, \alpha_k)$; note in fact that the first of (8.10) implies that (recall that $\rho_0 \equiv \xi_0/2$):

$$\rho_{k+1} \equiv \rho_k + \frac{\bar{\delta}_k}{(\gamma \Omega_k)^2} < \frac{\xi_0}{2} + \xi_0 \sigma^2 \sum_{k \geq 1} \frac{1}{k^4} \leq \xi_0, \qquad \forall k. \tag{8.11}$$

Now, if we denote by L either the identity operator, $\partial_{\theta^{(k+1)}}$, $[\theta^{(k+1)} \equiv (\theta_1, ..., \theta_{N_{k+1}})]$, or $\gamma D_{\omega_{k+1}}$, then the bounds in (7.10) and (8.10) yield:

$$\begin{aligned} \|Lu^{(k+1)}\|_{r_{k+1}} &\leq \|Lu^{(0)}\| + \sum_{h=0}^{k} \|L(u^{(h+1)} - (u^{(h)}, 0))\|_{r_{k+1}} \\ &\leq \|Lu^{(0)}\|_{r_0} + \sum_{k \geq 0} \frac{\delta_k}{\gamma \Omega_k} \\ &\leq \|Lu^{(0)}\|_{r_0} + \sigma \frac{\pi^2}{6}. \end{aligned} \tag{8.12}$$

This bound implies, in particular, that, for any $h > 0$,

$$\lim_{k \nearrow \infty} \sup_{\mathbb{T}^{N_{h+k}}} \|L[u^{(h+k)} - (u^{(k)}, \underbrace{0, ..., 0}_{h \text{ times}})]\| = 0 \tag{8.13}$$

(recall that $\|a\| \equiv \sum_{h=1}^{N} |a_h|$ for $a \in \mathbb{R}^N$). Thus we can *define two functions*, u, Du, on $\mathcal{T}_w$ as the uniform limits of, respectively, $u^{(k)}$, $D_{\omega_k} u^{(k)}$: here Du is just a symbol for a function and D must not be interpreted as a differential operator. The functions u, Du are continuous from $\mathcal{T}_w$ into $\mathcal{B}_w$ *for any weight* w. In fact, if $\theta_h \in \mathcal{T}_w$ converges (in the metric ρ_w) to θ, and if $\tilde{u}$ denotes here either u or Du, letting $\bar{w}$ denote $\sup_i w_i$, we see that

$$\begin{aligned} \frac{1}{\bar{w}} \sum_{i \in \mathbb{Z}^d} |\tilde{u}_i(\theta_h) - \tilde{u}_i(\theta)| \, w_i &\leq \sum_{i \in \mathbb{Z}^d} |\tilde{u}_i(\theta_h) - \tilde{u}_i(\theta)| \\ &\leq \sum_{i \in \mathbb{Z}^d} |\tilde{u}_i(\theta_h) - \tilde{u}_i^{(k)}(\theta_h)| + \sum_{i \in I_k} |\tilde{u}_i^{(k)}(\theta_h) - \tilde{u}_i^{(k)}(\theta)| \\ &\quad + \sum_{i \in \mathbb{Z}^d} |\tilde{u}_i^{(k)}(\theta) - \tilde{u}_i(\theta)|, \end{aligned} \tag{8.14}$$

a quantity that can be made as small as we please by taking, first, k large enough [recall (8.13)] and, then, considering h large enough. Define also, for each $i \in \mathbb{Z}^d$, a function $D^2 u_i$ as:

$$D^2 u_i(\theta) \equiv f_i(\theta + u(\theta)), \qquad \theta \in \mathcal{T}_w. \tag{8.15}$$

As above $D^2 u_i$ is just a symbol for a new function; note that the functions $D^2 u_i$ are continuous from $\mathcal{T}_w$ into $\mathbb{R}$ (but we do not consider the vector-valued function $\{D^2 u_i\}_{i \in \mathbb{Z}^d}$). Finally define, for any $\theta \in \mathcal{T}_w$,

$$x(t) \equiv x(t;\theta) \equiv [\theta + \omega t + u([\theta + \omega t])], \tag{8.16}$$

where

$$\omega_i \equiv \begin{cases} \omega_i^{(0)}, & i \in I_0 \\ \omega_{j_k} \equiv \alpha_k, & i \equiv j_k \notin I_0. \end{cases} \tag{8.17}$$

From the above construction follow immediately the following facts:

$$\text{(i)} \quad \frac{d}{dt} x_i(t) = \omega_i + D u_i([\theta + \omega t]), \qquad \frac{d^2}{dt^2} x_i(t) = D^2 u_i([\theta + \omega t]); \tag{8.18}$$

(ii) $x(t;\theta)$ is a solution of (2.6); and (iii) denoting, for $i \in I_0$, $x_i^{(0)}(t;\theta) \equiv \theta_i + \omega_i^{(0)} t + u_i^{(0)}(\theta^{(I_0)} + \omega^{(0)} t)$ [see (7.10), (8.12), (8.14)]:

$$\begin{aligned} &\sup_{\substack{t \in \mathbb{R} \\ p = 0, 1}} \left| \frac{d^p}{dt^p} (x_i(t;\theta) - x_i^{(0)}(t;\theta)) \right| \leqslant \sigma \frac{\pi^2}{6}, \qquad i \in I_0 \\ &\sup_{\substack{t \in \mathbb{R} \\ p = 0, 1}} \left| \frac{d^p}{dt^p} (x_i(t;\theta) - \omega_i t) \right| \leqslant \sigma \frac{\pi^2}{6}, \qquad i \notin I_0. \end{aligned} \tag{8.19}$$

The proof of Theorem 6.3 is completed by taking $\sigma \equiv \min\{6\varepsilon/\pi^2, 3\sqrt{5}/\pi^2\}$. ∎

Remark 8.1. (i) The above estimates imply that (choosing v in (8.7) close enough to 1) $\Omega_k \geqslant (\varepsilon\gamma)^{-1} k!^9$, for all k big enough (the exponent 9 comes from the fact that $\gamma\Omega_k \geqslant \delta_k(1+k)^2/\sigma > \delta_k \varepsilon^{-1}$, from the 6th power of the factorial in (7.9) and from the factor $(r - r')^{-[11(N+1)+1]} \sim k!^{11v}$). It is also easy to see that, for the examples (1.5) and (1.10) discussed in Section 1, it holds the upper bound $\Omega_k \leqslant (\varepsilon\gamma)^{-1} b^k k!^9$ for a suitable constant b (depending also on $x^{(0)}(t)$; see [Pe]); thus, in such examples

$$(\varepsilon\gamma)^{-1} k!^9 < \omega_{j_{k+1}} < (\varepsilon\gamma)^{-1} \bar{b}^k k!^9 \tag{8.20}$$

for k large enough and for a suitable constant $\bar{b}$ (recall that $\varepsilon > 0$ is arbitrary). As already noted, this fast growth is intimately related to the property of ω of being a Diophantine sequence; (obviously) 9 is far from optimal.

(ii) The regularity properties of the almost-periodic solutions, $x(t)$, constructed above are much stronger than just being continuous and having $C^2(\mathbb{R})$ components $x_i(t)$ (such properties are best reflected by the approximations via the real-analytic functions $u^{(k)}$). Note, however, that *$x_i(t)$ is not C^3*.

APPENDIX 1. Full Measure of Diophantine Vectors in $\mathbb{R}^N$

Let $\mathscr{D}_\tau$ denote the set of vectors ω in $\mathbb{R}^N$ which are (γ, τ)-Diophantine [see (5.3)] for some $\gamma > 0$ and let l be Lebesgue measure.

Proposition A1.1. *$l(\mathbb{R}^N \setminus \mathscr{D}_\tau) = 0$ provided $\tau > N-1$.*

Proof. It is enough to check that $\bar{\mathscr{D}}_{R,\tau} \equiv \{\omega : \|\omega\| \leqslant R,\ \omega \notin \mathscr{D}_\tau\}$ has Lebesgue measure zero for any $R > 0$. Now, if

$$\mathbb{C}_{R,\gamma} \equiv \left\{\omega : \|\omega\| \leqslant R \text{ and } \exists 0 \neq n \in \mathbb{Z}^N : |\omega \cdot n| < \frac{1}{\gamma\, |n|^\tau}\right\} \tag{A1.1}$$

(recall that $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^N$ while $|\cdot|$ is the sum of absolute values), then $\bar{\mathscr{D}}_{R,\tau} = \bigcap_{\gamma > 0} \mathbb{C}_{R,\gamma}$ and the claim follows from the estimate

$$\begin{aligned} l(\mathbb{C}_{R,\gamma}) &\leqslant \sum_{n \neq 0} l\left(\left\{\omega : \|\omega\| \leqslant R,\ |\omega \cdot n| < \frac{1}{\gamma\, |n|^\tau}\right\}\right) \\ &\leqslant \sum_{n \neq 0} \frac{b_1 R^{N-1}}{\gamma\, |n|^\tau\, \|n\|} < \frac{b_2 R^{N-1}}{\gamma}, \end{aligned} \tag{A1.2}$$

where b_1, b_2 are suitable (N, τ-dependent) positive constants. ∎

APPENDIX 2. A Short KAM Theory

Here we want to prove in all details Lemma 7.2.

As usual, for vectors in $\mathbb{C}^N$ (or $\mathbb{R}^N$ or $\mathbb{Z}^N$) we denote by $\|\cdot\|$ the Euclidean norm and by $|\cdot|$ the 1-norm (sum of absolute values of components); recall also the definition of norms of analytic functions given in Section 7 [(7.5)–(7.7)].

We start with a basic (elementary) tool (a "Cauchy estimate"):

LEMMA A2.1. *Let g be an analytic function from $\mathscr{D} \subset \mathbb{C}^N \to \mathbb{C}$ ($\partial\mathscr{D}$ smooth). Then for any subdomain $\mathscr{D}' \subset \mathscr{D}$ with $\delta \equiv \operatorname{dist}(\mathscr{D}', \partial\mathscr{D}) > 0$ and any $n \in \mathbb{N}^N$, one has:*

$$\|\partial_z^n g\|_{\mathscr{D}'} \equiv \sup_{\mathscr{D}'} \left| \frac{\partial^{|n|} g}{\partial_{z_1}^{n_1} \cdots \partial_{z_N}^{n_N}} \right| \leqslant |n|! \, \delta^{-|n|} \, \|g\|_{\mathscr{D}}. \tag{A2.1}$$

If g is analytic from $\mathscr{D}$ into $\mathscr{L}^p(\mathbb{C}^N)$, $p \in \mathbb{N}$, then for any $q \in \mathbb{Z}_+$, $\partial_z^q g \in \mathscr{L}^{p+q}(\mathbb{C}^N)$ and

$$\|\partial_z^q g\|_{\mathscr{D}'} \leqslant q! \, \delta^{-q} \, \|g\|_{\mathscr{D}}. \tag{A2.2}$$

The proof is a straightforward exercise based upon Cauchy's integral formula using the contour $|\zeta_h - z_h| = \delta$, $h = 1, ..., N$ ($z \in \mathscr{D}'$ fixed and $\zeta \in \mathscr{D}$ variable of integration. The exercise is carried out, e.g., in [CC2]).

As an application of Lemma A2.1 we prove a useful bound. If $r \in (0, 1)$, $p \in \mathbb{N}$ and $N \in \mathbb{Z}_+$, one has:

$$\Sigma_p^N(r) \equiv \sum_{n \in \mathbb{Z}^N} |n|^p e^{-r|n|} \leqslant p! \left(\frac{4}{r}\right)^{p+N}. \tag{A2.3}$$

Proof of (A2.3). For any $r > 0$

$$\Sigma_p^N = (-1)^p \, \partial_r^p \sum_{n \in \mathbb{Z}^N} e^{-r|n|} = (-1)^p \, \partial_r^p \left(\frac{e^r + 1}{e^r - 1}\right)^N \equiv (-1)^p \, \partial_r^p E(r)^N. \tag{A2.4}$$

Now, the function $g(x) \equiv (-1)^p E(x)^N$ is analytic and bounded in $B_{r-s}(r) \equiv \{x \in \mathbb{C} : |x - r| < r - s\}$, for any $0 < s < r$. Applying Lemma A2.1 with $\mathscr{D}' \equiv B_\varepsilon(r)$ $(0 < \varepsilon < r - s$ arbitrary), $\mathscr{D} \equiv B_{r-s}(r)$ and noting that $\|E\|_{\mathscr{D}} = E(s) < 1 + 2/s$, one gets:

$$\Sigma_p^N < \|\partial_x^p g\|_{B_\varepsilon(r)} \leqslant \frac{p!}{(r - s - \varepsilon)^p} \left(1 + \frac{2}{s}\right)^N. \tag{A2.5}$$

Inequality (A2.3) follows by taking $s = \frac{3}{4} r$ and using the arbitrariness of ε. ∎

An immediate consequence of this bound is:

LEMMA A2.2. *Let $p \in \mathbb{N}$; $N, M \in \mathbb{Z}_+$ and consider a function $g: \mathbb{T}^N \to \mathscr{L}^p(\mathbb{C}^M)$ with zero average and analytic on Δ_ξ^N for some $\xi < 1$. Let $\omega \in \mathbb{R}^N$ be a (γ, N)-Diophantine vector and denote by $D^{-1}g$ the unique solution with* *zero average of the equation* $Df = g$ $(D \equiv \sum_{h=1}^{N} \omega_h \partial_{\theta_h})$. *Then, for any* $0 < \delta < \xi$ *one has*:

$$\|D^{-1}g\|_{\xi-\delta} \leqslant \gamma 2^{4N} N!\, \delta^{-2N} \|g\|_{\xi}. \tag{A2.6}$$

Remark A2.4. Rüßmann [Rü] obtains (with a much more subtle proof) N as an exponent of δ^{-1}.

Proof. Using the standard bound on the Fourier coefficients of analytic functions, $|g_n| \leqslant \|g\|_{\xi} e^{-\xi|n|}$, by definition of (γ, N)-Diophantine and by (A2.3) one sees that:

$$\|D^{-1}g\|_{\xi-\delta} = \left\| \sum_{n\neq 0} \frac{g_n}{i\omega\cdot n} e^{in\cdot\theta} \right\|_{\xi-\delta} \leqslant \|g\|_{\xi} \sum_{n\neq 0} \frac{e^{-\delta|n|}}{|\omega\cdot n|}$$

$$\leqslant \gamma \|g\|_{\xi} \sum_{n\neq 0} |n|^N e^{-\delta|n|} \leqslant \gamma \|g\|_{\xi} 2^{4N} N!\, \delta^{-2N}. \quad ▮ \tag{A2.7}$$

Lemma 7.2 has an immediate corollary, which has been used in deriving Theorem 6.1 from Theorem 6.3:

COROLLARY A2.4. *Let* V *be as in Lemma* 7.2 *and let* ω_0 *be a* (γ, N)-*Diophantine vector. There exists a* $0 < \mu_0 < 1$ *such that, for any* $0 < \mu < \mu_0$, *the equation* $D_\omega^2 u = V_x(\theta + u)$, *with* $\omega \equiv \omega_0/\mu$, *admits a solution.*

Proof. First observe that, since $\mu < 1$, ω is (γ, N)-Diophantine. Now, take $v \equiv D_\omega^{-2} V_x(\theta) \equiv \mu^2 D_{\omega_0}^{-2} V_x(\theta)$. Then, taking, say, $\xi \equiv \xi_0/2$ one has

$$\|e\|_{\xi} \equiv \|D^2 v - V_x(\theta + v)\|_{\xi} = \|V_x(\theta) - V_x(\theta + v)\|_{\xi} \leqslant c\mu^2 \tag{A2.8}$$

for a suitable c (depending on ω_0 and V). It is then obvious that (7.12) can then be achieved by taking μ small enough. ▮

The proof of Lemma 7.2 is based on a "Newton scheme" obtained by iterating the construction of a new "approximate solution," v', (given a starting approximate solution v), producing a "quadratically smaller" error. Such a Newton scheme is summarized in the following

LEMMA A2.5. *Set* $\mathcal{M} \equiv (\mathrm{Id}_N + v_\theta)$, $e \equiv D^2 v - V_x(\theta + v)$ *and let* z *be a solution of*

$$D(\mathcal{M}^T \mathcal{M} Dz) = -\mathcal{M}^T e, \tag{A2.9}$$

where $(\cdot)^T$ *denotes matrix transposition. Then, setting*

$$w \equiv \mathcal{M} z, \qquad v' \equiv v + w, \tag{A2.10}$$

the equation

$$D^2v' - V_x(\theta + v') = e_\theta z + q_1 + q_2 \equiv e' \tag{A2.11}$$

holds, where

$$\begin{aligned} q_1 &\equiv -V_x(\theta + v + w) + V_x(\theta + v) + V_{xx}(\theta + v)\,w \\ q_2 &\equiv (\mathcal{M}^T)^{-1}\mathcal{A}Dz, \qquad \mathcal{A} \equiv \mathcal{M}^T D\mathcal{M} - (D\mathcal{M}^T)\,\mathcal{M}. \end{aligned} \tag{A2.12}$$

Furthermore the matrix-valued function $\mathcal{A}$ *satisfies*

$$D\mathcal{A} = \mathcal{M}^T e_\theta - e_\theta^T \mathcal{M}, \qquad \langle \mathcal{A} \rangle = 0. \tag{A2.13}$$

This is Lemma 1 of [CC1] (with f replaced here by $-V$; see also [M2, SZ]); however, for completeness we sketch the proof.

Proof of Lemma A2.5. First, using the definition of $e(\theta)$, one checks that (A2.9) makes sense (i.e., that the right hand side has zero mean value over $\mathbb{T}^N$). Then, (A2.11) follows easily from the definitions of q_1, q_2 and from

$$D^2v - V_x(\theta + v) = e, \qquad D^2\mathcal{M} - V_{xx}(\theta + v)\,\mathcal{M} = e_\theta, \tag{A2.14}$$

the first equation being the definition of e and the second one being obtained by taking the θ-gradient of the first equation. Finally equation (A2.13) follows easily from the second equation in (A2.14) and from its transposed; the vanishing of the average of $\mathcal{A}$ is checked using the definition of $\mathcal{A}$ and integration by parts. ▮

Remark A2.6. The general solution of (A2.9) depends upon an arbitrary constant $c \in \mathbb{R}^N$, which we shall fix by imposing $\langle w \rangle = 0$ (compare (2.6)–(2.8) of [CC1]).

Proof of Lemma 7.2. The idea is to estimate the objects in Lemma A2.5 and then to iterate the construction infinitely many times generating a sequence of approximations, v_j, ($v_0 \equiv v$, $v_1 \equiv v + w$, ...): Under the "convergence condition" (7.12), the "error function" at the jth step, $e_j \equiv D^2v_j - V_x(\theta + v_j)$, will be quadratically smaller than the preceding error function e_{j-1}, so that $\|e_j\| \searrow 0$ and v_j converges to the solution u.

We start by performing estimates on w, v', e' defined in Lemma A2.5 (but keep in mind that the corresponding estimates at the jth step of the procedure are basically identical, having replaced v with v_j, e with e_j and ξ with $\xi_j \searrow \xi'$ to be suitably defined below).

In the following bounds $1 \leqslant B_1 \leqslant B_2 \leqslant \cdots \leqslant B$ will denote suitable *universal constants.*

Fix $0<\delta<\xi$ and let z be the unique solution of (A2.9) such that $\langle w\rangle=0$ (see the above Remark A2.6); then, using Lemmas A2.1 and A2.2, one obtains, with suitable constants B_h,

$$\begin{aligned}\|Dz\|_{\xi-\delta/2} &\leqslant B_1 2^{6N} N!\,\delta^{-2N}\eta^2\bar{\eta}^4\gamma^{-1}\varepsilon\\ \|z\|_{\xi-\delta} &\leqslant B_2 2^{12N} N!^2\,\delta^{-4N}\eta^4\bar{\eta}^4\varepsilon\\ \|w\|_{\xi-\delta} &\leqslant B_3 2^{12N} N!^2\,\delta^{-4N}\eta^5\bar{\eta}^4\varepsilon\end{aligned} \tag{A2.15}$$

(to check the powers of η and $\bar{\eta}$, one needs to take into account the "constants of integration" coming from solving $Df=g$ with $\langle g\rangle=0$; compare (2.7), (2.8) of [CC1]). Now, observe that (7.12) implies that $\|w\|_{\xi-\delta}<\xi_0-\xi_0'$ so that e' is analytic on $\Delta^N_{\xi-\delta}$ and, using again Lemmas A2.1 and A2.2, one obtains the bounds

$$\begin{aligned}\|e_\theta z\|_{\xi-\delta} &\leqslant B_4 2^{12N} N!^2\,\delta^{-(4N+1)}(\eta\bar{\eta})^4\gamma^{-2}\varepsilon^2\\ \|q_1\|_{\xi-\delta} &\leqslant \|V_{xxx}\|_{\xi_0}\,\|w\|^2_{\xi-\delta}\\ &\leqslant B_5\,\|V_{xxx}\|_{\xi_0}\,2^{24N} N!^4\,\delta^{-8N}\eta^{10}\bar{\eta}^8\varepsilon^2\\ \|\mathscr{A}\|_{\xi-\delta} &\leqslant \|D^{-1}(\mathscr{M}^T e_\theta - e_\theta^T\mathscr{M})\|_{\xi-\delta}\\ &\leqslant B_6 2^{6N} N!\,\delta^{-(2N+1)}\eta\gamma^{-1}\varepsilon\\ \|q_2\|_{\xi-\delta} &\leqslant B_7 2^{10N} N!^2\,\delta^{-(4N+1)}\eta^3\bar{\eta}^5\gamma^{-2}\varepsilon^2,\end{aligned} \tag{A2.16}$$

where to bound $\|\mathscr{A}\|$ we used Lemma A2.2 with δ replaced by $\delta/2$ (so as to be able to bound subsequently $\|e_\theta\|$). Thus, observing that δ^{-1}, η, $\bar{\eta}$ are $\geqslant 1$, we get [see (7.11)]:

$$\varepsilon'\equiv\gamma^2\,\|e'\|_{\xi-\delta}\leqslant B_8 2^{24N} N!^4\,\delta^{-8N}\beta\eta^{10}\bar{\eta}^8\varepsilon^2. \tag{A2.17}$$

On $\Delta^N_{\xi-2\delta}$ we can bound immediately, by Lemma A2.1, the derivatives of w; hence we get:

$$\max\{\|w\|_{\xi-\delta},\,\|w_\theta\|_{\xi-2\delta}\}\leqslant B_3 2^{12N} N!^2\,\delta^{-(4N+1)}\eta^5\bar{\eta}^4\varepsilon. \tag{A2.18}$$

As already mentioned, we iterate now the above construction and relative estimates, setting $w_0\equiv v$, $v_j\equiv\sum_{i=0}^{j} w_i$, $\delta\equiv\delta_0$, and

$$\delta_j\equiv\frac{\xi-\xi'}{2^{j+2}},\qquad \hat{\xi}_j\equiv\hat{\xi}_{j-1}-2\delta_{j-1}\equiv\xi-2\sum_{i=0}^{j-1}\delta_i=\xi'+\frac{\xi-\xi'}{2^j} \tag{A2.19}$$

(note that $\hat{\xi}_0\equiv\xi$ and that $\hat{\xi}_j\searrow\xi'$; the choice of the "analyticity-loses" δ_j is rather arbitrary; however, the choice made here turns out to be particularly simple and also good for "sharp estimates," see [CC1]). We also set

$\eta_j \equiv \|\mathrm{Id}_N + \partial_\theta v_j\|_{\xi_j}$, $\bar{\eta}_j \equiv \|(\mathrm{Id}_N + \partial_\theta v_j)^{-1}\|_{\xi_j}$ and assume inductively that $\forall 0 \leqslant j \leqslant j_0$

$$\eta_j < 2\eta, \qquad \bar{\eta}_j < 2\bar{\eta}, \qquad \sum_{i=1}^{j} \|w_i\|_{\xi_i} < \xi_0 - \xi_0', \tag{A2.20}$$

the last inequality being necessary for the definition (and analyticity) of e_{j+1}. By the (analogous of the) above estimates (where $\varepsilon \to \varepsilon_j$, $\varepsilon' \to \varepsilon_{j+1}$, $\delta \to \delta_j$, $\eta \to \eta_j$, $\bar{\eta} \to \bar{\eta}_j$, $\xi \to \hat{\xi}_j$, $\xi - 2\delta \to \hat{\xi}_{j+1}$), we see that, setting

$$F \equiv B_9 2^{40N} N!^4 (\xi - \xi')^{-8N} \eta^{10} \bar{\eta}^8 \beta, \qquad G \equiv 2^{8N}, \tag{A2.21}$$

we obtain, for $j \leqslant j_0$,

$$\begin{aligned} \varepsilon_{j+1} &\equiv \gamma^2 \|e_{j+1}\|_{\xi_{j+1}} \equiv \gamma^2 \|D^2 v_j - V_x(\theta + v_j)\|_{\xi_{j+1}} \leqslant FG^j \varepsilon_j^2 \\ &\leqslant \varepsilon^{2^{j+1}} \prod_{i=0}^{j} (FG^{j-i})^{2^i} \\ &= [\varepsilon F^{(\sum_{i=1}^{j+1} 2^{-i})} G^{(\sum_{i=1}^{j+1} (i-1) 2^{-i})}]^{2^{j+1}} < (\varepsilon FG)^{2^{j+1}} \end{aligned} \tag{A2.22}$$

(note that both sums in the last line of the above formula converge to one as $j \to \infty$). Analogously,

$$\begin{aligned} \max\{\|w_{j+1}\|_{\xi_j - \delta_j}, \|\partial_\theta w_{j+1}\|_{\xi_{j+1}}\} &\leqslant B_{10} 2^{12N} N!^2 \left(\frac{2^{j+2}}{\xi - \xi'}\right)^{4N} \eta^5 \bar{\eta}^4 \varepsilon_j \\ &\equiv F_0 G_0^j \varepsilon_j \leqslant (\varepsilon FG)^{2^j}, \end{aligned} \tag{A2.23}$$

having used that $F_0 \leqslant F$, $G_0 \leqslant G$ (and that $\sum_{i \geqslant j+1} 2^{-i} = 2^{-j}$, $\sum_{i \geqslant j+1} (i-1) 2^{-i} = (j+1) 2^{-j}$). We check now the induction hypotheses (A2.20) for $j = j_0 + 1$. We find:

$$\eta_{j_0+1} \leqslant \eta + \sum_{j=0}^{j_0} \|\partial_\theta w_j\|_{\xi_j}, \qquad \bar{\eta}_{j_0+1} \leqslant \bar{\eta} \left(1 - \bar{\eta} \sum_{j=1}^{j_0} \|\partial_\theta w_j\|_{\xi_j}\right)^{-1}. \tag{A2.24}$$

Thus (recalling that $\xi_0 < 1$), all we need to complete the check is

$$\sum_{j=0}^{\infty} (\varepsilon FG)^{2^j} < \xi_0 - \xi_0', \tag{A2.25}$$

which is easily seen to be implied, for a suitable $B_{11} \geqslant 2B_{10}$, by

$$B_{11} N!^4 2^{40N} (\xi - \xi')^{-8N} \eta^{10} \bar{\eta}^8 \beta (\xi_0 - \xi_0')^{-1} \varepsilon < 1. \tag{A2.26}$$

This shows that the sequences $\partial_\theta^p v_j \equiv \sum_{i=0}^{j} \partial_\theta^p w_i$ $(p = 0, 1)$ converge uniformly on $\Delta_{\xi'}^N$ and, for $p = 0$, v_j converge to a function u, which satisfies (by construction) (7.13). Let us now estimate $\|D^p u\|_{\xi'}$. First, note that w_{j+1} and e_{j+1} are analytic and bounded on $\xi_j - \delta_j$ so that $D^2 w_{j+1}$, which is equal to $V_x(\theta + v_j + w_{j+1}) - V_x(\theta + v_j) + e_{j+1} - e_j$, can be bounded by

$$\gamma^2 \|D^2 w_{j+1}\|_{\xi_j - \delta_j} \leqslant B_{12} \beta 2^{12N} N!^2 \delta_j^{-4N} \eta^5 \bar{\eta}^4 \varepsilon_j, \tag{A2.27}$$

having used that $\eta_j < 2\eta$, $\bar{\eta}_j < 2\bar{\eta}$ and that $B_8 2^{12N} N!^2 \delta_j^{-4N} \eta^5 \bar{\eta}^4 \varepsilon_j \leqslant 1$. Then, by Lemma A2.2, it follows that

$$\gamma \|Dw_j\|_{\xi_j - 2\delta_j} = \gamma \|D^{-1} D^2 w_j\|_{\xi_{j+1}} \leqslant B_{12} \beta 2^{16N} N!^3 \delta_j^{-6N} \eta^5 \bar{\eta}^4 \varepsilon_j. \tag{A2.28}$$

The bounds (7.14) now follow easily. ∎

Acknowledgments

We are deeply indebted to C. Liverani for many helpful discussions and especially for suggesting the right form of the "approximate solution" in Proposition 7.1. We are also grateful to C. E. Wayne for useful remarks on a previous version of this paper. This work has been partially supported by Consiglio Nazionale delle Ricerche.

References

[A] V. I. Arnold (Ed.), "Dynamical Systems, III," Encyclopaedia of Mathematical Sciences, Vol. 3, Springer-Verlag, Berlin/New York, 1988.

[AN] V. I. Arnold and S. P. Novikov (Eds.), "Dynamical Systems, IV," Encyclopaedia of Mathematical Sciences, Vol. 4, Springer-Verlag, Berlin/New York, 1990.

[CC1] A. Celletti and L. Chierchia, Construction of analytic KAM surfaces and effective stability bounds, *Comm. Math. Phys.* **118** (1988), 119–161.

[CC2] A. Celletti and L. Chierchia, "A Constructive Theory of Lagrangian Tori and Computer-Assisted Applications," Carr Report, n. 27/90, II University of Rome "Tor Vergata," December, 1990.

[CZ] L. Chierchia and E. Zehnder, Asymptotic expansion of quasiperiodic solutions, *Ann. Scuola Norm. Sup. Pisa Cl. Sci. 4* **XVI** (1989), 245–258.

[FSW] J. Frölich, T. Spencer, and C. E. Wayne, Localization in disordered nonlinear dynamical systems, *J. Stat. Phys.* **42** (1986), 247–274.

[G] G. Gallavotti, "The Elements of Mechanics," Springer-Verlag, Berlin/New York, 1983.

[CW] W. Craig and C. E. Wayne, Newton's method and periodic solutions of nonlinear wave equations, preprint, 1992.

[Ha] P. R. Halmos, "Measure Theory," Springer-Verlag, Berlin/New York, 1974.

[Ka] Y. Katznelson, "An Introduction to Harmonic Analysis," Dover, New York, 1976.

[Ke] J. Kelley, "General Topology," Van Nostrand–Reinhold, Princeton, NJ, 1955.

[Ku] S. B. Kuksin, Perturbation of quasi-periodic solutions of infinite dimensional Hamiltonian systems, *Math. USSR-Izv.* **32**, No. 1 (1989), 39–62.

[M1] J. Moser, A rapidly convergent iteration method and nonlinear partial differential equations, *Ann. Scuola Norm. Sup. Pisa* **20** (1966), 265–315, 499–535.

[M2] J. Moser, Minimal solutions of variational problems on a torus, *Ann. Inst. H. Poincaré, Analyse Nonlinéaire* **3** (1986), 229–272.

[Pö] J. Pöschel, Small divisors with spatial structure in infinite dimensional Hamiltonian systems, *Comm. Math. Phys.* **127** (1990), 351–393.

[Pe] P. Perfetti, "Soluzioni Quasi-periodiche Massimali per Sistemi Lagrangiani Infinito Dimensionali," Ph.D. Thesis, Physics Department, University of Rome "La Sapienza," 1991.

[Rü] H. Rüßmann, Konvergente Reihenentwicklungen in der Störungstheorie der Himmelmechanik, *Selecta Mathematica V, Heidelberger Taschenbücher* **201** (1979), 93–260.

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

[VB] M. Vittot and J. Bellissard, "Invariant Tori for an Infinite Lattice of Coupled Rotators," Report Centre de Physique Théorique CNRS–Luminy, Marseille, 1985.

[Wa1] C. E. Wayne, The KAM theory of systems with short range interactions, I and II, *Comm. Math. Phys.* **96** (1984), 311–329, 331–344.

[Wa2] C. E. Wayne, Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory, *Comm. Math. Phys.* **127** (1990), 479–528.