KAM Lectures*

by

Luigi Chierchia

Department of Mathematics, Università ROMA TRE

Largo San L. Murialdo 1, I-00146 Roma, Italy

The aim of these lectures is to present, in a self contained way, the fundamentals of *KAM theory*, which, as well known, deals with the problem of constructing *quasi-periodic motions* in real-analytic or smooth conservative dynamical systems.

KAM theory is based upon quantitative techniques designed to overcome the so-called *small denominator* difficulties arising in the construction of quasi-periodic motions and works under rather stringent smallness and regularity assumptions.

For sake of presentation, we will consider only *second order Hamiltonian systems* with a finite number of degrees of freedom (periodic in the "space" variables), i.e., systems governed by Hamiltonian functions of of the form

$$H(y,x) = \frac{y^2}{2} + V(x) \ , \tag{1}$$

where y and x are *standard symplectic variables* $(y,x) \in \mathbb{R}^d \times \mathbb{T}^d$, and $V : \mathbb{T}^d \to \mathbb{R}$ is a (multi-periodic) smooth or real-analytic function; $y^2 := y \cdot y := \sum_{j=1}^d y_j^2$. Here, $\mathbb{T}^d$ denotes the standard flat d-torus $\mathbb{T}^d := \mathbb{R}^d/(2\pi\mathbb{Z}^d)$; the (standard) symplectic structure is: $dy \wedge dx = \sum_{j=1}^d dy_j \wedge dx_j$ and the Hamilton equations are

$$\dot y = -H_x \ , \qquad \dot x = H_y \ , \tag{2}$$

*Published in "Dynamical Systems. Part I: Hamiltonian Systems and Celestial Mechanics", Pubblicazioni della Classe di Scienze, Scuola Normale Superiore, Pisa. 1-56. Centro di Ricerca Matematica "Ennio De Giorgi" : Proceedings. (2003)

where H_y denotes the y-gradient $(H_{y_1}, ..., H_{y_d})$ and H_x denotes the x-gradient $(H_{x_1}, ..., H_{x_d})$; dot denotes time derivative.

The point of view taken up in these lectures is that of *non-linear functional analysis*, as we briefly proceed to explain. The problem of constructing (maximal[1]) quasi-periodic solutions is essentially equivalent to solve a non-linear partial differential equation on $\mathbb{T}^d$, $\mathcal{E}(u) = 0$, with real-analytic or C^k coefficients. If one is given an *approximate solution*, i.e. a function v for which $\mathcal{E}(v)$ is not zero but *small* (in suitable norms), then, under suitable conditions, it is possible to find a true near-by solution. The method we shall follow is based on a *Newton ("quadratic") scheme*, which allow to construct a sequence of better and better approximations (living in larger and larger Banach spaces) converging to a true solution. The loss of regularity (related to the inversion of non elliptic differential operators and to the above mentioned small denominator problems) arising in solving the associated linearized equation is overcome by the speed of convergence of the scheme.

The approach presented here - sometimes referred to as *KAM theory in configuration space* - avoids completely the use of symplectic transformations and needs less preparation than standard KAM theory.

The notes of the lectures are divided in two chapters:

In the first chapter a KAM theorem establishing the existence of quasi-periodic solutions (with prescribed "diophantine" frequencies), in real-analytic setting, is presented. The "potential" V in (1) is *not* assumed to be small; what allows to start up the perturbative procedure is the existence of a good enough *approximate solution*.

While no effort is put in trying to get "optimal estimates", a certain care is devoted to perform explicit estimates and also to discuss convenient norms (Fourier and complex sup-norms).

In the second chapter, we shall consider Hamiltonians H in (1) with $V \in C^l(\mathbb{T}^d)$, *which shall be assumed to be small in C^1 norm*. Then, assuming l big enough and using the approximation technique due to Bernstein, Jackson, Moser and Zehnder[2], we shall construct (using the real-analytic KAM theorem of the first chapter) a sequence of real-analytic approximate solutions converging to C^s quasi-periodic solutions; explicit estimates on l and s will be given.

[1]I.e., quasi periodic-solutions with d independent frequencies; for the definition of quasi-periodic solutions, see below.

[2]Such technique gives precise hypotheses in order to approximate C^l functions with real-analytic ones and, viceversa, to get C^l functions out of limits of real-analytic sequences.

The main references are:

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

[2] D. Salamon, *The Kolmogorov-Arnold-Moser theorem*, FIM-Preprint, ETH-Zurich, (1986), available on
http://www.math.ethz.ch/~salamon/PREPRINTS/KAM.htm

For the analytic part, see also:

[3] A. Celletti and L. Chierchia: *A constructive theory of Lagrangian tori and computer–assisted applications*, Dynamics reported, 60–130, Dynam. Report. Expositions Dynam. Systems (N.S.), 4, Springer, Berlin, 1995

Acknowledgments These lectures were delivered at SISSA/ISAS, Trieste (February 2001) and at SNS "Centro di Ricerca Matematica E. De Giorgi", Pisa (March 2002): I am grateful to, respectively, A. Ambrosetti and S. Marmi for their kind invitation. The students in Trieste and Pisa contributed non-trivially and in a stimulating way to the form and organization of these notes; in particular, I am indebted with N. Chtcherbakova, U. Serres and M. Sigalotti, who wrote a first draft of these notes, and with A. Sorrentino for various comments on the second chapter (obviously, responsibility for any mistake relies entirely upon the author). This work has been partially supported by the MIUR national project "Variational Methods and Nonlinear Differential Equations".

Contents

1 Analytic KAM Theory

1.1 Warm up: Newton scheme for the standard IFT

The aim of this section is to discuss a proof of the (standard) Implicit Function Theorem in $\mathbb{R}^n$ based on the "Newton method" with the purpose of illustrating, in a trivial case, the scheme of proof that we shall use to construct quasi-periodic motions for Hamiltonian systems.

Let $(\bar{y}, \bar{x}) \in \mathbb{R}^n \times \mathbb{R}^m$; denote by D^n_ρ the closed ball in $\mathbb{R}^n$ centered at $\bar{y}$ with radius ρ and by $D^m_{\bar{r}}$ the closed ball in $\mathbb{R}^m$ centered at $\bar{x}$ with radius $\bar{r}$; let

$X_{r,\rho}$ denote the Banach space, $C(D^m_r, D^n_\rho)$, of continuous function from D^m_r into D^n_ρ endowed with the sup-norm.

Theorem 1 *Let $F \in C(D^n_\rho \times D^m_{\bar r}, \mathbb{R}^n)$ be such that $y \mapsto F(y,x) \in C^2(D^n_\rho, \mathbb{R}^n)$ for all $x \in D^m_{\bar r}$ with F_y invertible on $D^n_\rho \times D^m_{\bar r}$. Let α and β be positive numbers such that*

$$\left\|(F_y)^{-1}\right\|_{\rho,\bar r} \le \alpha\ , \qquad \frac{\alpha^2}{2}\,\|F_{yy}\|_{\rho,\bar r} \le \beta\ , \tag{3}$$

$\|\cdot\|_{\rho,\bar r}$ being short for $\sup_{D^n_\rho \times D^m_{\bar r}} |\cdot|$. Suppose that, for some $0 < r \le \bar r$ and $0 < \sigma < 1$, there exists $u_0 \in X_{r,\rho}$ such that $\|u_0 - \bar y\|_r := \sup_{D^m_r} |u_0 - \bar y| < \rho$ and:

$$\|F(u_0(x),x)\|_r \le \min\left\{\frac{\sigma}{\beta}, \frac{1-\sigma}{\alpha}\,(\rho - \|u_0-\bar y\|_r)\right\}\ . \tag{4}$$

Then, there exists a unique function $u \in X_{r,\rho}$ such that:

$$F(u(x),x) = 0\ , \qquad \forall\ x \in D^m_r, \tag{5}$$

and

$$\|u - u_0\|_r \le \frac{\alpha}{1-\sigma}\|F(u_0(x),x)\|_r\ .$$

Remark 1 (i) The limiting case $\beta = 0$ corresponds to the linear case

$$F(y,x) = a(x) + A(x)y$$

(with A invertible), in which case the solution of $F(u,x) = 0$ is simply $u = -A^{-1}a$.

(ii) If $F(\bar y, \bar x) = 0$, one can obviously take $u_0(x) \equiv \bar y$ (choosing suitably r so as to meet condition (4)).

(iii) The function u_0 is called *an approximate solution* of (5); the function

$$\varepsilon_0(x) := F(u_0(x),x)\ , \tag{6}$$

is the associated *error function*. The inequality (4) should be interpreted as a *smallness condition* on the error function and the IFT can be rephrased by saying that *if the smallness condition is verified by the error function ε_0 associated to the approximate solution u_0, than there exists a (unique) true solution u, which is $\|\varepsilon_0\|$-close to the approximate solution u_0.*

Proof We first show how to construct out of u_0 a new approximate solution u_1 for which the associated error function $\varepsilon_1(x) := F(u_1(x), x)$ is quadratically smaller that ε_0.

Let ε_0 be as in (6) and define

$$w_0(x) = -\Big(F_y(u_0(x), x)\Big)^{-1}\varepsilon_0(x) \ , \qquad u_1 := u_0 + w_0 \ . \tag{7}$$

We claim that $u_1 \in X_{r,\rho}$ and that w_0 and $\varepsilon_1 := F(u_1, x)$ verify:

$$\|w_0\|_r \le \alpha\|\varepsilon_0\|_r \ , \qquad \|\varepsilon_1\|_r \le \beta\|\varepsilon_0\|_r^2 \ . \tag{8}$$

In fact, the first estimate in (8) is immediate consequence of the definitions of w_0 and α. To show that $u_1 : D_r^m \to D_\rho^n$, we compute:

$$\begin{aligned} \|u_1 - \bar y\|_r &:= \|u_0 + w_0 - \bar y\|_r \le \|u_0 - \bar y\|_r + \|w_0\|_r \\ &\le \|u_0 - \bar y\|_r + \alpha\|\varepsilon_0\|_r \le \|u_0 - \bar y\|_r + \alpha\frac{1-\sigma}{\alpha}\Big(\rho - \|u_0 - \bar y\|_r\Big) \\ &< \rho \ , \end{aligned}$$

where we have used the assumption (4) on $\varepsilon_0 := F(u_0, x)$. Observe that, by the definition of ε_1, w_0 and Taylor's formula, one gets:

$$\begin{aligned} \varepsilon_1 &:= F(u_1, x) := F(u_0 + w_0, x) = F(u_0, x) + F_y(u_0, x)w_0 + Q \\ &= \varepsilon_0 + F_y(u_0, x)w_0 + Q = Q \end{aligned} \tag{9}$$

where[3]

$$Q = \int_0^1 (1-t)F_{yy}(u_0 + tw_0, x)w_0 \cdot w_0 dt \ . \tag{10}$$

Thus, by the estimates on w_0 in (8) and the definition of β, we get

$$\|\varepsilon_1\|_r = \|Q\|_r \le \frac{1}{2}\|F_{yy}\|_{\rho,\bar r} \ \|w_0\|_r^2 \le \frac{\alpha^2}{2} \ \|F_{yy}\|_{\rho,\bar r} \ \|\varepsilon_0\|_r^2 \le \beta\|w_0\|_r^2 \ , \tag{11}$$

completing the proof of (8).

The idea is, now, to iterate such construction: Fix $k \ge 2$ and assume that u_1,...,u_{k-1} are given approximate solutions belonging to the Banach space $X_{r,\rho}$ and such that, if one defines

$$w_j := u_{j+1} - u_j \ , \qquad \varepsilon_j(x) := F(u_j(x), x) \ , \qquad \Big(0 \le j \le k-2\Big) \ , \tag{12}$$

[3] "$\cdot$" denotes, here, the standard inner product.

then the following inequalities hold for all $0 \le j \le k-2$:

$$\|w_j\|_r \le \alpha \|\varepsilon_j\|_r \ , \qquad \|\varepsilon_{j+1}\|_r \le \beta \|\varepsilon_j\|_r^2 \ . \tag{13}$$

Note that such inductive assumption has been verified for $k = 2$ with u_1 as in (7).

We claim that, under the inductive assumption (12) and (13), setting

$$\varepsilon_{k-1}(x) := F(u_{k-1}(x), x) \ , \qquad w_{k-1}(x) = -F_y(u_{k-1}(x), x)\varepsilon_{k-1}(x) \ , \tag{14}$$

then one has

$$u_k := u_{k-1} + w_{k-1} \in X_{r,\rho} \ , \tag{15}$$

and (13) holds also for $j = k-1$.

In fact, the estimate on $\|w_{k-1}\|_r$ follows at once (as above) from the definition of w_{k-1} and α (and the inductive assumption on u_{k-1}). Let us, now, show (15). Multiplying by β the second relation in (13) can be rewritten as

$$\beta \|\varepsilon_{j+1}\|_r \le (\beta \|\varepsilon_j\|_r)^2 \ , \tag{16}$$

which iterated leads to

$$\beta \|\varepsilon_j\|_r \le (\beta \|\varepsilon_0\|_r)^{2^j} \ , \qquad \forall \ 0 \le j \le k-1 \ . \tag{17}$$

Thus, by (12), (13) (first inequality), (17) and (4), one has

$$\begin{aligned}
\|u_k - \bar{y}\|_r &= \Big\| u_0 + \sum_{j=0}^{k-1} w_j - \bar{y} \Big\|_r \\
&\le \|u_0 - \bar{y}\|_r + \sum_{j=0}^{k-1} \|w_j\|_r \\
&\le \|u_0 - \bar{y}\|_r + \alpha \sum_{j=0}^{k-1} \|\varepsilon_j\|_r \\
&\le \|u_0 - \bar{y}\|_r + \frac{\alpha}{\beta} \sum_{j=0}^{k-1} (\beta \|\varepsilon_0\|_r)^{2^j} \\
&\le \|u_0 - \bar{y}\|_r + \frac{\alpha}{\beta} \sum_{j=1}^{\infty} (\beta \|\varepsilon_0\|_r)^{j}
\end{aligned}$$

$$\begin{aligned}
&= \|u_0 - \bar{y}\|_r + \alpha \frac{\|\varepsilon_0\|_r}{1 - \beta\|\varepsilon_0\|_r} \\
&\le \|u_0 - \bar{y}\|_r + \alpha \frac{\|\varepsilon_0\|_r}{1 - \sigma} \\
&\le \|u_0 - \bar{y}\|_r + \alpha \frac{1-\sigma}{\alpha} \frac{\rho - \|u_0 - \bar{y}\|_r}{1-\sigma} = \rho \ .
\end{aligned}$$

This shows (15). At this point, also the estimate on $\|\varepsilon_k\|_r$ follows: just replace ε_1, u_0 and w_0 in (9)÷(11) by, respectively, ε_k, u_{k-1} and w_{k-1}.

Thus, thanks to (4), the construction can be iterated indefinitely and $\{u_k\}$ will converge to a function $u \in X_{r,\rho}$. Clearly, since $\|\varepsilon_j\|_r \to 0$ (super-exponentially fast), one has

$$F(u, x) = \lim F(u_k, x) = \lim \varepsilon_k = 0 \ ,$$

showing (5).

Uniqueness is an obvious consequence of the invertibility of F_y. ∎

Remark 2 (i) The approximate solutions u_k's belong to the *same* Banach space $X_{r,\rho}$. This is so because w_{k-1} belongs to the same space of u_{k-1}. In the more complicate case of quasi-periodic solutions for Hamiltonian systems this will not be the case any more: the analogous of F_y^{-1} will be an unbounded operator (involving small divisors) and (the analogous of) w_{k-1} will lie, in general, in a *smaller* Banach space.

(ii) In fact, even formally, it will not be possible to solve the linearized equation[4] exactly but *only* up to quadratically small terms.

(iii) The argument to prove (local) uniqueness in the quasi-periodic case will be different (because of the lack of invertibility of F_y).

1.2 Quasi-periodic solutions (definitions)

Let, as above, $\mathbb{T}^d := \mathbb{R}^d/(2\pi\mathbb{Z}^d)$ be the standard d-dimensional flat torus and let Ω be a bounded domain in $\mathbb{R}^d$. Consider a smooth (say C^2) Hamiltonian

[4] I.e., the equation $F_y(u_0, x)w_0 + \varepsilon_0 = 0$.

$H(y,x)$ from $\Omega \times \mathbb{T}^d$ to $\mathbb{R}$ and the associated Hamiltonian equations

$$\begin{cases} \dot{y}_i = -\frac{\partial H}{\partial x_i} \\ \dot{x}_i = \frac{\partial H}{\partial y_i} \end{cases} \qquad i = 1, \ldots, d. \tag{18}$$

An interesting example is when the system is *nearly-integrable*, i.e., when H is of the form

$$H(y,x) = H_0(y) + \varepsilon H_1(y,x)$$

with ε a small parameter. The corresponding Hamiltonian equations become

$$\begin{cases} \dot{y}_i = -\varepsilon \frac{\partial H_1}{\partial x_i} \\ \dot{x}_i = \frac{\partial H_0}{\partial y_i} + \varepsilon \frac{\partial H_1}{\partial y_i} \end{cases} \qquad i = 1, \ldots, d. \tag{19}$$

When $\varepsilon = 0$ this system is *completely integrable* and all solutions,

$$\begin{cases} y(t) = y(0) \ , \\ x(t) = x(0) + \frac{\partial H_0}{\partial y}(y(0))\, t \ , \qquad (\bmod (2\pi, \ldots, 2\pi)) \ , \end{cases}$$

are *quasi-periodic*:

Definition 1 *A solution* $(y(t), x(t))$ *of* (18) *is said to be quasi-periodic, if there exist a vector* $\omega \in \mathbb{R}^d$ *(frequency vector) and two functions* u, $v \in C^2(\mathbb{T}^d, \mathbb{R}^d)$ *such that*

$$\begin{cases} y(t) = v(\omega t) \\ x(t) = \omega t + u(\omega t) \qquad (\bmod (2\pi, \ldots, 2\pi)). \end{cases} \tag{20}$$

for every t*. If the frequency vector is rationally independent (i.e.*[5] $\omega \cdot n \neq 0$ *for every* $n \in \mathbb{Z}^d \setminus \{0\}$*), then the solution* $(y(t), x(t))$ *is said to be maximal quasi-periodic.*

Remark 3 (i) Non-maximal quasi-periodic solutions include periodic solutions: this is the case when there exist $T > 0$ and $n \in \mathbb{Z}^d$ such that $\omega T = 2\pi n$; notice that in this case there exist $d-1$ linearly independent vectors $n_j \in \mathbb{Z}^d$ such that $\omega \cdot n_j = 0$ for[6] $j = 1, ..., d-1$. More in general, frequencies may be

[5] For vectors $a, b \in \mathbb{R}^d$ we denote $a \cdot b := \sum_{i=1}^d a_i b_i$

[6] In fact, if $\omega T = 2\pi n$ with $T > 0$ and $n \in \mathbb{Z}^d \setminus \{0\}$ then there are *exactly* $(d-1)$ independent vectors n_j s.t. $\omega \cdot n_j = 0$.

classified in terms of the number of relations $\omega \cdot m = 0$ satisfied by ω with *independent* vectors $m \in \mathbb{Z}^d$.

(ii) A maximal quasi-periodic solution is said to be *non-degenerate* if (ω is rationally independent and) the map $\theta \in \mathbb{T}^d \to \theta + u(\theta) \in \mathbb{T}^d$ is a *diffeomorphism* of $\mathbb{T}^d$ so that the map $\theta \in \mathbb{T}^d \to (v(\theta), \theta + u(\theta)) \in \mathbb{R}^d \times \mathbb{T}^d$ yields an embedding of the d-dimensional torus into the phase space $\Omega \times \mathbb{T}^d$. The relation (20) says that *non-degenerate maximal quasi-periodic solutions correspond to d-dimensional invariant tori on which the H-flow is conjugate to the linear flow $\theta \to \theta + \omega t$.*

(iii) In these lectures *we shall consider only non-degenerate maximal quasi-periodic solutions* and, hereafter, "quasi-periodic solution" will be used as synonymous of "non-degenerate maximal quasi-periodic solutions". In particular, the frequency vector ω is always assumed to be rationally independent.

Consider a quasi-periodic solution $(y(t), x(t))$ as in (20). Differentiating it with respect to t we get

$$\begin{cases} \dot{y}(t) = Dv(\omega t) \\ \dot{x}(t) = \omega + Du(\omega t)\,, \end{cases}$$

where

$$D := D_\omega := \sum_{i=1}^{d} \omega_i \frac{\partial}{\partial \theta_i}\,.$$

Since (y, x) is a solution of (18), we have

$$\begin{cases} Dv_i(\omega t) = -\frac{\partial H}{\partial x_i}(v(\omega t), \omega t + u(\omega t)) \\ \omega_i + Du_i(\omega t) = \frac{\partial H}{\partial y_i}(v(\omega t), \omega t + u(\omega t)) \end{cases} \qquad i = 1, \ldots, d\,,$$

which, by density of the trajectory $t \mapsto \omega t$ on[7] $\mathbb{T}^d$, are equivalent to

$$\begin{cases} Dv(\theta) = -H_x(v(\theta), \theta + u(\theta)) \\ \omega + Du(\theta) = H_y(v(\theta), \theta + u(\theta)) \end{cases} \qquad \theta \in \mathbb{T}^d. \tag{21}$$

[7] As well known, $\theta \in \mathbb{T}^d \to \omega t \in \mathbb{T}^d$ *is dense if and only if ω is rationally independent*; see, e.g., [V.I. Arnold, *Mathematical methods of classical mechanics*, Springer-Verlag, 1989].

Hereafter, (for simplicity) we shall consider only Hamiltonians H of the form

$$H(y,x) = \frac{y^2}{2} + V(x) := \frac{1}{2}\sum_{j=1}^{d} y_j^2 + V(x) \ .$$

In this special case (18) takes the form

$$\begin{cases} \dot{y} = -V_x \\ \dot{x} = y \end{cases}$$

or, equivalently,

$$\ddot{x} = -V_x.$$

Notice that, in such a case, the second equation in (21) becomes simply

$$v(\theta) = \omega + Du(\theta) \ , \tag{22}$$

so that the system (21) becomes the following single (vector) equation for u:

$$D^2 u(\theta) = -V_x(\theta + u(\theta)) \ . \tag{23}$$

The lectures are devoted to discuss solutions of (23).

It is clear that an important rôle in the study of (21) or (23) is played by the *linear equation*

$$Du = f \ ,$$

with f a given function on $\mathbb{T}^d$. Proceeding formally, we expand both sides in Fourier series getting

$$\sum_{n\in\mathbb{Z}^d} f_n e^{in\cdot\theta} = \sum_{n\in\mathbb{Z}^d} i(\omega\cdot n) u_n e^{in\cdot\theta} \ .$$

Equating Fourier coefficient, we get, for $n = 0$, the compatibility condition[8]

$$f_0 = \langle f \rangle = 0 \ , \tag{24}$$

and, for $n \neq 0$,

$$u_n = \frac{f_n}{i(\omega\cdot n)} \ , \qquad \big(n \in \mathbb{Z}^d \setminus \{0\}\big) \ . \tag{25}$$

[8] We denote $\langle \cdot \rangle := {\int\!\!\!\!\!-}_{\mathbb{T}^d} \cdot \, d\theta := (2\pi)^{-d} \int_{\mathbb{T}^d} \cdot \, d\theta$.

The denominator $(\omega \cdot n)$ in (25), even though never vanishes, might become arbitrarily small making doubtful the convergence of the Fourier series

$$\sum_{n\in\mathbb{Z}^d} u_n e^{in\cdot\theta} \, . \tag{26}$$

Definition 2 *We say that* $\omega \in \mathbb{R}^d$ *is* (γ, τ)*-diophantine if* γ*,* τ *are positive constants such that*

$$|\omega{\cdot}n| \geq \frac{\gamma}{|n|^\tau} \qquad \text{for every} \quad n \in \mathbb{Z}^d \setminus \{0\} \, . \tag{27}$$

Remark 4 For $\tau > d-1$ fixed, the set of diophantine vectors is of full measure (**exercise**). For $\tau < d-1$ (27) is never satisfied (Liouville).

Suppose now that ω is (γ, τ)-diophantine and f is a smooth enough function with vanishing mean value, $\langle f \rangle = 0$. Then (26)-(25) actually define the function u, solution of $Du = f$, up to an additive constant (the average of u); the unique solution of the system:

$$Du = f \, , \qquad \langle u \rangle = 0 \, ,$$

will be denoted by $D^{-1}f$.

Exercise Find a lower bound on k so that if $f \in C^k(\mathbb{T}^d)$ then $D^{-1}f$ has an absolutely convergent Fourier series expansion.

Remark 5 The analysis described in these lectures could be easily extended to the *non-autonomous case*, i.e, the case when the potential $V = V(x,t)$ depends also explicitly (and periodically) on time t, $V : \mathbb{T}^{d+1} \to \mathbb{R}$. In such a case D_ω has to be replaced by

$$D = \sum_{i=1}^{d} \omega_i \frac{\partial}{\partial x_i} + \frac{\partial}{\partial t}$$

with $(\omega, 1) \in \mathbb{R}^{d+1}$ rationally independent and equation (23) becomes

$$D^2 u(\theta, t) = -V_x(\theta + u(\theta, t), t) \, .$$

1.3 Newton scheme for Quasi-periodic solutions

In this section we describe the Newton scheme on which the construction of solutions of the functional equation (23) will be based.

The strategy that we shall follow mimics the proof of Theorem 1: we shall start from an *approximate solution* v *of* (23), i.e., a (smooth) function v such that the associated *error function*

$$\varepsilon := \mathcal{E}(v) := D^2 v + V_x(\theta + v) \ , \tag{28}$$

is "small" and try to construct a "better" approximate solution

$$v' := v + w \ , \tag{29}$$

whose associated error function

$$\varepsilon' := \mathcal{E}(v') := D^2 v' + V_x(\theta + v') \tag{30}$$

is "quadratically smaller" than the error function associated to v.

Remark 6 The discussion in this section will be *algebraic in character* and the necessary estimates will be discussed later (§ 1.5). Therefore, words such as "small" or "quadratically smaller" are used, here, in a somewhat formal way[9]. Roughly speaking, the idea is to look for $w \sim \varepsilon$ (i.e., "of the same order of ε") so that $\mathcal{E}(v+w) \sim \varepsilon^2$. However, as clarified also in Remark 7 below, the reader can also disregard any reference to "smallness" following only the algebraic identities involved.

Define Q_1 as

$$Q_1 := V_x(\theta + v + w) - V_x(\theta + v) - V_{xx}(\theta + v)w. \tag{31}$$

and note that, by Taylor's formula, Q_1 is quadratic in[10] w. Expanding $V_x(\theta + v + w)$ we find:

$$\varepsilon' := \mathcal{E}(v') \ \ := \ \ D^2 v + D^2 w + V_x(\theta + v + w)$$

[9] At an intuitive level, one should think to substitute the error function ε with $\mu\varepsilon$ thinking μ as a small real parameter: the terms appearing with a μ in front will be thought of as "small" terms and of the same "order" of the error function, terms with a μ^2 in front will be thought of as "quadratically smaller terms", etc.

[10] Here and in what follows the symbol Q_j's stand for terms "quadratic in ε".

$$
\begin{aligned}
&= D^2 v + V_x(\theta + v) + D^2 w + V_{xx}(\theta + v) w + Q_1 \\
&=: \mathcal{E}(v) + D^2 w + V_{xx}(\theta + v) w + Q_1 \\
&=: \varepsilon + D^2 w + V_{xx}(\theta + v) w + Q_1 \ . \qquad (32)
\end{aligned}
$$

The perfect analogue of the Newton scheme described in the proof of the standard IFT given in § 1 would consist in finding an "explicit" solution of the the following PDE on $\mathbb{T}^d$

$$\varepsilon + D^2 w + V_{xx}(\theta + v) w = 0 \, . \qquad (33)$$

However this is not so easy and, in fact, *we shall be able to solve* (33) *only up to quadratic terms in* ε.

To proceed further, we look at the variation equation for (23), i.e., the equation

$$\varepsilon_\theta = D^2 v_\theta + V_{xx}(\theta + v)(I + v_\theta) \ , \qquad (34)$$

which is gotten by differentiating with respect to θ the system (23); here, for a given function $u : \mathbb{T}^d \to \mathbb{R}^d$, u_θ denotes the Jacobian matrix

$$u_\theta = \left(\frac{\partial u_i}{\partial \theta_j}(\theta) \right)_{i,j=1,\ldots,d} \ ,$$

and $I := I_d$ denotes the unit $(d \times d)$ matrix.

Setting

$$M := I + v_\theta \qquad (35)$$

we can rewrite (34) in the form

$$\varepsilon_\theta = D^2 M + V_{xx}(\theta + v) M. \qquad (36)$$

Assume that $M(\theta)$ *is invertible for all* $\theta \in \mathbb{T}^d$. From (36) we get

$$V_{xx}(\theta + v) = (\varepsilon_\theta - D^2 M) M^{-1}$$

and plugging this equality in (32), we find

$$
\begin{aligned}
\varepsilon' &= \varepsilon + D^2 w + (\varepsilon_\theta - D^2 M) M^{-1} w + Q_1 \\
&= \varepsilon + D^2 w - (D^2 M) M^{-1} w + \varepsilon_\theta M^{-1} w + Q_1 \\
&=: \varepsilon + D^2 w - (D^2 M) M^{-1} w + Q_2 \qquad (37)
\end{aligned}
$$

with

$$Q_2 = Q_1 + \varepsilon_\theta M^{-1} w \ . \tag{38}$$

Setting

$$z := M^{-1} w \ , \tag{39}$$

we get:

$$\begin{aligned} \varepsilon' &= \varepsilon + D^2(Mz) - (D^2 M)z + Q_2 \\ &= \varepsilon + D(MDz) + D(DMz) - (D^2 M)z + Q_2 \\ &= \varepsilon + D(MDz) + (DM)(Dz) + Q_2. \end{aligned} \tag{40}$$

Denote by M^T the transpose of the matrix M and let $M^{-T} := (M^T)^{-1}$. Then:

$$\begin{aligned} \varepsilon' &= M^{-T}\Big(M^T\varepsilon + M^T D(MDz) + M^T(DM)(Dz)\Big) + Q_2 \\ &= M^{-T}\Big(M^T\varepsilon + D(M^T MDz) - (DM^T)(MDz) + M^T(DM)(Dz)\Big) + Q_2 \\ &=: M^{-T}\Big(M^T\varepsilon + D(M^T MDz)\Big) + g + Q_2 \ , \end{aligned} \tag{41}$$

with

$$g := M^{-T}(M^T DM - (DM^T)M)Dz \ . \tag{42}$$

We claim that g is quadratic in ε. To check this, we, first, remark that

$$\begin{aligned} \langle M^T DM - (DM^T)M\rangle &= \langle M^T Dv_\theta - (Dv_\theta^T)M\rangle \\ &= \langle (I + v_\theta^T)Dv_\theta - (Dv_\theta^T)(I + v_\theta)\rangle \\ &= \langle v_\theta^T Dv_\theta - Dv_\theta^T v_\theta\rangle \ , \end{aligned} \tag{43}$$

(since $\langle Du\rangle = 0$ for any periodic function u). Integrating by parts,

$$\begin{aligned} \langle v_\theta^T Dv_\theta - Dv_\theta^T v_\theta\rangle_{ij} &= \sum_{k=1}^d \ ⨍_{\mathbb{T}^d} \left(\frac{\partial v_k}{\partial \theta_i}\left(D\frac{\partial v_k}{\partial \theta_j}\right) - \left(D\frac{\partial v_k}{\partial \theta_i}\right)\frac{\partial v_k}{\partial \theta_j}\right) \\ &= -\sum_{k=1}^d \ ⨍_{\mathbb{T}^d} \left(\frac{\partial^2 v_k}{\partial \theta_j \partial \theta_i}(Dv_k) - (Dv_k)\frac{\partial^2 v_k}{\partial \theta_i \partial \theta_j}\right) \\ &= 0 \ , \end{aligned}$$

showing that

$$\langle M^T DM - (DM^T)M\rangle = 0 \, . \tag{44}$$

Thus, we can write:

$$M^T DM - (DM^T)M = D^{-1}\left[D\left(M^T DM - (DM^T)M\right)\right] \ .$$

But, by (36),

$$\begin{aligned} D\left(M^T DM - (DM^T)M\right) &= M^T D^2 M - (D^2 M^T)M \\ &= -M^T V_{xx} M + M^T \varepsilon_\theta + M^T V_{xx} M - \varepsilon_\theta^T M \\ &= M^T \varepsilon_\theta - \varepsilon_\theta^T M \ , \end{aligned}$$

showing that

$$\langle M^T \varepsilon_\theta - \varepsilon_\theta^T M\rangle = 0 \ , \tag{45}$$

and that

$$M^T DM - (DM^T)M = D^{-1}(M^T \varepsilon_\theta - \varepsilon_\theta^T M) \ .$$

Thus

$$g = M^{-T}(D^{-1}(M^T \varepsilon_\theta - \varepsilon_\theta^T M))Dz \tag{46}$$

is quadratic in ε. Furthermore (41) can be rewritten as

$$\varepsilon' = M^{-T}\left(M^T \varepsilon + D(M^T MDz)\right) + Q_3 \ , \tag{47}$$

with

$$Q_3 := Q_2 + M^{-T}(D^{-1}(M^T \varepsilon_\theta - \varepsilon_\theta^T M))Dz. \tag{48}$$

We can now show that the equation

$$M^T \varepsilon + D(M^T MDz) = 0 \tag{49}$$

can be explicitly solved.

We have already studied the inversion of the differential operator D and therefore we know that a necessary condition to solve our equation is that the average $\langle M^T \varepsilon\rangle$ of $M^T \varepsilon$ over $\mathbb{T}^d$ is equal to 0. This is indeed the case, as we proceed to show.

First, by the definitions of M and ε,

$$\begin{aligned} \langle M^T \varepsilon\rangle &= \langle (I + v_\theta^T)(D^2 v + V_x(\theta + v))\rangle \\ &= \langle v_\theta^T D^2 v\rangle + \langle M^T V_x(\theta + v)\rangle \end{aligned} \tag{50}$$

where the latter equality follows since $\langle D^2 v\rangle = 0$. Let us compute the i-th component of $\langle v_\theta^T D^2 v\rangle$. Integrating by parts

$$\begin{aligned}
\langle v_\theta^T D^2 v\rangle_i &= \sum_{k=1}^d \fint_{\mathbb{T}^d} \frac{\partial v_k}{\partial \theta_i} D^2 v_k d\theta \\
&= (-1)\sum_{k=1}^d \fint_{\mathbb{T}^d} v_k \frac{\partial}{\partial \theta_i}(D^2 v_k) d\theta \\
&= (-1)^2\sum_{k=1}^d \fint_{\mathbb{T}^d} (Dv_k) D\frac{\partial v_k}{\partial \theta_i} d\theta \\
&= (-1)^3\sum_{k=1}^d \fint_{\mathbb{T}^d} (D^2 v_k) \frac{\partial v_k}{\partial \theta_i} d\theta \\
&= -\langle v_\theta^T D^2 v\rangle_i \ .
\end{aligned} \tag{51}$$

Thus $\langle v_\theta^T D^2 v\rangle = 0$ and, in view of (50), it remains to check that

$$\langle M^T V_x(\theta + v)\rangle = 0\,.$$

By the chain rule:

$$\langle M^T V_x(\theta+v)\rangle_i = \langle \sum_{k=1}^d \frac{\partial(\theta + v(\theta))_k}{\partial \theta_i} V_{x_k}(\theta+v)\rangle = \langle \frac{\partial}{\partial\theta_i} V(\theta+v)\rangle = 0\,,$$

showing that

$$\langle M^T \varepsilon\rangle = 0 \ . \tag{52}$$

Inverting D in (49) we find that

$$M^T M D z = -D^{-1}(M^T\varepsilon) + c\,, \tag{53}$$

where c is a suitable constant vector that we shall shortly identify.

Let

$$P = M^T M \tag{54}$$

and notice that $P = P(\theta)$ is, for $\theta \in \mathbb{T}^d$, a strictly positive defined matrix: $P > 0$. Thus P is invertible and $P^{-1} > 0$. Rephrasing (53) in terms of P:

$$Dz = -P^{-1}D^{-1}(M^T\varepsilon) + P^{-1}c. \tag{55}$$

By the positiveness of P^{-1} and its integrability over $\mathbb{T}^d$, we have that also $\langle P^{-1}\rangle$ is positive and in particular invertible. By taking the average on both sides of (55), we see that in order for (55) to make sense we have to choose:

$$c := \langle P^{-1}\rangle^{-1}\langle P^{-1}D^{-1}(M^T\varepsilon)\rangle. \tag{56}$$

We can now solve for z obtaining:

$$\begin{aligned} z &= b + D^{-1}(-P^{-1}D^{-1}(M^T\varepsilon) + P^{-1}c) \\ &=: b + \hat{z} \end{aligned} \tag{57}$$

having defined $\hat{z}$ as

$$\hat{z} = D^{-1}(-P^{-1}D^{-1}(M^T\varepsilon) + P^{-1}c), \tag{58}$$

and b denotes the arbitrary average of z. We fix this ambiguity by *requiring that*

$$\langle v'\rangle = \langle v\rangle \ , \tag{59}$$

which is equivalent to

$$0 = \langle w\rangle = \langle Mz\rangle = \langle Mb\rangle + \langle M\hat{z}\rangle = \langle M\rangle b + \langle M\hat{z}\rangle = b + \langle M\hat{z}\rangle \ ,$$

i.e.,

$$b = -\langle M\hat{z}\rangle. \tag{60}$$

The above analysis may be summarized in the following

Lemma 2 (KAM scheme) *Let $V : \mathbb{T}^d \to \mathbb{R}$ be smooth enough and let $\omega \in \mathbb{R}^d$ be a diophantine vector. Assume that a smooth enough function $v : \mathbb{T}^d \to \mathbb{R}^d$ is given so that*

$$M = I + v_\theta$$

is an invertible matrix on $\mathbb{T}^d$ and define

$$\varepsilon(\theta) := D^2 v + V_x(\theta + v) \ , \qquad \Big(D := \sum_{i=1}^d \omega_i \frac{\partial}{\partial\theta_i}\Big) \ .$$

Then:

$$\langle M^T \varepsilon \rangle = 0 \qquad \text{and} \qquad \langle M^T \varepsilon_\theta - \varepsilon_\theta^T M \rangle = 0\,. \tag{61}$$

Furthermore, if we let:

$$\begin{aligned}
P &:= M^T M \\
c &:= \langle P \rangle^{-1} \langle P^{-1} D^{-1} (M^T \varepsilon) \rangle \\
\hat{z} &:= D^{-1}(-P^{-1} D^{-1}(M^T \varepsilon) + P^{-1} c) \\
b &:= -\langle M \hat{z} \rangle \\
z &:= b + \hat{z} \\
w &:= Mz \\
Q_1 &:= V_x(\theta + v + w) - V_x(\theta + v) - V_{xx}(\theta + v) w \\
Q_2 &:= Q_1 + \varepsilon_\theta z \\
Q_3 &:= Q_2 + M^{-T}(D^{-1}(M^T \varepsilon_\theta - \varepsilon_\theta^T M)) D z \\
v' &:= v + w \\
\varepsilon'(\theta) &:= D^2 v' + V_x(\theta + v')\ ,
\end{aligned}$$

then:

$$\varepsilon' = Q_3 \qquad \text{and} \qquad \langle v' \rangle = \langle v \rangle. \tag{62}$$

Remark 7 (i) The above lemma *does not contain any quantitative statement*, nor its proof uses in any way the fact that ε should be a "small" function.

(ii) The **proof** of Lemma 2 is based upon a series of identities: (50)÷(52) and (43)÷(45) [proof of (61)]; (32), (37), (40), (41), (46), (48), (49), (59) [proof of (62)].

(iii) At this level, the above KAM scheme is purely "algebraic" and it will be only after having equipped it with quantitative estimates that it will be possible to iterate the scheme and to actually construct solutions of (23).

1.4 Banach spaces of analytic functions and technical lemmata

In this section we introduce "monotone families" of Banach spaces of real-analytic functions on $\mathbb{T}^d$; such families will depend upon a parameter $\xi \geq 0$

and "monotone" means that a space parameterized by $\xi > \xi'$ is smaller than the space parameterized by ξ'.

Usually in KAM theory one works either with *complex sup-norms* or with Fourier norms. In connection with smooth theory (chapter two below) sup-norms are more convenient, while for the extension of KAM theory to infinite dimensions Fourier norms are more suited. In this section we shall discuss sup-norms and for completeness we present the analogous technical results also for Fourier norms in Appendix B.

In these lectures we use the following standard notation: for $n \in \mathbb{Z}^d$, $\alpha \in \mathbb{N}^d$ and $x \in \mathbb{C}^d$, we let

$$|n| = \sum_{i=1}^{d} |n_i| \ , \quad |\alpha| = \sum_{i=1}^{d} \alpha_i \ , \quad \partial^\alpha f := \frac{\partial^{\alpha_1+\ldots+\alpha_d} f}{\partial x_1^{\alpha_1} \ldots \partial x_d^{\alpha_d}} \ , \quad x^\alpha := x_1^{\alpha_1} \ldots x_d^{\alpha_d} \ ; \tag{63}$$

Fourier coefficients of a periodic function will be denoted f_n; denote, also, by Δ^d_ξ the complex strip

$$\Delta^d_\xi := \{x \in \mathbb{C}^d : \ |\operatorname{Im} x_j| < \xi, \ j = 1, \ldots, d\} \ . \tag{64}$$

For $\xi \geq 0$ we define[11]

$\mathcal{R}_\xi(\mathbb{T}^d, \mathbb{R}^N) := \Big\{ f \in C(\mathbb{T}^d, \mathbb{R}^N)$ with bounded analytic extension on $\Delta^d_\xi \Big\}$.

$\mathcal{R}_\xi(\mathbb{T}^d, \mathbb{R}^N)$ endowed with the sup-norm

$$\|f\|_\xi := \sup_{\Delta^d_\xi} |f|$$

is a Banach space.

Remark 8 In the following, we will consider function $f \in \mathcal{R}_\xi(\mathbb{T}^d, X)$ with values in a matrix or tensor space X; in such cases the definition of the norm will be adapted in the obvious way[12].

[11] $\mathcal{R}_0$ denotes simply $C(\mathbb{T}^d, \mathbb{R}^N)$ endowed with the sup-norm.

[12] For example, if $X = \mathrm{Mat}(n \times n)$, then $f_n \in X$ and, in the definition of the norm $\|f\|_\xi$, the expression $|f_n|$ denotes the standard "operator norm" $\sup_{|c|=1} |f_n c|$.

We proceed to discuss, from a quantitative point of view, the equation $Du = f$ (for f with $\langle f\rangle = 0$) for ω diophantine. While it is elementary to get a bound of the form[13]

$$\|u\|_{\xi-\delta} \le c(d,\tau)\frac{1}{\gamma\delta^a}\|f\|_\xi$$

(with $0 < \delta \le \xi$) for some $a > 0$, to get the *optimal* dependence on the "analyticity loss" δ (i.e., the best a) is a subtle matter, which was solved by H. Rüssmann. We present a version of Rüssmann's result due to J. Moser (compare also Salamon's paper [2]).

Lemma 3 (Rüssmann, Moser) *Let $d \ge 2$, $\gamma > 0$ and $\tau \ge d-1$; let $\omega \in \mathbb{R}^d$ be (γ,τ)-diophantine and let $f \in \mathcal{R}_\xi$ be such that $\langle f\rangle = 0$. Denote (as above) by $u := D^{-1}f$ the unique solution of $Du = f$ with zero average. Then, there exists a constant $c = c(\tau,d) > 0$ such that for every*

$$0 < \delta \le \min\{1,\xi\}\ ,$$

one has

$$\|u\|_{\xi-\delta} = \|D^{-1}f\|_{\xi-\delta} \le \frac{c}{\gamma\delta^\tau}\|f\|_{L^2,\,\Delta^d_\xi} \tag{65}$$

where

$$\|f\|_{L^2,\,\Delta^d_\xi} := \sup\left\{\left(\int_{\mathbb{T}^d}|f(u+iv)|^2du\right)^{\frac12},\quad |v|<\xi\right\}.$$

Remark 9 Clearly $\|f\|_{L^2,\,\Delta^d_\xi} \le \|f\|_\xi$.

In order to prove the above lemma, we shall make use of the following general estimates, the proof of which are deferred to the Appendix A.

Lemma 4 *Let $f \in \mathcal{R}_\xi$. Then, for every $n \in \mathbb{Z}^d$,*

$$\begin{aligned} |f_n| &\le \|f\|_{L^2,\,\Delta^d_\xi}\; e^{-|n|\xi} \\ &\le \|f\|_\xi\; e^{-|n|\xi}\ . \end{aligned} \tag{66}$$

Furthermore, there exists a constant $c_0 = c_0(d) > 0$ such that, for every positive number $\delta < \min\{1,\xi\}$ and for every $x \in \Delta^d_{\xi-\delta}$, one has:

$$|f(x)| \le \sum_{n\in\mathbb{Z}^d}|f_n|e^{-(n\cdot\,\mathrm{Im}\,x)} \le \frac{c_0}{\delta^{\frac d2}}\|f\|_{L^2,\,\Delta^d_\xi}\ . \tag{67}$$

[13] **Exercise.**

Proof of Lemma 3 As already discussed above, the unique analytic solutions with vanishing mean value of $Du = f$ is given by

$$u(\theta) := (D^{-1}f)(\theta) = \sum_{n\in\mathbb{Z}^d\setminus\{0\}} u_n e^{i(n\cdot\theta)} \ , \qquad u_n := \frac{f_n}{i(n\cdot\omega)} \ .$$

In order to establish the inequality (65) we first single out the subset

$$J_0 := \Big\{ n \in \mathbb{Z}^d \setminus \{0\} \ : \ |n\cdot\omega| \ge \frac{\gamma}{2} \Big\}$$

and define

$$u^0(x) = \sum_{n\in J_0} u_n e^{i(n\cdot x)} \ .$$

By Lemma 4, we get that, for $|\operatorname{Im} x| < \xi - \delta$:

$$\begin{aligned} |u^0(x)| &\le \sum_{n\in J_0} |n\cdot\omega|^{-1} |f_n| e^{-(n\cdot \operatorname{Im} x)} \\ &\le \frac{2}{\gamma} \sum_{n\in Z^d} |f_n| e^{-(n\cdot \operatorname{Im} x)} \\ &\le \frac{c_1}{\gamma\delta^\tau} \|f\|_{L^2,\Delta_\xi^d} \end{aligned} \tag{68}$$

where $c_1 := 2c_0$ and we have used that $\tau \ge d-1 \ge \frac{d}{2}$.

The more delicate part of the estimate concerns the integer vectors in $\mathbb{Z}^d \setminus J_0$. First of all, let us assume, without loss of generality, that

$$|\omega_k| < |\omega_d| \ , \qquad \forall \ 1 \le k \le d-1 \ ; \tag{69}$$

let us also introduce the following notation: if $y = (y_1, ..., y_d)$ is a vector with d components, we denote by $\hat{y} := (y_1, ..., y_{d-1})$ the vector formed by the first $(d-1)$ components of y. Let, now, $K \ge 1$ be a fixed number and for $\nu = 1,\ 2\ldots$, define

$$\begin{aligned} J(\nu, K) &:= \{ n \in \mathbb{Z}^d,\ 0 < |n| \le K : \quad 2^\nu\gamma^{-1} < |n\cdot\omega|^{-1} \le 2^{\nu+1}\gamma^{-1} \} \\ &= \{ n \in \mathbb{Z}^d,\ 0 < |n| \le K : \quad 2^{-(\nu+1)}\gamma \le |n\cdot\omega| < 2^{-\nu}\gamma \} \ . \end{aligned}$$

Here is a list of **properties of** ω **and** $J(\nu, K)$:

(i) $|\omega_d| > |\omega_k| \ge \gamma$;

(ii) *if* $n \in J(\nu, K)$ *then* $\hat{n} \neq 0$*;*

(iii) *if* $n, n' \in J(\nu, K)$ *and* $\hat{n} = \hat{n}'$ *then* $n = n'$*;*

(iv) *if* $n \in \mathbb{Z}^d$ *is such that* $\hat{n} \neq 0$ *then*

$$|n \cdot \omega| \geq \frac{\gamma}{(3|\hat{n}|)^\tau} \ . \tag{70}$$

(v) *if* $n, n' \in J(\nu, K)$ *and* $n \neq n'$ *then* $|\hat{n} - \hat{n}'| \geq \frac{2^{\frac{\nu-1}{\tau}}}{3}$*;*

(vi) *there exists a constants* $c_2 = c_2(d) > 0$ *such that*

$$\text{Card}\, J(\nu, K) \leq c_2 K^{d-1} 2^{-\frac{\nu(d-1)}{\tau}} \ ;$$

(vii) $J(\nu, K) = \emptyset$ *when* $2^{\frac{\nu}{\tau}} \geq K$*;*

(viii) *there exists a constant* $c_3 = c_3(d) > 0$ *such that the following holds. If* $J(K)$ *denotes the set*

$$J(K) := \{n \in \mathbb{Z}^d : 0 < |n| \leq K \text{ and } |n \cdot \omega| < \gamma/2\} = \bigcup_{\nu \geq 1} J(\nu, K) \ ,$$

then

$$\sum_{n \in J(K)} \frac{1}{|n \cdot \omega|} \leq c_3 \ \frac{K^\tau}{\gamma} \ . \tag{71}$$

Proof of properties (i)÷(viii)

(i): The first inequality is (69). The second inequality follows from the Diophantine property (27) by taking $n = e_k$ (the unit versor in $\mathbb{Z}^d$).

(ii): If $n \in J(\nu, K)$, then

$$|n \cdot \omega| < 2^{-\nu}\gamma \leq \gamma/2 \ ; \tag{72}$$

thus from $\hat{n} = 0$ it would follow, by (i), that $|n \cdot \omega| = |n_d \omega_d| \geq |\omega_d| > \gamma$, which would contradict (72).

(iii): Assume (by contradiction) that $n, n' \in J(\nu, K)$ with $\hat{n} = \hat{n}'$ and $n_d \neq n'_d$. Then, by (i) and (72):

$$\gamma < |\omega_d| \leq |n_d - n'_d| \ |\omega_d| = |n \cdot \omega - n' \cdot \omega| \leq |n \cdot \omega| + |n' \cdot \omega| < \frac{\gamma}{2} + \frac{\gamma}{2} = \gamma \ ,$$

which is impossible.

(iv) Fix $\hat{n}\neq 0$ and choose $n_d\in\mathbb{Z}$ so that $n_{\min}:=(\hat{n},n_d)$ minimizes $|n\cdot\omega|$. Clearly, $n_{\min}$ minimizes also $|\hat{n}\cdot\frac{\hat{\omega}}{\omega_d}+n_d|=|n\cdot\omega|/|\omega_d|$. Thus, $|\hat{n}\cdot\frac{\hat{\omega}}{\omega_d}+n_d|\leq 1$. Therefore, by (69), $|n_d|\leq 1+|\hat{n}\cdot\frac{\hat{\omega}}{\omega_d}|\leq 1+|\hat{n}|\leq 2|\hat{n}|$, which implies that $|n_{\min}|=|\hat{n}|+|n_d|\leq 3|\hat{n}|$. In conclusion, by (27), and the above estimates,

$$|n\cdot\omega|\geq|n_{\min}\cdot\omega|\geq\frac{\gamma}{|n_{\min}|^\tau}\geq\frac{\gamma}{(3|\hat{n}|)^\tau}\,.$$

(v): By (iii), $\hat{n}\neq\hat{n}'$. Thus, by (iv) (applied to the difference $n-n'$) and by the definition of $J(\nu,K)$, we find $|\hat{n}-\hat{n}'|^{-\tau}\leq\frac{3^\tau}{\gamma}|(n-n')\cdot\omega|\leq\frac{3^\tau}{\gamma}(|n\cdot\omega|+|n'\cdot\omega|)\leq 3^\tau 2^{-(\nu-1)}$, proving the claim.

(vi): By (iii), $J(\nu,K)$ is in a one-to-one correspondence with $\hat{J}(\nu,K):=\{\hat{n}\in\mathbb{Z}^{d-1}: n\in J(\nu,K)\}$. By the estimate in (v), the distance between two points in $\hat{J}(\nu,K)$ is at least $\frac{2^{\frac{\nu-1}{\tau}}}{3}$. Thus a simple geometrical argument yields the desired upper bound on the cardinality of $\hat{J}(\nu,K)$ and hence on the cardinality of $J(\nu,K)$.

(vii): If $n\in J(\nu,K)$, by definition $|n\cdot\omega|<\gamma 2^{-\nu}$; on the other hand the Diophantine property (27) implies that $|n\cdot\omega|\geq\gamma|n|^{-\tau}\geq\gamma K^{-\tau}$ implying that $K>2^{\nu/\tau}$, which is equivalent to the claim.

(viii): In view of (vii), $J(K)=\bigcup_{\nu=1}^{\nu_*}J(\nu,K)$ where ν_* denotes the integer part of $\tau\log K/\log 2$ (i.e., ν_* is the maximal integer ν for which $2^{\nu/\tau}\leq K$: for $\nu>\nu_*$, $J(\nu,K)=\emptyset$). Thus, by (vii), the definition of $J(\nu,K)$, (vi), one finds

$$\begin{aligned}\sum_{n\in J(K)}\frac{1}{|n\cdot\omega|} &\leq \sum_{\nu=1}^{\nu_*}\sum_{n\in J(\nu,K)}\frac{1}{|n\cdot\omega|}\\ &\leq \sum_{\nu=1}^{\nu_*}\sum_{n\in J(\nu,K)}\frac{2^{\nu+1}}{\gamma}=\sum_{\nu=1}^{\nu_*}\frac{2^{\nu+1}}{\gamma}\,\mathrm{Card}\,J(\nu,K)\\ &\leq \frac{2c_2K^{d-1}}{\gamma}\sum_{\nu=1}^{\nu_*}2^{\nu\frac{(\tau+1-d)}{\tau}}\\ &\leq \frac{c_3}{\gamma}K^\tau.\end{aligned}$$

We are now ready to conclude the proof of (65).

$$\begin{aligned}\|u-u^0\|_{\xi-\delta} &\leq \sum_{n\notin J_0}|f_n||n\cdot\omega|^{-1}e^{|n|(\xi-\delta)}\\ &\overset{(*)}{\leq} \|f\|_{L^2,\Delta_\xi^d}\sum_{n\notin J_0}|n\cdot\omega|^{-1}e^{-|n|\delta}\\ &= \|f\|_{L^2,\Delta_\xi^d}\sum_{k=1}^\infty\sum_{n\notin J_0,\,|n|^2=k}|n\cdot\omega|^{-1}e^{-\sqrt{k}\delta}\end{aligned}$$

$$
\begin{aligned}
&= \|f\|_{L^2,\Delta^d_\xi} \sum_{k=1}^{\infty} \sum_{J(\sqrt{k})\setminus J(\sqrt{k-1})} |n\cdot\omega|^{-1} e^{-\sqrt{k}\delta} \\
&= \|f\|_{L^2,\Delta^d_\xi} \sum_{k=1}^{\infty} \sum_{n\in J(\sqrt{k})} |n\cdot\omega|^{-1} \left(e^{-\sqrt{k}\delta} - e^{-\sqrt{k+1}\delta}\right) \\
&\overset{(**)}{\leq} \|f\|_{L^2,\Delta^d_\xi} \Big(\sum_{k=1}^{\infty} \frac{\delta}{2\sqrt{k}} e^{-\sqrt{k}\delta}\Big)\Big(\sum_{n\in J(\sqrt{k})} |n\cdot\omega|^{-1}\Big) \\
&\overset{\dagger}{\leq} \frac{c_3}{2\gamma} \|f\|_{L^2,\Delta^d_\xi} \sum_{k=1}^{\infty} k^{\frac{\tau-1}{2}} \delta e^{-\sqrt{k}\delta} \\
&\leq \frac{c_3}{2\gamma}\,\frac{1}{\delta^\tau}\,\|f\|_{L^2,\Delta^d_\xi} \sup_{0<\lambda\leq 1} \sum_{k=1}^{\infty} \lambda^{\tau+1} k^{\frac{\tau-1}{2}} e^{-\sqrt{k}\lambda} \\
&\overset{\ddagger}{\leq} \frac{c_4}{\gamma\delta^\tau}\,\|f\|_{L^2,\Delta^d_\xi}\,,
\end{aligned}
$$

where: $(*)$ is by (66); $(**)$ follows from the elementary bounds $e^{-s} - e^{-s-\varepsilon} \leq \varepsilon e^{-s}$ (any $s > 0$, $\varepsilon > 0$) and $\sqrt{t+1} - \sqrt{t} \leq (2\sqrt{t})^{-1}$ (any $t > 0$); $(\dagger)$ is by property (viii) above; $(\ddagger)$ holds for a suitable constant $c_4 = c_4(d,\tau) > 0$ since $\sup_{0<\lambda\leq 1}\sum_{k=1}^{\infty} \lambda^{\tau+1} k^{\frac{\tau-1}{2}} e^{-\sqrt{k}\lambda} < \infty$. The proof is completed if one takes $c = \max\{c_1, c_4\}$. ∎

Another fundamental tool are the so-called *Cauchy estimates*, i.e., the estimates of the sup-norm of derivatives of $f \in \mathcal{R}_\xi$ in $\Delta^d_{\xi-\delta}$.

Denote by $D^d_r(x)$ the complex polydisc

$$
D^d_r(x) = \{\theta \in \mathbb{C}^d \ : \ \ |\theta_i - x_i| \leq r,\ \forall i\}\,,
$$

and let $r > 0$ be such that[14] $D^d_r(x) \subset \Delta^d_\xi$. Then, by Cauchy Integral Formula,

$$
\partial^\alpha f(x) = \frac{\alpha!}{(2\pi i)^d} \int_{\partial D^d_r(x)} \frac{f(\theta)}{(\theta_1 - x_1)^{\alpha_1+1}\ldots(\theta_d - x_d)^{\alpha_d+1}} d\theta\,.
$$

In particular, taking $r = \delta$, we have for every $x \in \Delta^d_{\xi-\delta}$:

[14]Notice that f is bounded on Δ^d_ξ so f cannot have singularities on the boundary of the strip.

$$\begin{aligned} |\partial^\alpha f(x)| &\le \frac{\alpha!}{(2\pi)^d}\int_{\partial D_\delta^d}\frac{|f(\theta)|}{|\theta_1-x_1|^{\alpha_1+1}\dots|\theta_d-x_d|^{\alpha_d+1}}d\theta \\ &\le \frac{\alpha!}{(2\pi)^d}\frac{\|f\|_\xi}{\delta^{|\alpha|+d}}\int_{\partial D_\delta^d}d\theta \\ &= \frac{\alpha!}{\delta^{|\alpha|}}\|f\|_\xi\,. \end{aligned}$$

Thus, the following *Cauchy estimate* holds

$$\|\partial^\alpha f\|_{\xi-\delta}\le \alpha!\delta^{-|\alpha|}\|f\|_\xi\,. \tag{73}$$

Combining the above estimates one gets easily the following

Lemma 5 *Let $f\in\mathcal{R}_\xi(\mathbb{T}^d,X)$, let $p\in\mathbb{N}$, $\alpha\in\mathbb{N}^d$ and assume $\langle f\rangle=0$ when $\alpha=0$. Let $0<\delta\le\xi$. Then, there exist $C(p,\alpha,\tau)>0$ such that*

$$\|D^{-p}\partial^\alpha f\|_{\xi-\delta}\le\frac{C(p,\alpha,\tau)}{\gamma^p\delta^{p\tau+|\alpha|}}\|f\|_\xi\,,$$

Exercise Give an explicit estimate of $C(p,\alpha,\tau)$.

Remark 10 We shall use also the following trivial facts:

(i) *Let $f\in\mathcal{R}_\xi(\mathbb{T}^d,X)$, $g\in\mathcal{R}_\xi(\mathbb{T}^d,Y)$, with X and Y tensor spaces ($\mathbb{R}^N$, matrices or higher dimensional tensors) and assume that the product fg is well defined. Then*

$$\|f\,g\|_\xi\le\|f\|_\xi\|g\|_\xi\,.$$

(ii) *Let $V\in\mathcal{R}_{\xi_*}(\mathbb{T}^d,\mathbb{R})$, $g\in\mathcal{R}_\xi(\mathbb{T}^d,\mathbb{R}^d)$, with $\xi_*>\xi$. If $\|\operatorname{Im}g_j\|_\xi\le\xi_*-\xi$ for all $j=1,\dots,d$, then $\|V(x+g(x))\|_\xi\le\|V\|_{\xi_*}$.*

Remark 11 The same estimates could be done also for the non-autonomous case. In this case we consider a function $f(x,t)$ defined on the $(d+1)$-dimensional torus and we require ω to be $(\gamma\,,\tau)$-diophantine in the sense that

$$|\omega\cdot n+m|\ge\frac{\gamma}{|n|^\tau}$$

for some $\tau\ge d$, $\gamma>0$ and every $(n,m)\in\mathbb{Z}^d\times\mathbb{Z}$ with $n\neq0$.

1.5 An analytic KAM theorem

In this section we prove a KAM theorem in the real-analytic setting.

Theorem 6 *Fix* $0 < \bar\xi < \xi < \xi_* \le 1$. *Let* $V \in \mathcal{R}_{\xi_*}(\mathbb{T}^d, \mathbb{R})$, *let* $v \in \mathcal{R}_\xi(\mathbb{T}^d, \mathbb{R}^d)$ *be such that*

$$\|\operatorname{Im} v\|_\xi \le \xi_* - \xi \ , \tag{74}$$

and let ω *be* (γ,τ)*-diophantine*[15]. *Let, also,* λ, η *and* α *be numbers greater or equal than one such that*

$$\lambda \ge \|I + v_\theta\|_\xi \,, \qquad \eta \ge \|(I+v_\theta)^{-1}\|_\xi \,, \qquad \alpha \ge \frac{\|V_{xxx}\|_{\xi_*}}{\gamma^2} \,.$$

There exists a constant $C = C(\tau, d) > 1$, *such that if*[16]

$$E := \frac{C}{\gamma^2}\,\|\mathcal{E}(v)\|_\xi \ \ \alpha\,(\lambda\eta)^{10}\,(\xi-\bar\xi)^{-(4\tau+2)} \le 1 \,, \tag{75}$$

then there exists $u \in \mathcal{R}_{\bar\xi}(\mathbb{T}^d, \mathbb{R}^d)$ *with* $\langle u\rangle = \langle v\rangle$, *which solves the Euler equation*

$$D^2 u + V_x(\theta+u) =: \mathcal{E}(u) = 0 \,. \tag{76}$$

Furthermore, there exists a constant $K = K(\tau,d) > 0$ *such that*

$$\max\Big\{\|u-v\|_{\bar\xi}\ ,\ \|\partial_\theta u - \partial_\theta v\|_{\bar\xi}\Big\} \le K\,E\,. \tag{77}$$

Proof As a *first step*, we equip the KAM scheme described in Lemma 2 with analytical estimates. The *second step* will be to iterate the procedure controlling the convergence.

Remark 12 Here and below for simplicity we will use the notation "const" to denote finite (different) constants, which depend only on τ and d.

---

[15] Recall Definition 2.

[16] Recall that $\mathcal{E}$ is the differential operator defined as $\mathcal{E}(v) := D^2 v + V_x(\theta+v)$ where $D := \sum_{i=1}^d \omega_i \partial_{\theta_i}$; notice, also, that, since $\|\operatorname{Im} v_j\|_\xi \le \xi_* - \xi$, then $\theta + v \in \Delta^d_{\xi_*}$ whenever $\theta \in \Delta^d_\xi$.

Define

$$\varepsilon(\theta) := D^2 v + V_x(\theta + v) \, .$$

and let $\mu > 0$ be such that

$$\|\varepsilon\|_\xi \le \mu \, .$$

Recalling the definition of $M = I + v_\theta$, we have, by hypothesis,

$$\|M\|_\xi \le \lambda \, , \qquad \|M^{-1}\|_\xi \le \eta \, .$$

Fix[17] $\bar{\xi} < \xi' < \xi$ and let $\delta = (\xi - \xi')/2$:

$$\xi' =: \xi - 2\delta \, , \qquad \delta = \frac{\xi - \xi'}{2} \, , \tag{78}$$

and let us denote, as above,

$$P = M^T M \, .$$

Then

$$\|P\|_\xi \le \lambda^2 \, , \qquad \|P^{-1}\|_\xi \le \eta^2 \, . \tag{79}$$

We start by estimating c in (56). To estimate it, we will use the fact[18] that for each positive symmetric matrix $T : \mathbb{T}^d \to \mathrm{Mat}(d \times d)$,

$$\|\langle T \rangle^{-1}\| \le \sup_{\theta \in \mathbb{T}^d} \|T^{-1}\| \, . \tag{80}$$

Therefore, by (80) and (79),

$$\|\langle P^{-1} \rangle^{-1}\| \le \|P\|_0 \le \lambda^2 \, .$$

By Lemma 5 we get

$$\|\langle P^{-1} D^{-1}(M^T \varepsilon) \rangle\|_0 \le \|P^{-1}\|_0 \|D^{-1}(M^T \varepsilon)\|_0 \le \frac{\mathrm{const}}{\gamma \xi^\tau} \eta^2 \|M^T \varepsilon\|_\xi \, ,$$

which leads to

$$|c| \le \mathrm{const} \, \frac{\lambda^3 \eta^2}{\gamma \xi^\tau} \mu \, . \tag{81}$$

[17] Later we shall make a specific (somewhat arbitrary) choice.

[18] For the proof see Lemma 12 in Appendix A.

We proceed to estimate $\hat{z}$ (see (58)). Applying Lemma 5 twice, we have

$$\begin{aligned}
\|\hat{z}\|_{\xi'} &\leq \text{const}\,\frac{1}{\gamma\delta^\tau}\| - P^{-1}D^{-1}(M^T\varepsilon) + P^{-1}c\|_{\xi-\delta} \\
&\leq \text{const}\,\frac{\eta^2}{\gamma\delta^\tau}\left[\|D^{-1}(M^T\xi)\|_{\xi-\delta} + \frac{\lambda^3\eta^2\mu}{\gamma\xi^\tau}\right] \\
&\leq \text{const}\,\frac{\eta^2}{\gamma\delta^\tau}\left[\frac{\lambda^3\eta^2\mu}{\gamma\xi^\tau} + \frac{\lambda\mu}{\gamma\delta^\tau}\right] \\
&\leq \text{const}\,\frac{\lambda^3\eta^4\mu}{\gamma^2\delta^{2\tau}}\,.
\end{aligned}$$

Thus,

$$\|\hat{z}\|_{\xi'} \leq \text{const}\,\frac{\lambda^3\eta^4}{\gamma^2\delta^{2\tau}}\mu\,. \tag{82}$$

Now, since $z = b + \hat{z}$ and $b = -\langle M\hat{z}\rangle$, we have

$$|b| = |\langle M\hat{z}\rangle| \leq \lambda\|\hat{z}\|_{\xi'}\ .$$

Therefore,

$$\|z\|_{\xi'} \leq \text{const}\,\frac{(\lambda\eta)^4}{\gamma^2\delta^{2\tau}}\mu\,, \qquad \|w\|_{\xi'} \leq \text{const}\,\frac{\lambda^5\eta^4}{\gamma^2\delta^{2\tau}}\mu\,. \tag{83}$$

Let us estimate now the remainder Q_1. Using again the standard formula for the remainder of the Taylor expansion and applying Remark 10 and the definition of α, we have

$$\|Q_1\|_{\xi'} \leq \frac{1}{2}\|V_{xxx}\|_{\xi_*}\|w\|^2_{\xi'} \leq \text{const}\|V_{xxx}\|_{\xi_*}\frac{\lambda^{10}\eta^8\mu^2}{\gamma^4\delta^{4\tau}} \leq \text{const}\,\frac{\alpha\lambda^{10}\eta^8}{\gamma^2\delta^{4\tau}}\mu^2\ , \tag{84}$$

provided (compare Remark 10)

$$\|\operatorname{Im}(v + sw)\|_{\xi'} \leq \xi_* - \xi'$$

for every $s \in [0, 1]$. But, by (74) and (83),

$$\begin{aligned}
\|\operatorname{Im}(v + sw)\|_{\xi'} &\leq \|\operatorname{Im} v\|_\xi + \|w\|_{\xi'} \leq \xi_* - \xi + \|w\|_{\xi'} \\
&\leq \xi_* - \xi + \text{const}\,\frac{\lambda^5\eta^4}{\gamma^2\delta^{2\tau}}\mu \\
&\leq \xi_* - \xi'\ ,
\end{aligned}$$

which is implied if we assume that

$$\text{const}\,\frac{(\lambda\eta)^5}{\gamma^2\delta^{2\tau+1}}\mu \leq 1\,. \tag{85}$$

Since $Q_2 = Q_1 + \varepsilon_\theta z$, by Lemma 5 and Remark 10, we find

$$\|\varepsilon_\theta\|_{\xi-\delta} \leq \frac{\mu}{\delta}\,, \tag{86}$$

and

$$\|Q_2\|_{\xi'} \leq \|Q_1\|_{\xi'} + \|\varepsilon_\theta z\|_{\xi'} \leq \|Q_1\|_{\xi'} + \|\varepsilon_\theta\|_{\xi'}\|z\|_{\xi'} \leq \|Q_1\|_{\xi'} + \frac{\mu}{2\delta}\|z\|_{\xi'}\,.$$

Thus, by (83) and (84),

$$\|Q_2\|_{\xi'} \leq \text{const}\,\frac{\alpha\lambda^{10}\eta^8}{\gamma^2\delta^{4\tau}}\mu^2\,. \tag{87}$$

Here we implicitly used the fact that $\tau \geq 1/2$ and $\delta \leq 1$.

We turn to the estimate of the norm of Q_3 (defined in (48)): we will use again Lemma 5. First, observe that $Dz = D\hat{z}$ and that

$$D\hat{z} = -P^{-1}D^{-1}(M^T\varepsilon) + P^{-1}c\,.$$

Thus

$$\|D\hat{z}\|_{\xi'} = \|-P^{-1}D^{-1}(M^T\varepsilon) + P^{-1}c\|_{\xi'} \leq \text{const}\,\frac{\lambda^3\eta^4}{\gamma\delta^\tau}\mu\,.$$

Recalling (86), we find

$$\|D^{-1}(M^T\varepsilon_\theta - \varepsilon_\theta^T M)\|_{\xi'} \leq \text{ const}\,\frac{1}{\gamma\delta^\tau}\|M^T\varepsilon_\theta\|_{\xi-\delta} \leq \text{const}\,\frac{\lambda}{\gamma\delta^{\tau+1}}\mu\,.$$

As for the second term in the formula for Q_3, we find

$$\|M^{-T}(D^{-1}(M^T\varepsilon_\theta - \varepsilon_\theta^T M))Dz\|_{\xi'} \leq \text{const}\,\frac{\lambda\eta\mu}{\gamma\delta^{\tau+1}}\,\frac{\lambda^3\eta^4\mu}{\gamma\delta^\tau} \leq \text{const}\,\frac{\lambda^4\eta^5}{\gamma^2\delta^{2\tau+1}}\mu^2\,.$$

Finally, recalling (87), we get the following bound

$$\|Q_3\|_{\xi'} \leq \text{const}\,\frac{\alpha\lambda^{10}\eta^8}{\gamma^2\delta^{4\tau}}\mu^2 \leq \text{const}\,\frac{\alpha(\lambda\eta)^{10}}{\gamma^2\delta^{4\tau}}\mu^2\,. \tag{88}$$

Our next step will be to estimate v', $M' := M + w_\theta$ and ε' in terms of v, M and ε.

Let us start with w_θ. Here again, as we have already done for $\hat{z}$, we will apply Lemma 5 twice. In fact,

$$\|w_\theta\|_{\xi'} \leq \frac{1}{\delta}\|w\|_{\xi-\delta}\,.$$

Then, since $\xi' = \xi - \delta = \xi - \delta/2 - \delta/2$, using twice Lemma 5, we get:

$$\begin{aligned}
\|\hat{z}\|_{\xi-\delta} &\leq \text{const}\frac{1}{\gamma\delta^\tau}\| - P^{-1}D^{-1}(M^T\varepsilon) + P^{-1}c\|_{\xi-\delta/2} \\
&\leq \text{const}\,\frac{\eta^2}{\gamma\delta^\tau}\Big[\frac{\lambda^3\eta^2\mu}{\gamma\xi^\tau} + \|D^{-1}(M^T\varepsilon)\|_{\xi-\delta/2}\Big] \\
&\leq \text{const}\,\frac{\eta^2}{\gamma\delta^\tau}\Big[\frac{\lambda^3\eta^2\mu}{\gamma\xi^\tau} + \frac{\lambda\mu}{\gamma(\delta/2)^\tau}\Big] \\
&\leq \text{const}\,\frac{\eta^4\lambda^3}{\gamma^2\delta^{2\tau}}\mu\,.
\end{aligned}$$

Since $z = b + \hat{z}$ and $w = Mz$, we get the estimate

$$\|w_\theta\|_{\xi'} \leq \text{const}\,\frac{\lambda^5\eta^4}{\gamma^2\delta^{2\tau+1}}\mu\,. \tag{89}$$

Next, using (89), we easily get

$$\begin{aligned}
\|M'\|_{\xi'} &\leq \|M + w_\theta\|_{\xi'} \leq \lambda + \text{const}\,\frac{\lambda^5\eta^4\mu}{\gamma^2\delta^{2\tau+1}} \\
&= \lambda\Big(1 + \text{const}\frac{(\lambda\eta)^4\mu}{\gamma^2\delta^{2\tau+1}}\Big) =: \lambda'\,.
\end{aligned} \tag{90}$$

As for the inverse matrix

$$\begin{aligned}
\|(M')^{-1}\|_{\xi'} &\leq \|(M + w_\theta)^{-1}\|_{\xi'} = \|(I + M^{-1}w_\theta)^{-1}M^{-1}\|_{\xi'} \\
&\leq \eta\left(1 - \text{const}\,\frac{(\lambda\eta)^5\mu}{\gamma^2\delta^{2\tau+1}}\right)^{-1}\,,
\end{aligned}$$

provided

$$\text{const}\,\frac{(\lambda\eta)^5\mu}{\gamma^2\delta^{2\tau+1}} < 1\,.$$

In fact, assuming that

$$\text{const}\,\frac{(\lambda\eta)^5}{\gamma^2\delta^{2\tau+1}}\mu \le \frac{1}{2}\,, \tag{91}$$

one finds[19]

$$\|(M')^{-1}\|_{\xi'} \le \eta\Big(1+\text{const}\,\frac{(\lambda\eta)^5}{\gamma^2\delta^{2\tau+1}}\mu\Big) =: \eta'\,. \tag{92}$$

Note that (91) is the same condition (up to the constant) as (85). Since, by definition, $\varepsilon' = Q_3$, we have obtained

$$\|\varepsilon'\|_{\xi'} \le \text{const}\,\frac{\alpha(\lambda\eta)^{10}}{\gamma^2\delta^{4\tau}}\mu^2 =: \mu'\,. \tag{93}$$

It will be useful to introduce also a "dimensionless" parameter

$$\bar\mu = \mu\gamma^{-2}\ , \tag{94}$$

in terms of which (93) may be rewritten as

$$\|\varepsilon'\|_{\xi'}\gamma^{-2} \le \Big(\text{const}\,\frac{\sqrt{\alpha}(\lambda\eta)^5\bar\mu}{\delta^{2\tau}}\Big)^2 \le \Big(\text{const}\,\frac{\sqrt{\alpha}(\lambda\eta)^5\bar\mu}{\delta^{2\tau+1}}\Big)^2 =: \bar\mu'\ . \tag{95}$$

From now on we will replace (91) by the stronger condition

$$\text{const}\,\frac{\sqrt{\alpha}(\lambda\eta)^5\bar\mu}{\delta^{2\tau+1}} \le 1\,. \tag{96}$$

Let us now turn to the *second step*, i.e., to the control of the convergence of the iteration process. For $i \ge 0$, let the input data v, ε, ξ correspond to the i-th step of iteration, and let the output v', ε' and ξ' correspond to $(i+1)$-th step: in particular the function v and the parameter ξ in the statement of Theorem 6 will be denoted, respectively, v_0, ξ_0. Thus,

$$v_i = v_0 + \sum_{j=0}^{i-1} w_j\,,$$

[19]This follows from the following inequality: $(1-x)^{-1} \le 1+2x$ valid for $0 \le x \le 1/2$.

and our aim is to show that this sequence converges to some real-analytic function u, which solves the Euler equation

$$D^2 u + V_x(\theta + u(t)) = 0\,.$$

We fix a sequence $\{\xi_i\}$ *as follows*

$$\xi_i = \bar{\xi} + \frac{\xi - \bar{\xi}}{2^i}\,.$$

So, $\{\xi_i\}$ is a decreasing sequence, which tends to $\bar{\xi}$. In view of the above definition of δ, *we fix also*

$$\delta_i := \frac{\xi_i - \xi_{i+1}}{2} = \frac{\xi - \bar{\xi}}{2^{i+2}}\,.$$

If (96) holds at each each step of the iteration, i.e., if

$$\text{const}\,\frac{\sqrt{\alpha}(\lambda_j \eta_j)^5 \bar{\mu}_j}{\delta_j^{2\tau+1}} \le 1\,, \qquad\qquad j = 0, 1, \ldots, i\,, \tag{97}$$

then, in particular, we see that (compare (90) and (92) attaching the indices i and $i+1$ in the obvious way)

$$\lambda_j \le 2^j \lambda_0\,, \quad \eta_j \le 2^j \eta_0\,, \qquad\qquad j = 0, 1, \ldots, i\,. \tag{98}$$

In terms of λ_0 and η_0 condition (97) can be rewritten in the form

$$C_0\,\frac{\sqrt{\alpha}(\lambda_0 \eta_0)^5 (2^{2\tau+11})^i \bar{\mu}_i}{(\xi - \bar{\xi})^{2\tau+1}} \le 1\,,$$

where $C_0 = C_0(d, \tau)$ denotes the largest constant " const " occurred until now. Denoting,

$$A := \Big(C_0 \frac{\sqrt{\alpha}(\lambda_0 \eta_0)^5}{(\xi - \bar{\xi})^{2\tau+1}}\Big)^2\,, \qquad B := 2^{2\tau+11}\,,$$

we see that (95) yields

$$\bar{\mu}_{i+1} \le A B^i \bar{\mu}_i^2 \le 1\,. \tag{99}$$

Such relation may be rewritten as

$$\hat{\mu}_{i+1} \le \hat{\mu}_i^2 \,, \qquad \hat{\mu}_i := AB^{i+1}\bar{\mu}_i \,,$$

which, iterated, leads to

$$\bar{\mu}_i \le \frac{(AB\bar{\mu}_0)^{2^i}}{AB^{i+1}} \,. \tag{100}$$

In particular, one can conclude that the iteration process converges if

$$AB\bar{\mu}_0 < 1 \,, \qquad \bar{\mu}_0 \ge \frac{\|\varepsilon_0\|_{\xi_0}}{\gamma^2} \,,$$

showing, in particular, that $v + \sum_{j=0}^{\infty} w_j$ converges uniformly on the complex strip of width $\bar{\xi}$ to the real-analytic function

$$u := v + \sum_{j=0}^{\infty} w_j \,,$$

which (since $\varepsilon_j \to 0$ uniformly) will satisfy the Euler equation (76).

Finally, we prove (77). First of all note that in fact E is nothing else but

$$E = AB\bar{\mu}_0 \,.$$

Now, (compare (98)),

$$\begin{aligned} \|w_j\|_{\bar{\xi}} &\le \text{const} \frac{2^{2\tau(j+2)}\lambda_j^5 \eta_j^4 \bar{\mu}_j}{(\xi - \bar{\xi})^{2\tau}} = \text{const}(4^{\tau j}\lambda_j^5 \eta_j^4 \bar{\mu}_j) \\ &\le \Big(\text{const}(\lambda_0\eta_0)^5\Big) 2^{10j+2\tau j} \bar{\mu}^j =: \tilde{C} 2^{10j+2\tau j} \bar{\mu}_j \,, \end{aligned}$$

where $\tilde{C} :=$ const $(\lambda_0\eta_0)^5$. Since by (100)

$$\bar{\mu}_i < (AB\mu_0)^{2^i} \,,$$

we have

$$\sum_{j=0}^{\infty} 2^{10j+2\tau j} \bar{\mu}_j \le \sum_{j=0}^{\infty} 2^{2(5+\tau)j} (AB\bar{\mu}_0)^{2^j} = \sum_{j=0}^{\infty} (C_1^2)^j E^{2^j} \,,$$

where we have denoted $C_1 = 2^{5+\tau}$. Continuing the last inequality, we get

$$\begin{aligned}\sum_{j=0}^{\infty} C_1^{2j} E^{2^j} &< \sum_{j=0}^{\infty} C_1^{2^j} E^{2^j} = \sum_{j=0}^{\infty} (C_1 E)^{2^j} \\ &\leq \sum_{j=1}^{\infty} (C_1 E)^j = \frac{C_1 E}{1 - C_1 E} \\ &\leq C_1 E (1 + 2C_1 E) \leq 2C_1 E\,,\end{aligned}$$

provided $0 < C_1 E \leq 1/2$. This last assumption can be always satisfied by the right choice of the constant. ∎

We conclude this section with an immediate application of Theorem 6 to the "nearly-integrable" case.

Corollary 1 *If*

$$\|V_x\|_\xi \leq \frac{\gamma^2}{C} \frac{(\xi - \bar{\xi})^{4\tau+2}}{\max\{1, \|V_{xxx}\|_{\xi_*} \gamma^{-2}\}}\,, \tag{101}$$

then there exists a function $u \in \mathcal{R}_\xi(\mathbb{T}^d, \mathbb{R}^d)$ *such that* $\langle u \rangle = 0$ *which solves the Euler equation* $D^2 u + V_x(\theta + u) = 0$ *with*

$$\|u\|_{\bar{\xi}} \leq \frac{C}{\gamma^2} \|V_x\|_\xi \frac{\max\{1, \|V_{xxx}\|_{\xi_*} \gamma^{-2}\}}{(\xi - \bar{\xi})^{4\tau+2}}\,.$$

Proof Take as initial approximate solution the function $v \equiv 0$. Then $\mathcal{E}(v) = \mathcal{E}(0) = V_x(\theta)$ and one can take $\lambda = \eta = 1$ so that (101) is recognized to be (75) . ∎

Remark 13 Let $\|V_x\|_\xi$, $\|V_{xxx}\|_\xi \leq \varepsilon$. From the properties of diophantine numbers it follows that, if we denote

$$\Omega_r = \left\{ \omega \in B_r^d : \quad |\omega \cdot n| \geq \frac{\gamma}{|n|^\tau} \quad \forall n \in \mathbb{Z}^d \setminus \{0\} \right\},$$

then

$$\mathrm{meas}(B_r^d \setminus \Omega_r) \leq \mathrm{const}\, \mathrm{meas}(B_r^d) \gamma\,.$$

Now, condition (101) can be met by taking $\gamma = \sqrt{\varepsilon}\hat{C}$ with $\hat{C}$ big enough, showing that the set of ω's for which we can find simultaneously a solution for the Euler equation fills (as $\varepsilon \to 0$) a ball of radius r up to a set of measure at most const$\sqrt{\varepsilon}$.

1.6 Local uniqueness

In this section we formulate a sufficient condition which provides "local" uniqueness for the solution of Euler equation. First we remark that if u verifies

$$D^2 u + V_x(\theta + u) = 0 \tag{102}$$

then also

$$\bar{u} : (\theta) \mapsto c + u(\theta + c)$$

is a solution of the same equation, for every constant $c \in \mathbb{R}^d$. Since $\langle \bar{u} \rangle = \langle u \rangle + c$, it is natural to investigate local uniqueness of solutions with prescribed average.

Proposition 7 *Let $\omega \in \mathbb{R}^d$ be (γ, τ)-diophantine. Let $V \in \mathcal{R}_{\xi_*}(\mathbb{T}^d, \mathbb{R})$ and u, $\bar{u} \in \mathcal{R}_\xi(\mathbb{T}^d, \mathbb{R}^d)$. Assume that u and $\bar{u}$ are two solutions of (102) such that $\langle u \rangle = \langle \bar{u} \rangle$. Assume moreover that $I + u_\theta$ is invertible everywhere on $\mathbb{T}^d$ and that*

$$\begin{aligned} \|u\|_\xi, \|\bar{u}\|_\xi &\leq \xi_* - \xi \\ \|(I + u_\theta)^{-1}\|_\xi &\leq \eta < +\infty \\ \|I + u_\theta\|_\xi &\leq \lambda < +\infty. \end{aligned}$$

Define

$$c := c_0 \frac{\gamma^2 \xi^{2\tau}}{\lambda^5 \eta^4 \|V_{xxx}\|_{\xi_*}}, \tag{103}$$

where $c_0 = c_0(d, \tau) \geq 1$ is a suitable constant. Then, if $\|u - \bar{u}\|_\xi < c$ one has that $u \equiv \bar{u}$.

Proof Let

$$w := \bar{u} - u$$

and notice that $\langle w \rangle = 0$. Since $\bar{u}$ and u are solutions of (102), we have

$$\begin{aligned} 0 &= D^2\bar{u} + V_x(\theta + \bar{u}) \\ &= D^2 w + D^2 u + V_x(\theta + w + u) \\ &= D^2 w + V_x(\theta + w + u) - V_x(\theta + u) \\ &= D^2 w + V_{xx}(\theta + u)w + Q \end{aligned} \tag{104}$$

where

$$Q = \int_0^1 (1-s) V_{xxx}(\theta + u + sw) ww \, ds.$$

From the expression of Q and Lemma 15 [Remark 10] we easily get

$$\|Q\|_{\xi'} \leq \frac{1}{2} \|V_{xxx}\|_{\xi_*} \|w\|_{\xi'}^2 \tag{105}$$

for every $\xi' \in [0, \xi]$.

Let $M = I + u_\theta$. Differentiating with respect to θ equation (102) we get the equality

$$V_{xx}(\theta + u) = -(D^2 M) M^{-1} \tag{106}$$

that we can plug in (104) obtaining

$$0 = D^2 w - (D^2 M) M^{-1} w + Q.$$

Letting

$$z = M^{-1} w$$

we have

$$\begin{aligned} 0 &= D^2(Mz) - (D^2 M)z + Q \\ &= D(MDz) + D(DMz) - (D^2 M)z + Q \\ &= D(MDz) + (DM)(Dz) + Q\,, \end{aligned}$$

which can be rewritten as

$$0 = M^{-T}(M^T D(MDz) + M^T(DM)(Dz)) + Q. \tag{107}$$

Moreover from (106), we have that the matrix $(D^2M)M^{-1}$ is symmetric and so

$$\begin{aligned} 0 &= M^T(D^2M) - (D^2M)^TM \\ &= D(M^TDM - (DM^T)M); \end{aligned}$$

in particular, since $D^{-1}0 = 0$ we have

$$M^TDM - (DM^T)M = \langle M^TDM - (DM^T)M\rangle.$$

By equation (44) we already know, however, that

$$\langle M^TDM - (DM^T)M\rangle = 0$$

for every matrix M of the form $I + u_\theta$. Thus

$$M^TDM - DM^TM = 0,$$

i.e.

$$M^TDM = DM^TM.$$

From (107) it follows:

$$\begin{aligned} 0 &= M^{-T}(M^TD(MDz) + (DM^T)MDz) + Q \\ &= M^{-T}D(M^TMDz) + Q \end{aligned}$$

which means, setting $P = M^TM$, that $PDz = -D^{-1}(M^TQ) + c_1$ for a suitable constant vector c_1. Thus

$$Dz = -P^{-1}D^{-1}(M^TQ) + P^{-1}c_1 \tag{108}$$

and

$$w = MD^{-1}(-P^{-1}D^{-1}(M^TQ) + P^{-1}c_1) + Mc_2 \tag{109}$$

for a suitable constant vector c_2. Taking averages in (108) and (109) we obtain the following expressions for c_1 and c_2:

$$c_1 = \langle P^{-1}\rangle^{-1}\langle P^{-1}D^{-1}(M^TQ)\rangle \tag{110}$$

$$c_2 = -\langle MD^{-1}(-P^{-1}D^{-1}(M^TQ) + P^{-1}c_1)\rangle. \tag{111}$$

Let us now define, for every $j \in \mathbb{N}$, $\xi_j = 2^{-j}\xi$. For every $j \in \mathbb{N}$, by estimates similar to the ones already seen in the previous section we get from (110) and (111) the following inequalities:

$$\begin{aligned} |c_1| &\leq \text{const } \lambda^3\eta^2 \frac{2^{j\tau}}{\gamma\xi^\tau}\|Q\|_{\xi_j} \\ |c_2| &\leq \text{const } \lambda^4\eta^4 \frac{4^{(j+1)\tau}}{\gamma^2\xi^{2\tau}}\|Q\|_{\xi_j} \end{aligned}$$

which can be inserted in (109) obtaining

$$\text{const } \|w\|_{\xi_{j+1}} \leq \lambda^5\eta^4 \frac{4^{2j\tau}}{\gamma^2\xi^{2\tau}}\|Q\|_{\xi_j} \ . \tag{112}$$

Letting

$$k = \max\Big\{1 \ , \ \ \text{const } \lambda^5\eta^4 \frac{1}{\gamma^2\xi^{2\tau}}\|V_{xxx}\|_{\xi_*}\Big\}$$

we obtain from (112) and (105)

$$\|w\|_{\xi_{j+1}} \leq k4^{2j\tau}\|w\|_{\xi_j}$$

and, iterating as done above (compare (99), (100)), we get

$$\|w\|_0 \leq \|w\|_{\xi_{j+1}} \leq (k4^{2\tau}\|w\|_\xi)^{2^j} \ ,$$

showing that $\|w\|_0 = 0$ (and hence by analyticity $w \equiv 0$) whenever

$$k4^{2\tau}\|w\|_\xi < 1 \ . \qquad \blacksquare$$

2 Smooth KAM Theory

The aim of this chapter is to exhibit a result of existence of quasi-periodic solutions for systems that are no more required to be analytic but just smooth enough. We will heavily use the previous results, passing through analytic approximations of smooth functions.

2.1 Approximation Theory

Here we prove the necessary technical approximation results.

We start by introducing Hölder norms. First of all, for every $l_0 \in \mathbb{N}$ and for every $f \in C^{l_0}(\mathbb{R}^m)$, we define

$$|f|_{C^{l_0}} = \sup_{|\alpha|\le l_0} \sup_{\mathbb{R}^m} |\partial^\alpha f| \,.$$

If $l = l_0 + \mu$ with $l_0 \in \mathbb{N}$ and $\mu \in (0,1)$, we set

$$|f|_{C^l} := |f|_{C^{l_0}} + \sup_{|\alpha|=l_0} \sup_{0<|x-y|<1} \frac{|\partial^\alpha f(x) - \partial^\alpha f(y)|}{|x-y|^\mu} \,.$$

For every $l \geq 0$ we define

$$C^l(\mathbb{R}^d, \mathbb{R}^m) = \{ f : \mathbb{R}^d \to \mathbb{R}^m : \quad |f|_{C^l} < +\infty \} \,.$$

The space $C^l(\mathbb{R}^d, \mathbb{R}^m)$ endowed with the norm $|\cdot|_{C^l}$ is a Banach space; the subspace of $C^l(\mathbb{R}^d, \mathbb{R}^m)$ made of functions which are 2π-periodic in each variable will be denoted $C^l(\mathbb{T}^d, \mathbb{R}^m)$.

Remark 14 In this section it is convenient to work with Euclidean norms on vectors and the associated operator norms on matrices and tensors.

Proposition 8 (Jackson, Moser, Zehnder) *Let $l \geq 0$, $d \in \mathbb{Z}_+$ and $f \in C^l(\mathbb{R}^d)$. There exists a constant $c = c(l,d) > 0$ such that for every $0 < r \leq 1$ there exists a real analytic function f_r on Δ^d_r which satisfies*[20]

$$\left| \partial^\alpha f_r(x) - \sum_{|\beta|\le l-|\alpha|} \partial^{\alpha+\beta} f(\operatorname{Re} x) \frac{(i \operatorname{Im} x)^\beta}{\beta!} \right| \leq c\, |f|_{C^l}\, r^{l-|\alpha|} \,, \qquad \forall\; x \in \Delta^d_r \tag{113}$$

for all α such that $|\alpha| \leq l$.

[20] We will use the following notations:

$$\operatorname{Re} x := (\operatorname{Re} x_1, \ldots, \operatorname{Re} x_d) \qquad \text{and} \qquad \operatorname{Im} x := (\operatorname{Im} x_1, \ldots, \operatorname{Im} x_d) \,.$$

In fact, the analytic extension f_r may be defined as follows. Let ϕ_1 be an even function in $C_0^\infty(\mathbb{R})$ with support $[-1,1]$, increasing in $[-1,0]$ and such that

$$\phi_1(0) = 1 \ , \qquad \partial^n \phi_1(0) = 0 \quad (\forall\ n \geq 1) \ ;$$

for $\xi \in \mathbb{R}^d$, let $\phi(\xi) = \phi_1(|\xi|^2)$ and let K be the anti-Fourier transform of ϕ:

$$K(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \phi(\xi) e^{ix\cdot\xi} d\xi \ ;$$

then f_r can be taken to be

$$\begin{aligned} f_r(x) &:= (\mathcal{S}_r f)(x) := r^{-d} \int_{\mathbb{R}^d} K\left(\frac{x-y}{r}\right) f(y) dy \\ &= \int_{\mathbb{R}^d} K\left(\frac{x}{r} - \xi\right) f(r\xi) d\xi = \int_{\mathbb{R}^d} K(\eta) f(x - r\eta) d\eta. \end{aligned} \tag{114}$$

Proof Since ϕ is a real smooth function with compact support, K is real analytic on $\mathbb{C}^d$. Some properties of K are collected in the following

Lemma 9 *The derivatives of K satisfy*

$$\forall p \in \mathbb{N}\,, \quad \exists\ c_p : \ \left|\partial^\beta K(x)\right| \leq c_p \frac{e^{|\mathrm{Im} x|}}{(1+|x|)^p} \ , \quad \forall\ |\beta| \leq p \ ; \tag{115}$$

$$\sup_{x \in \mathbb{R}^d} \sup_{\beta \in \mathbb{N}^d} \left|\partial^\beta K(x)\right| \leq \frac{1}{(2\pi)^d} \|\phi\|_{L^1} \ . \tag{116}$$

Furthermore, if α, $\beta \in \mathbb{N}^d$ and $x = u + iv \in \mathbb{C}^d$, then[21]

$$I_{\alpha,\beta} := \int_{\mathbb{R}^d} u^\beta \partial^\alpha K(u+iv) du = \begin{cases} (-1)^{|\beta|} \frac{\beta!}{(\beta-\alpha)!} (iv)^{\beta-\alpha} \ , & \text{if } \alpha \leq \beta \ , \\ 0 \ , & \text{otherwise} \, . \end{cases} \tag{117}$$

[21] For vectors α, $\beta \in \mathbb{N}^d$, we denote

$$\alpha \leq \beta \quad \Longleftrightarrow \quad \alpha_i \leq \beta_i \quad \forall\ i = 1, \ldots, d\,.$$

Proof First of all, remark that if $u \in \operatorname{supp}\phi = B_1(0)$, then

$$\left|e^{ix\cdot u}\right| \;=\; e^{-\mathrm{Im}x\cdot u} \le e^{|\mathrm{Im}x|} .$$

Let us denote $\phi_\beta(u) = u^\beta\phi(u)$. We have

$$\begin{aligned}\partial^\beta K(x) &:= \partial^\beta\left(\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\phi(u)e^{ix\cdot u}du\right)\\ &= \frac{i^{|\beta|}}{(2\pi)^d}\int_{\mathbb{R}^d}\phi_\beta(u)e^{ix\cdot u}du ,\end{aligned}$$

and, for any multi-index $\alpha \in \mathbb{N}^d$, $|\alpha|$ integrations by part give

$$x^\alpha\partial^\beta K(x) = \frac{i^{|\alpha|+|\beta|}}{(2\pi)^d}\int_{\mathbb{R}^d}\partial^\alpha\phi_\beta(u)e^{ix\cdot u}du .$$

Hence

$$|x^\alpha||\partial^\beta K(x)| \le \frac{1}{(2\pi)^d}\left|\partial^\alpha\phi_\beta\right|_{L^1(\mathbb{R}^d)} e^{|\mathrm{Im}x|} .$$

Now remark that for any $p \in \mathbb{N}$,

$$(1+|x|)^p \le (1+|x_1|+|x_2|+\cdots+|x_d|)^p = \sum_{|\alpha|\le p}\frac{p!}{(p-|\alpha|)!\alpha!}|x^\alpha| .$$

Hence for $|\beta| \le p$, one finds

$$\begin{aligned}(1+|x|)^p|\partial^\beta K| &\le \frac{1}{(2\pi)^d}\sum_{|\alpha|\le p}\frac{p!}{(p-|\alpha|)!\alpha!}\left|\partial^\alpha\phi_\beta\right|_{L^1(\mathbb{R}^d)} e^{|\mathrm{Im}x|}\\ &\le \frac{1}{(2\pi)^d}\left(\sum_{|\alpha|\le p}\frac{p!}{(p-|\alpha|)!\alpha!}\sup_{|\beta|\le p}\left\{\left|\partial^\alpha\phi_\beta\right|_{L^1(\mathbb{R}^d)}\right\}\right)e^{|\mathrm{Im}x|} .\end{aligned}$$

Thus

$$|\partial^\beta K(x)| \le c_p\frac{e^{|\mathrm{Im}x|}}{(1+|x|)^p} ,$$

proving (115). Moreover, if $x \in \mathbb{R}^d$ and $\beta \in \mathbb{N}^d$, we have:

$$\begin{aligned} \left|\partial^\beta K(x)\right| &\leq \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \left|u^\beta \phi(u)\right| du \\ &\leq \frac{1}{(2\pi)^d} \|\phi\|_{L^1} . \end{aligned}$$

which is (116).

The relation (115) shows that the integral $I_{\alpha,\beta}$ is well defined. Moreover, notice that if $\alpha_i > \beta_i$ for some i, α_i integrations by part show that $I_{\alpha,\beta} = 0$ since $\partial^{\alpha_i} u^\beta = 0$. So we can assume that $\alpha_i \leq \beta_i$ for all i (i.e. $\alpha \leq \beta$). By integration by part, we see that

$$I_{\alpha,\beta} = (-1)^{|\alpha|} \frac{\beta!}{(\beta-\alpha)!} \int_{\mathbb{R}^d} u^{\beta-\alpha} K(u+iv) du \, .$$

Notice also that

$$\begin{aligned} K(u+iv) &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \phi(\xi) e^{i\,\xi\cdot(u+iv)}\, d\xi \\ &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \phi(\xi) e^{-v\cdot\xi} e^{i\,\xi\cdot u}\, d\xi =: \tilde{K}_v(u) \end{aligned}$$

that we can think to be the anti-Fourier transform of

$$\tilde{\phi}_v(\xi) = \phi(\xi) e^{-v\cdot\xi} \in C_0^\infty(\mathbb{R}^d)$$

for each fixed v.

Hence:

$$\begin{aligned} \int_{\mathbb{R}^d} u^{\beta-\alpha} K(u+iv)\, du &= \frac{1}{(-i)^{|\beta-\alpha|}} \left.\frac{\partial^{|\beta-\alpha|}}{\partial\, \xi^{\beta-\alpha}}\right|_{\xi=0} \int_{\mathbb{R}^d} K(u+iv) e^{-iu\xi}\, du \\ &= \frac{1}{(-i)^{|\beta-\alpha|}} \left.\frac{\partial^{|\beta-\alpha|}}{\partial\, \xi^{\beta-\alpha}}\right|_{\xi=0} \int_{\mathbb{R}^d} \tilde{K}_v(u) e^{-iu\xi}\, du \\ &= \frac{1}{(-i)^{|\beta-\alpha|}} \left.\frac{\partial^{|\beta-\alpha|}}{\partial\, \xi^{\beta-\alpha}}\right|_{\xi=0} \left(\phi(\xi) e^{-v\xi}\right) \end{aligned}$$

where, in the last equalities, we used the fact that the right hand-side integral is the Fourier transform of $\tilde{K}_v$ (i.e. $\tilde{\phi}_v$). Now, using (8), we obtain:

$$\begin{aligned}\int_{\mathbb{R}^d} u^{\beta-\alpha} K(u+iv)\,du &= \frac{1}{(-i)^{|\beta-\alpha|}}\phi(0)(-v)^{\beta-\alpha} \\ &= (-iv)^{\beta-\alpha} \;.\end{aligned}$$

From the properties above, we can conclude:

$$I_{\alpha,\beta} := \int_{\mathbb{R}^d} u^\beta \partial^\alpha K(u+iv)\,du = \begin{cases} (-1)^{|\beta|}\frac{\beta!}{(\beta-\alpha)!}(iv)^{\beta-\alpha} \;, & \text{if } \alpha \le \beta \;, \\ 0 \;, & \text{otherwise} \;. \end{cases} \qquad \blacksquare$$

We proceed with the proof of Proposition 8. Recall the definition of $f_r = \mathcal{S}_r$ given in (114) and notice that (117) implies $\mathcal{S}_r P = P$ for any polynomial P:

$$\begin{aligned}(\mathcal{S}_r P)(x) &= \int_{\mathbb{R}^d} K(\eta) P(x - r\eta) d\eta \\ &:= \int_{\mathbb{R}^d} K(\eta) \left(\sum_{|k| \le \deg P} a_k(x,r)\eta^k \right) d\eta \\ &= \sum_{|k|\le \deg P} a_k(x,r) \int_{\mathbb{R}^d} K(\eta)\eta^k d\eta \\ &= a_0(x,r) = P(x) \;.\end{aligned}$$

We claim also that for every $l \in \mathbb{R}_+$ and $f \in C^l$, then there exists a constant $c_0(l,d)$ such that

$$|f(x+y) - P_k(x,y)| \;\le\; c_0(l,d)\, |f|_{C^l}\, |y|^l \qquad (118)$$

where

$$P_k(x,y) = \sum_{|\alpha|\le k} \frac{\partial^\alpha f(x)}{\alpha!} y^\alpha \quad \text{and} \quad k = [\, l\,] \;.$$

In fact, if l is an integer (118) comes immediately from Taylor's formula (and, actually, one could get $c_0(l,d)$ as small as one wants); if $l = k + \mu$ with $\mu \in (0,1)$, we get

$$\begin{aligned}|f(x+y) - P_k(x,y)| &\le \left| \int_0^1 \frac{(1-t)^{k-1}}{(k-1)!} \left(\partial^k f(x+ty) y^{(k)} - \partial^k f(x) y^{(k)} \right) dt \right| \\ &\le \int_0^1 \frac{(1-t)^{k-1}}{(k-1)!} \frac{|\partial^k f(x+ty) - \partial^k f(x)|}{|ty|^\mu} |y|^{k+\mu}\, dt \\ &\le \frac{|f|_{C^l}}{k!} |y|^l \,.\end{aligned}$$

Let $x = u + iv$, $\eta = \dfrac{u-y}{r}$ where u, v, y belong to $\mathbb{R}^d$. Then

$$\begin{aligned}
\partial^\alpha f_r(x) &= \partial^\alpha \left(\frac{1}{r^d} \int_{\mathbb{R}^d} K\left(\frac{x-y}{r}\right) f(y)dy \right) \\
&= \frac{1}{r^d} \int_{\mathbb{R}^d} \frac{1}{r^{|\alpha|}} \partial^\alpha K \left(\frac{u-y}{r} + i\frac{v}{r} \right) f(y)dy \\
&= \frac{1}{r^{|\alpha|}} \int_{\mathbb{R}^d} \frac{1}{r^d} \partial^\alpha K \left(\eta + i\frac{v}{r} \right) f(u-r\eta)d(u-r\eta) \\
&= \frac{1}{r^{|\alpha|}} \int_{\mathbb{R}^d} \partial^\alpha K \left(\eta + i\frac{v}{r} \right) f(u-r\eta)d\eta \;.
\end{aligned}$$

Let us consider now $|\beta| \le l - |\alpha|$; in view of (117):

$$\begin{aligned}
\frac{\partial^{\alpha+\beta} f(u)}{\beta!}(iv)^\beta &= \frac{\partial^{\alpha+\beta} f(u)}{\beta!} \frac{(iv)^\beta}{(-1)^{|\alpha|} \frac{(\beta+\alpha)!}{\beta!} \left(-i\frac{v}{r}\right)^\beta} \int_{\mathbb{R}^d} \partial^\alpha K(\eta + i\frac{v}{r})\eta^{\alpha+\beta} d\eta \\
&= \int_{\mathbb{R}^d} \partial^\alpha K(\eta + i\frac{v}{r}) \frac{\partial^{\alpha+\beta} f(u)}{(\beta+\alpha)!} (-1)^{|\alpha+\beta|} r^{|\beta|} \eta^{\alpha+\beta} d\eta \\
&= \frac{1}{r^{|\alpha|}} \int_{\mathbb{R}^d} \partial^\alpha K(\eta + i\frac{v}{r}) \frac{\partial^{\alpha+\beta} f(u)}{(\beta+\alpha)!} (-1)^{|\alpha+\beta|} (r\eta)^{\alpha+\beta} d\eta \;.
\end{aligned}$$

Hence, if we denote $k = [l]$ and apply again (117), we obtain:

$$\begin{aligned}
&\sum_{|\beta| \le l - |\alpha|} \frac{\partial^{\alpha+\beta} f(u)}{\beta!} (iv)^\beta \\
&\quad = \frac{1}{r^{|\alpha|}} \int_{\mathbb{R}^d} \partial^\alpha K(\eta + i\frac{v}{r}) \sum_{|\beta| \le l-|\alpha|} \left(\frac{\partial^{\alpha+\beta} f(u)}{(\beta+\alpha)!} (-1)^{|\alpha+\beta|} (r\eta)^{\alpha+\beta} \right) d\eta \\
&\quad = \frac{1}{r^{|\alpha|}} \int_{\mathbb{R}^d} \partial^\alpha K(\eta + i\frac{v}{r}) P_k(u - r\eta)\, d\eta \,.
\end{aligned}$$

In view of the above calculations and of the inequality (118), we have:

$$\begin{aligned}
&\left| \partial^\alpha f_r(x) - \sum_{|\beta| \le l-|\alpha|} \partial^{\alpha+\beta} f(u) \, \frac{(iv)^\beta}{\beta!} \right| \\
&\qquad \le \frac{1}{r^{|\alpha|}} \int_{\mathbb{R}^d} \left| \partial^\alpha K \left(\eta + i\frac{v}{r} \right) \right| \left| f(u-r\eta) - P_k(u-r\eta) \right| d\eta
\end{aligned}$$

$$\leq c_0 |f|_{C^l} r^{l-|\alpha|} \int_{\mathbb{R}^d} \left| \partial^\alpha K \left(\eta + i \frac{v}{r} \right) \right| |\eta|^l \, d\eta$$
$$\leq c_0 |f|_{C^l} r^{l-|\alpha|} \int_{\mathbb{R}^d} \left| \partial^\alpha K \left(\eta + i \frac{v}{r} \right) \right| (1 + |\eta|)^l \, d\eta \ .$$

Applying (117) for $p > l$ big enough (i.e., $p > l + d$), we get

$$\left| \partial^\alpha f_r(x) - \sum_{|\beta| \leq l - |\alpha|} \partial^{\alpha+\beta} f(u) \, \frac{(iv)^\beta}{\beta!} \right| \leq c_1 |f|_{C^l} r^{l-|\alpha|} \int_{R^d} (1 + |\eta|)^{l-p} \, d\eta$$
$$\leq c(l,d) |f|_{C^l} r^{l-|\alpha|}$$

which completes the proof. ∎

Remark 15 (i) In particular, (113) (with $x \in \mathbb{R}^d$) implies, for $s \leq l$ (s integer),

$$|f_r - f|_{C^s} \leq c' |f|_{C^l} \ r^{l-s} \ , \qquad (s \leq l) \ , \tag{119}$$

for a suitable $c' = c'(l,d)$.

(ii) Also, (113) with $l = 0$, yields, for every $f \in C^0$ and any $r > 0$,

$$\sup_{\Delta_r^d} |f_r| \leq (c+1) \ |f|_{C^0} \ . \tag{120}$$

(iii) If x is real, the definition of f_r implies immediately that, if $f \in C^l$ and $|\alpha| \leq l$,

$$\partial^\alpha f_r(x) = (\partial^\alpha f)_r(x) \ . \tag{121}$$

By analyticity, such relation is seen to hold for any $x \in \mathbb{C}^d$.

(iv) Using the observation in (iii) and Cauchy estimates, one can give the following bound on the derivatives of f_r with $f \in C^l$:

$$|f_r|_{C^s} \leq c'' \ |f|_{C^l} r^{l-s} \ , \qquad (s \geq l \text{ integers}) \ , \tag{122}$$

$c'' = c''(l,d)$ being a suitable positive constant.

Proof If $s = l$, (122) comes from (119). Let $s > l$. For any multi-index α such that $|\alpha| = s$ we can find β and α_0 such that $\alpha = \beta + \alpha_0$ with $|\alpha_0| = l$ and $|\beta| = s - l$. Then, by (121), Cauchy estimates and (120), denoting by " const " (possibly different) constants depending on l and d, we find

$$\begin{aligned} |\partial^\alpha f_r|_{C^0} &= |\partial^\beta (\partial^{\alpha_0} f)_r|_{C^0} \le \text{ const } \sup_{\Delta^d_r} |(\partial^{\alpha_0} f)_r| \frac{1}{r^{s-l}} \\ &\le \text{ const } |\partial^{\alpha_0} f|_{C^0} \frac{1}{r^{s-l}} \le \text{ const } |f|_{C^l} r^{l-s} \ . \qquad \blacksquare \end{aligned}$$

(v) (*Convexity estimates*) Let $l > 0$, let $f \in C^l$ and let k, m be integers such that $0 \le k \le m \le l$. Then, there exist a constant $\hat{c} = \hat{c}(l, d) > 0$ such that

$$|f|_{C^m} \le \hat{c}\ |f|^{\frac{l-m}{l-k}}_{C^k}\ |f|^{\frac{m-k}{l-k}}_{C^l} \ . \tag{123}$$

Proof Define $r := \left(\frac{|f|_{C^k}}{|f|_{C^l}}\right)^{\frac{1}{l-k}}$. Then $r \le 1$ and by (119) and (122), we get

$$|f|_{C^m} \le |f_r - f|_{C^m} + |f_r|_{C^m} \le \text{ const } \Big(|f|_{C^l} r^{l-m} + |f|_{C^k} r^{k-m}\Big) = 2 \text{ const } |f|^{\frac{l-m}{l-k}}_{C^k}\ |f|^{\frac{m-k}{l-k}}_{C^l} \ . \qquad \blacksquare$$

(vi) If f is periodic, f_r is obviously periodic. Moreover, if f belongs to $C^0(\mathbb{T}^d)$, then

$$\begin{aligned} (\mathcal{S}_r f)_n &= \int\!\!\!\!\!\!-_{\mathbb{T}^d} \mathcal{S}_r f(x) e^{-ix\cdot n} dx \\ &= \int\!\!\!\!\!\!-_{\mathbb{T}^d} \Big(\int_{\mathbb{R}^d} K(\eta) f(x - r\eta) d\eta\Big) e^{-ix\cdot n} dx \\ &= \int_{\mathbb{R}^d} K(\eta) e^{-irn\cdot\eta} \Big(\int\!\!\!\!\!\!-_{\mathbb{T}^d} f(x - r\eta) e^{-i(x - r\eta)\cdot n} dx\Big) d\eta \\ &= f_n \phi(rn) \ , \end{aligned}$$

showing that f_r is a *trigonometric polynomial.*

Proposition 10 (Bernstein, Moser) *Let $l > 0$ and $d \in \mathbb{Z}_+$. Let $f_0 = 0$ and for each j in $\mathbb{Z}_+$, let f_j be a real analytic function on $\Delta^d_{r_j} \subset \mathbb{C}^d$ where $r_j = r_0/2^j$ for some $0 < r_0 \le 1$. Assume that*

$$|f_j - f_{j-1}|_{r_j} \le A r_j^l \tag{124}$$

for every $j \ge 1$ and some constant A.
Then, f_j tends to f uniformly on $\mathbb{R}^d$ and $f \in C^s(\mathbb{R}^d)$ for every non integer

$s \leq l$. Furthermore, there exists a constant $C = C(l,d)$ such that:

$$|f|_{C^s(\mathbb{R}^d)} \leq \frac{CA}{\mu(1-\mu)} r_0^{l-s} \tag{125}$$

where $\mu = s - [s]$. Finally, if the f_i's are periodic in each variable x_j then so is f.

Remark 16 (i) If $f_0 \neq 0$ and (124) holds for all $j \geq 1$, we can apply the proposition to $\tilde{f}_j := f_j - f_0$ getting that $\tilde{f}_j$ tends uniformly to $\tilde{f} \in C^s$ so that f_j tends uniformly to $f := f_0 + \tilde{f} \in C^s$. Moreover,

$$|f - f_0|_{C^s(\mathbb{R}^d)} \leq \frac{CA}{\mu(1-\mu)} r_0^{l-s} \,. \tag{126}$$

(ii) It is enough to prove Proposition 10 in the particular case where $l \in (0,1)$ and $s = l = \mu$ as we proceed to check.

Proof of point (ii) of Remark 16 Let us consider the three following claims:

- (a) "Proposition 10 holds true for $0 < s < l = 1$" .
- (b) "Proposition 10 holds true for $0 < s < l < 1$" .
- (c) "Proposition 10 holds true for $0 < s = l < 1$" .

We will show that

$$(c) \Rightarrow (b) \Rightarrow (a)$$

and finally that (a) implies the general case.
$(b) \Rightarrow (a)$: To prove (a), we assume (124) with $l = 1$ and fix $0 < s < 1$. Then, for every $s < l < 1$,

$$|f_j - f_{j-1}|_{r_j} \leq A r_j \leq A r_j^l$$

which shows that f_j satisfies the hypothesis (124) of (b). Applying (b), we get

$$|f|_{C^s} \leq \frac{CA}{s(1-s)} r_0^{l-s}$$

and taking the infimum in the above expression over $l < 1$, we get

$$|f|_{C^s} \leq \frac{CA}{s(1-s)} r_0^{1-s}.$$

$(c) \Rightarrow (b)$: We have

$$|f_j - f_{j-1}|_{r_j} \leq A r_j^l = A r_j^{l-s} r_j^s \leq A r_0^{l-s} r_j^s$$

which shows that $\tilde{f}_j = f_j / r_0^{l-s}$ satisfies the hypothesis of (c). Then, by (c), the uniform limit $\tilde{f}$ of $\tilde{f}_j$ belongs to C^s and

$$|\tilde{f}|_{C^s} \leq \frac{CA}{s(1-s)},$$

which is equivalent to

$$|f|_{C^s} \leq \frac{CA}{s(1-s)} r_0^{l-s} \ .$$

which proves (b).

Now, let us show that the claim (c) implies the general case. We prove by induction on k ($k \geq 1$) that

$(\mathcal{P}_k)$ "Proposition 10 holds true for $0 < l \leq k$" .

First of all, notice that $(\mathcal{P}_1)$ holds true since (c) implies (a) and (b). Let s be a non integer such that $0 < s \leq l \leq k+1$. We can assume that $k < l \leq k+1$(if not, then $0 < l \leq k$ and we can apply the inductive hypothesis). By assumption, we have

$$|f_j - f_{j-1}|_{r_j} \leq A r_j^l \ .$$

Using Cauchy estimates (Lemma 5) we have for every $\alpha \in \mathbb{N}^d$ such that $|\alpha| = 1$,

$$|\partial^\alpha f_j - \partial^\alpha f_{j-1}|_{\frac{r_j}{2}} \leq |f_j - f_{j-1}|_{r_j} \left(\frac{r_j}{2}\right)^{-1} \leq 2A r_j^{l-1} \ .$$

Then by $(\mathcal{P}_k)$, f_j converges uniformly to $f \in C^s$ for any $s \leq l-1$ and

$$|f|_{C^s} \leq \frac{C(l-1)A}{\mu(1-\mu)} r_0^{l-1-s}. \quad \blacksquare$$

Proof of Proposition 10 In view of point (ii) of Remark 16, we may suppose, without loss of generality, that

$$0 < l = s = \mu < 1 \ . \tag{127}$$

In this case, we have to prove that f_j converges uniformly on $\mathbb{R}^d$ to f and that

$$|f|_{C^\mu} \ = \ |f|_{C^0} + \sup_{0<|x-y|\leq 1} \frac{|f(x)-f(y)|}{|x-y|^\mu} \leq \frac{CA}{\mu(1-\mu)} \ .$$

Set $g_j = f_j - f_{j-1}$. First of all, let us prove that f_j converges uniformly on $\mathbb{R}^d$. For any $1 \le n \le N$, one has (recall that $r_j = r_0/2^j$)

$$\begin{aligned} \Big|\sum_{j=n}^{N} g_j\Big|_{C^0} &\le \sum_{j=n}^{N} |g_j|_{C^0} = \sum_{j=n}^{N} |f_j - f_{j-1}|_{C^0} \\ &= \sum_{j=n}^{\infty} A\left(\frac{r_0}{2^j}\right)^{\mu} = \frac{Ar_0^{\mu}}{2^{\mu n}}\frac{1}{1-2^{-\mu}} \end{aligned} \tag{128}$$

which converges to zero as n goes to $+\infty$; thus f_j converges to $f = \sum_{j=1}^{\infty} g_j$ uniformly on $\mathbb{R}^d$ and (setting $n = 1$ and $N = +\infty$ in the above estimates) we have

$$|f|_{C^0} \le Ar_0^{\mu}\frac{2^{-\mu}}{1-2^{-\mu}}.$$

Since $1 - 2^{-\mu} \ge \mu/2$ for each $\mu \in [0, 1]$, we get:

$$|f|_{C^0} \le \frac{2Ar_0^{\mu}}{\mu} < \frac{2A}{\mu(1-\mu)}r_0^{\mu} \le \frac{2A}{\mu(1-\mu)}. \tag{129}$$

In order to estimate the second part of $|f|_{C^{\mu}}$ we need to distinguish two cases according to whether $r_0 < |x - y| \le 1$ or $|x - y| \le r_0$.

First case: $r_0 < |x - y| \le 1$. Then using the second inequality in (129), we get

$$|f(x) - f(y)| \le 2\,|f|_{C^0} \le \frac{4A}{\mu(1-\mu)}r_0^{\mu} \le \frac{4A}{\mu(1-\mu)}\,|x-y|^{\mu}. \tag{130}$$

Second case: $0 < |x - y| \le r_0$. Then there exists a N in $\mathbb{N}$ such that:

$$\frac{r_0}{2^{N+1}} \le |x-y| \le \frac{r_0}{2^N}. \tag{131}$$

The second inequality in (129) is equivalent to

$$\left(2^N r_0^{-1}\right)^{1-\mu} \le |x-y|^{\mu-1}. \tag{132}$$

Now,

$$\begin{aligned} |f(x) - f(y)| &\le \sum_{j=1}^{\infty} |g_j(x) - g_j(y)| \\ &= \sum_{j=1}^{N} |g_j(x) - g_j(y)| + \sum_{j=N+1}^{\infty} |g_j(x) - g_j(y)|. \end{aligned}$$

Let us estimate separately the two sums. Using Cauchy estimates, we have (recall that, by hypothesis, $|g_j|_{r_j} \le A r_j^{\mu}$):

$$|\partial_x g_j|_{\frac{r_j}{2}} \le |g_j|_{r_j} \left(\frac{r_j}{2}\right)^{-1} \le |g_j|_{C^0} \left(\frac{r_j}{2}\right)^{-1} \le 2A r_j^{\mu-1}.$$

Hence,

$$\begin{aligned} \sum_{j=1}^{N} |g_j(x) - g_j(y)| &\le 2A\,|x-y| \sum_{j=1}^{N} \left(\frac{2^j}{r_0}\right)^{1-\mu} \\ &= 2A\,|x-y| \left((r_0^{-1})^{1-\mu}\right) \frac{2^{(N+1)(1-\mu)} - 1}{2^{1-\mu} - 1}. \end{aligned}$$

Since $2^t - 1 \ge t/2$ for any $t \ge 0$ (and since $2^{1-\mu} \le 2$), by (132) we get

$$\begin{aligned} \sum_{j=1}^{N} |g_j(x) - g_j(y)| &\le 2A\,|x-y| \left((r_0^{-1})^{1-\mu}\right) \left(4 \frac{2^{(N+1)(1-\mu)}}{1-\mu}\right) \\ &\le 16A \frac{|x-y|^{\mu}}{1-\mu}. \end{aligned} \tag{133}$$

Next (using again $1 - 2^{-\mu} \ge \mu/2$ for $\mu \in [0,1]$ and (129))

$$\begin{aligned} \sum_{j=N+1}^{\infty} |g_j(x) - g_j(y)| &\le 2 \sum_{j=N+1}^{\infty} |g_j|_{C^0} \le 2A \sum_{j=N+1}^{\infty} \left(\frac{r_0}{2^j}\right)^{\mu} \\ &\le 2A \left(\frac{r_0}{2^{N+1}}\right)^{\mu} \frac{1}{1-2^{-\mu}} \le 4A \frac{|x-y|^{\mu}}{\mu}. \end{aligned} \tag{134}$$

Putting (133) and (134) together, we get:

$$\begin{aligned} |f(x) - f(y)| &\le \frac{16A}{1-\mu} |x-y|^{\mu} + \frac{4A}{\mu} |x-y|^{\mu} \\ &\le \frac{16A}{\mu(1-\mu)} |x-y|^{\mu} \ . \end{aligned} \tag{135}$$

Thus, by (129) and (135), we get

$$|f|_{C^{\mu}} \le \frac{CA}{\mu(1-\mu)} \ ,$$

with $C = 18$. ∎

2.2 A KAM theorem in C^k category

In this section we extend KAM theory to the finitely differentiable case.

For simplicity, we shall discuss only the nearly integrable case: in particular we prove the following generalization of Corollary 1.

Theorem 11 *Let $\omega \in \mathbb{R}^d$ be (γ, τ)-diophantine, let $l > l_0 := 4\tau + 3$, let $V \in C^l(\mathbb{T}^d)$ and let $M > 0$ be such that $|V|_{C^l} \le M$. There exists a constant $\kappa = \kappa(l, d, \tau, \gamma, M) > 1$ such that if*

$$\kappa \, (|V_x|_{C^0})^{\frac{l-l_0}{l-1}} \le 1, \tag{136}$$

then there exists a function $u : \mathbb{T}^d \to \mathbb{R}^d$, which belongs to C^s for all $s \le l - l_0$ not integer, satisfying

$$D^2 u + V_x(\theta + u) = 0 \ , \tag{137}$$

and

$$|u|_{C^s} \le \frac{\kappa}{\mu(1-\mu)} \, (|V_x|_{C^0})^{\frac{l-l_0-s}{l-1}} \ , \qquad \mu := s - [s] \ . \tag{138}$$

If $s < 2$, from the proof given below it follows easily (as relation (137) suggests) that the double directional derivative $D^2 u$ exists and is a $C^s(\mathbb{T}^d)$ function (**exercise**).

Proof Let

$$\varepsilon := (|V_x|_{C^0})^{\frac{1}{l-1}} \ , \qquad \xi_j := \frac{\varepsilon}{2^j} \ , \qquad \hat{\xi}_j := \frac{\xi_{j+1}}{2} = \frac{\xi_j}{4} \ . \tag{139}$$

Notice that (136) implies that $\varepsilon < 1$.

By Proposition 8, the real-analytic functions $V_j := \mathcal{S}_{\xi_j} V \in \mathcal{R}_{\xi_j}$ satisfy

$$\left| \partial^\alpha V_j(x) - \sum_{|\beta| \le l - |\alpha|} \frac{\partial^{\beta+\alpha} V(\,\mathrm{Re}\, x)}{\beta!} (i \,\mathrm{Im}\, x)^\beta \right| \le c \, |V|_{C^l} \xi_j^{l-|\alpha|} \ , \tag{140}$$

for every $x \in \Delta^d_{\xi_j}$ and $|\alpha| \le l$. Denote by $\mathcal{E}_j$ the differential operator

$$\mathcal{E}_j : v \to \mathcal{E}_j(v) := D^2 v + \partial_x V_j(\theta + v) \ .$$

The strategy is to construct a sequence of real-analytic functions $u_j \in \mathcal{R}_{\hat{\xi}_j}$, satisfying $\mathcal{E}_j(u_j) = 0$ and to obtain, by Proposition 10, the solution u as uniform limit of the $u_j's$.

For the purpose of this proof, we denote by " const " (possibly different) constants depending on l, d and τ and by κ_i suitable constants depending on l, d, τ, γ, M. *The constant* κ *in* (136) *is assumed to be such that*

$$\kappa \geq \kappa_i \ , \qquad \forall \ i \ . \tag{141}$$

As a preliminary remark, we observe that, for any $|\alpha| \leq 3$ and for any $j \geq 0$,

$$\sup_{\Delta^d_{\xi_j}} |\partial^\alpha V_j| \leq \text{ const } M \ , \tag{142}$$

as it follows from (140) and the fact that $\varepsilon < 1$:

$$\begin{aligned} |\partial^\alpha V_j(x)| \ &\leq \ \Big|\partial^\alpha V_j(x) - \sum_{|\beta|\leq l-3} \frac{\partial^{\beta+\alpha} V(\,\mathrm{Re}\, x)}{\beta!}(i\,\mathrm{Im}\, x)^\beta\Big| \\ &\qquad + \Big|\sum_{|\beta|\leq l-3} \frac{\partial^{\beta+\alpha} V(\,\mathrm{Re}\, x)}{\beta!}(i\,\mathrm{Im}\, x)^\beta\Big| \\ &\leq \ \text{const } \Big(M\xi_j^{l-3} + |V|_{C^3} + M\xi_j\Big) \\ &\leq \ \text{const } M \ . \end{aligned}$$

We proceed in three steps: construction of u_0; inductive construction of u_j ($j \geq 1$); construction of u as $\lim u_j$.

Step 1: construction of u_0. We want to apply the KAM Theorem 6 with $v \equiv 0$, $\xi_* = \xi_0$, $\xi = \xi_0/2 := \varepsilon/2$, $\bar{\xi} = \hat{\xi}_0 = \varepsilon/4$. We start by estimating $\mathcal{E}_0(v) = \mathcal{E}_0(0)$. Let $\theta \in \Delta^d_\xi$, i.e., $|\,\mathrm{Im}\,\theta_k| \leq \varepsilon/2$. Then

$$\begin{aligned} |\mathcal{E}_0(0)(\theta)| \ &:= \ |\partial_x V_0(\theta)| \\ &\leq \ \Big|\partial_x V_0(\theta) - \sum_{|\beta|\leq l-1} \frac{\partial^\beta \partial_x V(\,\mathrm{Re}\,\theta)}{\beta!}(i\,\mathrm{Im}\,\theta)^\beta\Big| \\ &\qquad + \Big|\sum_{|\beta|\leq l-1} \frac{\partial^\beta \partial_x V(\,\mathrm{Re}\,\theta)}{\beta!}(i\,\mathrm{Im}\,\theta)^\beta\Big| \end{aligned}$$

$$
\begin{aligned}
&\overset{(*)}{\leq} \quad \text{const}\Big(|V|_{C^l}\varepsilon^{l-1} + \sum_{j\leq l-1}|V_x|_{C^j}\varepsilon^j\Big)\\
&\overset{(**)}{\leq} \quad \text{const}\Big(|V|_{C^l}\varepsilon^{l-1} + \sum_{j\leq l-1}|V_x|_{C^0}^{\frac{l-1-j}{l-1}}\ |V|_{C^l}^{\frac{j}{l-1}}\ \varepsilon^j\Big)\\
&\overset{(\dagger)}{=} \quad \text{const}\Big(|V|_{C^l} + \sum_{j\leq l-1}|V|_{C^l}^{\frac{j}{l-1}}\Big)\varepsilon^{l-1}\\
&\leq \quad \text{const}\ \ \max\{1, M\}\ \varepsilon^{l-1}\\
&=: \quad \kappa_1\ \varepsilon^{l-1}\ , \qquad (143)
\end{aligned}
$$

where: $(*)$ is implied by (140); $(**)$ is the convexity estimate (123) with $k=0$, $m=j$ and l replaced by $l-1$; $(\dagger)$ is the definition of ε. Thus, recalling the notations in Theorem 6, and observing that $\alpha \leq \max\{1,\ \text{const}\ M/\gamma^2\}$ by (142), that $\lambda = \eta = 1$ and recalling the definition of $\hat{\xi}_j$, we see that (75) is implied, in our case, by

$$\kappa_2\ \varepsilon^{l-l_0} \leq 1\ , \qquad (144)$$

for a suitable $\kappa_2 > 1$. Such condition, in view of (141) and of the definition of ε, is implied by (136). Therefore, by Theorem 6, there exists a function $u_0 \in \mathcal{R}_{\hat{\xi}_0}$ such that

$$\mathcal{E}_0(u_0) = 0$$

and such that

$$\|u_0\|_{\hat{\xi}_0}\ ,\ \|\partial_\theta u_0\|_{\hat{\xi}_0} \leq \kappa_3\ \varepsilon^{l-l_0} \leq 1\ , \qquad (145)$$

where $\kappa_3 = K\kappa_2$, K being the constant in (77); the second inequality holds because of (141) and (136). The first step is completed.

Step 2: construction of $\{u_j\}$. We proceed inductively constructing u_{j+1}, for $j \geq 0$, via Theorem 6 by taking $v = u_j$ as approximate solution. We also take $\xi_* = \xi_{j+1}$, $\xi = \hat{\xi}_j$ and $\bar{\xi} = \hat{\xi}_{j+1}$. The parameter α, in view of (142), is uniformly bounded by $\max\{1,\ \text{const}\ M/\gamma^2\}$.

We, now, *assume that, for* $0 \leq k \leq j$, *there exist functions* $u_k \in \mathcal{R}_{\hat{\xi}_k}$ *such that*

$$\mathcal{E}_k(u_k) = 0$$

and such that[22]

$$\|u_k - u_{k-1}\|_{\hat{\xi}_k}\ ,\ \|\partial_\theta(u_k - u_{k-1})\|_{\hat{\xi}_k} \leq \kappa_4\ \Big(\frac{\varepsilon}{2^k}\Big)^{l-l_0}\ , \qquad (1 \leq k \leq j) \qquad (146)$$

[22]For $j = 0$ (146) is obviously replaced by the already proven (145).

for a suitable $\kappa_4 \geq \kappa_3$ specified below; finally we assume that $0 \leq k \leq j$:

$$\|\partial_\theta u_k\|_{\hat{\xi}_k} \leq 1 \;, \qquad \|(I + \partial_\theta u_k)^{-1}\|_{\hat{\xi}_k} \leq 2 \;. \tag{147}$$

Notice that $\|\partial_\theta u_k\|_{\hat{\xi}_k} \leq 1$ implies[23] $\|I + \partial_\theta u_k\|_{\hat{\xi}_k} \leq 2$ so that, if (147) holds, then in Theorem 6 one can take $\lambda = \eta = 2$. The inductive assumption (147) and the definitions in (139) imply that (74) is satisfied: in fact, if $\theta \in \Delta^d_{\hat{\xi}_j}$, then

$$\begin{aligned} |\operatorname{Im} u_j(\theta)| &= |\operatorname{Im}(u_j(\theta) - u_j(\operatorname{Re}\theta))| \\ &\leq |u_j(\theta) - u_j(\operatorname{Re}\theta)| \\ &\leq \|\partial_\theta u_j\|_{\hat{\xi}_j} \hat{\xi}_j \\ &\leq \hat{\xi}_j := \xi_{j+1} - \hat{\xi}_j \;. \end{aligned}$$

We need, now, to estimate $\mathcal{E}_{j+1}(u_j)$. Since, by the inductive assumption, $\mathcal{E}_j(u_j) = 0$, we find, for $\theta \in \Delta^d_{\hat{\xi}_j}$ and because of (140),

$$\begin{aligned} |\mathcal{E}_{j+1}(u_j)(\theta)| &:= |D^2 u_j(\theta) + \partial_x V_{j+1}(\theta + u_j)| \\ &= |\partial_x V_{j+1}(\theta + u_j) - \partial_x V_j(\theta + u_j)| \\ &\leq \Big|\partial_x V_{j+1}(\theta + u_j) - \sum_{|\beta| \leq l-1} \frac{\partial_x^\beta \partial_x V(\operatorname{Re}(\theta + u_j))}{\beta!} (i \operatorname{Im}(\theta + u_j))^\beta \Big| \\ &\quad + \Big| \sum_{|\beta| \leq l-1} \frac{\partial_x^\beta \partial_x V(\operatorname{Re}(\theta + u_j))}{\beta!} (i \operatorname{Im}(\theta + u_j))^\beta - \partial_x V_j(\theta + u_j) \Big| \\ &\leq \text{const}\, |V|_{C^l} \left(\xi_{j+1}^{l-1} + \xi_j^{l-1} \right) | \\ &\leq \text{const}\, M \xi_j^{l-1} \,. \end{aligned} \tag{148}$$

Thus, (75) becomes, in the present case,

$$\kappa_5 \left(\frac{\varepsilon}{2^{j+1}} \right)^{l - l_0} \leq 1 \;, \tag{149}$$

for a suitable $\kappa_5 > 1$. We now define[24] κ_4 as

$$\kappa_4 := \max\{\kappa_3 \,,\, K\kappa_5\} \;. \tag{150}$$

[23] We are choosing norms for which $\|I\| = 1$.

[24] This is well defined since in the computations leading to the definition of κ_5 the inductive hypotheses (146) have not been used.

Notice that condition (149) is again implied by (136). Thus, by Theorem 6, there exists a function $u_{j+1} \in \mathcal{R}_{\hat{\xi}_{j+1}}$ such that

$$\mathcal{E}_{j+1}(u_{j+1}) = 0$$

and such that

$$\|u_{j+1} - u_j\|_{\hat{\xi}_{j+1}}, \|\partial_\theta (u_{j+1} - u_j)\|_{\hat{\xi}_{j+1}} \leq \kappa_4 \left(\frac{\varepsilon}{2^{j+1}}\right)^{l-l_0}, \qquad (151)$$

which is exactly (146) with $k = j + 1$. The bounds (151) together with the condition (136) easily implies that the inductive assumptions (147) are satisfied also for[25] $k = j + 1$, allowing to iterate the inductive procedure indefinitely. The second step is completed.

Step 3: construction of u. At this point we can apply Proposition 10 (see also Remark 16) with: l replaced by $l - l_0$; $f_j = u_j - u_0$; $r_j = \hat{\xi}_j := \varepsilon/2^{j+2}$ (so that $r_0 = \varepsilon/4$); $A = \kappa_4 4^{l-2}$ (compare (124), (151) and the choice of r_j). The thesis of the theorem now follows at once from Proposition 10. ∎

Exercise Discuss the C^∞ case.

Exercise* Extend Theorem 6 to the differentiable case.

[25]**Exercise** Fill in the details.

3 Appendix A

Lemma 12 *If $T = T(\theta)$ is a strictly positive and symmetric real matrix for each $\theta \in \mathbb{T}^d$, then*

$$\|\langle T \rangle\|^{-1} \le \sup_{\theta \in \mathbb{T}^d} \left\| T^{-1} \right\| .$$

Proof By hypotheses there exists an orthogonal matrix P such that $P^T T P$ is diagonal. Let $\{\lambda_i : i = 1, \ldots, d\}$ be the spectrum of T and y a vector whose coordinates are y_i for $i = 1, \ldots, d$ in the basis where T is diagonal. Then, we have:

$$\begin{aligned} (Ty) \cdot y &= \sum_{i=1}^{d} \lambda_i y_i^2 \\ &\ge \min_{i \in \{1,\ldots,d\}} \{\lambda_i\} \|y\|^2 . \end{aligned}$$

But, $\min_{i \in \{1,\ldots,d\}} \{\lambda_i\} = \|T^{-1}\|^{-1}$, thus,

$$\begin{aligned} (Ty) \cdot y &\ge \frac{\|y\|^2}{\|T^{-1}\|} \\ &\ge \frac{\|y\|^2}{\sup_{\theta \in \mathbb{T}^d} \|T^{-1}\|} . \end{aligned}$$

Set $y = \langle T \rangle^{-1} x$. Taking the average of the last expression, we get

$$\left\langle \left(T \langle T \rangle^{-1} x \right) \cdot y \right\rangle \ge \left\langle \frac{\|y\|^2}{\sup_{\theta \in \mathbb{T}^d} \|T^{-1}\|} \right\rangle ,$$

i.e.,

$$\left(\langle T \rangle \langle T \rangle^{-1} x \right) \cdot y \ge \frac{\|y\|^2}{\sup_{\theta \in \mathbb{T}^d} \|T^{-1}\|} .$$

Finally, using Schwarz's inequality in the left-hand side, and dividing by $\|y\|$, we get

$$\|x\| \ge \frac{\|y\|}{\sup_{\theta \in \mathbb{T}^d} \|T^{-1}\|} ,$$

i.e.,

$$\sup_{\theta\in\mathbb{T}^d}\left\|T^{-1}\right\| \geq \frac{\|y\|}{\|x\|} = \frac{\|\langle T\rangle^{-1}x\|}{\|x\|} \ .$$

Since this is true for all x, the result follows. ∎

Proof of Lemma 4 To prove (66), we observe that by Parseval identity, if $v\in\mathbb{C}^d$ is such that $|\operatorname{Im} v| < \xi$, then

$$\sum_{n\in\mathbb{Z}^d} |f_n|^2 e^{-2(n\cdot v)} = \int_{\mathbb{T}^d} |f(u+iv)|^2 \ du \leq \|f\|^2_{L^2,\Delta^d_\xi} \ . \tag{152}$$

Thus, for any $n\in\mathbb{Z}^d$ and any v as above,

$$e^{-n\cdot v}|f_n| \leq \|f\|_{L^2,\Delta^d_\xi} \ ;$$

and choosing[26] $v = -(sign\, n_1, ..., sign\, n_d)(\xi-\varepsilon)$ we get

$$e^{|n|\xi} \ |f_n| \leq \|f\|_{L^2,\Delta^d_\xi} \ e^{\varepsilon|n|} \ ;$$

letting $\varepsilon\to 0$, (66) follows.

Let us turn to (67). For the purpose of the following argument we let $|n|$ denote the Euclidean norm also for integer vectors.

The first inequality in (67) is obvious; in order to establish the second one for a fixed vector $x\in\Delta^d_{\xi-\delta}$ we define the set

$$I_0 = \{n\in\mathbb{Z}^d : \quad (n\cdot \operatorname{Im} x) \leq -\frac{|n|(\xi-\delta)}{2}\} \,,$$

and let

$$\mu := \frac{\xi}{\xi-\delta} \ .$$

Observe that $(n\cdot \operatorname{Im} x) \leq -\frac{|n|(\xi-\delta)}{2}$ is equivalent to

$$-n\cdot \operatorname{Im} x \leq -n\cdot\mu \operatorname{Im} x - \frac{|n|\delta}{2} \ . \tag{153}$$

[26] Here, we let $sign\, a$ be 1 if $a\geq 0$ and (-1) otherwise.

Thus, by (153), Schwarz inequality and (152) (since $\mu x \in \Delta^d_\xi$), we get

$$\begin{aligned}\sum_{n\in I_0} |f_n| e^{-(n\cdot \operatorname{Im} x)} &\le \sum_{n\in I_0} |f_n| e^{-(n\cdot\mu \operatorname{Im} x)} e^{-|n|\delta} \\ &\le \left(\sum_{n\in\mathbb{Z}^d} |f_n|^2 e^{-2(n\cdot\mu\operatorname{Im} x)}\right)^{\frac12} \left(\sum_{n\in\mathbb{Z}^d} e^{-|n|\delta}\right)^{\frac12} \\ &\le \frac{c_1^{\frac12}}{\delta^{\frac d2}} \|f\|_{L^2,\,\Delta^d_\xi}\,,\end{aligned}$$

where[27]

$$c_1 = c_1(d) = \sup_{0<\lambda\le 1} \sum_{n\in\mathbb{Z}^d} \lambda^d e^{-|n|\lambda} < \infty\,.$$

It is easy to see that there exist an integer $s = s(d)$ and a collection of s unit vectors $e_1, \ldots, e_s$ in $\mathbb{R}^d$, such that for ever $y \in \mathbb{R}^d$ there exists $\sigma \in \{1,\ldots,s\}$ with[28]

$$(y\cdot e_\sigma) > \frac{|y|}{2}\,.$$

Now, every nonzero integer vector outside I_0, lies in one the sets I_σ defined as

$$I_\sigma := \left\{ n\in\mathbb{Z}^d : \quad (n\cdot\operatorname{Im} x) > -\frac{|n|(\xi-\delta)}{2}, \quad (n\cdot e_\sigma) > \frac{|n|}{2} \right\}\,.$$

But, (using again Schwarz inequality and (152)),

$$\begin{aligned}\sum_{n\in I_\sigma} |f_n| e^{-(n\cdot\operatorname{Im} x)} &\le \sum_{n\in I_\sigma} |f_n| e^{|n|\frac{(\xi-\delta)}{2}} \\ &\le \left(\sum_{n\in I_\sigma} |f_n|^2 e^{|n|\xi}\right)^{\frac12} \left(\sum_{n\in I_\sigma} e^{-|n|\delta}\right)^{\frac12}\end{aligned}$$

27

$$\sum e^{-|n|\delta} = \frac{1}{\delta^d}\Big(\delta^d \sum e^{-|n|\delta}\Big) \le \frac{1}{\delta^d} \sup_{0<\lambda<1} \Big(\lambda^d \sum e^{-|n|\lambda}\Big).$$

28 Given $v\in S^{d-1} := \{y\in\mathbb{R}^d : |y| = 1\}$, let $C_v := \{w\in S^{d-1} : w\cdot v > \frac12\}$. Then C_v is an open (in the relative topology) neighborhood of v and $\{C_v : v\in S^{d-1}\}$ is an open cover of the compact set S^{d-1}. Thus there exist unit vectors $v_1 =: e_1,\ldots,v_s =: e_s$ such that $S^{d-1}\subset \cup_{j=1}^s C_{e_j}$: this is equivalent to the claim.

$$\leq \left(\sum_{n\in\mathbb{Z}^d} |f_n|^2 e^{-2(n\cdot\xi e_\sigma)}\right)^{\frac{1}{2}} \left(\sum_{n\in\mathbb{Z}^d} e^{-|n|\delta}\right)^{\frac{1}{2}}$$

$$\leq \frac{c_1^{\frac{1}{2}}}{\delta^{\frac{d}{2}}} \|f\|_{L^2,\Delta_\xi^d} .$$

Inequality (67) now follows and one can take $c_0 = (s+1)\sqrt{c_1}$. ∎

4 Appendix B (Fourier Norms)

For $\xi \geq 0$ let us define the space

$$\hat{\mathcal{R}}_\xi(\mathbb{T}^d, \mathbb{R}^N) := \{f \in C(\mathbb{T}^d, \mathbb{R}^N) \text{ s. t. } \|f\|_\xi^{\hat{}} := \sum_{n\in\mathbb{Z}^d} |f_n| e^{|n|\xi} < \infty\} .$$

The space $\hat{\mathcal{R}}_\xi(\mathbb{T}^d, \mathbb{R}^N)$ is a Banach space with respect to the norm $\|f\|_\xi^{\hat{}}$. Moreover, since

$$|f_n| \leq \|f\|_\xi^{\hat{}} e^{-|n|\xi} ,$$

(when $\xi > 0$) the function f has a holomorphic extension to the complex strip Δ_ξ^d. Notice that

$$\sup_{\Delta_\xi^d} |f| = \sup_{\Delta_\xi^d} \Big| \sum_{n\in\mathbb{Z}^d} f_n e^{inx} \Big| \leq \sum_{n\in\mathbb{Z}^d} |f_n| e^{|n|\xi} = \|f\|_\xi^{\hat{}} , \tag{154}$$

and if $0 \leq \xi' < \xi$, then $\hat{\mathcal{R}}_\xi \subsetneqq \hat{\mathcal{R}}_{\xi'}$ (**exercise**). In particular, (154) shows that $f \in \hat{\mathcal{R}}_\xi$ admits a holomorphic and *bounded* extension to Δ_ξ^d.

Lemma 13 *Let $f \in \hat{\mathcal{R}}_\xi(\mathbb{T}^d, X)$, $p \in \mathbb{Z}$, $\alpha \in \mathbb{N}^d$ be such that $|p| + |\alpha| > 0$. If $p > 0$, assume either $|\alpha| > 0$ or $\langle f \rangle = 0$. Let $0 < \delta \leq \xi$. Then*

$$\|D^{-p}\partial^\alpha f\|_{\xi-\delta}^{\hat{}} \leq C_{p,\alpha}(\omega) \|f\|_\xi^{\hat{}} ,$$

where

$$C_{p,\alpha}(\omega) := \sup_{n\neq 0} \frac{|n^\alpha| e^{-\delta|n|}}{|\omega\cdot n|^p} .$$

If ω is (γ,τ)-diophantine, then

$$C_{p,\alpha}(\omega) \le \begin{cases} \dfrac{(p\tau+|\alpha|)!}{\gamma^p \delta^{p\tau+|\alpha|}}\,, & \text{for } p \ge 0\ , \\ \\ (\sup|\omega_i|)^{|p|}\,\dfrac{(|p|+|\alpha|)!}{\delta^{|p|+|\alpha|}}\,, & \text{for } p < 0\ . \end{cases}$$

Proof

$$\|D^{-p}\partial^\alpha f\|^{\hat{}}_{\xi-\delta} = \Big\|\sum_{n\neq 0}\frac{n^\alpha f_n}{(\omega\cdot n)^p}e^{inx}\Big\|^{\hat{}}_{\xi-\delta} = \sum_{n\neq 0}\frac{|f_n||n^\alpha|}{|\omega\cdot n|^p}e^{|n|(\xi-\delta)} \le C_{p,\alpha}(\omega)\|f_n\|^{\hat{}}_\xi\ .$$

If $p \ge 0$ and ω is (γ,τ)-diophantine, then

$$C_{p,\alpha}(\omega) \le \sup_{n\neq 0}\frac{|n|^{|\alpha|+\tau p}}{\gamma^p}e^{-|n|\delta}\ .$$

The function on the right is of the form $g(t) = t^a e^{-\delta|t|}$, $t > 0$, $a \ge 0$, and has a maximum at the point $t_m = \frac{a}{\delta}$ such that $g(t_m) = a^a(e\delta)^{-a} \le a!\delta^{-a}$, where if a is not an integer, we define[29] $a! = ([a]+1)!$

If now $p < 0$, one can repeat the previous arguments using the fact that in this case $|\omega\cdot n|^{|p|} \le (\sup|\omega_i|)^{|p|}|n|^{|p|}$. ∎

In particular, if $\langle f\rangle = 0$, then

$$\|D^{-1}f\|^{\hat{}}_{\xi-\delta} \le \|f\|^{\hat{}}_\xi\frac{\tau!}{\gamma\delta^\tau}\ . \tag{155}$$

Lemma 14 *Let $f \in \hat{\mathcal{R}}_\xi(\mathbb{T}^d, X)$, $g \in \hat{\mathcal{R}}_\xi(\mathbb{T}^d, Y)$, with X and Y tensor spaces ($\mathbb{R}^N$, matrices or higher dimensional tensors) and assume that the product fg is well defined. Then*

$$\|f\,g\|^{\hat{}}_\xi \le \|f\|^{\hat{}}_\xi\|g\|^{\hat{}}_\xi\ .$$

[29] If $a \in \mathbb{N}$, then $\left(\frac{a}{e}\right)^a = \frac{a^a}{1+a+\frac{a^2}{2}+\ldots+\frac{a^a}{a!}+\ldots} \le \frac{a^a}{\frac{a^a}{a!}} = a!$. If a is non-integer, then we can repeat the same argument eliminating all terms in the Taylor expansion for the exponent except $1 + a^{[a]+1}/([a]+1)!$.

Proof Indeed,

$$\begin{aligned}\|f\,g\|^{\hat{}}_{\xi} &= \sum_n |(fg)_n| e^{|n|\xi} = \sum_n \left| \sum_m f_{n-m} g_m \right| e^{|n|\xi} \\ &\leq \sum_{n,m} |f_{n-m}||g_m| e^{|n-m|\xi} e^{|m|\xi} = \|f\|^{\hat{}}_{\xi} \|g\|^{\hat{}}_{\xi}. \quad \blacksquare\end{aligned}$$

Lemma 15 *Let* $V \in \hat{\mathcal{R}}_{\xi_*}(\mathbb{T}^d, \mathbb{R})$, $g \in \hat{\mathcal{R}}_{\xi}(\mathbb{T}^d, \mathbb{R}^d)$, *with* $\xi_* > \xi$. *If* $\|g_i\|^{\hat{}}_{\xi} \leq \xi_* - \xi$ *for all* $i = 1, \ldots, d$, *then* $\|V(x + g(x))\|^{\hat{}}_{\xi} \leq \|V\|^{\hat{}}_{\xi_*}$.

Proof Using Lemma 14, the fact that $\|\cdot\|^{\hat{}}_{\xi}$ is a norm, one finds

$$\begin{aligned}\|V(x + g(x))\|^{\hat{}}_{\xi} &= \sum_n |(V(x + g(x))_n| e^{|n|\xi} \\ &= \sum_n \Big| \Big(\sum_m V_m e^{im\cdot(x+g(x))} \Big)_n \Big| e^{|n|\xi} \\ &= \sum_n \Big| \sum_m V_m (e^{im\cdot g(x)})_{n-m} \Big| e^{|n|\xi} \\ &\leq \sum_{n,m} |V_m| |(e^{im\cdot g(x)})_{n-m}| e^{|n|\xi} \\ &\leq \sum_{n,m} |V_m| \left| \sum_{j\geq 0} \left[\frac{(im\cdot g(x))^j}{j!} \right]_{n-m} \right| e^{|n-m|\xi} e^{|m|\xi} \\ &\leq \sum_{m,j} \frac{|V_m|}{j!} e^{|m|\xi} \sum_n \left| \left[(im\cdot g(x))^j \right]_{n-m} \right| e^{|n-m|\xi} \\ &\leq \sum_{m,j} \frac{|V_m|}{j!} e^{|m|\xi} (\|m\cdot g\|^{\hat{}}_{\xi})^j \\ &\leq \sum_{m,j} \frac{|V_m|}{j!} e^{|m|\xi} (\sum_{i=1}^d m_i \|g_i\|^{\hat{}}_{\xi})^j \\ &\leq \sum_{m,j} \frac{|V_m|}{j!} e^{|m|\xi} (\sup_i \|g_i\|^{\hat{}\,j}_{\xi}) |m|^j \\ &\leq \sum_{m,j} \frac{|V_m|}{j!} e^{|m|\xi} (\xi_* - \xi)^j |m|^j \\ &= \sum_m |V_m| e^{|m|\xi} e^{(\xi_*-\xi)|m|} = \|V\|^{\hat{}}_{\xi_*}. \quad \blacksquare\end{aligned}$$

The Fourier norm $\|\cdot\|_\xi^{\hat{}}$ and the sup-norm $\|\cdot\|_\xi$ are *not* equivalent (**exercise**); however they are strictly related. We have already seen (compare (154)) that

$$\|f\|_\xi \le \|f\|_\xi^{\hat{}} \ , \tag{156}$$

which implies immediately[30]

$$\hat{\mathcal{R}}_\xi(\mathbb{T}^d) \subset \mathcal{R}_\xi(\mathbb{T}^d) \ . \tag{157}$$

We now prove that a weaker version of the converse of (156) is true. Let $\xi' > \xi > 0$ and assume that $f \in \mathcal{R}_{\xi'}$. Since, for every $n \in \mathbb{Z}^d$,

$$|f_n| \le \|f\|_{\xi'} e^{-|n|\xi'} \ ,$$

we have

$$\|f\|_\xi^{\hat{}} = \sum_n |f_n| e^{|n|\xi} \le \|f\|_{\xi'} \sum_n e^{-|n|(\xi'-\xi)} \ .$$

But (for suitable positive constants $c(d), C(d)$)

$$\sum_{n\in\mathbb{Z}^d} e^{-|n|(\xi'-\xi)} \le c(d) \int_{\mathbb{R}^d} e^{-|x|(\xi'-\xi)} dx = \frac{c(d)}{(\xi'-\xi)^d} \int_{\mathbb{R}^d} e^{-|y|} dy = \frac{C(d)}{(\xi'-\xi)^d} \ ,$$

so that

$$\|f\|_\xi^{\hat{}} \le C(d) \ \frac{\|f\|_{\xi'}}{(\xi'-\xi)^d} \ . \tag{158}$$

This relation shows, in particular, that

$$\mathcal{R}_{\xi'}(\mathbb{T}^d) \subset \hat{\mathcal{R}}_\xi(\mathbb{T}^d) \ , \qquad \forall \ \xi' > \xi \ . \tag{159}$$

Exercise Give explicit upper bounds on $c(d)$ and $C(d)$.

[30] Actually $\hat{\mathcal{R}}_\xi(\mathbb{T}^d) \subsetneq \mathcal{R}_\xi(\mathbb{T}^d)$ (**exercise**).