

# Optimal linearization of vector fields on the torus in non-analytic Gevrey classes

Abed Bounemoura

**Abstract.** We study linear and non-linear small divisors problems in analytic and non-analytic regularity. We observe that the Bruno arithmetic condition, which is usually attached to non-linear analytic problems, can also be characterized as the optimal condition to solve the linear problem in some fixed non-quasi-analytic class. Based on this observation, it is natural to conjecture that the optimal arithmetic condition for the linear problem is also optimal for non-linear small divisors problems in any reasonable non-quasi-analytic classes. Our main result proves this conjecture in a representative non-linear problem, which is the linearization of vector fields on the torus, in the most natural non-quasi-analytic class, which is the Gevrey class. The proof follows Moser's argument of approximation by analytic functions, and uses works of Popov, Rüssmann and Pöschel in an essential way.

## 1. Introduction

The motivation of this work is to try to understand the discrepancy between an elementary remark and a deep theorem. Let  $\mathbb{T} := \mathbb{R}/\mathbb{Z}$ ,  $\alpha \in \mathbb{T}$  irrational and  $(q_n)_{n \in \mathbb{N}}$  be the sequence of denominators of the “best” rational approximations of  $\alpha$  (such approximations are given by expansion into continued fractions). Let us say that  $\alpha \in \mathcal{R}$  (for Rüssmann, see [32]) if  $\ln(q_{n+1}) = o(q_n)$  as  $n$  goes to infinity and  $\alpha \in \mathcal{B}$  (for Bruno, see [12]) if the sequence  $q_n^{-1} \ln(q_{n+1})$  is summable; obviously  $\mathcal{B} \subsetneq \mathcal{R}$ . The remark is that  $\alpha \in \mathcal{R}$  if and only if for any real-analytic function  $f: \mathbb{T} \rightarrow \mathbb{R}$  with zero average, there exists a real-analytic function  $g: \mathbb{T} \rightarrow \mathbb{R}$  satisfying the linear equation

$$g \circ R_\alpha - g = f, \tag{1.1}$$

where  $R_\alpha: \mathbb{T} \rightarrow \mathbb{T}$  denotes rotation by  $\alpha$ . The deep theorem, due to Yoccoz (see [37]), states that  $\alpha \in \mathcal{B}$  if and only if for any real-analytic orientation-preserving circle diffeomorphism  $F: \mathbb{T} \rightarrow \mathbb{T}$  with rotation number  $\alpha$  and sufficiently close to  $R_\alpha$ , there exists a real-analytic circle diffeomorphism  $\phi: \mathbb{T} \rightarrow \mathbb{T}$  close to the identity satisfying the conjugacy equation

$$\phi^{-1} \circ F \circ \phi = R_\alpha. \tag{1.2}$$

The linear equation (1.1) appears as a “linearization” of the non-linear equation (1.2), hence it may be surprising that the arithmetic condition needed to solve (1.2) is stronger than the one needed to solve (1.1); this reflects the failure of any kind of inverse function theorem (either the classical version or the Nash–Moser version) for this particular problem. Such a discrepancy does not appear in the smooth category; in this case (1.1) and (1.2) are solvable if and only if  $\alpha \in \mathcal{D}$  (for Diophantine) which can be expressed by the asymptotic condition  $\ln(q_{n+1}) = O(\ln(q_n))$ . In higher dimensions, results in the smooth case are exactly the same (and the proofs as well) but not in the analytic case: Yoccoz’s proof relies on a geometric construction which deeply uses the theory of holomorphic functions in one variable and continued fractions for which there are no known good analogues in higher dimensions. Yet one can still define a Bruno condition  $\mathcal{B}$  in any dimension and Rüssmann (see [34]) proved that it is a sufficient condition to solve the higher-dimensional analogue of (1.2); it is unknown whether such a condition is necessary.<sup>1</sup>

The purpose of this paper is to study these linear and non-linear small divisors problems in any dimension for regularities intermediate between smooth and analytic. Those regularities fall into two basic classes: the quasi-analytic classes, which are made of functions completely determined by their Taylor expansion at just one point exactly like analytic functions, and the others, the non-quasi-analytic classes, and we will be mainly concerned with these latter classes. At that point we should stress that even if one is interested only in analytic problems, non-quasi-analytic classes appear naturally in problems like (1.1) or (1.2). As we shall see below in Section 2.4, the Bruno condition  $\alpha \in \mathcal{B}$ , which can be characterized by the solvability of (1.2) in the analytic case, can also be characterized if one looks at (1.1) not in the analytic case but in some arbitrary yet fixed non-quasi-analytic class; apparently this has not been noticed before, even though this is elementary. Another way in which non-quasi-analytic classes show up is more classical: for various reasons one may be interested in the regularity of  $g = g(\alpha)$  in (1.1) and  $\phi = \phi(\alpha)$  in (1.2) as functions of  $\alpha$  (with  $\alpha$  varying in a closed set, so regularity has to be understood in the sense of Whitney). In general such a dependence is not analytic even if the data is analytic, it is always smooth (see [31]) and, most probably, it always belongs to some non-quasi-analytic class which depends on the arithmetic properties of  $\alpha$ ; in the best-case scenario (when  $\alpha$  is Diophantine) such a dependence belongs to some Gevrey non-analytic class ([24, 28]), which are the most studied non-quasi-analytic classes.

Our main result deals with those Gevrey classes, so let us informally state what it implies in the particular case of (1.1) and (1.2). Let  $f_n \in \mathbb{C}$ ,  $n \in \mathbb{Z}$  be the Fourier coefficients of an integrable function on  $\mathbb{T}$ ; it is well known that  $f$  is analytic if and only if  $|f_n| \leq e^{-s|n|}$  for some  $s > 0$  and for all  $n$  large enough; we shall say it belongs to the  $a$ -Gevrey class,  $0 < a \leq 1$ , if we have instead the asymptotic inequalities  $|f_n| \leq e^{-s|n|^a}$ . So 1-Gevrey is analytic, but  $a$ -Gevrey functions for  $0 < a < 1$  are non-quasi-analytic (indeed, for  $0 < a < 1$  it is not hard to construct explicitly a non-zero  $a$ -Gevrey function with an arbitrarily small support). It is easy to observe that (1.1) can be solved if and only

---

<sup>1</sup>According to Yoccoz, it is not (personal communication).

if  $\ln(q_{n+1}) = o(q_n^a)$ , while a consequence of the results in [9] and [23] is that (1.2) can be solved if  $q_n^{-a} \ln(q_{n+1})$  is summable. The latter condition is an adapted Bruno-type condition, which is optimal in the analytic case  $a = 1$  in view of the work of Yoccoz, and one may ask whether this is the case when  $0 < a < 1$ . The answer is negative, as the following informal theorem shows.

**Theorem 1.** *In the  $a$ -Gevrey class for  $0 < a < 1$ , we have  $\ln(q_{n+1}) = o(q_n^a)$  if and only if (1.1) can be solved if and only if (1.2) can be solved.*

This is a particular case of Theorem C that will be stated in Section 2.2, and is a corollary of our main result Theorem A which is a more quantitative statement, valid in any dimension and in the continuous setting (the discrete setting can be recovered by the usual suspension-section argument, which is rather straightforward in our context). So for Gevrey non-analytic classes, there is no Bruno-type condition; actually this should be true for any reasonable non-quasi-analytic class (where by reasonable we mean that such a class should be stable with respect to basic non-linear operations, which, unlike (1.1), seems needed if one wants to study (1.2)). For quasi-analytic classes, the problem is more subtle.

## 2. The linear problem

### 2.1. The cohomological equation

The continuous version of equation (1.1) in any dimension  $n \geq 2$  is as follows. Let  $\omega \in \mathbb{R}^n$  be a non-resonant vector, meaning that for all non-zero  $k \in \mathbb{Z}^n$ , the Euclidean inner product  $k \cdot \omega$  is non-zero. We denote by  $\mathcal{F}$  the space of real formal Fourier series, whose elements are of the form  $f = \sum_{k \in \mathbb{Z}^n} f_k e_k$  with  $f_k \in \mathbb{C}$  satisfying  $\bar{f}_k = f_{-k}$ , and where  $e_k(x) = e^{2\pi i k \cdot x}$ . The problem is to find  $g = \sum_{k \in \mathbb{Z}^n} g_k e_k \in \mathcal{F}$  and  $c \in \mathbb{R}$  such that

$$X_\omega g = f - c, \tag{L}$$

where  $X_\omega = \omega$  is the constant vector field on  $\mathbb{T}^n$ , acting formally on  $\mathcal{F}$  by derivation in the direction  $\omega$ . Equation (L), which is usually called the cohomological equation, is easily solved because it is a linear equation: both  $c = c(f)$  and  $g = g(f)$  depend linearly on  $f$ . Necessarily,

$$c = f_0 = \int_{\mathbb{T}^n} f \, dx,$$

where  $dx$  is the Haar measure on  $\mathbb{T}^n$ , and solutions of (L) are such that  $g_0 \in \mathbb{R}$  is arbitrary, while for non-zero  $k \in \mathbb{Z}^n$ ,

$$g_k = (i2\pi k \cdot \omega)^{-1} f_k. \tag{2.1}$$

The solution  $g$  is thus unique if we normalize the value of  $g_0$ : we shall always choose  $g_0 = 0$  and refer to  $g$  defined this way as the solution of (L). Now if  $f$  is regular, in order for  $g$  to be regular there will be competition between the decay of  $|f_k|$  and the inevitable

growth of  $|k \cdot \omega|$  as the norm of  $k$  goes to infinity. We shall quantify this by introducing a weight as follows.

**Definition 1.** A weight  $\varphi: [0, +\infty) \rightarrow [0, +\infty)$  is a continuous function, normalized by  $\varphi(0) = 0$ , which is non-decreasing and satisfies  $\ln(1 + t) = O(\varphi(t))$ .

One should keep in mind the following three main examples of weights:

- the “smooth” weight  $\varphi_0(t) := \ln(1 + t)$ ;
- the “Gevrey” weight  $\varphi_a(t) := t^a/a$ , for  $0 < a < 1$ ;
- the “analytic” weight  $\varphi_1(t) := t$ .

We have  $\varphi_a(t) \rightarrow \varphi_1(t)$  as  $a \rightarrow 1$ , whereas for  $t \geq 1$ ,  $\varphi_a(t) - 1/a = (t^a - 1)/a \rightarrow \ln(t)$  as  $a \rightarrow 0$ , which is equivalent to  $\varphi_0(t)$ . As we shall see below, our minimal growth requirement on  $\varphi$ , which can be written again as  $\varphi_0(t) = O(\varphi(t))$ , is no loss of generality for the problems we are interested in. To a given weight function we will associate a scale of regularity classes and a scale of arithmetic classes, where we shall always use the norm

$$|x| := \max_{1 \leq i \leq n} |x_i|, \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n.$$

**Definition 2.** Given a weight  $\varphi$  and a regularity parameter  $r \in \mathbb{R}$ , we define a regularity class

$$\mathcal{F}_r^\varphi := \{f \in \mathcal{F} \mid \|f\|_r := \sup_{k \in \mathbb{Z}^n} |f_k| e^{r\varphi(|k|)} < +\infty\}.$$

Observe that  $\mathcal{F}_r^\varphi$  is always a Banach space, and we have compact inclusions  $\mathcal{F}_{r_2}^\varphi \subseteq \mathcal{F}_{r_1}^\varphi$  whenever  $r_1 \leq r_2$ . Let us make some technical comment on the choice of the “Fourier-based”  $l_\infty$ -norm  $\|\cdot\|_r$  we used to define the space  $\mathcal{F}_r^\varphi$ . We could have easily chosen any  $l_p$ -norm for  $p \in [1, +\infty]$ :

$$\|f\|_r^p := \left( \sum_{k \in \mathbb{Z}^n} |f_k|^p e^{p r \varphi(|k|)} \right)^{1/p}.$$

These norms are non-equivalent and lead to different spaces, yet they are “comparable” so the (projective and inductive) limits we will consider below are the same. For the linear problem that we shall consider here, the precise choice of such a norm does not make any difference, but this is not the case for non-linear problems. The  $l_1$ -norm has the advantage that it defines a Banach algebra provided the weight  $\varphi$  is sub-additive (that is,  $\varphi(t + s) \leq \varphi(t) + \varphi(s)$ ) and thus leads to very simple product estimates, while the  $l_2$ -norm has the advantage that it defines a Hilbert space, and the norm is equivalent to a “space-based”  $L_2$ -norm which is characterized by the growth of the derivatives, provided  $f$  is a smooth function. We choose the  $l_\infty$ -norm since it will be more convenient for the (almost) characterization of such spaces (under assumptions on the weight  $\varphi$ ) by almost analytic extension and analytic approximation (see Section 4.2).

For the smooth weight  $\varphi_0(t) = \ln(1 + t)$ , we shall write  $\mathcal{F}_r^{\varphi_0} = \mathcal{F}_r^0$  and this space can be compared to the space  $C^r(\mathbb{T}^n)$  of Hölder 1-periodic functions on  $\mathbb{R}^n$  as follows:

$$\mathcal{F}_{r+d+\varepsilon}^0 \subseteq C^r(\mathbb{T}^n) \subset \mathcal{F}_r^0$$

for any  $\varepsilon > 0$ , hence  $\mathcal{F}_r^0$  contains sufficiently smooth functions when  $r$  is sufficiently large, and so the same remains true in general for  $\mathcal{F}_r^\varphi$  according to our assumption  $\varphi_0(t) = O(\varphi(t))$ .

**Definition 3.** Given a weight  $\varphi$  and an arithmetic parameter  $\tau > 0$ , we define an arithmetic class

$$\mathcal{A}_\tau^\varphi := \left\{ \omega \in \mathbb{R}^n \mid \gamma_\tau^{-1}(\omega) := \sup_{k \in \mathbb{Z}^n \setminus \{0\}} |2\pi k \cdot \omega|^{-1} e^{-\tau\varphi(|k|)} < +\infty \right\}.$$

Observe that  $\mathcal{A}_\tau^\varphi$  might be the empty set; however, the set  $\mathcal{A}_\tau^{\varphi_0} = \mathcal{A}_\tau^0$  is precisely the set of Diophantine vectors with exponent  $\tau$ , which is non-empty for  $\tau \geq n - 1$  and of full Lebesgue measure for  $\tau > n - 1$ . Hence our assumption  $\varphi_0(t) = O(\varphi(t))$  ensures that  $\mathcal{A}_\tau^\varphi$  is non-empty and of full measure for all  $\tau$  large enough, and this explains why our minimal growth assumption is no loss of generality. The following lemma is obvious.

**Lemma 2.** For any  $\omega \in \mathcal{A}_\tau^\varphi \neq \emptyset$  and any  $r \in \mathbb{R}$ , we have

$$\|g\|_{r-\tau} \leq \gamma_\tau^{-1}(\omega) \|f\|_r,$$

hence  $g \in \mathcal{F}_{r-\tau}^\varphi$  provided  $f \in \mathcal{F}_r^\varphi$ .

As a side remark, the above lemma holds true (up to a multiplicative constant which depends on the dimension) with  $\mathcal{F}_r^0$  replaced by the space  $C^r(\mathbb{T}^n)$ ,  $r > 0$  and  $r \notin \mathbb{N}$  (this follows from a Paley–Littlewood decomposition, and holds true actually for general Besov spaces allowing  $r \in \mathbb{R}$ ).

### 2.2. The projective limit

Next we look at the projective limit of the scale of Banach spaces  $(\mathcal{F}_r^\varphi)_{r>0}$  and its associated arithmetic class

$$\mathcal{F}_\infty^\varphi := \bigcap_{r>0} \mathcal{F}_r^\varphi, \quad \mathcal{A}_\infty^\varphi := \bigcup_{\tau>0} \mathcal{A}_\tau^\varphi,$$

which are also characterized by

$$f \in \mathcal{F}_\infty^\varphi \Leftrightarrow \varphi(|k|) = o(\ln(|f_k|^{-1})), \quad \omega \in \mathcal{A}_\infty^\varphi \Leftrightarrow \ln(|2\pi k \cdot \omega|^{-1}) = O(\varphi(|k|)).$$

Endowed with its natural projective limit topology,  $\mathcal{F}_\infty^\varphi$  is a Fréchet space. We shall say that a weight  $\varphi$  dominates another weight  $\psi$  if

$$\psi(t) = O(\varphi(t)) \Leftrightarrow \psi \preceq \varphi$$

and we obviously have

$$\psi \preceq \varphi \Rightarrow \mathcal{F}_\infty^\varphi \subseteq \mathcal{F}_\infty^\psi, \quad \mathcal{A}_\infty^\psi \subseteq \mathcal{A}_\infty^\varphi.$$

The converse holds true under some restrictions (for sub-additive weights for instance). By definition, an arbitrary weight dominates the smooth weight  $\varphi_0$  and since

$$\mathcal{F}_\infty^0 = C^\infty(\mathbb{T}^n), \quad \mathcal{A}_\infty^0 = \mathcal{D},$$

where  $\mathcal{D}$  is the set of all Diophantine vectors, for an arbitrary weight  $\varphi$  we have

$$\mathcal{F}_\infty^\varphi \subseteq C^\infty(\mathbb{T}^n), \quad \mathcal{D} \subseteq \mathcal{A}_\infty^\varphi.$$

so  $\mathcal{A}_\infty^\varphi$  is always of full measure. If  $\psi \preceq \varphi$  and  $\varphi \preceq \psi$  the weights are said to be equivalent and they define, in the above limits, the same regularity and arithmetic classes. Finally, for the analytic weight  $\varphi_1(t) = t$ ,  $\mathcal{F}_\infty^{\varphi_1} = \mathcal{F}_\infty^1$  identifies with entire periodic functions. The following lemma is well known.

**Lemma 3.** *The vector  $\omega \in \mathcal{A}_\infty^\varphi$  if and only if equation (L) can be solved in  $\mathcal{F}_\infty^\varphi$ . More precisely, if  $\omega \in \mathcal{A}_\infty^\varphi$ , then for any  $f \in \mathcal{F}_\infty^\varphi$  we have  $g \in \mathcal{F}_\infty^\varphi$  and if  $\omega \notin \mathcal{A}_\infty^\varphi$ , there exists  $f \in \mathcal{F}_\infty^\varphi$  such that  $g$  does not belong to*

$$\mathcal{F}_{-\infty}^\varphi := \bigcup_{r < 0} \mathcal{F}_r^\varphi.$$

So the statement says that either we can always solve equation (L) in the same regularity class or the solution may lose all regularity (observe that  $\mathcal{F}_{-\infty}^0 = \mathcal{F}_{-\infty}^{\varphi_0}$  is precisely the space of periodic distributions). When  $\omega \in \mathcal{A}_\infty^\varphi$ , the assertion follows directly from Lemma 2. When  $\omega \notin \mathcal{A}_\infty^\varphi$ , we have

$$\limsup_{k \in \mathbb{Z}^n \setminus \{0\}} \frac{\ln(|2\pi k \cdot \omega|^{-1})}{\varphi(|k|)} = +\infty,$$

hence for any positive sequence  $\tau_j \rightarrow +\infty$  one can find a sequence of integer vectors  $k_j \in \mathbb{Z}^n \setminus \{0\}$  with  $|k_j| \rightarrow +\infty$  such that  $|2\pi k_j \cdot \omega|^{-1} \geq e^{\tau_j \varphi(|k_j|)}$ . If we choose

$$f = \sum_j f_{k_j} e_{k_j}, \quad f_{k_j} := (i2\pi k_j \cdot \omega) |2\pi k_j \cdot \omega|^{-1/2}$$

then  $f \in \mathcal{F}$  and the infinitely many non-zero Fourier coefficients of  $f$  and  $g$  satisfy

$$|f_{k_j}| = |2\pi k_j \cdot \omega|^{1/2} \leq e^{-(\tau_j/2)\varphi(|k_j|)}, \quad g_{k_j} = |2\pi k_j \cdot \omega|^{-1/2} \geq e^{(\tau_j/2)\varphi(|k_j|)}$$

and the conclusion follows.

### 2.3. The inductive limit

Now we look at the inductive limit and its associated arithmetic class

$$\mathcal{F}_+^\varphi := \bigcup_{r > 0} \mathcal{F}_r^\varphi, \quad \mathcal{A}_+^\varphi := \bigcap_{\tau > 0} \mathcal{A}_\tau^\varphi$$

which admit the “dual” characterizations

$$f \in \mathcal{F}_+^\varphi \Leftrightarrow \varphi(|k|) = O(\ln(|f_k|^{-1})), \quad \omega \in \mathcal{A}_+^\varphi \Leftrightarrow \ln(|2\pi k \cdot \omega|^{-1}) = o(\varphi(|k|)),$$

and which are again well defined up to equivalent weights. The inductive limit topology on  $\mathcal{F}_+^\varphi$  is more complicated, but it is still a complete locally convex topological vector space (yet not metrizable). We obviously have the strict inclusions

$$\mathcal{F}_\infty^\varphi \subsetneq \mathcal{F}_+^\varphi, \quad \mathcal{A}_+^\varphi \subsetneq \mathcal{A}_\infty^\varphi.$$

Now we shall say that a weight  $\varphi$  strictly dominates another weight  $\psi$  if

$$\psi(t) = o(\varphi(t)) \Leftrightarrow \psi < \varphi$$

and we have

$$\psi < \varphi \Rightarrow \mathcal{F}_+^\varphi \subset \mathcal{F}_+^\psi, \quad \mathcal{A}_\infty^\psi \subset \mathcal{A}_+^\varphi.$$

In particular, if  $\varphi_0(t) = o(\varphi(t))$ , then  $\mathcal{A}_+^\varphi$  always contains the set of Diophantine vectors, and therefore it is of full measure (but in full generality it might be empty, as is the case for  $\varphi = \varphi_0$ ). For the analytic weight  $\varphi_1(t) = t$ ,  $\mathcal{F}_+^{\varphi_1} = \mathcal{F}_+^1$  is precisely the space  $\mathcal{O}(\mathbb{T}^n)$  of real-analytic functions on  $\mathbb{T}^n$  and the condition  $\mathcal{A}_+^{\varphi_1} = \mathcal{A}_+^1$  is exactly the Rüssmann condition  $\mathcal{R}$  as introduced in [32]. For the Gevrey weight  $\varphi_a(t) = t^a/a$ ,  $\mathcal{F}_+^{\varphi_a} = \mathcal{F}_+^a$  is the so-called space of Gevrey functions.

**Lemma 4.** *The vector  $\omega \in \mathcal{A}_+^\varphi \neq \emptyset$  if and only if equation (L) can be solved in  $\mathcal{F}_+^\varphi$ . More precisely, if  $\omega \in \mathcal{A}_+^\varphi$ , then for any  $f \in \mathcal{F}_+^\varphi$  we have  $g \in \mathcal{F}_+^\varphi$  and if  $\omega \notin \mathcal{A}_+^\varphi$ , there exists  $f \in \mathcal{F}_+^\varphi$  such that  $g$  does not belong to*

$$\mathcal{F}_-^\varphi := \bigcap_{r < 0} \mathcal{F}_r^\varphi.$$

Again, either we can always solve equation (L) in the same regularity class or the solution may lose all positive regularity (observe that  $\mathcal{F}_-^1 = \mathcal{F}_-^{\varphi_1}$  is precisely the space of periodic hyperfunctions in the sense of Kato). The proof is completely analogous to the proof of Lemma 3: when  $\omega \in \mathcal{A}_+^\varphi$  this follows from Lemma 2 and when  $\omega \notin \mathcal{A}_+^\varphi$  we have

$$\limsup_{k \in \mathbb{Z}^n \setminus \{0\}} \frac{\ln(|2\pi k \cdot \omega|^{-1})}{\varphi(|k|)} > 0$$

and one can use the same example as before, replacing the positive sequence  $\tau_j \rightarrow +\infty$  by some  $\tau > 0$ .

### 2.4. The Bruno condition

Finally, we would like to discuss the so-called Bruno condition, which was first introduced by Bruno in [12] in a linearization problem in the vicinity of a singular point, and then in various other equivalent forms for other linearization problems on a torus by Rüssmann (see for instance [33–35]). This condition appears in a non-linear version of Lemma 4 in the analytic case, and for  $n = 2$  it can be actually characterized this way (recall problem (1.2) in Section 1); we shall explain how it can also be characterized if one looks at Lemma 3 in an arbitrary non-quasi-analytic class, at the projective limit. We should say that a weight is non-quasi-analytic if

$$\int_1^{+\infty} \frac{\varphi(t)}{t^2} dt < \infty$$

and we denote by NQ the set of all non-quasi-analytic weights; the terminology is justified by the famous Denjoy–Carleman theorem which identifies the above condition as

the necessary and sufficient condition for non-quasi-analyticity (or even more precisely, for the existence of functions with arbitrarily small support). Now following the definition of [35], a vector satisfies the Bruno condition if and only if it belongs to  $\mathcal{A}_1^\varphi$  for some quasi-analytic weight  $\varphi$ ; indeed the “approximating function” in the sense of Rüssmann is nothing but the exponential of a quasi-analytic weight (we point out that Rüssmann does not require  $\ln(1+t) = O(\varphi(t))$ , but without such an assumption  $\mathcal{A}_1^\varphi$  is always empty so the definition remains unchanged). Using the fact that a weight  $\varphi$  is quasi-analytic if and only if  $\tau\varphi$  is quasi-analytic for some  $\tau > 0$ , we arrive at the following representation of Bruno vectors:

$$\mathcal{B} = \bigcup_{\varphi \in \text{NQ}} \mathcal{A}_1^\varphi = \bigcup_{\varphi \in \text{NQ}} \bigcup_{\tau > 0} \mathcal{A}_\tau^\varphi = \bigcup_{\varphi \in \text{NQ}} \mathcal{A}_\infty^\varphi. \tag{2.2}$$

For  $n = 2$ ,  $\omega = (1, \alpha)$ , the elementary equivalence of  $\omega \in \mathcal{B}$  and  $\alpha \in \mathcal{B}$ , as defined in Section 1, is proved in [35] for instance. Lemma 3 allows the following characterization of  $\mathcal{B}$ .

**Lemma 5.** *The vector  $\omega \in \mathcal{B}$  if and only if there exists  $\varphi \in \text{NQ}$  such that equation (L) can be solved in  $\mathcal{F}_\infty^\varphi$ .*

The discrepancy between  $\mathcal{B} \subsetneq \mathcal{R}$  seems to be related somehow to the existence of non-trivial quasi-analytic classes  $\mathcal{F}_\infty^\psi$  which strictly contains the analytic class  $\mathcal{O}(\mathbb{T}^n)$ . Indeed, first observe that when  $\varphi$  is non-quasi-analytic, we obviously have  $\varphi(t) = o(t)$ , therefore we have the following inclusions, which are moreover strict:

$$\mathcal{F}_\infty^\varphi \supsetneq \mathcal{O}(\mathbb{T}^n), \quad \mathcal{A}_\infty^\varphi \subsetneq \mathcal{R}.$$

Taking respectively the intersection and the union over all non-quasi-analytic classes one actually obtains

$$\bigcap_{\varphi \in \text{NQ}} \mathcal{F}_\infty^\varphi = \mathcal{O}(\mathbb{T}^n), \quad \mathcal{B} = \bigcup_{\varphi \in \text{NQ}} \mathcal{A}_\infty^\varphi \subsetneq \mathcal{R}.$$

The first equality claims that the intersection of all non-quasi-analytic classes is precisely the space of real-analytic functions: this is a rather surprising result first proved in [4] in a different setting (see [6] for more general results and see also [5, 11, 13] for results adapted to our setting). On the other hand, the set  $\mathcal{B}$ , which is the union of  $\mathcal{A}_\infty^\varphi$  over all non-quasi-analytic classes, is not equal to  $\mathcal{R}$ ; indeed, take any quasi-analytic weight  $\psi(t) = o(t)$  for which  $\mathcal{O}(\mathbb{T}^n) \subsetneq \mathcal{F}_\infty^\psi$ . For instance, one may choose  $\psi = \psi_1$  with

$$\psi_1(t) := t(\log(1+t))^{-1}; \tag{2.3}$$

then any vector in  $\mathcal{A}_\infty^\psi$  (which is non-empty) is in  $\mathcal{R}$  but not in  $\mathcal{B}$ .

As a final remark, we would like to point out that in his first attempts to obtain a linearization result with a Bruno condition in any dimension, Rüssmann (see [33]) required not only the weight to be non-quasi-analytic but also that  $t^{-1}\varphi(t)$  decrease monotonically to zero; the extra condition is the monotonicity requirement but such an (a priori mild) condition prevents the identification of vectors  $\omega = (1, \alpha) \in \mathbb{R}^2$  satisfying this condition



with Bruno numbers  $\alpha$ . In fact, when  $t^{-1}\varphi(t)$  decreases monotonically to zero, it is easy to show that  $\varphi$  is sub-additive; if we let NQS be the set of all non-quasi-analytic sub-additive weights, then we have another Bruno-type condition,

$$\mathcal{B}^1 := \bigcup_{\varphi \in \text{NQS}} \mathcal{A}_\infty^\varphi \subsetneq \mathcal{B}. \tag{2.4}$$

It is proved in [5] that for  $\varphi \in \text{NQS}$ , not only do we have  $\varphi(t) = o(t)$  but also  $\varphi(t) = o(\psi_1(t))$  with  $\psi_1$  as in (2.3) and therefore (see [5, 6]) the intersection of all non-quasi-analytic sub-additive classes is no longer the real-analytic class, but the larger quasi-analytic class associated to  $\psi_1$ ,

$$\bigcap_{\varphi \in \text{NQS}} \mathcal{F}_\infty^\varphi = \mathcal{F}_+^{\psi_1} \supsetneq \mathcal{O}(\mathbb{T}^n). \tag{2.5}$$

One should point out that in (2.4) and (2.5), one could replace NQS by quasi-analytic weights which are either strictly concave or satisfy  $t^{-1}\varphi(t) \searrow 0$ : the strict concavity implies the monotonicity condition which implies sub-additivity, and quasi-analytic sub-additive weights are equivalent to strict concave weights (see [5, 27] for instance). For non-linear problems, one may largely speculate that in the quasi-analytic class  $\mathcal{F}_+^{\psi_1}$ , vectors in  $\mathcal{B}^1$  could play the same role as Bruno vectors in the analytic class.

Finally, let us briefly explain how, in the sequel we will use the observation that Bruno vectors can be represented as (2.2). For a given  $\varphi \in \text{NQ}$ , this observation implies that non-linear analytic problems (such as (1.2) described in Section 1) can be solved if  $\omega \in \mathcal{A}_\tau^\varphi$  for an arbitrary  $\tau$ ; if one has a suitable control on how the analytic problem can be solved, and if one can characterize functions in  $\mathcal{F}_r^\varphi$  in terms of approximation by real-analytic functions, then one may expect that  $\omega \in \mathcal{A}_\tau^\varphi$  also allows us to solve the problem in the larger space  $\mathcal{F}_r^\varphi$ ; clearly there will be competition in this approach but it does work in the smooth case  $\varphi = \varphi_0$ , as was proved by Moser in [25]. The competition in the Gevrey case  $\varphi = \varphi_a$  is more subtle (since it has to become singular for  $a = 1$ ) but we will show that it works too, and the proof also shows that it should be true for a large class of non-quasi-analytic weights (the Gevrey weights allow for a simpler approach and more explicit computations).

### 3. A non-linear problem and the main result

#### 3.1. Linearization of vector fields on the torus

So now we finally look at the continuous version of problem (1.2) in any dimension  $n \geq 2$ . We consider a smooth vector field  $X = X_\omega + F$  on  $\mathbb{T}^n$ , where  $F$  is sufficiently close to zero with respect to a suitable topology, and we wish to conjugate  $X$  to  $X_\omega$  by a diffeomorphism close to the identity. This is clearly not possible in general as there are obvious topological restrictions, so to circumvent them the question we ask is whether we can find

a constant vector field  $X_\lambda = \lambda \in \mathbb{R}^n$  close to zero and a diffeomorphism  $\Phi: \mathbb{T}^n \rightarrow \mathbb{T}^n$  close to the identity such that

$$\Phi^*(X - X_\lambda) = X_\omega; \tag{V}$$

that is, we wish to conjugate the modified vector field  $X - X_\lambda$ , with an unknown  $\lambda \in \mathbb{R}^n$ , to  $X_\omega$ . So given  $X = X_\omega + F$ , a solution of (V) is a couple  $(\Phi, \lambda)$  which satisfy equation (V) and which is close to  $(\text{Id}, 0)$ : upon normalization (for instance fixing the average of  $\Phi - \text{Id}$  to be zero) it can be shown that such a solution, if it exists, is unique. Such a “modifying term” formulation was introduced by Arn’old (see [2]) in this setting, and then generalized by Moser (see [26]). If we require the rotation set of  $X$  to contain  $\omega$ , then it is not hard to show that necessarily  $\lambda = 0$  and this is precisely the generalization of (1.2) in any dimension, but we prefer to keep the modifying term since it gives a more flexible formulation which allows us to consider arbitrary vector fields  $X$  close to  $X_\omega$ . In general problem (V) is non-linear, in the sense that both  $\lambda = \lambda(F)$  and  $\Phi - \text{Id} = (\Phi - \text{Id})(F)$  depend on  $F$  in a non-linear fashion. There is, however, one simple case in which (V) is actually equivalent to the linear problem (L) we studied in Section 2: this is the case where  $X$  is proportional to  $X_\omega$ , which can be written

$$X = X_\omega + P = \frac{1}{f} X_\omega \tag{3.1}$$

for some nowhere-vanishing smooth function  $f: \mathbb{T}^n \rightarrow \mathbb{R}$  (one may assume  $f$  to be close to 1 to consider this problem as perturbative, but this is not necessary since it is not a perturbative problem). The flow associated to  $X$  is nothing but a time-reparametrization of the flow of  $X_\omega$ , and it is not hard to prove that the unique (upon normalization) solution  $(\Phi, \lambda)$  of (V) has to be of the form

$$\Phi - \text{Id} = g\omega, \quad \lambda = (1 - c)\omega, \quad c = \int_{\mathbb{T}^n} f \, dx, \tag{3.2}$$

where  $g: \mathbb{T}^n \rightarrow \mathbb{R}$  is the solution of (L). This is essentially due to Kolmogorov (see [20]) in an implicit form and Herman (see [17]) in this more explicit form. It follows that the necessary and sufficient conditions on  $\omega$  which were given respectively in Lemmas 3 and 4 to solve (L) respectively in  $\mathcal{F}_\infty^\varphi$  and  $\mathcal{F}_+^\varphi$  are automatically necessary (but a priori not sufficient) conditions to solve (V) in those regularity classes. In general, since we look at (V) in a perturbative setting, linearizing the conjugacy equation at  $P = 0$ ,  $\Phi = \text{Id}$  yields a linear equation between vector fields, which is nothing but a vector-valued version of (L). The basic question is whether the optimal arithmetic conditions introduced for the linear problem, which we know are necessary for the non-linear problem, are also sufficient.

For the general smooth case, the answer is yes: the vector  $\omega \in \mathcal{D}$  if and only if one can solve the problem (V), which is the exact analogue of Lemma 3 (observe that the analogue of Lemma 4 does not make any sense here since the corresponding arithmetic class is empty). In fact, a much more precise statement holds true: if  $\omega \in \mathcal{D}^\tau$  for some  $\tau \geq n - 1$ , then it suffices for  $P$  to be of class  $C^r$  with  $r > \tau + 1$  ( $r \notin \mathbb{N}$ ) for  $\Phi$  to be of

class  $C^{r-\tau-\varepsilon}$  for any  $\varepsilon > 0$ . This is almost a non-linear analogue of Lemma 2, except for the fact that one needs  $r > \tau + 1$  instead of  $r > \tau$  (for  $\Phi$  to be at least  $C^1$ ), but the loss of regularity is essentially the same, namely one loses any  $\tau' > \tau$  instead of  $\tau$ .

For the analytic case, the answer is no in general. As we have already explained, the Bruno condition  $\omega \in \mathcal{B}$  which is stronger than the Rüssmann condition  $\mathcal{R}$ , is known to be sufficient to solve (V) for any  $n \geq 2$  ([34]) and it is known that it is also necessary for  $n = 2$  ([37]), hence there is no analogue of Lemma 4 for the analytic weight  $\varphi_1(t) = t$  when  $n = 2$  (and this is most probably the case for any  $n \geq 2$ ). Here again, a more precise statement is true: Yoccoz’s example shows in fact that there is no analogue of Lemma 2 since the Bruno condition is optimal for (V) no matter how large is the regularity parameter  $r > 0$  (which here is essentially a width of analyticity). The question of whether a non-linear analogue of Lemma 3 holds true (the regularity class in this case corresponds to entire functions) is, to our knowledge, open.

The main result of this article deals with Gevrey regularity, which is associated to the weight  $\varphi_a(t) = t^a/a$  for  $0 < a < 1$ , and we shall prove that in this case the answer is again positive. Actually, we shall prove a non-linear perturbative version of Lemma 2, which is Theorem A below, with however a significant quantitative difference: the loss of derivatives will be essentially  $c(a)\tau$  with  $1 < c(a) < +\infty$ ; at the formal limit  $a \rightarrow 0$  then  $c(a) \rightarrow 1$  in concordance with the fact that in this case functions tend to “behave” as general smooth functions (beware though that there are many smooth functions that do not belong to any Gevrey class) and at the formal limit  $a \rightarrow 1$  then  $c(a) \rightarrow +\infty$  in concordance with the fact that such a result cannot hold true in the analytic case (beware that other constants, such as the threshold of applicability, will have a singular limit as  $a \rightarrow 1$  so the fact that  $c(a) \rightarrow +\infty$  is not necessarily unavoidable). Regardless of this, such a statement is still enough to guarantee we have exact non-linear analogues of Lemma 3 and Lemma 4.

### 3.2. Main results: the Gevrey case

To state our result properly for the Gevrey weight  $\varphi_a(t) = t^a/a$  for  $0 < a < 1$ , we recall that

$$\mathcal{F}_r^a := \mathcal{F}_r^{\varphi_a} = \{f \in \mathcal{F} \mid \|f\|_r = \sup_{k \in \mathbb{Z}^n} |u_k| e^{r\varphi_a(|k|)} < +\infty\} \tag{3.3}$$

and

$$\mathcal{A}_\tau^a := \mathcal{A}_\tau^{\varphi_a} = \{\omega \in \mathbb{R}^n \mid \gamma_\tau^{-1}(\omega) = \sup_{k \in \mathbb{Z}^n \setminus \{0\}} |2\pi k \cdot \omega|^{-1} e^{-\tau\varphi_a(|k|)} < +\infty\}. \tag{3.4}$$

In particular,  $\mathcal{F}_r^a$  is strictly contained in  $C^\infty(\mathbb{T}^n)$  for any  $r > 0$  and  $\mathcal{A}_\tau^a$  strictly contains  $\mathcal{D}$  for any  $\tau > 0$ . For any smooth vector field  $F = (f_1, \dots, f_n): \mathbb{T}^n \rightarrow \mathbb{R}^n$ , slightly abusing notation we shall still say that  $F \in \mathcal{F}_r^a$  if

$$\|F\|_r := \max_{1 \leq i \leq p} \|f_i\|_r < +\infty.$$

This is the main result.

**Theorem A.** *Let  $0 < a < 1$ ,  $r_0 > r > 0$ ,  $\tau > \tau_0 > 0$ ,  $1 < \kappa < 2$ , assume that*

$$r > c\tau, \quad c = c(a, \kappa) := 32^{a/(1-a)}(\kappa - 1)(\kappa^{1-a} - 1)^{-1/(1-a)} \tag{3.5}$$

*and let  $\rho := r - c\tau > 0$  and  $\nu := 8^{-a/(1-a)}\rho > 0$ . Then there exist a small positive constant  $\varepsilon^*$  and a large positive constant  $C$  that depend on  $a, r_0, r, \tau, \tau_0, \kappa$  and  $n$  such that for any*

$$X = X_\omega + F, \quad \omega \in \mathcal{A}_{\tau_0}^a, \quad \|F\|_{r_0} = \varepsilon \leq \gamma_{\tau_0}(\omega)\varepsilon^*$$

*there exists a vector  $\lambda \in \mathbb{R}^n$  and a diffeomorphism  $\Phi \in \mathcal{F}_\nu^a$ , with  $\nu := \nu/2$ , which solve (V) with the estimates*

$$|\lambda| \leq C\varepsilon, \quad \|\Phi - \text{Id}\|_\nu \leq C\left(\frac{\varepsilon}{\gamma_{\tau_0}(\omega)}\right)^\iota, \quad \iota := 8^{-a/(1-a)}(2r)^{-1\nu}.$$

We shall discuss later this rather technical statement, especially the conditions  $r_0 > r$ ,  $\tau > \tau_0$  and  $r > c\tau$ , but first we would like to observe that these conditions disappear at the limits, that is, when we look at the spaces

$$\mathcal{F}_\infty^a := \bigcap_{r>0} \mathcal{F}_r^a, \quad \mathcal{F}_+^a := \bigcup_{r>0} \mathcal{F}_r^a$$

and their associated arithmetic classes

$$\mathcal{A}_\infty^a := \bigcup_{\tau>0} \mathcal{A}_\tau^a, \quad \mathcal{A}_+^a := \bigcap_{\tau>0} \mathcal{F}_\tau^a$$

and lead to the following non-technical statements, which are non-linear perturbative analogues of Lemmas 3 and 4.

**Theorem B.** *The vector  $\omega \in \mathcal{A}_\infty^a$  if and only if equation (V) can be solved in  $\mathcal{F}_\infty^a$ .*

**Theorem C.** *The vector  $\omega \in \mathcal{A}_+^a$  if and only if equation (V) can be solved in  $\mathcal{F}_+^a$ .*

The direct implications for both statements follow from Theorem A. It is important to observe that the requirement that  $F$  be close to zero need not be formulated in the limit topology; one only need assume the norm of  $F$  to be small in some fixed space  $\mathcal{F}_{r_0}^a$ : in Theorem B, once  $\tau_0$  is given (possibly large), it suffices to choose  $r_0 = r_0(\tau_0)$  large enough to satisfy (3.5), and in Theorem C, once  $r_0$  is given (possibly small), it suffices to choose  $\tau_0 = \tau_0(r_0)$  small enough to satisfy (3.5). That this is so is clear for Theorem C, but not so obvious for Theorem B, yet this is a classical matter. (One has to verify that we can impose additional smoothness on  $F$  in Theorem A to get additional smoothness on  $\Phi$  with the same smallness condition. See for example [36, Corollary 1] in the smooth case; the argument in our case is similar.) The converse implications of both statements follow from the converse implications in Lemmas 3 and 4 respectively, together with the case of reparametrized constant vector fields described in (3.1) and (3.2).

### 3.3. Some comments on Theorem A

Let us now comment on the technical conditions  $r_0 > r$ ,  $\tau > \tau_0$  and  $r > c\tau$  in Theorem A. The fact that the initial arithmetic parameter  $\tau_0$  deteriorates to  $\tau > \tau_0$  is in a sense unavoidable: this comes from the fact that for the non-linear problem (V), the small divisors are not the quantities  $|k \cdot \omega|$  but rather  $|k|^{-1}|k \cdot \omega|$ , or equivalently, the small parameter is not the size of  $F$  but rather the size of its differential (of course one could still modify the space (3.3) or the set (3.4) to take this into account, but we have decided not to do so). In the smooth case, this has the effect of replacing  $\tau_0$  by  $\tau = \tau_0 + 1$  but in the Gevrey case, any  $\tau > \tau_0$  is sufficient. Then the initial regularity parameter  $r_0$  also deteriorates to  $r < r_0$  but this, however, is artificial: this will be used in order to compensate the fact that, throughout the proof, we shall use various non-equivalent norms and consequently, the constants  $\varepsilon$  and  $C$  becomes singular at the limit  $r \rightarrow r_0$ . One can in principle work with the same norm throughout the proof and allow  $r = r_0$ , but for our non-linear problem this leads to technical difficulties (changing norms will allow us to somehow avoid these difficulties, at the expense of introducing a singular limit  $r \rightarrow r_0$ ). For some other simpler non-linear problems (that will be mentioned below) one can indeed work with the same norm and reach  $r = r_0$  (and also  $\tau = \tau_0$ ). Finally, let us discuss our assumption  $r > c(a, \kappa)\tau$  with

$$c(a, \kappa) = 32^{a/(1-a)}(\kappa - 1)(\kappa^{1-a} - 1)^{-1/(1-a)}.$$

As we have already mentioned, when  $a \rightarrow 0$  then  $c(a, \kappa) \rightarrow 1$  and this is the best one can hope for, in view of the linear analogue given by Lemma 2. But when  $a \rightarrow 1$ , then  $c(a, \kappa) \rightarrow +\infty$  at variance with Lemma 2: here we do not know whether this is artificial or not.

To discuss this last issue, let us recall that the smooth analogue of Theorem A, which we have already mentioned, can be proved either by analytic approximations or by using Nash–Moser theory, and even though these techniques are clearly related, they are not the same (though they may be equivalent in some sense): in Nash–Moser theory a crucial use is made of the so-called “tame estimates”, whereas they are not used at all if one uses analytic approximations. Our proof of Theorem A will use analytic approximations, together with the application of an analytic KAM theorem at each step, and it is the analytic KAM theorem which introduces this large constant  $c = c(a, \kappa)$ . One could, and perhaps should, replace the analytic KAM theorem by an analytic KAM step since this is clearly what is needed; to apply a KAM step one essentially has to require  $r > \tau$  but the convergence argument becomes quite technical and heuristic considerations suggest that one needs to ask for much more than  $r > \tau$  (applying a KAM theorem instead of a KAM step makes the convergence proof elementary). In principle we could also use Nash–Moser theory, albeit in a “generalized” sense since we only have “generalized” tame estimates, and we would like to point out that with this approach one faces the same singular limit when  $a \rightarrow 1$ . First recall that for the Fourier  $l_1$ -norm

$$\|f\|_r^1 = \sum_{k \in \mathbb{Z}^n} |f_k| e^{r\varphi(|k|)}$$

associated to any sub-additive weight (so in particular the smooth weight  $\varphi_0$ , the Gevrey weight  $\varphi_a$  and the analytic weight  $\varphi_1$ ), one has the general product estimate

$$\|fg\|_r^1 \leq \|f\|_r^1 \|g\|_r^1.$$

Now for the smooth weight  $\varphi_0$ , we have the better tame estimates

$$\|fg\|_r^1 \leq C(r)(\|f\|_r^1 \|g\|_0^1 + \|f\|_0^1 \|g\|_r^1), \quad \varphi_0(t) = \ln(1 + t), \tag{3.6}$$

which are fundamental in Nash–Moser theory; these product estimates are the simplest non-linear tame estimates. For the Gevrey weight  $\varphi_a$ , one can only prove weaker tame estimates

$$\|fg\|_r^1 \leq C(\|f\|_r^1 \|g\|_{\chi(a)r}^1 + \|f\|_{\chi(a)r}^1 \|g\|_r^1), \quad \varphi_a(t) = \frac{t^a}{a} \tag{3.7}$$

with  $0 < \chi(a) < 1$  such that  $\chi(a) \rightarrow 0$  as  $a \rightarrow 0$  and  $\chi(a) \rightarrow 1$  as  $a \rightarrow 1$ ; in general these estimates are the best possible and for the analytic weight  $a = 1$ , one cannot do better than the general product estimate. Even though estimates (3.7) are weaker than those in (3.6), a generalized Nash–Moser argument applies: this was done in [22] for instance, and their result leads to exactly the same singular behavior as  $a \rightarrow +\infty$ . More precisely, in the language of Nash–Moser theory we have a “non-constant loss of derivatives” for the non-linear tame estimates (we still have nice linear tame estimates as Lemma 2 shows) and the result of [22] requires  $r > c'(a)\tau$  with  $c'(a) = O_2(1/(1 - \chi(a)))$  (there  $\kappa = 3/2$  is fixed). Of course product estimates are not sufficient to deal with the non-linear problem (V), as one also needs composition estimates which are usually harder to obtain, and this is why we eventually used analytic approximations, which, in our opinion, turn out to be simpler (even though they require the use of non-equivalent norms and induce an artificial small loss of regularity).

### 3.4. Some comments on other (non-)quasi-analytic classes

Let us now discuss the case of other weights, starting with non-quasi-analytic ones. Again, to deal with non-linear problems such as (V), one should require further properties that ensure stability with respect to non-linear operations such as products and compositions: a general property that guarantees such stability properties is the sub-additivity of the weight. A statement such as Theorem A seems hard to obtain for an abstract weight, since it depends strongly on the weight and on the choice of a norm for a given regularity parameter: this is not the case for statements such as Theorems B or C which depend only on the equivalence class of the weight and on the topology of the limit spaces, so we shall discuss only those statements. Let us say that a weight  $\varphi$  is of moderate growth if there exists  $H > 1$  such that

$$\liminf_{t \rightarrow +\infty} \frac{\varphi(Ht)}{\varphi(t)} > 1.$$

The following conjecture seems reasonable.

**Conjecture 1.** Theorems B and C hold true for any non-quasi-analytic sub-additive weight of moderate growth.

Indeed, for any non-quasi-analytic sub-additive weights, the result of Section 4.2, which deals with characterization of functions in  $\mathcal{F}_r^\varphi$  by their approximation with real-analytic functions, and the result of Section 4.3, which deals with the analytic KAM theorem for vectors in  $\mathcal{A}_r^\varphi$ , apply (for the results of Section 4.2, one needs to change to some equivalent weight and somehow lose precise control of the regularity parameter; see [27] or [16]). To be more accurate, the approximation by real-analytic functions and the application of an analytic KAM step (and not KAM theorem) are both governed by the same function, which is the Young conjugate of  $\varphi$  (see Section 4.2 for a definition): the case of Gevrey weight is simpler since its Young conjugate has an explicit expression, and, more importantly, it also governs the application of the analytic KAM theorem (up to a large factor, a point which we have already discussed). For more general weights, simple examples show that this last point is no longer true, so to solve the above conjecture one probably has to replace the analytic KAM theorem by an analytic KAM step, but heuristic considerations suggest that one can still obtain a convergent scheme (which has to be more involved than what is done in Section 4.4). The moderate growth condition, which also frequently appears in the literature, is clearly used in the proof in Section 4.4, but we do not know whether this is necessary: again Gevrey weights are homogeneous and this moderate growth property is then obvious and very explicit. Apart from the Gevrey weight, a specific family of weights satisfying the requirements of the above conjecture is given by

$$\psi_b(t) = t(\ln(1 + t))^{-b}, \quad b > 1.$$

The quasi-analytic case is more subtle, and we should only discuss the analogue of Theorem C in the most representative case of the weight

$$\psi_1(t) = t(\ln(1 + t))^{-1}$$

which has already appeared in Section 2.4. We can ask the following question.

**Question 1.** What is the necessary and sufficient condition on  $\omega$  to solve (V) in  $\mathcal{F}_+^{\psi_1}$ ?

The condition  $\omega \in \mathcal{A}_+^{\psi_1}$ , defined in Section 2.3, is clearly a necessary condition but we do not know whether it is also a sufficient condition. The results of Section 4.2 do apply as well, but not the results of Section 4.3 since vectors in  $\mathcal{A}_1^{\psi_1}$  are not Bruno: but exactly as before, one could replace the analytic KAM theorem by an analytic KAM step (which only requires  $\varphi(t) = o(t)$ , hence applies to  $\varphi = \psi_1$ ) but unlike non-quasi-analytic classes, heuristic considerations do not suggest that such a scheme could converge.

For this quasi-analytic problem, it may be the case that a Bruno-type condition appears, like they do for the analytic problem. Results in [8] do apply here and give the following sufficient condition to solve (V) in  $\mathcal{F}_+^{\psi_1}$ :

$$\omega \in \mathcal{A}_1^\varphi, \quad \int_1^{+\infty} \frac{\ln(1 + t)\varphi(t)}{t^2} < +\infty.$$

One may still improve this condition; perhaps the weaker condition  $\omega \in \mathcal{B}^1$ , which was introduced in Section 2.4 and reads  $\omega \in \mathcal{A}_1^\varphi$  for some non-quasi-analytic sub-additive weight (equivalently,  $\varphi$  is non-quasi-analytic and satisfies  $t^{-1}\varphi(t) \searrow 0$ ), could play a role here.

### 3.5. Some comments on other non-linear problems

Finally, we claim that Theorem A (and hence Theorems B and C) also holds true for other non-linear problems on the torus: we choose the case of perturbation of constant vector fields since it is the simplest case in which non-trivial arithmetic conditions are known to be necessary. But as will be clear from the proof (which will be described in Section 4.1), everything also works for the persistence of a Lagrangian quasi-periodic invariant torus in Hamiltonian systems as pioneered by Kolmogorov ([21]), or, even more generally, for the quasi-periodic solutions constructed in [26] upon introducing modifying terms: considering this more general setting only introduces further non-essential technicalities that we have decided to avoid.

In particular, everything applies as well to the problem of reducibility of elliptic quasi-periodic cocycles, as studied in [14, 33] for instance: with modifying terms this is nothing but a “linear” particular case of [25] in which the frequency  $\omega$  is fixed under perturbation (but some other “elliptic” frequency moves in a non-linear fashion). This problem is technically simpler as only product estimates are needed (see for instance [7]) and the proof of the equivalent of Theorem A in this case greatly simplifies: using a “generalized” Nash–Moser argument that we described in Section 3.3, we can work with the same norm and the statement of Theorem A holds true with  $r = r_0$  and even  $\tau = \tau_0$ . But more seems to be true: indeed it was proved in [19] (see [3] for the discrete setting) that for  $n = 2$ , the analogue of Theorem A also holds true even for the analytic case  $a = 1$  (with a proper arithmetic condition on the elliptic component): hence there are no Bruno-type conditions as far as the “base” frequency  $\omega$  is concerned. It is reasonable to expect that there should be no Bruno-type condition on  $\omega$  for any  $n \geq 2$ , yet this is an open problem.

## 4. Proof of the main result

### 4.1. Strategy of the proof, following Moser

The proof follows Moser’s argument of approximation by real-analytic functions, as described for instance in [25] in the smooth case. The proof in the smooth case relies on the following two principles. First, a finitely differentiable function can be characterized by the rate of approximation by real-analytic functions; if  $r$  is the regularity and  $s_j$  is the sequence (converging to zero) of analytic widths associated to the analytic approximations, then the optimal rate of convergence is of order  $s_j^r$ . It is important to observe that those analytic approximations can actually be chosen to be much more than analytic, namely one can choose entire functions of exponential type (which, in the periodic case we are considering, are nothing but trigonometric polynomials) and they are easily obtained



by convolution. Second, for an analytic perturbation, with analytic width  $s_j$ , of a constant Diophantine vector field with exponent  $\tau > 0$ , the threshold of applicability of the analytic KAM theorem is easily seen to be of order  $s_j^{\tau+1}$ . Combining these two principles one obtains a result in regularity  $r > \tau + 1$ .

In the  $a$ -Gevrey case for  $0 < a < 1$ , we will follow the same principles but they are somehow more complicated in this situation. First, as was observed in [28], approximation of Gevrey functions by real-analytic functions through entire functions of exponential type do not lead to an optimal rate of approximation: if  $r$  is the regularity and  $s_j$  is the sequence of analytic widths, then the rate of convergence is of order  $\exp(-r(1/s_j)^a)$  and does not characterize  $a$ -Gevrey functions: one could argue that the space of entire functions of exponential type is too small and does not allow us to discriminate between the case  $0 < a < 1$  and the analytic case  $a = 1$ , whereas one obviously would like the rate of approximation to tend to infinity as  $a$  approaches 1. The idea of Popov (see [28]) is to obtain a sequence of analytic approximations in a much more precise way through almost analytic extensions, which are an extension to the complex domain for which the  $\bar{\partial}$ -operator does not vanish (so the extension is not analytic) but vanishes asymptotically with a precise rate as the imaginary part tends to zero. Once we have almost analytic extensions, one can further approximate them to solve the  $\bar{\partial}$ -operator and this leads to a sequence of real-analytic approximations with an optimal rate of order  $\exp(-r(r/s_j)^b)$  with  $b = a/(1 - a)$ , which does tend to infinity as  $a$  tends to 1 and which does characterize  $a$ -Gevrey functions. Second, the threshold of applicability of the analytic KAM theorem with a Bruno frequency vector (recall that frequency vectors  $\omega \in \mathcal{A}_\tau^a$  are Bruno) is more subtle, but very precise statements have been given by Rüssmann (see [33–35]). In fact, results in [34, 35] do apply to an arbitrary Bruno vector (see also [10, 30]) and it turns out that they are not very well suited to the special Bruno vectors  $\omega \in \mathcal{A}_\tau^a$  we are considering (they lead to an artificial singular limit when  $a$  approaches 0). Consequently, we shall rely on a result of Pöschel ([29]) which follows [33] and gives a quite precise statement for vectors  $\omega \in \mathcal{A}_\tau^a$ : the threshold is of order  $\exp(-c\tau(\tau/s_j)^b)$  with  $c = c(a) > 1$ . Again, combining these two principles one obtains a result in regularity  $r > c\tau$ .

#### 4.2. Analytic approximation, following Popov

In this section we recall the results of Popov (see [28] and also [18]) on the approximation of Gevrey functions by real-analytic functions through almost analytic extension. Some little modifications are required as far as the almost analytic extensions are concerned, since we use a Fourier-based  $l_\infty$ -norm whereas the results in [28] use a space-based  $L_\infty$ -norm (that characterizes the regularity of a function through the growth of the sequence of its derivatives). One could simply convert our Fourier norm into a proper space norm, as this would only involve an arbitrarily small loss of regularity parameter (up to some polynomially large factor that can be absorbed by the exponentially small error term that will come into play; anyway such a procedure will be used several times in the sequel). However, those results are known in greater generality and they are valid for a large class of weights (even some quasi-analytic weights). The first general result on the existence

of almost analytic extension is due to Dynkin (see [15] for a survey) still using space-based norms; for Fourier norms we shall rely on the results contained in [1, 27]: results in [27] are stated for functions with one variable but since our weights are “radial” (in the sense that  $\varphi(|k|)$  for  $k \in \mathbb{Z}^n$  depends only on  $|k| \in \mathbb{N}$ ) these results apply as well, but alternatively the results of [1] are stated in any dimension and for general (not necessarily radial) weights. Finally, there are also recent results in [16] which apply to both space-based or Fourier-based norms.

Given  $v > 0$ , we consider the complex domain

$$\mathbb{T}_v^n := \{\theta = x + iy \in \mathbb{C}^n / \mathbb{Z}^n \mid |y| < v\}$$

and for any bounded function  $g: \mathbb{T}_v^n \rightarrow \mathbb{C}$ , we consider the sup-norm over the complex domain, as well as the sup-norm over the real domain

$$|g|_v := \sup_{\theta \in \mathbb{T}_v^n} |g(\theta)|, \quad |g|_0 := \sup_{\theta \in \mathbb{T}^n} |g(\theta)|.$$

Recall that a  $C^1$  function  $\tilde{f}: \mathbb{T}_v^n \rightarrow \mathbb{C}$  is analytic if and only if it is holomorphic, that is it satisfies a system of Cauchy–Riemann equations

$$\bar{\partial}_l \tilde{f}(\theta) := \frac{1}{2}(\partial_{x_l} \tilde{f}(\theta) + i \partial_{y_l} \tilde{f}(\theta)) = 0, \quad 1 \leq l \leq n,$$

for any  $\theta = (x_1 + iy_1, \dots, x_n + iy_n) \in \mathbb{T}_v^n$ . Now a real-analytic function  $f \in \mathcal{O}(\mathbb{T}^n) = \bigcap_{s>0} \mathcal{F}_s^1$  possesses a (unique) analytic extension  $\tilde{f}$  to  $\mathbb{T}_v^n$  for some  $v > 0$ : more precisely, one can choose any  $v < s$  provided  $f \in \mathcal{F}_s^1$ . This is no longer true for  $f \in \mathcal{F}_r^a$  with  $0 < a < 1$ , but a (non-unique) “almost” analytic extension always exists (and one can take  $v$  as large as one wants, though only the case where  $v$  is small will be of interest).

To state the result, we shall introduce the Young conjugate of the weight  $\varphi_a(t) = t^a/a$ , which is defined by

$$\varphi_a^*(\xi) := \sup_{t \geq 0} \{\varphi_a(t) - \xi t\}, \quad \xi > 0. \tag{4.1}$$

It is always finite since  $\varphi_a(t) = o(t)$ , convex, decreasing and we also have

$$\varphi_a(t) = \inf_{\xi > 0} \{\varphi_a^*(\xi) + \xi t\}, \quad t > 0$$

since  $\varphi_a$  is concave. Moreover, for any  $r > 0$  we have

$$r\varphi_a^*(\xi/r) = \sup_{t \geq 0} \{r\varphi_a(t) - \xi t\}, \quad \xi > 0. \tag{4.2}$$

In fact, the form of the Gevrey weight allows for a simple computation of  $\varphi_a^*$ : define the conjugate exponent  $b$  by

$$0 < b := \frac{a}{1-a} < +\infty, \quad \frac{1}{a} - \frac{1}{b} = 1; \tag{4.3}$$

then it is easy to check that

$$\varphi_a^*(\xi) = \frac{1}{b\xi^b} = \varphi_b(1/\xi), \quad \xi > 0. \tag{4.4}$$

We have the following statement.

**Proposition 6.** *Let  $f \in \mathcal{F}_{r_0}^a$  and  $0 < r < r_0$ . Then  $f$  admits a  $C^1$  extension  $\tilde{f}: \mathbb{T}_{r_0}^n \rightarrow \mathbb{C}$  such that for any  $0 < v \leq r_0$ ,*

$$|\tilde{f}|_v \leq C_1 \|f\|_{r_0}, \quad |\bar{\partial}_l \tilde{f}|_v \leq C_1 \|f\|_{r_0} e^{-r\varphi_b(r_0/v)}, \quad 1 \leq l \leq n$$

with a positive constant  $C_1 = C_1(a, r, r_0, n)$ .

This is a consequence of [27, Theorem 2.4] (see also [1, Theorems 2.2, 4.1]). Actually, the quality of approximation is more precisely an exponentially small factor  $e^{-r_0\varphi_b(r_0/v)}$  up to some polynomially large factor  $(r_0/v)^d$  with  $d = d(a, n)$ ; we simply decreased  $r < r_0$  to absorb this last factor.

Now, for the sequence  $v_j = 2^{-j}v_0$ ,  $j \in \mathbb{N}$  with  $v_0 \leq r_0$ , the proposition gives a sequence  $\tilde{f}_j: \mathbb{T}_{v_j}^n \rightarrow \mathbb{C}$  which is almost analytic in the sense that  $|\bar{\partial}_l \tilde{f}_j|_{v_j}$  for any  $1 \leq l \leq n$  decreases to zero with a stretched exponential speed (in  $v_j$ ) with exponent  $b$ ; in view of (4.3) we have  $b \rightarrow +\infty$  as  $a \rightarrow 1$  which agrees with the fact that for  $a = 1$ , the extension can be chosen so that  $|\bar{\partial}_l \tilde{f}_j|_{v_j}$  is identically zero.

Proposition 6 constitutes one half of Popov’s approximation lemma by real-analytic functions; once we have almost analytic extensions, [28, Proposition 3.1] (see also [18, Proposition 2.1]) yields a sequence of real-analytic approximations  $f_j$  with the same stretched exponential speed. However, the approximations  $f_j$  have to be constructed on a domain  $\mathbb{T}_{u_j}^n$  which is slightly smaller than the domain  $\mathbb{T}_{v_j}^n$  on which one has estimates for  $\tilde{f}_j$ ; it suffices to take  $u_j < v_j$  (in [28] the author chooses  $u_j = v_j/2$  for simplicity) and so again we use our absorbing factor  $r < r_0$  to choose  $u_j = (r/r_0)v_j$ . Proposition 6, together with Popov’s approximation lemma, leads to the following statement.

**Proposition 7.** *Let  $f \in \mathcal{F}_{r_0}^a$ ,  $0 < u_0 \leq r < r_0$  and  $u_j = 2^{-j}u_0$  for  $j \in \mathbb{N}$ . There exists a sequence of real-analytic functions  $f_j: \mathbb{T}_{u_j}^n \rightarrow \mathbb{C}$  such that*

$$\begin{cases} |f_0|_{u_0} \leq C_2 \|f\|_{r_0}, \\ |f_{j+1} - f_j|_{u_{j+1}} \leq C_2 \|f\|_{r_0} e^{-r\varphi_b(r/u_j)}, \\ |f_j - f|_0 \leq C_2 \|f\|_{r_0} e^{-r\varphi_b(r/u_j)} \end{cases}$$

with a positive constant  $C_2 = C_2(a, r, r_0, n)$ .

Finally, as in [18], we shall prove that the above proposition admits a partial converse; we shall state it in a way adapted to the proof of Theorem A.

**Proposition 8.** *Given  $r > 0$ , consider for  $j \in \mathbb{N}$  a geometric sequence  $w_{j-1} = 2^{-j}w_{-1} < r$  and a sequence of real-analytic functions  $f^{j-1}: \mathbb{T}_{w_{j-1}}^n \rightarrow \mathbb{C}$  such that  $f^{-1} = 0$ , and assume that*

$$|f^j - f^{j-1}|_{w_{j-1}} \leq e^{-v\varphi_b(r/w_{j-1})}, \quad w_{j-1} < r, \quad j \in \mathbb{N}$$

for some constant  $0 < v < r$ . Then for any  $0 < v < v$ , if

$$e^{-(v-v)(2^b-1)\varphi_b(r/w_{-1})} \leq \frac{1}{2}, \tag{4.5}$$

the sequence  $f^{j-1}$  converges in  $\mathcal{F}_v^a$  and the limit  $f$  satisfies

$$\|f\|_v \leq 2e^{-(v-v)\varphi_b(r/w_{-1})}.$$

*Proof.* Let us define  $h^j = f^j - f^{j-1}$  and expand it into Fourier series  $h^j = \sum_{k \in \mathbb{Z}^n} h_k^j e_k$ : since  $h^j$  is analytic we have

$$|h_k^j| \leq |h_j|_{w_{j-1}} e^{-w_{j-1}|k|} \leq e^{-v\varphi_b(r/w_{j-1})} e^{-w_{j-1}|k|}$$

so that

$$\sum_{j \geq 0} \|h^j\|_v = \sum_{j \geq 0} \sup_{k \in \mathbb{Z}^n} |h_k^j| e^{v\varphi_a(|k|)} \leq \sum_{j \geq 0} e^{-v\varphi_b(r/w_{j-1})} \sup_{t \geq 0} \{e^{v\varphi_a(t) - w_{j-1}t}\};$$

therefore, from (4.2) and (4.4) we have

$$\sum_{j \geq 0} \|h^j\|_v \leq \sum_{j \geq 0} e^{-v\varphi_b(r/w_{j-1})} e^{v\varphi_b(v/w_{j-1})} \leq \sum_{j \geq 0} e^{-(v-v)\varphi_b(r/w_{j-1})} < 2e^{-(v-v)\varphi_b(r/w_{-1})},$$

where we used the fact that  $\varphi_b(v/w_{j-1}) \leq \varphi_b(r/w_{j-1})$  since  $0 < v < v < r$  and  $r > w_{j-1}$ , and (4.5) which allows us to bound the last series by a geometric series. It follows that the sum of the  $h^j$  converges normally in  $\mathcal{F}_v^a$ , and since the latter is a Banach space, the sum converges to some  $f$  which is necessarily the limit of  $f^{j-1}$  and from the last inequality and the fact that  $f^{-1} = 0$ , at the limit one has the wanted estimate. ■

### 4.3. Analytic KAM theorem, following Rüssmann

In this section we shall state an analytic KAM theorem adapted to a frequency  $\omega \in \mathcal{A}_{\tau_0}^a$ , which by definition (recall (3.4)) satisfies the inequalities

$$|2\pi k \cdot \omega| \geq \gamma_{\tau_0}(\omega) e^{-\tau_0 \varphi_a(|k|)}, \quad k \in \mathbb{Z}^n \setminus \{0\},$$

which can be written as

$$|k \cdot \omega| \geq \frac{\alpha}{\Delta(|k|)}, \quad \Delta(t) := e^{\tau_0 \varphi_a(t)}, \quad \alpha := \frac{\gamma_{\tau_0}(\omega)}{2\pi}, \quad k \in \mathbb{Z}^n \setminus \{0\}. \tag{4.6}$$

Vectors which satisfy (4.6) are clearly Bruno vectors, hence the results of Rüssmann (see [34, 35], see also [10, 30]) apply. However, those results apply to any Bruno vector  $\omega$  and as we will explain below, for the special vectors satisfying (4.6) they do not give the best quantitative result.<sup>2</sup> The difference is that in order to reach a statement valid for all Bruno vectors, it seems that one has to avoid using a “superlinear” scheme of convergence (Newton method) but rather use a scheme whose “speed” depends on the arithmetic property of  $\omega$ . Now, for vectors satisfying (4.6) (or more generally for the set of vectors

---

<sup>2</sup>One should point out that, unlike [35] for instance, the threshold of the main theorem in [34] is not correct: this comes from a slight mistake in Lemma 3.2 in that reference.

$\mathcal{B}^1$  we mentioned in Sections 2.4 and 4.4), a superlinear scheme is possible and does give a better quantitative result. Such an analytic KAM theorem with a superlinear scheme is contained in [33] where the sup-norm is used and a variant of this scheme (in a much more general setting) is contained in [29] where the Fourier  $l_1$ -norm is used; it turns out that using the Fourier norm, and consequently the results in [29], is much more practical (this will also be explained below).

So to state the main result of [29] (in the simple setting we are considering), recall that  $\Delta$  was defined in (4.6), and we now define

$$\Gamma(\sigma) := \sup_{t \geq 0} \{(1+t)\Delta(t)e^{-\sigma t}\}, \quad \sigma > 0, \tag{4.7}$$

and for a given  $1 < \kappa < 2$ , we set

$$\Psi_\kappa(\sigma) := \inf \prod_{j=0}^{+\infty} \Gamma(\sigma_j)^{\kappa_j}, \quad \kappa_j := (\kappa - 1)\kappa^{-(j+1)}, \tag{4.8}$$

where the infimum (which can be shown to be a minimum) is taken over all non-increasing sequences  $(\sigma_j)_{j \in \mathbb{N}}$  whose sum is less than or equal to  $\sigma$ . The fact that (4.7) and (4.8) are indeed finite will be verified later; we shall actually need explicit estimates for them. Given  $s > 0$  and  $f: \mathbb{T}^n \rightarrow \mathbb{R}$  we recall that

$$\|f\|_s^1 := \sum_{k \in \mathbb{Z}^n} |f_k| e^{s|k|}$$

and the assumption that  $\|f\|_s^1 < +\infty$  implies that  $f$  extends to a holomorphic function on  $\mathbb{T}_s^n$ . We then extend the definition of the above norm for vector fields on  $\mathbb{T}^n$  by taking the maximum norm of each component. Here is the analytic KAM theorem we shall rely on.

**Proposition 9.** *Let  $0 < a < 1$ ,  $1 < \kappa < 2$ ,  $0 < 2\sigma < s$  and for  $X = X_\omega + F$  with  $\omega$  satisfying (4.6), assume that*

$$C_4 \alpha^{-1} \Psi_\kappa(\sigma) \|F\|_s^1 \leq 1$$

for some positive numerical constant  $C_4$ . Then there exist  $\lambda \in \mathbb{R}^n$  and  $\Phi: \mathbb{T}_{s-2\sigma}^n \rightarrow \mathbb{T}_s^n$  such that

$$\Phi^*(X - \lambda) = X_\omega \tag{4.9}$$

with the estimates

$$|\lambda| \leq C_4 \|F\|_s^1, \quad \max\{\|\Phi - \text{Id}\|_{s-2\sigma}^1, \|D\Phi - \text{Id}\|_{s-2\sigma}^1\} \leq C_4 \alpha^{-1} \Psi_\kappa(\sigma) \|F\|_s^1.$$

This is a direct consequence of [29, Theorem A] and [29, Estimates of Theorem A] in the very special case where the Hamiltonian is linear with respect to the action variables, there are no elliptic variables (one can put  $M = 0$  in the above reference) and the Cantor

set of frequencies is reduced to a single point. We shall now make several modifications to this result in order to have a statement which will be more convenient in this sequel.

First we shall actually use Proposition 9 in the case where  $s$ , and thus  $\sigma < s/2$ , is a small parameter and in this case, we shall obtain explicit estimates for  $\Gamma(\sigma)$  in (4.7) and  $\Psi_\kappa(\sigma)$  in (4.8). Observe that

$$\Gamma(\sigma) = \sup_{t \geq 0} \{(1+t)e^{\tau_0 \varphi_a(t) - \sigma t}\}$$

and the supremum is reached at a value  $t_\sigma \rightarrow +\infty$  as  $\sigma \rightarrow 0$  and that within this limit, the polynomially large factor  $(1+t)$  is dominated by the exponentially large factor  $e^{\tau_0 \varphi_a(t)}$ . Hence given  $\tau > \tau_0$ , there exists  $s^* = s^*(a, \tau, \tau_0) > 0$  such that for  $0 < 2\sigma < s \leq s^*$ , we have

$$\Gamma(\sigma) \leq \sup_{t \geq 0} \{e^{\tau \varphi_a(t) - \sigma t}\}$$

and recalling the definition of  $\varphi_a^*$  in (4.1), together with relations (4.2) and (4.4), this gives

$$\Gamma(\sigma) \leq e^{\tau \varphi_b(\tau/\sigma)}.$$

If we had used the sup-norm (as in [33]) instead of the Fourier  $l_1$ -norm, the supremum in the definition of  $\Gamma(\sigma)$  in (4.7) would have to be replaced by a sum (or an integral) and the computations in this case are less explicit: this is the reason why we choose the Fourier  $l_1$ -norm (sup-norms will be converted into Fourier  $l_1$ -norms below). Next we shall estimate  $\Psi_\kappa(\sigma)$  in (4.8), and actually, we have nothing to do since it is explicitly done in [29, Lemma 6]: if we optimize among geometric sequences one is lead to consider

$$\sigma_j := \sigma(\kappa^{1-a} - 1)\kappa^{-(1-a)(j+1)}, \quad j \in \mathbb{N}$$

which gives the estimate

$$\Psi_\kappa(\sigma) \leq e^{\delta \tau \varphi_b(\tau/\sigma)}, \quad \delta = \delta(a, \kappa) := (\kappa - 1)(\kappa^{1-a} - 1)^{-1/(1-a)}. \tag{4.10}$$

Observe that as  $a \rightarrow 1$ , then  $\delta \rightarrow +\infty$ , and as  $a \rightarrow 0$ , then  $\delta \rightarrow 1$  which is the best one can expect: if we had used results valid for any Bruno vectors such as [35], we would have found another constant  $\delta'$  for which  $\delta' \rightarrow +\infty$  as  $a \rightarrow 1$  but also as  $a \rightarrow 0$ , which is clearly not natural.

Next we convert sup-norms into Fourier  $l_1$ -norms by using the well-known relations

$$\|f\|_s \leq \|f\|_s^1, \quad \|f\|_{s-2\sigma}^1 \leq (\coth \sigma)^n \|f\|_s, \quad 0 < 2\sigma < s.$$

Using again  $\tau > \tau_0$ , the polynomial large factor  $(\coth \sigma)^n$  can be absorbed by the exponential large factor in (4.10) (strictly speaking, we should introduced yet another  $\tau' > \tau > \tau_0$  but clearly one can replace  $\tau$  by  $\tau_0 + (\tau - \tau_0)/2$  and then take  $\tau' = \tau$ ) and it follows that Proposition 9 holds true with (4.10) if we replace the Fourier  $l_1$ -norm by the sup-norm, require  $0 < 4\sigma < s$  instead of  $0 < 2\sigma < s$  and allow the constant  $C_4$  to depend now on  $\tau, \tau_0, a, \kappa$  and  $n$ .

Finally, as observed by Moser in [25], the modifying term  $\lambda$  in (4.9) need not be constant and it can be replaced by a non-constant modifying term of the form  $\Theta^*\lambda$ , that is,

$$\Theta^*\lambda(x) := (D_x\Theta)^{-1}\lambda$$

with

$$\Theta: \mathbb{T}_s^n \rightarrow \mathbb{C}^n/\mathbb{Z}^n, \quad |D\Theta - \text{Id}|_s \leq \frac{1}{3}. \tag{4.11}$$

Indeed, the last inequality implies  $|(D\Theta)^{-1} - \text{Id}|_s \leq 1/2$  and therefore, for any  $\lambda \in \mathbb{R}^n$ , we have

$$\frac{|\lambda|}{2} \leq |\Phi^*\lambda|_s \leq \frac{3|\lambda|}{2}.$$

We refer to [25] for the reduction of this seemingly more general statement to the case where  $\Theta = \text{Id}$ ; clearly this only changes  $C_4$  by a numerical factor. With all these modifications, we can finally state a more convenient version of Proposition 9.

**Proposition 10.** *Let  $0 < a < 1$ ,  $\tau > \tau_0 > 0$ ,  $1 < \kappa < 2$ . There exists  $s^* = s^*(a, \tau, \tau_0)$  such that for any  $0 < 4\sigma < s \leq s^*$ , the following holds true. Given  $X = X_\omega + F$  with  $s$  satisfying (4.6) and  $\Theta$  satisfying (4.11), if we assume that*

$$C_4\alpha^{-1}e^{\delta\tau\varphi_b(\tau/\sigma)}|F|_s \leq 1, \quad \delta = (\kappa - 1)(\kappa^{1-a} - 1)^{-1/(1-a)} \tag{4.12}$$

for some positive constant  $C_4 = C_4(a, \kappa, \tau, \tau_0, n)$ , then there exist  $\lambda \in \mathbb{R}^n$  and  $\Phi: \mathbb{T}_{s-4\sigma}^n \rightarrow \mathbb{T}_s^n$  such that

$$\Phi^*(X - \Theta^*\lambda) = X_\omega$$

with the estimates and

$$|\lambda| \leq C_4|F|_s, \quad \max\{|\Phi - \text{Id}|_{s-4\sigma}, |D\Phi - \text{Id}|_{s-4\sigma}\} \leq C_4\alpha^{-1}|F|_s e^{\delta\tau\varphi_b(\tau/\sigma)}.$$

#### 4.4. Proof of Theorem A

This section is entirely devoted to the proof of Theorem A. We recall that we are given  $0 < a < 1$ ,  $r_0 > r > 0$ ,  $\tau > \tau_0 > 0$ ,  $1 < \kappa < 2$ , such that

$$r > c\tau, \quad c = c(a, \kappa) := 32^b\delta = 32^b(\kappa - 1)(\kappa^{1-a} - 1)^{-1/(1-a)}. \tag{4.13}$$

Our assumption is that  $\omega \in \mathcal{A}_{\tau_0}^a$  and

$$\|F\|_{r_0} := \varepsilon_0$$

will be required to be sufficiently small. We now define  $u_0 = u_0(\varepsilon_0) > 0$  by the equality

$$e^{-r\varphi_b(r/(2u_0))} := \frac{C_2C_4\varepsilon_0}{\alpha} \leq \frac{1}{2}, \tag{4.14}$$

where  $\alpha = \gamma_{\tau_0}(\omega)/2\pi$  and  $C_2$  and  $C_4$  are the constants appearing in Propositions 7 and 10 respectively. Then we set

$$\sigma_0 = \frac{s_0}{8} = \frac{u_0}{16}, \tag{4.15}$$

and we define geometric sequences converging to zero,

$$\sigma_j := 2^{-j} \sigma_0, \quad s_j := 2^{-j} s_0, \quad u_j := 2^{-j} u_0. \tag{4.16}$$

We have already assumed that  $\varepsilon_0$  is small enough so that (4.14) is less than 1/2 (we will require much more than that in the sequel), and in view of (4.15), any further smallness condition on  $\varepsilon_0$  is equivalent to a smallness condition on  $u_0$  or  $\sigma_0$  or  $s_0$ . Thus we may assume  $2u_0 < r$ , so in particular Proposition 7 applies, with the sequence  $u_j$  defined in (4.16), to each component of  $F$  and yields a sequence of analytic vector fields  $F_j: \mathbb{T}_{u_j}^n \rightarrow \mathbb{C}$  such that

$$\begin{cases} |F_0|_{u_0} \leq C_2 \varepsilon_0, \\ |F_{j+1} - F_j|_{u_{j+1}} \leq C_2 \varepsilon_0 e^{-r\varphi_b(r/u_j)}, \\ |F_j - F|_0 \leq C_2 \varepsilon_0 e^{-r\varphi_b(r/u_j)}. \end{cases} \tag{4.17}$$

We wish to apply Proposition 10 to the analytic vector field  $X_0 = X_\omega + F_0$  on the domain  $\mathbb{T}_{s_0}^n$ , which makes sense since  $F_0$  is defined on  $\mathbb{T}_{u_0}^n$  and  $s_0 < u_0$ , and with  $\Theta = \text{Id}$ . To do so, first observe that  $s_0 \leq s^*$  is yet another smallness condition,  $4\sigma_0 = s_0/2 < s_0$  so that we only need to verify (4.12), which is implied by

$$e^{\delta\tau\varphi_b(\tau/\sigma_0)} \frac{C_2 C_4 \varepsilon_0}{\alpha} = e^{32^b \delta\tau\varphi_b(\tau/2u_0)} e^{-r\varphi_b(r/(2u_0))} \leq e^{-(r-c\tau)\varphi_b(r/(2u_0))} \leq 1, \tag{4.18}$$

where we used the equality in (4.14), the fact that  $\sigma_0 = u_0/16 = 2u_0/32$ ,  $2u_0 < r$  and the definition of  $c = 32^b \delta$ . But our assumption (4.13) is precisely that  $r > c\tau$ , therefore the last inequality in (4.18) holds true because of the inequality in (4.14). Consequently, Proposition 10 applies and, since  $s_0 - 4\sigma_0 = s_1$ , we obtain  $\lambda^0 \in \mathbb{R}^n$  and  $\Phi^0: \mathbb{T}_{s_1}^n \rightarrow \mathbb{T}_{s_0}^n$  such that

$$(\Phi^0)^*(X_0 - \lambda^0) = X_\omega,$$

and setting  $\rho = r - c\tau > 0$  we have the estimates

$$|\lambda^0| \leq \alpha e^{-r\varphi_b(r/(2u_0))}, \quad \max\{|\Phi^0 - \text{Id}|_{s_1}, |D\Phi^0 - \text{Id}|_{s_1}\} \leq e^{-\rho\varphi_b(r/(2u_0))}.$$

Moreover, since  $s_0 < u_1$ , we have  $\Phi^0: \mathbb{T}_{s_1}^n \rightarrow \mathbb{T}_{u_1}^n$ .

*Claim.* Set  $\lambda^{-1} = 0 \in \mathbb{R}^n$ ,  $\Phi^{-1} = \text{Id}$  and  $u_{-1} = 2u_0$ . We claim that for any  $j \in \mathbb{N}$ , there exist  $\Phi^j: \mathbb{T}_{s_{j+1}}^n \rightarrow \mathbb{T}_{u_{j+1}}^n$  and  $\lambda^j \in \mathbb{R}^n$  such that for  $X_j = X_\omega + F_j$  we have

$$(\Phi^j)^*(X_j - \lambda^j) = X_\omega \tag{4.19}$$

and

$$\begin{cases} |\lambda^j - \lambda^{j-1}| \leq \alpha e^{-r\varphi_b(r/u_{j-1})}, \\ \max\{|\Phi^j - \Phi^{j-1}|_{s_{j+1}}, |D\Phi^j - D\Phi^{j-1}|_{s_{j+1}}\} \leq e^{-\rho\varphi_b(r/u_{j-1})}. \end{cases} \tag{4.20}$$



*Proof of the claim.* We have already shown the claim to be true for  $j = 0$ , so we may proceed by induction and we assume the statement to hold true for  $0 \leq i \leq j$  with  $j \in \mathbb{N}$ , and we need to show that it remains true for  $j + 1$ . First observe that from (4.20), for  $u_0$  small enough one has

$$\max\{|\Phi^j - \text{Id}|_{s_{j+1}}, |D\Phi^j - \text{Id}|_{s_{j+1}}\} \leq 2e^{-\rho\varphi_b(r/(2u_0))} \leq \frac{1}{3}. \quad (4.21)$$

Then we write

$$X_{j+1} = X_\omega + F_{j+1} = X_j + (F_{j+1} - F_j)$$

and we apply our inductive assumption to get

$$(\Phi^j)^*(X_{j+1} - \lambda^j) = (\Phi^j)^*(X_j - \lambda^j) + (\Phi^j)^*(F_{j+1} - F_j) = X_\omega + G_j, \quad (4.22)$$

where  $G_j = (\Phi^j)^*(F_{j+1} - F_j)$  is well defined since  $\Phi^j$  maps  $\mathbb{T}_{s_{j+1}}^n$  into  $\mathbb{T}_{u_{j+1}}^n$  which is the domain of definition of  $F_{j+1} - F_j$ ; it follows from (4.21) and (4.17) that

$$|G_j|_{s_{j+1}} \leq |D\Phi^j|_{s_{j+1}} |F_{j+1} - F_j|_{u_{j+1}} \leq 2C_2\varepsilon_0 e^{-r\varphi_b(r/u_j)}. \quad (4.23)$$

We wish to apply Proposition 10 to the vector field (4.22), with  $\Theta = \Phi^j$ : it follows from (4.21) that (4.11) holds true; obviously  $s_{j+1} \leq s^*$  since this is the case for  $j = -1$ . Also, in view of (4.23), to verify (4.12) one needs to check that

$$\begin{aligned} e^{-\delta\tau\varphi_b(\tau/\sigma_{j+1})} \left( \frac{2C_2C_4\varepsilon_0}{\alpha} \right) e^{-r\varphi_b(r/u_j)} &\leq e^{-32^b\delta\tau\varphi_b(\tau/u_j)} e^{-r\varphi_b(r/u_j)} \\ &\leq e^{-\rho\varphi_b(r/u_j)} \leq 1, \end{aligned} \quad (4.24)$$

where we used the inequality in (4.14), the fact that  $\sigma_{j+1} = u_{j+1}/16 = u_j/32$  and the definition of  $\rho = r - 32^b\delta\tau$ . Again, (4.24) holds true since it holds true for  $j = -1$  in view of (4.18). Hence Proposition 10 applies (with  $\Theta = \Phi^j$ ) and gives a vector  $\lambda_{j+1}$  and a transformation  $\Phi_{j+1}$  with the estimates

$$|\lambda_{j+1}| \leq \alpha e^{-r\varphi_b(r/u_j)}, \quad \max\{|\Phi_{j+1} - \text{Id}|_{s_{j+2}}, |D\Phi_{j+1} - \text{Id}|_{s_{j+2}}\} \leq e^{-\rho\varphi_b(r/u_j)} \quad (4.25)$$

so that

$$(\Phi_{j+1})^*((\Phi^j)^*(X_{j+1} - \lambda^j) - (\Phi^j)^*\lambda_{j+1}) = X_\omega.$$

Hence if we set  $\Phi^{j+1} = \Phi^j \circ \Phi_{j+1}$ ,  $\lambda^{j+1} = \lambda^j + \lambda_{j+1}$  then

$$(\Phi^{j+1})^*(X_{j+1} - \lambda^{j+1}) = X_\omega$$

and the estimates (4.20), with  $j$  replaced by  $j + 1$ , follow from (4.25) and (4.21). Finally, since (4.20) holds true with  $j$  replaced by  $j + 1$ , then (4.21) as well holds true with  $j$  replaced by  $j + 1$ , and since the transformations are real, for  $\theta \in \mathbb{T}_{s_{j+2}}^n$  we have

$$|\text{Im}(\Phi^{j+1}(\theta))| \leq |D\Phi^{j+1}|_{s_{j+2}} |\text{Im}(\theta)| \leq 2s_{j+2} = u_{j+2}$$

and therefore  $\Phi^{j+1}$  maps  $\mathbb{T}_{s_{j+2}}^n$  into  $\mathbb{T}_{u_{j+2}}^n$ , hence the claim is proved.  $\blacksquare$

To conclude the proof of Theorem A, we define  $w_{j-1} = s_{j+1} = u_{j-1}/8$  for  $j \in \mathbb{N}$  so that (4.20) gives

$$\max\{|\Phi^j - \Phi^{j-1}|_{w_{j-1}}, |D\Phi^j - D\Phi^{j-1}|_{w_{j-1}}\} \leq e^{-v\varphi_b(r/(w_{j-1}))}, \quad v := 8^{-b}\rho.$$

Since  $w_{-1} = u_0/4 < r$ , if we fix  $0 < v = v/2 < v$ , the assumptions of Proposition 8 are satisfied because (4.5) is yet another smallness condition, so we can apply this proposition to each component of  $\Phi^{j-1} - \text{Id}$  and  $D\Phi^{j-1} - \text{Id}$  and consequently  $\Phi^{j-1}$  converges to some map  $\Phi \in \mathcal{F}_v^a$ , which is necessarily a diffeomorphism in view of the convergence of  $D\Phi^{j-1}$ , and satisfies the estimate

$$\|\Phi - \text{Id}\|_v \leq 2e^{-(v-v)\varphi_b(r/w_{-1})} = 2e^{-8^{-b}(v-v)\varphi_b(r/2u_0)} = 2\left(\frac{C_2C_4\varepsilon_0}{\alpha}\right)^\iota \tag{4.26}$$

with  $\iota = 8^{-b}(2r)^{-1}v$ , where we used the definition in (4.14), whereas for the  $\lambda^{j-1}$ , it follows directly from (4.20) that it converges to some  $\lambda \in \mathbb{R}$  which satisfies

$$|\lambda| \leq 2\alpha e^{-r\varphi_b(r/u_{-1})} = 2\alpha e^{-r\varphi_b(r/(2u_0))} = 2C_2C_4\varepsilon_0. \tag{4.27}$$

Since  $F_j$  converges uniformly to  $F$  by (4.17),  $X_j$  converges uniformly to  $X = X_\omega + F$  and going to the limit  $j \rightarrow +\infty$  in (4.19) we find

$$\Phi^*(X - \lambda) = X_\omega$$

and the wanted estimates follows from (4.26) and (4.27).

**Funding.** The author has benefited from partial funding from the ANR project Beyond KAM (ANR-15-CE40-0001).

## References

- [1] M. Andersson and B. Berndtsson, Almost holomorphic extensions of ultradifferentiable functions. *J. Anal. Math.* **89** (2003), 337–365 Zbl [1061.30028](#) MR [1981924](#)
- [2] V. I. Arnol'd, Small denominators. I. Mapping the circle onto itself. *Izv. Akad. Nauk SSSR Ser. Mat.* **25** (1961), 21–86 Zbl [0135.42603](#) MR [0140699](#)
- [3] A. Avila, B. Fayad, and R. Krikorian, A KAM scheme for  $\text{SL}(2, \mathbb{R})$  cocycles with Liouvillean frequencies. *Geom. Funct. Anal.* **21** (2011), no. 5, 1001–1019 Zbl [1277.37089](#) MR [2846380](#)
- [4] T. Bang, *Om quasi-analytiske Funktioner*. Thesis, University of Copenhagen, 1946 MR [0017782](#)
- [5] G. Björck, Linear partial differential operators and generalized distributions. *Ark. Mat.* **6** (1966), 351–407 Zbl [0166.36501](#) MR [203201](#)
- [6] J. Boman, On the intersection of classes of infinitely differentiable functions. *Ark. Mat.* **5** (1963/65), 301–309 (1963/65) Zbl [0135.26104](#) MR [220049](#)
- [7] A. Bounemoura, C. Chavaudret, and S. Liang, Reducibility of ultra-differentiable quasiperiodic cocycles under an adapted arithmetic condition. *Proc. Amer. Math. Soc.* **149** (2021), no. 7, 2999–3012 Zbl [07352298](#) MR [4257810](#)

- [8] A. Bounemoura and J. Féjzo, Hamiltonian perturbation theory for ultra-differentiable functions. *Mem. Amer. Math. Soc.* **270** (2021), no. 1319, v+89 Zbl [07403464](#) MR [4273633](#)
- [9] A. Bounemoura and J. Féjzo, KAM,  $\alpha$ -Gevrey regularity and the  $\alpha$ -Bruno-Rüssmann condition. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **19** (2019), no. 4, 1225–1279 Zbl [07224201](#) MR [4050197](#)
- [10] A. Bounemoura and S. Fischler, A Diophantine duality applied to the KAM and Nekhoroshev theorems. *Math. Z.* **275** (2013), no. 3-4, 1135–1167 Zbl [1308.37028](#) MR [3127051](#)
- [11] R. W. Braun, R. Meise, and B. A. Taylor, Ultradifferentiable functions and Fourier analysis. *Results Math.* **17** (1990), no. 3-4, 206–237 Zbl [0735.46022](#) MR [1052587](#)
- [12] A. D. Brjuno, Analytic form of differential equations. I. *Trudy Moskov. Mat. Obšč.* **25** (1971), 119–262 Zbl [0263.34003](#) MR [0377192](#)
- [13] I. Cioranescu and L. Zsidó,  $\omega$ -ultradistributions and their application to operator theory. In *Spectral theory (Warsaw, 1977)*, pp. 77–220, Banach Center Publ. 8, PWN, Warsaw, 1982 Zbl [0497.46026](#) MR [738280](#)
- [14] E. I. Dinaburg and J. G. Sinaĭ, The one-dimensional Schrödinger equation with quasiperiodic potential. *Funkcional. Anal. i Priložen.* **9** (1975), no. 4, 8–21 Zbl [0333.34014](#) MR [0470318](#)
- [15] E. M. Dyn'kin, The pseudoanalytic extension. *J. Anal. Math.* **60** (1993), 45–70 Zbl [0795.30034](#) MR [1253229](#)
- [16] S. Fördös, D. N. Nanning, A. Rainer, and G. Schindl, Almost analytic extensions of ultradifferentiable functions with applications to microlocal analysis. *J. Math. Anal. Appl.* **481** (2020), no. 1, 123451, 51 Zbl [1427.32009](#) MR [4002151](#)
- [17] M.-R. Herman, Exemples de flots hamiltoniens dont aucune perturbation en topologie  $C^\infty$  n'a d'orbites périodiques sur un ouvert de surfaces d'énergies. *C. R. Acad. Sci. Paris Sér. I Math.* **312** (1991), no. 13, 989–994 Zbl [0759.58016](#) MR [1113091](#)
- [18] X. Hou and G. Popov, Rigidity of reducibility of Gevrey quasi-periodic cocycles on  $U(n)$ . *Bull. Soc. Math. France* **144** (2016), no. 1, 1–52 Zbl [1375.37009](#) MR [3481260](#)
- [19] X. Hou and J. You, Almost reducibility and non-perturbative reducibility of quasi-periodic linear systems. *Invent. Math.* **190** (2012), no. 1, 209–260 Zbl [1294.37027](#) MR [2969277](#)
- [20] A. N. Kolmogorov, On dynamical systems with an integral invariant on the torus. *Doklady Akad. Nauk SSSR (N.S.)* **93** (1953), 763–766 Zbl [0052.31904](#) MR [0062892](#)
- [21] A. N. Kolmogorov, On conservation of conditionally periodic motions for a small change in Hamilton's function. *Dokl. Akad. Nauk SSSR (N.S.)* **98** (1954), 527–530 Zbl [0056.31502](#) MR [0068687](#)
- [22] S. Łojasiewicz, Jr. and E. Zehnder, An inverse function theorem in Fréchet-spaces. *J. Functional Analysis* **33** (1979), no. 2, 165–174 Zbl [0431.46032](#) MR [546504](#)
- [23] J. Lopes Dias and J. P. Gaivão, Linearization of Gevrey flows on  $\mathbb{T}^d$  with a Brjuno type arithmetical condition. *J. Differential Equations* **267** (2019), no. 12, 7167–7212 Zbl [1441.37066](#) MR [4011043](#)
- [24] S. Marmi and D. Sauzin, Quasianalytic monogenic solutions of a cohomological equation. *Mem. Amer. Math. Soc.* **164** (2003), no. 780, vi+83 Zbl [1054.30046](#) MR [1982715](#)
- [25] J. Moser, A rapidly convergent iteration method and non-linear differential equations. II. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3)* **20** (1966), 499–535 Zbl [0144.18202](#) MR [206461](#)
- [26] J. Moser, Convergent series expansions for quasi-periodic motions. *Math. Ann.* **169** (1967), 136–176 Zbl [0149.29903](#) MR [208078](#)
- [27] H.-J. Petzsche and D. Vogt, Almost analytic extension of ultradifferentiable functions and the boundary values of holomorphic functions. *Math. Ann.* **267** (1984), no. 1, 17–35 Zbl [0517.30028](#) MR [737333](#)

- [28] G. Popov, KAM theorem for Gevrey Hamiltonians. *Ergodic Theory Dynam. Systems* **24** (2004), no. 5, 1753–1786 Zbl [1088.37030](#) MR [2104602](#)
- [29] J. Pöschel, On elliptic lower-dimensional tori in Hamiltonian systems. *Math. Z.* **202** (1989), no. 4, 559–608 Zbl [0662.58037](#) MR [1022821](#)
- [30] J. Pöschel, KAM à la R. *Regul. Chaotic Dyn.* **16** (2011), no. 1-2, 17–23 Zbl [1219.37044](#) MR [2774375](#)
- [31] E. Risler, Linéarisation des perturbations holomorphes des rotations et applications. *Mém. Soc. Math. Fr. (N.S.)* (1999), no. 77, viii+102 Zbl [0929.37017](#) MR [1779976](#)
- [32] H. Rüssmann, On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus. In *Dynamical systems, theory and applications (Rencontres, Battelle Res. Inst., Seattle, Wash., 1974)*, pp. 598–624, Lecture Notes in Phys. 38, Springer, Berlin, 1975 Zbl [0319.35017](#) MR [0467824](#)
- [33] H. Rüssmann, On the one-dimensional Schrödinger equation with a quasiperiodic potential. In *Nonlinear dynamics (Internat. Conf., New York, 1979)*, pp. 90–107, Ann. New York Acad. Sci. 357, New York Acad. Sci., New York, 1980 Zbl [0477.34007](#) MR [612810](#)
- [34] H. Rüssmann, On the frequencies of quasi-periodic solutions of analytic nearly integrable Hamiltonian systems. In *Seminar on Dynamical Systems (St. Petersburg, 1991)*, edited by S. Kuksin, V. Lazutkin, and J. Pöschel, pp. 160–183, Progr. Nonlinear Differential Equations Appl. 12, Birkhäuser, Basel, 1994 Zbl [0797.58028](#) MR [1279397](#)
- [35] H. Rüssmann, Invariant tori in non-degenerate nearly integrable Hamiltonian systems. *Regul. Chaotic Dyn.* **6** (2001), no. 2, 119–204 Zbl [0992.37050](#) MR [1843664](#)
- [36] D. A. Salamon, The Kolmogorov-Arnold-Moser theorem. *Math. Phys. Electron. J.* **10** (2004), Paper 3, 37 Zbl [1136.37348](#) MR [2111297](#)
- [37] J.-C. Yoccoz, Analytic linearization of circle diffeomorphisms. In *Dynamical systems and small divisors (Cetraro, 1998)*, pp. 125–173, Lecture Notes in Math. 1784, Springer, Berlin, 2002 Zbl [1417.37153](#) MR [1924912](#)

Received 12 June 2020; revised 4 May 2021; accepted 21 June 2021.

### **Abed Bounemoura**

CEREMADE, CNRS, PSL Research University, Université Paris-Dauphine and Observatoire de Paris, Place du Maréchal de Lattre de Tassigny, 75775 Paris, France;  
[bounemoura@ceremade.dauphine.fr](mailto:bounemoura@ceremade.dauphine.fr)