KAM, α-Gevrey regularity and the α-Bruno-Rüssmann condition

Abed Bounemoura and Jacques Féjoz

February 8, 2018

Abstract

We prove a new invariant torus theorem, for α-Gevrey smooth Hamiltonian systems, under an arithmetic assumption which we call the α-Bruno-Rüssmann condition, and which reduces to the classical Bruno-Rüssmann condition in the analytic category. Our proof is direct in the sense that, for analytic Hamiltonians, we avoid the use of complex extensions and, for non-analytic Hamiltonians, we do not use analytic approximation nor smoothing operators. Following Bessi, we also show that if a slightly weaker arithmetic condition is not satisfied, the invariant torus may be destroyed. Crucial to this work are new functional estimates in the Gevrey class.

1 Introduction

1.1 The general question

We consider small perturbations of an integrable Hamiltonian system, defined by

$$\dot{q} = \nabla_p H(q,p), \quad \dot{p} = -\nabla_q H(q,p)$$

where H is a Hamiltonian of the form

$$H(q,p) = h(p) + \epsilon f(q,p), \quad (q,p) \in \mathbb{T}^n \times \mathbb{R}^n, \quad 0 \leq \epsilon < 1$$

where $n \geq 2$, $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$, $\omega_0 = \nabla h(0) \in \mathbb{R}^n$, and $\nabla^2 h(0) \in M_n(\mathbb{R})$ is non-degenerate. When $\epsilon = 0$, the torus $\mathcal{T}_0$ of equation $p = 0$ is invariant and quasi-periodic of frequency ω_0. The general question we are interested in is the persistence of this torus for $\epsilon \neq 0$ sufficiently small : does there exist a torus $\mathcal{T}_\epsilon$ which is invariant and quasi-periodic of frequency ω_0 and which converges (in a suitable sense) to $\mathcal{T}_0$ as ϵ goes to zero?

This question was answered positively by Kolmogorov in his foundational paper [Kol54] under the assumption that H is real-analytic and ω_0 is a τ-Diophantine vector ($\tau \geq n-1$): there exists $\gamma > 0$ such that for all $k \in \mathbb{Z}^n \setminus \{0\}$, $|k \cdot \omega_0| \geq \gamma |k|^{-\tau}$. As a conclusion, the perturbed torus is real-analytic. It became clear that a regularity assumption on H and an

arithmetic condition on ω_0 were necessary, and then further works investigate the interplay between the analysis and the arithmetic.

It was certainly a remarkable contribution of Moser (see [Mos62]) to realize that the question can also be answered for Hamiltonians which are only finitely differentiable. More precisely (see [Sal04]), if ω_0 is τ-Diophantine and if H is of class C^r, with $r > 2(\tau + 1)$, then the torus persists and it is of class $C^{r'+\tau+1}$ for any $r' < r - 2(\tau + 1)$. If H is smooth, that is C^∞, there is no restriction on τ and the perturbed torus is smooth. It follows from a recent result of Cheng-Wang [CW13] (which uses an idea of Bessi [Bes00]) that the result is false if H is of class C^r, with $r < 2(\tau + 1)$. Thus in the finitely differentiable or smooth case, one may consider this Diophantine condition as essentially optimal.

In the real-analytic setting, the Diophantine condition is not necessary. Indeed, it is sufficient to assume that ω_0 satisfies the weaker Bruno-Rüssmann condition (see §2.2 for a definition), as was first proved by Rüssmann in [Rüs01]; an equivalent condition was actually introduced earlier by Bruno [Bru71], [Bru72] in a different but related small divisors problem, the Siegel linearization problem. The necessity of this condition turns out to be a more subtle problem. In the Siegel problem, it is optimal in dimension one (this is a celebrated result of Yoccoz [Yoc88], [Yoc95]) but in higher dimension it is unknown. In the Hamiltonian problem we are considering here, the only general result we are aware of is due to Bessi [Bes00] (extending an earlier result of Forni [For94] for twist maps of the annulus) in which a torus with a frequency not satisfying a slightly weaker condition can be destroyed by an arbitrary small analytic perturbation. This leaves open the possibility of slightly improving the Bruno-Rüssmann condition.

1.2   Main results of the paper

Real-analytic functions are characterized by a growth of their derivatives of order $s^{-|k|}k!$ for some analyticity width $s > 0$; in the periodic case, this is equivalent to a decay of Fourier coefficients of order $e^{-s|k|}$. Given a real parameter $\alpha \geq 1$, allowing a growth of the derivatives of order $s^{-|k|}k!^\alpha$ or, equivalently, a decay of Fourier coefficients of order $e^{-\alpha s|k|^{1/\alpha}}$, one is lead to consider α-Gevrey functions, which thus corresponds to real-analytic functions when $\alpha = 1$. Since the introduction by Gevrey of the class of functions now baring his name ([Gev18]), there has been a huge amount of works on Gevrey functions, mainly for PDEs, but also more recently in other fields, including dynamical systems (see §1.3 for some related works in dynamical systems dealing with Gevrey regularity).

In this paper, we study the persistence when ϵ is small of the torus $\mathcal{T}_0$, as a Gevrey quasiperiodic invariant torus $\mathcal{T}_\varepsilon$, under the assumptions that H itself has Gevrey regularity. The only general result so far is due to Popov [Pop04] who proved that the latter holds true if ω satisfies a Diophantine condition. This result, the proof of which uses analytic approximation, extends the result of Kolmogorov when $\alpha = 1$ but not the one of Rüssmann: clearly one would expect an arithmetic condition which does depend on α and that reduces to the Bruno-Rüssmann condition when $\alpha = 1$.

The main result of the paper is to solve this persistency problem, assuming that the frequency ω_0 satisfies some arithmetic condition which we call the α-Bruno-Rüssmann

condition, which is weaker than the Diophantine condition and agrees with the Bruno-Rüssmann condition when $\alpha = 1$. This is the content of Theorem A; Theorem B and Theorem C deal respectively with the iso-energetic and time-periodic versions. We will also state and prove a Gevrey analogue of Arnold's normal form theorem for vector fields on the torus (Theorem E). Theorem H is a more precise, quantitative statement, with parameters, which does not require non-degeneracy, and from which Theorem A and Theorem E follow. We also notice that Bessi's ideas [Bes00] may be adapted to the Gevrey setting, to provide a necessary arithmetic condition for the invariant torus to persist (Theorem D). The so-obtained condition fails to agree with the sufficient condition of Theorem A and, as in the analytic case, it remains open to determine the optimal condition. Finally, we will also give discrete versions of Theorem A and Theorem E, which are, respectively, Theorem F and Theorem G.

When a Hamiltonian is not real analytic, it is often the case that there is still some control on its derivatives and that it has Gevrey regularity. This may happen for example for the restriction of an analytic Hamiltonian restricted to a Gevrey, symplectic, central manifold. Technically, Gevrey regularity luckily extends the well-behaved analytic regularity in KAM theory: the effect of small denominators in Fourier series reduces to decreasing the "Gevrey width" s, the analogue of the analyticity width. This makes it possible to adapt Kolmogorov's proof of his invariant torus theorem without using analytic approximations or smoothing operators as in the smooth setting. Yet there are two issues one needs to solve.

The first and main issue is that the estimates needed in the general problem of perturbation theory were missing. This is why we provide an appendix with an adequate choice of norms and spaces, together with the estimates needed in our proof. In particular, Proposition 20 provides a "geometric" estimate of the composition of two Gevrey functions, in which the loss of Gevrey width is arbitrarily small when composing a function to the right by a diffeomorphism close to the identity, in continuity with the real-analytic setting. Starting with the work of Gevrey itself [Gev18], there have been many results concerning the composition of Gevrey functions (see, for instance, Yamanaka [Yam89], Marco-Sauzin [MS02], Cadeddu-Gramchev [CG03], Popov [Pop04]) but none of them allowed an arbitrarily small loss of width except in some particular cases (the one-dimensional case and the analytic case). To our knowledge, our composition result is new and may be of independent interest.

The second and minor issue is that to reach a weak arithmetic condition, it is usually better not to solve exactly the cohomological equation but an approximate version of it, and hence one cannot proceed as in Kolmogorov's proof. The strategy of Rüssmann, that we could have tried to pursue here, consists in solving this equation not for the original perturbation but for a polynomial approximation of it. We will rather adopt the strategy of [BF13], [BF14] in which periodic approximations of the frequency are used and only cohomological equations associated to periodic vectors need to be solved: estimates on the solution are straightforward in this case, unlike the cohomological equation associated to a non-resonant vector.

As a further remark concerning the proof, in invariant tori problems derivatives in the

3

angle and action directions do not play the same role: in the analytic case it is customary to introduce anisotropic norms. However, as Theorem H and its proof show, we can still get good estimates if we keep track separately of the sizes of various terms in the expansion of the Hamiltonian with respect to the actions: this turns out simpler than using anisotropic Gevrey norms. Such a feature is not present in dealing with linearization problem such as in Theorem E; a direct proof of the latter result would have been much simpler.

1.3 Related results

Apart from the work of Popov that we have already mentioned, there have been several works dealing with Gevrey regularity in a related context.

The first setting is the so-called Siegel-Sternberg linearization problem. Under a non-resonance condition, a formal solution to the conjugacy problem always exists and Sternberg proved that the solution is in fact smooth. In the analytic case, under the Bruno-Rüssmann condition the conjugacy is analytic; this arithmetic condition is thus sufficient but also necessary in (complex) dimension one (a result of Yoccoz we already mentioned). In the Gevrey setting, still under the Bruno-Rüssmann condition, Carletti-Marmi [CM00] and Carletti [Car03] have shown that the formal solution still has Gevrey growth (with the same Gevrey exponent); an interesting feature of their result is that allowing a worse Gevrey exponent for the formal solution, one can relax accordingly the arithmetic condition. All these results are actually valid for a class of ultra-differentiable functions that includes analytic and Gevrey functions. It was then proved by Stolovitch [Sto13] that this formal Gevrey solution actually give rise to a Gevrey smooth solution, and recently, Pöschel [Pös17] gave a very general version of the Siegel-Sternberg theorem for ultra-differentiable functions that contains all the previous results (the smooth, analytic, Gevrey and ultra-differentiable cases). Let us mention that all these results do use stability by composition, but a precise composition result is not needed as they do not require to keep track of the width.

In the analytic setting, the Siegel problem and the problem of the linearization of circle diffeomorphisms are solved under the same arithmetic condition [PM97]. But this may well be incidental, and, to our knowledge, it may well not be true in the Gevrey setting. The only result concerning Gevrey circle diffeomorphism we are aware of is due to Gramchev-Yoshino [GY99]: they proved the linearization theorem under a condition which is weaker than the Diophantine condition but stronger than the α-Bruno-Rüssmann condition (they actually introduce a condition equivalent to our α-Bruno-Rüssmann condition and conjecture that the result should hold under this condition). To prove such a result, they use a composition result but in one dimension only; in this special case, as we already pointed out above, good composition estimates are known (see, for instance, [MS02]). As a matter of fact, Theorem G (the discrete version of Theorem E) gives linearization of Gevrey torus diffeomorphism close to a translation under the α-Bruno-Rüssmann condition, extending the result in [GY99] (and giving a positive answer to their conjecture).

1.4 Further results

Let us describe some further results that could be achieved using the techniques of this paper. The literature on KAM theory is enormous and so there are many potential applications; we will only describe here some of those that may have some interest.

First, and more importantly, the technical estimates we derive in Appendix B for Gevrey functions actually hold true for a larger class of ultra-differentiable functions that includes Gevrey (and thus analytic) functions as a particular case. This not only leads to a further extension of the KAM theorems we state and prove here, but also allows us to generalize other perturbative results such as the Nekhoroshev theorem (extending the result of [MS02] in the convex case and [Bou11] in the steep case). To keep this paper to a reasonable length, all these results will be derived in a subsequent article [BF17].

Then, our main result Theorem A deals with the persistence of Lagrangian tori; KAM theory also deals with lower-dimensional tori (see, for instance, [Rüs01] for a comprehensive treatment in the analytic case), and one may expect that our result extend to such a setting.

Finally, one may consider the problem of reducibility of quasi-periodic cocycles close to constant. In the analytic case, the Bruno-Rüssmann condition is sufficient, as was shown in [CM12]; in the α-Gevrey case, the α-Bruno-Rüssmann condition is sufficient. In fact, this setting is simpler from a technical point of view and our Gevrey estimates are not necessary to obtain such a result; one simply needs to go through the proof of [CM12]. A possible explanation for this is that for quasi-periodic cocycles, composition occur in a linear Lie group, thus only estimates for linear composition (product of matrices) are necessary and so everything boils down to good estimates for the product of two functions.

1.5 Plan of the paper

The plan of the paper is as follows.

In Section 2, we describe precisely the setting, namely we properly define the Gevrey norms we will use and the α-Bruno-Rüssmann condition. In Section 3 we state our main results:

- Theorem A about the persistence of a torus in a non-degenerate Hamiltonian system under the α-Bruno-Rüssmann condition;

- Theorem B, the iso-energetic version of Theorem A;

- Theorem C, the non-autonomous time-periodic version;

- Theorem D about the destruction of a torus in the same context not assuming a condition weaker that the α-Bruno-Rüssmann condition;

- Theorem E about linearization of vector fields on the torus close to constant (we will also discuss necessary arithmetic conditions here, albeit in a restricted context);

- Theorem F, the discrete version of Theorem A, about the persistence of a torus in a non-degenerate exact-symplectic map;

- Theorem G, the discrete version of Theorem E, about the linearization of diffeomorphisms of the torus close to a translation.

In Section 4 we state Theorem H, the main technical result of this paper, which is a KAM theorem which do not require non-degeneracy but depends on parameters. In Section 5, we give the proof of Theorems A and E, assuming Theorem H. Section 6 contains the proof of Theorem H. Section 7 contains the proof of Theorem D, a straightforward extension of the work of Bessi [Bes00].

Finally, two appendices contain technical results. Appendix A provides various characterizations of the α-Bruno-Rüssmann condition. Appendix B, which is absolutely crucial in this work, provides estimates on Gevrey functions (and in particular our composition result Proposition 20) which are use throughout the paper.

2 Setting

2.1 Gevrey Hamiltonians

Recall that $n \geq 1$ is an integer, $\mathbb{T}^n = \mathbb{R}^n / \mathbb{Z}^n$ and let $B \subseteq \mathbb{R}^n$ be a bounded open domain containing the origin. For a small parameter $\epsilon \geq 0$, we consider a Hamiltonian function $H : \mathbb{T}^n \times B \to \mathbb{R}$ of the form

$$\begin{cases} H(q,p) = h(p) + \epsilon f(q,p), \\ \nabla h(0) := \omega_0 \in \mathbb{R}^n. \end{cases} \quad (*)$$

The Hamiltonian h is *non-degenerate* at the origin if the matrix $\nabla^2 h(0)$ itself is non-degenerate. We shall assume that the Hamiltonian H is α-*Gevrey* on $\mathbb{T}^n \times \bar{B}$, with $\alpha \geq 1$ and where $\bar{B}$ denotes the closure of B in $\mathbb{R}^n$: H is smooth on a open neighborhood of $\mathbb{T}^n \times \bar{B}$ in $\mathbb{T}^n \times \mathbb{R}^n$ and there exists $s_0 > 0$ such that, using multi-indices notation (see Appendix B) and $|k| = \sum_{i=1}^{2n} k_i$,

$$|H|_{\alpha,s_0} := c \sup_{(\theta,I) \in \mathbb{T}^n \times \bar{B}} \left(\sup_{k \in \mathbb{N}^{2n}} \frac{(|k|+1)^2 s_0^{\alpha|k|} |\partial^k H(\theta,I)|}{|k|!^\alpha} \right) < \infty, \quad c := 4\pi^2/3. \quad (1)$$

This definition can be extended to vector-valued function $X : \mathbb{T}^n \times \bar{B} \to \mathbb{R}^p$ by setting

$$|X|_{\alpha,s_0} := c \sup_{(\theta,I) \in \mathbb{T}^n \times \bar{B}} \left(\sup_{k \in \mathbb{N}^{2n}} \frac{(|k|+1)^2 s_0^{\alpha|k|} |\partial^k X(\theta,I)|_1}{|k|!^\alpha} \right) < \infty \quad (2)$$

where $|\,.\,|_1$ is the l_1-norm of vectors in $\mathbb{R}^p$, or the sum of the absolute values of the components. As a rule, we will use the l_1-norm for vectors, so for simplicity we shall write $|\,.\,|_1 = |\,.\,|$. To emphasize the role of the "Gevrey width" s_0, we shall also say that H is (α, s_0)-Gevrey if (1) holds. Observe that a function is 1-Gevrey if and only it is real-analytic, in which case the parameter $s_0 > 0$ is the width of analyticity.

Properties of these Gevrey norms are described in Appendix B; in particular we explain there the (inessential) role of the factor $(|k|+1)^2$ and the normalizing constant $c > 0$ in (1).

2.2 The α-Bruno-Rüssmann condition

Given $\omega_0 \in \mathbb{R}^n$, define the function

$$\Psi_{\omega_0} : [1, +\infty) \to [\Psi_{\omega_0}(1), +\infty], \quad Q \mapsto \max\{|k \cdot \omega_0|^{-1} \mid k \in \mathbb{Z}^n,\ 0 < |k| \leq Q\}. \quad (3)$$

This function Ψ_{ω_0} measures the size of the so-called small denominators which will come into play in our computations. Call BR the set of vectors ω_0 satisfying the so-called *Bruno-Rüssmann condition*,

$$\int_1^{+\infty} \frac{\ln(\Psi_{\omega_0}(Q))}{Q^2} dQ < \infty \quad \text{(BR)}$$

and, given $\alpha \geq 1$, call BR_α the set of vectors ω_0 satisfying the α-*Bruno-Rüssmann condition*, which we define as

$$\int_1^{+\infty} \frac{\ln(\Psi_{\omega_0}(Q))}{Q^{1+\frac{1}{\alpha}}} dQ < \infty. \quad \text{(BR}_\alpha\text{)}$$

These conditions prevent Ψ_{ω_0} from growing too fast at infinity. If $\omega_0 \in \text{BR} = \text{BR}_1$, in particular $\Psi_{\omega_0}(Q)$ is finite for all Q, i.e. ω_0 is non-resonant. Besides, the set BR_α decreases with respect to α. For example, if $\Psi_{\omega_0}(Q) = \exp(Q^\beta)$ then $\omega_0 \in \text{BR}_\alpha$ if and only if $\beta < 1/\alpha$ (we let the reader check, using continued fractions if $n = 2$, that the set of vectors ω_0 having such function Ψ_{ω_0} is not empty).

Let D_τ be the set of τ-Diophantine vectors ($\tau \geq n-1$), i.e. for which there exists $\gamma > 0$ such that $\Psi(Q) \leq Q^\tau/\gamma$ for all $Q \geq 1$. D_τ is non-empty and has full measure if $\tau > n-1$ [Rüs75]. As definitions show, for all $\alpha \geq 1$, we have $D_\tau \subset \text{BR}_\alpha$. Thus, as Example 10 shows,

$$\cap_{\alpha \geq 1}\text{BR}_\alpha \setminus \cup_{\tau \geq n-1}D_\tau$$

has zero-measure but is non-empty.

Now assume that ω_0 is non-resonant. The function Ψ_{ω_0} is non-decreasing, piecewise constant, and has a countable number of discontinuities. In the sequel, it will be more convenient to work with a continuous version of Ψ_{ω_0}: it is not hard to prove (see, for instance, Appendix A of [BF13]) that one can find a continuous non-decreasing function $\Psi : [1, \infty) \to [\Psi(1), +\infty)$ such that $\Psi(1) = \Psi_{\omega_0}(1)$ and

$$\Psi_{\omega_0}(Q) \leq \Psi(Q) \leq \Psi_{\omega_0}(Q+1), \quad Q \geq 1. \quad (4)$$

For all $k \in \mathbb{Z}^n \setminus \{0\}$, we still have

$$|k \cdot \omega_0| \geq 1/\Psi(|k|)$$

and in the condition (BR_α) (which defines $\omega_0 \in \text{BR}_\alpha$), one may use Ψ instead of Ψ_{ω_0}.

Let us now define the function

$$\Delta : [1, +\infty) \to [\Psi(1), +\infty), \quad Q \mapsto Q\Psi(Q).$$

It is continuous and increasing, and thus is a homeomorphism, whose functional inverse is

$$\Delta^{-1} : [\Psi(1), +\infty) \to [1, +\infty), \quad \Delta^{-1} \circ \Delta = \Delta \circ \Delta^{-1} = \mathrm{Id}.$$

In Appendix A we show that the set BR_α agrees with the set A_α defined by the condition

$$\int_{\Delta(1)}^{+\infty} \frac{dx}{x(\Delta^{-1}(x))^{\frac{1}{\alpha}}} < \infty. \tag{A_α}$$

3 Main results

3.1 KAM theorem for non-degenerate integrable Hamiltonians

The image of the map $\Theta_0 : \mathbb{T}^n \to \mathbb{T}^n \times B$, $q \mapsto (q, 0)$, is an embedded torus invariant by the flow of h carrying a quasi-periodic flow with frequency ω_0. We shall prove that this quasi-periodic invariant Gevrey-smooth embedded torus is preserved by an arbitrary small perturbation, provided h is non-degenerate, H is α-Gevrey and ω_0 satisfies the α-Bruno-Rüssmann condition.

Theorem A. Let H be as in $(*)$, where H is (α, s_0)-Gevrey, $\omega_0 \in \mathrm{BR}_\alpha$ and h is non-degenerate. Then there exists $0 < s_0' < s_0$ such that for ϵ small enough, there exists an (α, s_0')-Gevrey torus embedding $\Theta_{\omega_0} : \mathbb{T}^n \to \mathbb{T}^n \times B$ such that $\Theta_{\omega_0}(\mathbb{T}^n)$ is invariant by the Hamiltonian flow of H and quasi-periodic with frequency ω_0. Moreover, Θ_{ω_0} is close to Θ_0 in the sense that

$$|\Theta_{\omega_0} - \Theta_0|_{\alpha, s_0'} \leq c\sqrt{\epsilon}$$

for some constant $c > 0$ independent of ϵ.

Theorem A will be deduced from a KAM theorem for a Hamiltonian with parameters, for which a quantitative statement is given in §4. Let us also state the corresponding iso-energetic and non-autonomous time-periodic versions.

We say that the integrable Hamiltonian h is *iso-energetically non-degenerate* at 0 if the so-called bordered Hessian of h,

$$\begin{pmatrix} \nabla^2 h(0) & {}^t\nabla h(0) \\ \nabla h(0) & 0 \end{pmatrix},$$

has a non-zero determinant. Under this assumption, the unperturbed torus $p = 0$, with energy $h(0)$, can be continued to a torus with the same energy but with a frequency of the form $\lambda\omega_0$ for λ close to one.

Theorem B. Let H be as in $(*)$, where H is (α, s_0)-Gevrey, $\omega_0 \in \mathrm{BR}_\alpha$ and h is iso-energetically non-degenerate. Then there exists $0 < s_0' < s_0$ such that for ϵ small enough, there exist $\lambda \in \mathbb{R}^*$ and an (α, s_0')-Gevrey torus embedding $\Theta_{\omega_0} : \mathbb{T}^n \to \mathbb{T}^n \times B$ such that

$\Theta_{\omega_0}(\mathbb{T}^n)$ is invariant by the Hamiltonian flow of H, contained in $H^{-1}(h(0))$ and quasi-periodic with frequency $\lambda\omega_0$. Moreover, λ is close to one and Θ_{ω_0} is close to Θ_0 in the sense that

$$|\lambda - 1| \leq c\sqrt{\epsilon}, \quad |\Theta_{\omega_0} - \Theta_0|_{\alpha,s_0'} \leq c\sqrt{\epsilon}$$

for some constant $c > 0$ independent of ϵ.

We can also look at the non-autonomous time-periodic version; we consider a slightly different setting by looking at a Hamiltonian function $\tilde{H} : \mathbb{T}^n \times B \times \mathbb{T} \to \mathbb{R}$ of the form

$$\begin{cases} \tilde{H}(q,p) = h(p) + \epsilon f(q,p,t), \\ \nabla h(0) := \omega_0 \in \mathbb{R}^n. \end{cases} \tag{$\tilde{*}$}$$

It is better to consider the unperturbed torus $p = 0$ as an invariant torus for the integrable Hamiltonian $\tilde{h} : B \times \mathbb{R}$ defined by $\tilde{h}(p,e) := h(p) + e$: it is then quasi-periodic with frequency $\tilde{\omega}_0 := (\omega_0, 1)$, has dimension $n + 1$ and is the image of the trivial embedding $\tilde{\Theta}_0 : \mathbb{T}^n \times \mathbb{T} \to \mathbb{T}^n \times B \times \mathbb{T}$.

Theorem C. Let $\tilde{H}$ be as in $(\tilde{*})$, where $\tilde{H}$ is (α, s_0)-Gevrey, $\omega_0 \in \mathrm{BR}_\alpha$ and h is non-degenerate. Then there exists $0 < s_0' < s_0$ such that for ϵ small enough, there exists an (α, s_0')-Gevrey torus embedding $\tilde{\Theta}_{\omega_0} : \mathbb{T}^n \times \mathbb{T} \to \mathbb{T}^n \times B \times \mathbb{T}$ such that $\tilde{\Theta}_{\omega_0}(\mathbb{T}^n \times \mathbb{T})$ is invariant by the Hamiltonian flow of $\tilde{H}$ and quasi-periodic with frequency $\tilde{\omega}_0$. Moreover, $\tilde{\Theta}_{\omega_0}$ is close to $\tilde{\Theta}_0$ in the sense that

$$|\tilde{\Theta}_{\omega_0} - \tilde{\Theta}_0|_{\alpha,s_0'} \leq c\sqrt{\epsilon}$$

for some constant $c > 0$ independent of ϵ.

Theorem B and Theorem C are essentially equivalent statements and can be easily deduced from Theorem A; in the analytic case details are given in [TZ10], Chapter 2, but it is plain to observe that the arguments still work in the Gevrey case.

3.2   Destruction of invariant tori

According to Theorem A, the α-Bruno-Rüssmann condition is sufficient for the preservation of an invariant torus under an α-Gevrey perturbation. A natural question is: is it necessary? To this question, here we only bring a partial answer, which circumscribes the optimal arithmetic condition, if any. Following Bessi [Bes00], one can show that if $\omega = \omega_0$ satisfies a condition (the condition (B_α) defined below), the torus can be destroyed. In particular, this shows that the exponent $1 + 1/\alpha$ in (BR_α) cannot be replaced by a strictly larger exponent. As a matter of fact, the example of Bessi already shows this in the analytic case $\alpha = 1$; our observation here is that Bessi's example gives a similar result for any $\alpha \geq 1$.

Theorem D. Given $\alpha \geq 1$, assume that the vector $\omega \in \mathbb{R}^n$ satisfies the following condition:

$$\limsup_{Q \to +\infty} \frac{\ln(\Psi_\omega(Q))}{Q^{1/\alpha}} > 0. \tag{B_α}$$

Then an invariant torus with frequency ω can be destroyed by an arbitrarily small α-Gevrey perturbation.

Thus the condition that ω_0 does not satisfy (B_α), namely

$$\lim_{Q \to +\infty} \frac{\ln(\Psi_\omega(Q))}{Q^{1/\alpha}} = 0, \tag{R_α}$$

is a necessary condition for the conclusion of Theorem A to hold true. For $\alpha = 1$, this condition (R_α) is actually a sufficient (and most probably necessary) condition to solve the cohomological equation associated to ω (see [Rüs75]); in the general case $\alpha \geq 1$ this should also be true but we couldn't find a reference. Let us also note that (R_α) is implied by (but clearly not equivalent to) the condition that $\omega \in \mathrm{BR}_\alpha$, see Remark 1 in Appendix A.

For a more precise statement and how this follows from [Bes00], we refer to Theorem 7 in Section 7. It is likely that one could improve this result for $\alpha > 1$ by using perturbations with compact support as in [CW13].

Observe that for any $\alpha \geq 1$ and any $0 < \beta < \alpha$, vectors $\omega \in \mathbb{R}^n$ for which

$$\Psi_\omega(Q) \sim e^{Q^{1/\alpha}}$$

satisfies (B_α) but also the β-Bruno-Rüssmann condition. (That such vectors do exist is a classical matter in number theory.) The following corollary is then obvious.

Corollary 1. *For any $\alpha \geq 1$ and any $0 < \beta < \alpha$, there exist invariant tori with frequency vectors $\omega \in \mathrm{BR}_\beta$ which can be destroyed by an arbitrary small α-Gevrey perturbation. In particular, there exist invariant tori with frequency vectors $\omega \in \mathrm{BR}$ which can be destroyed by an arbitrary small Gevrey non-analytic perturbation.*

3.3   KAM theorem for constant vector fields on the torus

Now we state a Gevrey version of Arnold's normal form theorem for vector fields on the torus.

Theorem E. Let $\omega_0 \in \mathrm{BR}_\alpha$ and $X \in G_{\alpha,s}(\mathbb{T}^n, \mathbb{R}^n)$ a vector field on $\mathbb{T}^n$ of the form

$$X = \omega_0 + B, \quad |B|_{\alpha,s} \leq \mu.$$

Then, for μ sufficiently small, there exist a vector $\omega_0^* \in \mathbb{R}^n$ and an $(\alpha, s/2)$-Gevrey diffeomorphism $\Xi : \mathbb{T}^n \to \mathbb{T}^n$ such that $X + \omega_0^* - \omega_0$ is conjugate to ω_0 via Ξ:

$$\Xi^*(X + \omega_0^* - \omega_0) = \omega_0.$$

Moreover, we have the estimate

$$|\omega_0^* - \omega_0| \le c\mu, \quad |\Xi - \mathrm{Id}|_{\alpha,s/2} \le c\mu$$

for some constant $c \ge 1$ independent of μ.

Observe that because of the shift of frequency $\omega_0^* - \omega_0$, in general this result does not give any information on the vector field X. Under some further assumption (for instance, if ω_0 belongs to the rotation set of X, see [Kar16]), then this shift vanishes and Theorem E implies that X is conjugated to ω_0.

An even more restricted setting is when X is proportional to ω_0 (so that the flow of X is a re-parametrization of the linear flow of frequency ω_0 and thus ω_0 is the unique rotation vector of X); Theorem E applies in this case to give a conjugacy to ω_0, assuming that $\omega_0 \in \mathrm{BR}_\alpha$, but the proof is actually much simpler in this case (it boils down to solve only once the cohomological equation) and should require the weaker condition that ω_0 satisfies (R_α), as it is stated in the case $\alpha = 1$ in [Fay02]. Still in [Fay02], it is proved that for $\alpha = 1$ (there are also versions in the C^r case), if ω_0 satisfies (B_α), then there is a dense set of reparametrized linear flow which are weak-mixing (and so cannot be conjugated to the linear flow); thus a necessary condition for Theorem E to hold true is that ω_0 satisfies (R_α) (and this is also a sufficient condition if we impose that X is proportional to ω_0)[1]. Clearly, this should extend to the general case $\alpha \ge 1$ and thus the condition that ω_0 does not satisfy (B_α) is a necessary condition for Theorem E to hold true, as in Theorem A.

3.4   KAM theorem for maps

In this section, we give the statement of discrete versions of Theorem A and Theorem E.

Let us start with the discrete analogue of Theorem A. Given a function $h : \bar{B} \to \mathbb{R}$, we define the exact-symplectic map

$$F_h : \mathbb{T}^n \times \bar{B} \to \mathbb{T}^n \times \bar{B}, \quad (q,p) \mapsto (q + \nabla h(p), p).$$

As before, let us fix $\alpha \ge 1$ and $s_0 > 0$.

Theorem F. Let $F : \mathbb{T}^n \times \bar{B} \to \mathbb{T}^n \times \bar{B}$ be an (α, s_0)-Gevrey exact symplectic map with

$$|F - F_h|_{\alpha,s_0} \le \epsilon.$$

Assume that $\omega_0 = \nabla h(0) \in \mathrm{BR}_\alpha$ and that h is non-degenerate. Then there exists $0 < s_0' < s_0$ such that for ϵ small enough, there exists an (α, s_0')-Gevrey torus embedding $\Theta_{\omega_0} : \mathbb{T}^n \to \mathbb{T}^n \times B$ such that $\Theta_{\omega_0}(\mathbb{T}^n)$ is invariant by F and Θ_{ω_0} gives a conjugacy between the translation of vector ω_0 on $\mathbb{T}^n$ and the restriction of F to $\Theta_{\omega_0}(\mathbb{T}^n)$. Moreover, Θ_{ω_0} is close to Θ_0 in the sense that

$$|\Theta_{\omega_0} - \Theta_0|_{\alpha,s_0'} \le c\sqrt{\epsilon}$$

for some constant $c > 0$ independent of ϵ.

---

[1]We would like to thank B. Fayad for a discussion on this topic.

Theorem F follows at once from Theorem B (or Theorem C) provided one has a suitable quantitative "suspension" result; in the analytic case $\alpha = 1$ this was proved in [KP94] and in the Gevrey non-analytic case $\alpha > 1$ this is contained in [LMS16].

In the same way, we have the following discrete analogue of Theorem E. Given $\omega_0 \in \mathbb{R}^n$, let T_{ω_0} be the translation of $\mathbb{T}^n$ of vector ω_0:

$$T_{\omega_0} : \mathbb{T}^n \to \mathbb{T}^n, \quad \theta \mapsto \theta + \omega_0.$$

Let $\alpha \geq 1$ and $s > 0$.

Theorem G. Let $\omega_0 \in \mathrm{BR}_\alpha$ and $T \in G_{\alpha,s}(\mathbb{T}^n, \mathbb{T}^n)$ a diffeomorphism of $\mathbb{T}^n$ of the form

$$T = T_{\omega_0} + B, \quad |B|_{\alpha,s} \leq \mu.$$

Then, for μ sufficiently small, there exist a vector $\omega_0^* \in \mathbb{R}^n$ and an $(\alpha, s/2)$-Gevrey diffeomorphism $\Xi : \mathbb{T}^n \to \mathbb{T}^n$ such that $T + \omega_0^* - \omega_0$ is conjugate to T_{ω_0} via Ξ:

$$\Xi^{-1} \circ (T + \omega_0^* - \omega_0) \circ \Xi = T_{\omega_0}.$$

Moreover, we have the estimate

$$|\omega_0^* - \omega_0| \leq c\mu, \quad |\Xi - \mathrm{Id}|_{\alpha,s/2} \leq c\mu$$

for some constant $c \geq 1$ independent of μ.

4   Statement of the KAM theorem with parameters

Let us now consider the following setting. Fix $\omega_0 \in \mathbb{R}^n \setminus \{0\}$. Re-ordering the components of ω_0 and re-scaling the Hamiltonian allow us to assume without loss of generality that

$$\omega_0 = (1, \bar{\omega}_0) \in \mathbb{R}^n, \quad \bar{\omega}_0 \in [-1, 1]^{n-1}.$$

Given real numbers $r > 0$ and $h > 0$, we let

$$D_r := \{I \in \mathbb{R}^n \mid |I| \leq r\}, \quad D_h^{\omega_0} := \{\omega \in \mathbb{R}^n \mid |\omega - \omega_0| \leq h\}, \quad D_{r,h} := D_r \times D_h^{\omega_0}.$$

Our Hamiltonians will be defined on $\mathbb{T}^n \times D_{r,h}$, a neighborhood of $\mathbb{T}^n \times \{0\} \times \{\omega_0\}$ in $\mathbb{T}^n \times \mathbb{R}^n \times \mathbb{R}^n$.

Let $\alpha \geq 1$, $s > 0$, $\eta \geq 0$ a fixed parameter, $\varepsilon \geq 0$ and $\mu \geq 0$ two small parameters. We consider a function $H \in G_{\alpha,s}(\mathbb{T}^n \times D_{r,h})$ of the form

$$\begin{cases} H(\theta, I, \omega) = \underbrace{e(\omega) + \omega \cdot I}_{N(I,\omega)} + \underbrace{A(\theta, \omega) + B(\theta, \omega) \cdot I}_{P(\theta,I,\omega)} + \underbrace{M(\theta, I, \omega) \cdot I^2}_{R(\theta,I,\omega)} \\ |A|_{\alpha,s} \leq \varepsilon, \quad |B|_{\alpha,s} \leq \mu, \quad |\nabla_I^2 R|_{\alpha,s} \leq \eta \end{cases} \qquad (**)$$

where the notation $M(\theta, I, \omega) \cdot I^2$ stands for the vector I given twice as an argument to the symmetric bilinear form $M(\theta, I, \omega)$. Observe that $A : \mathbb{T}^n \times D_h^{\omega_0} \to \mathbb{R}$, $B : \mathbb{T}^n \times D_h^{\omega_0} \to \mathbb{R}^n$ whereas $M : \mathbb{T}^n \times D_{r,h} \to M_n(\mathbb{R})$ with $M_n(\mathbb{R})$ the ring of real square matrices of size n. Observe that we do not assume $\varepsilon = \mu$ because these two small parameters play different roles in applications (in Theorem A we will have $\mu = \sqrt{\varepsilon}$ while in Theorem E, $\varepsilon = 0$ and μ will be the only small parameter).

The function H in $(**)$ should be considered as a Gevrey Hamiltonian on $\mathbb{T}^n \times D_r$, depending on a parameter $\omega \in D_h^{\omega_0}$; for a fixed parameter $\omega \in D_h^{\omega_0}$, when convenient, we will write

$$H_\omega(I, \theta) = H(I, \theta, \omega), \quad N_\omega(I) = N(I, \omega), \quad P_\omega(I, \theta) = P(I, \theta, \omega), \quad R_\omega(I, \theta) = R(I, \theta, \omega).$$

The image of the map $\Phi_0 : \mathbb{T}^n \to \mathbb{T}^n \times D_r$, $\theta \mapsto (\theta, 0)$ is a smooth embedded torus in $\mathbb{T}^n \times D_r$, invariant by the Hamiltonian flow of $N_{\omega_0} + R_{\omega_0}$ and quasi-periodic with frequency ω_0. The next theorem asserts that this quasi-periodic torus will persist, being only slightly deformed, as an invariant torus not for the Hamiltonian flow of H_{ω_0} but for the Hamiltonian flow of $H_{\omega_0^*}$, where ω_0^* is a parameter close to ω_0, provided ε and μ are sufficiently small and ω_0 satisfies the α-Bruno-Rüssmann condition. Here is the precise statement.

Theorem H. Let H be as in $(**)$, with $\omega_0 \in \mathrm{BR}_\alpha$. Then there exist positive constants $c_1 \leq 1$, $c_2 \leq 1$ and $c_3 \geq 1$ depending only on n and α such that if

$$\sqrt{\varepsilon} \leq \mu \leq h/2, \quad \sqrt{\varepsilon} \leq r, \quad h \leq c_1(Q_0 \Psi(Q_0))^{-1} \tag{5}$$

where $Q_0 \geq n + 2$ is sufficiently large so that

$$Q_0^{-\frac{1}{\alpha}} + (\ln 2)^{-1} \int_{\Delta(Q_0)}^{+\infty} \frac{dx}{x(\Delta^{-1}(x))^{\frac{1}{\alpha}}} \leq c_2(1 + \eta)^{-1/\alpha} s, \tag{6}$$

there exist a vector $\omega_0^* \in \mathbb{R}^n$ and an $(\alpha, s/2)$-Gevrey embedding

$$\Phi_{\omega_0}^* : \mathbb{T}^n \times D_{r/2} \to \mathbb{T}^n \times D_r$$

of the form

$$\Phi_{\omega_0}^*(\theta, I) = (\theta + E^*(\theta), I + F^*(\theta) \cdot I + G^*(\theta))$$

with the estimates

$$|\omega_0^* - \omega_0| \leq c_3 \mu, \quad |E^*|_{\alpha, s/2} \leq c_3 \Psi(Q_0)\mu, \quad |F^*|_{\alpha, s/2} \leq c_3 \Delta(Q_0)\mu, \quad |G^*|_{\alpha, s/2} \leq c_3 \Delta(Q_0)\varepsilon \tag{7}$$

and such that

$$H_{\omega_0^*} \circ \Phi_{\omega_0}^*(\theta, I) = e_0^* + \omega_0 \cdot I + R^*(\theta, I), \quad R^*(\theta, I) = M^*(\theta, I) \cdot I^2,$$

with the estimates

$$|e_0^* - e_{\omega_0^*}| \leq c_3 \varepsilon, \quad |\nabla_I^2 R^* - \nabla_I^2 R_{\omega_0^*}|_{\alpha, s/2} \leq c_3 \eta \Delta(Q_0)\mu.$$

Theorem A follows quite directly from Theorem H, introducing the frequencies $\omega = \nabla h(p)$ as independent parameters, taking $\mu = \sqrt{\varepsilon}$, and tuning the shift of frequency $\omega_0^* - \omega_0$ using the non-degeneracy assumption on the unperturbed Hamiltonian. Theorems E follows also from Theorem H by realizing X as the restriction of a Hamiltonian vector field on an invariant torus, setting $\varepsilon = \eta = 0$ and letting μ be the only small parameter. These arguments are made precise in Section 5.

5 Proofs of Theorems A and E, assuming Theorem H

5.1 Proof of Theorem A

In this section, we assume Theorem H and we show how it implies Theorem A, following [Pös01] (in the analytic case) and [Pop04] (in the Gevrey case).

Proof of Theorem A. As noticed at the beginning of Section 4, we may assume that ω_0 is of the form
$$\omega_0 = (1, \bar{\omega}_0) \in \mathbb{R}^n, \quad \bar{\omega}_0 \in [-1, 1]^{n-1}.$$

For $p_0 \in B$, we expand h in a small neighborhood of p_0: writing $p = p_0 + I$ for I close to zero, we get

$$h(p) = h(p_0) + \nabla_p h(p_0) \cdot I + \int_0^1 (1 - t) \nabla_p^2 h(p_0 + tI) \cdot I^2 \, dt.$$

Similarly, we expand ϵf with respect to p, in a small neighborhood of p_0:

$$\epsilon f(q, p) = \epsilon f(q, p_0) + \epsilon \nabla_p f(q, p_0) \cdot I + \epsilon \int_0^1 (1 - t) \nabla_p^2 f(q, p_0 + tI) \cdot I^2 \, dt.$$

Since $\nabla_p h : B \to \Omega$ is a diffeomorphism, instead of p_0 we can use $\omega = \nabla_p h(p_0)$ as a new variable, and letting $\nabla_\omega g := (\nabla h)^{-1}$, we write

$$h(p) = e(\omega) + \omega \cdot I + R_h(I, \omega)$$

with
$$e(\omega) := h(\nabla_\omega g(\omega)), \quad R_h(I, \omega) := \int_0^1 (1 - t) \nabla_p^2 h(\nabla_\omega g(\omega) + tI) \cdot I^2 \, dt$$

and also, letting $\theta = q$,

$$\epsilon f(q, p) = \epsilon \tilde{A}(\theta, \omega) + \epsilon \tilde{B}(\theta, \omega) \cdot I + \epsilon R_f(\theta, I, \omega)$$

with
$$\tilde{A}(\theta, \omega) := f(\theta, \nabla_\omega g(\omega)), \quad \tilde{B}(\theta, \omega) := \nabla_p f(\theta, \nabla_\omega g(\omega))$$

and
$$R_f(\theta, I, \omega) := \epsilon \int_0^1 (1 - t) \nabla_p^2 f(\theta, \nabla_\omega g(\omega) + tI) \cdot I^2 \, dt.$$

Finally, we can set

$$A := \epsilon \tilde{A}, \quad B := \epsilon \tilde{B}, \quad R := R_h + \epsilon R_f = M(\theta, I, \omega) \cdot I^2,$$

so that $h + \epsilon f$ can be written as

$$H(\theta, I, \omega) = e(\omega) + \omega \cdot I + A(\theta, \omega) + B(\theta, \omega) \cdot I + R(\theta, I, \omega),$$

and we have

$$\nabla_I^2 R(\theta, I, \omega) = \nabla_I^2 h(\nabla_\omega g(\omega) + I) + \epsilon \nabla_I^2 f(\theta, \nabla_\omega g(\omega) + I).$$

By assumption, h and f are (α, s_0)-Gevrey on $\mathbb{T}^n \times \bar{B}$, and since the space of Gevrey functions is closed under taking derivatives, products, composition and inversion (up to restricting the parameter s_0, see Appendix B for the relevant estimates), we claim that we can find[2] $s > 0$, $r > 0$, $h > 0$ and $\tilde{c} > 0$ which are independent of ϵ such that H is (α, s)-Gevrey on the domain $\mathbb{T}^n \times D_{r,h}$ with the estimates

$$|A|_{\alpha,s} \le \tilde{c}\epsilon, \quad |B|_{\alpha,s} \le \tilde{c}\epsilon, \quad |\nabla_I^2 R|_{\alpha,s} \le \tilde{c}.$$

We may set

$$\varepsilon := \tilde{c}\epsilon, \quad \mu := \sqrt{\varepsilon}, \quad \eta := \tilde{c}$$

and assuming ϵ small enough, we have $\tilde{c}\epsilon \le \mu = \sqrt{\varepsilon}$. Thus we have

$$|A|_{\alpha,s} \le \varepsilon, \quad |B|_{\alpha,s} \le \mu, \quad |\nabla_I^2 R|_{\alpha,s} \le \eta.$$

Having fixed $s > 0$ and $r > 0$, we may choose Q_0 sufficiently large so that (6) holds true, and then by further restricting first h, and then ϵ, we may assume that the condition (5) is satisfied. Theorem H applies: there exist an $(\alpha, s/2)$-Gevrey embedding $\Upsilon_{\omega_0} : \mathbb{T}^n \to \mathbb{T}^n \times D_r$, defined by

$$\Upsilon_{\omega_0}(\theta) := \Phi_{\omega_0}(\theta, 0) = (\theta + E^*(\theta), G^*(\theta))$$

where Φ_{ω_0} is given by Theorem H, and a vector $\omega_0^* \in \mathbb{R}^n$ such that $\Upsilon_{\omega_0}(\mathbb{T}^n)$ is invariant by the Hamiltonian flow of $H_{\omega_0^*}$ and quasi-periodic with frequency ω_0. Moreover, ω_0^* and Υ_{ω_0} satisfy the estimates

$$|\omega_0^* - \omega_0| \le c\mu, \quad |E^*|_{\alpha,s/2} \le c\Psi(Q_0)\mu, \quad |G^*|_{\alpha,s/2} \le cQ_0\Psi(Q_0)\varepsilon$$

for some large constant $c > 1$. Since h is non-degenerate, there exists p_0^* such that $\nabla h(p_0^*) = \omega_0^*$ and, up to taking $c > 1$ larger and recalling that $\mu = \sqrt{\varepsilon}$, the above estimates imply

$$|p_0^*| \le c\sqrt{\varepsilon}, \quad |E^*|_{\alpha,s/2} \le c\Psi(Q_0)\sqrt{\varepsilon}, \quad |G^*|_{\alpha,s/2} \le cQ_0\Psi(Q_0)\varepsilon. \tag{8}$$

---

[2]There is a local clash of notations here, since one should not mix up the integrable Hamiltonian h of theorem A and the width parameter h of theorem H.

Now observe that an orbit $(\theta(t), I(t))$ for the Hamiltonian $H_{\omega_0^*}$ corresponds to an orbit $(q(t), p(t)) = (\theta(t), I(t) + p_0^*)$ for our original Hamiltonian. Hence, if we define $T : \mathbb{T}^n \times \mathbb{R}^n \to \mathbb{T}^n \times \mathbb{R}^n$ by $T(\theta, I) = (\theta, I + p_0^*)$ and

$$\Theta_{\omega_0} = T \circ \Upsilon_{\omega_0} : \mathbb{T}^n \to \mathbb{T}^n \times \mathbb{R}^n, \quad \Theta_{\omega_0}(\theta) = (\theta + E^*(\theta), G^*(\theta) + p_0^*)$$

then Θ_{ω_0} is an $(\alpha, s/2)$-Gevrey torus embedding such that $\Theta_{\omega_0}(\mathbb{T}^n)$ is invariant by the Hamiltonian flow of H and quasi-periodic with frequency ω_0. The estimates on the distance between Θ_{ω_0} and the trivial embedding Θ_0 follows directly from (8), which finishes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

5.2 Proof of Theorem E

Now we show how Theorem E follows from Theorem H.

Proof of Theorem E. Consider the vector field $X = \omega_0 + B \in G_{\alpha,s}(\mathbb{T}^n, \mathbb{R}^n)$ as in the statement. It can be trivially included into a parameter-depending vector field: given $h > 0$, let $\hat{X} \in G_{\alpha,s}(\mathbb{T}^n \times D_h^{\omega_0}, \mathbb{R}^n)$ be such that

$$\hat{X}(\theta, \omega) = \hat{X}_\omega(\theta) = \omega + B(\theta), \quad \omega \in D_h^{\omega_0}, \quad \hat{X}_{\omega_0} = X.$$

Now given any $r > 0$, consider the Hamiltonian H defined on $\mathbb{T}^n \times D_{r,h}$ by

$$H(\theta, I, \omega) = H_\omega(\theta, I) := \omega \cdot I + B(\theta) \cdot I. \tag{9}$$

Clearly, for any parameter ω, the torus $\mathbb{T}^n \times \{0\}$ is invariant by the Hamiltonian vector field X_{H_ω}, and, upon identifying $\mathbb{T}^n \times \{0\}$ with $\mathbb{T}^n$, the restriction of X_{H_ω} to this torus coincides with $\hat{X}_\omega$.

Now the Hamiltonian H defined in (9) is of the form $(**)$ with $\varepsilon = \eta = 0$ (and $e = 0$) and therefore for μ sufficiently small, Theorem H applies: there exist a vector $\omega_0^* \in \mathbb{R}^n$ and an $(\alpha, s/2)$-Gevrey embedding

$$\Phi_{\omega_0}^* : \mathbb{T}^n \times D_{r/2} \to \mathbb{T}^n \times D_r$$

here of the form

$$\Phi_{\omega_0}^*(\theta, I) = (\theta + E^*(\theta), I + F^*(\theta) \cdot I)$$

with the estimates

$$|\omega_0^* - \omega_0| \le c_3 \mu, \quad |E^*|_{\alpha,s/2} \le c_3 \Psi(Q_0)\mu, \quad |F^*|_{\alpha,s/2} \le c_3 \Delta(Q_0)\mu$$

and such that

$$H_{\omega_0^*} \circ \Phi_{\omega_0}^*(\theta, I) = \omega_0 \cdot I. \tag{10}$$

The embedding $\Phi_{\omega_0}^*$ clearly leaves invariant the torus $\mathbb{T}^n \times \{0\}$ and induces a diffeomorphism of this torus that can be identified to $\Xi := \text{Id} + E^*$. Writing the equality (10) in terms

of Hamiltonian vector fields, we have, upon restriction to the invariant torus and recalling that the restriction of X_{H_ω} coincides with $\hat{X}_\omega$,

$$\Xi^*(\hat{X}_{\omega_0^*}) = \omega_0.$$

But $\hat{X}_{\omega_0^*} = \hat{X}_{\omega_0} + \omega_0^* - \omega_0 = X + \omega_0^* - \omega_0$ and therefore

$$\Xi^*(X + \omega_0^* - \omega_0) = \omega_0$$

which, together with the estimates on ω_0^* and $\Xi - \mathrm{Id} = E^*$, was the statement we wanted to prove. $\qquad\square$

6 Proof of Theorem H

This section is devoted to the proof of Theorem H, in which we will construct, by an iterative procedure, a vector ω_0^* close to ω_0 and a Gevrey-smooth torus embedding $\Phi_{\omega_0}^*$ whose image is invariant by the Hamiltonian flow of $H_{\omega_0^*}$. We start, in §6.1, by recalling the Diophantine result of [BF13] which will be used in our approach. Then, in §6.2, we describe an elementary step of our iterative procedure, and finally, in §6.3, we will show that infinitely many steps may be carried out, to converge towards a solution.

In this paper, we do not pay attention to how constants depend on the dimension n or the Gevrey-exponent α, both being fixed. Hence in this section, we shall write

$$u \lessdot v \quad (\text{respectively } u \cdot\!\!\lessdot v)$$

if, for some constant $C \geq 1$ depending only on n and α, we have $u \leq Cv$ (respectively $Cu \leq v$). In particular, $u \cdot\!\!\lessdot v$ is stronger than $u \lessdot v$.

6.1 Approximation by rational vectors

Recall that we have written $\omega_0 = (1, \bar{\omega}_0) \in \mathbb{R}^n$ with $\bar{\omega}_0 \in [-1, 1]^{n-1}$. For a given $Q \geq 1$, it is always possible to find a rational vector $v = (1, p/q) \in \mathbb{Q}^n$, with $p \in \mathbb{Z}^{n-1}$ and $q \in \mathbb{N}$, which is a Q-approximation in the sense that $|q\omega_0 - qv| \leq Q^{-1}$, and for which the denominator q satisfies the upper bound $q \leq Q^{n-1}$: this is essentially the content of Dirichlet's theorem on simultaneous rational approximations, and it holds true without any assumption on ω_0. In our situation, since we have assumed that ω_0 is non-resonant, there exist not only one, but n linearly independent rational vectors in $\mathbb{Q}^n$ which are Q-approximations. Moreover, one can obtain not only linearly independent vectors, but rational vectors $v_1, \ldots, v_n$ of denominators $q_1, \ldots, q_n$ such that the associated integer vectors $q_1 v_1, \ldots, q_n v_n$ form a $\mathbb{Z}$-basis of $\mathbb{Z}^n$. However, the upper bound on the corresponding denominators $q_1, \ldots, q_n$ is necessarily larger than Q^{n-1}, and is given by a function of Q that we can call here Ψ'_{ω_0} (see [BF13] for more precise and general information, but note that in this reference, Ψ'_{ω_0} was denoted by Ψ_{ω_0} and Ψ_{ω_0}, which we defined in (3), was denoted by Ψ'_{ω_0}). A consequence of the main Diophantine result of [BF13] is that this function Ψ'_{ω_0} is in fact essentially equivalent to the function Ψ_{ω_0}.

Proposition 2. *Let $\omega_0 = (1, \bar{\omega}_0) \in \mathbb{R}^n$ be a non-resonant vector, with $\bar{\omega}_0 \in [-1, 1]^{n-1}$. For any $Q \geq n + 2$, there exist n rational vectors $v_1, \ldots, v_n$, of denominators $q_1, \ldots, q_n$, such that $q_1 v_1, \ldots, q_n v_n$ form a $\mathbb{Z}$-basis of $\mathbb{Z}^n$ and for $j \in \{1, \ldots, n\}$,*

$$|\omega_0 - v_j| \lessdot (q_j Q)^{-1}, \quad 1 \leq q_j \lessdot \Psi(Q).$$

For a proof of the above proposition with Ψ_{ω_0} instead of Ψ, we refer to [BF13], Theorem 2.1 and Proposition 2.3; now by (4), $\Psi_{\omega_0} \leq \Psi$ and so one may replace Ψ_{ω_0} by Ψ.

Now given a q-rational vector v and a smooth function H defined on $\mathbb{T}^n \times D_{r,h}$, we define

$$[H]_v(\theta, I, \omega) = \int_0^1 H(\theta + tqv, I, \omega)dt. \tag{11}$$

Given n rational vectors $v_1, \ldots, v_n$, we let $[H]_{v_1,\ldots,v_d} = [\cdots [H]_{v_1} \cdots]_{v_d}$. Finally we define

$$[H](I, \omega) = \int_{\mathbb{T}^n} H(\theta, I, \omega)d\theta. \tag{12}$$

The following proposition is a consequence of the fact that the vectors $q_1 v_1, \ldots, q_n v_n$ form a $\mathbb{Z}$-basis of $\mathbb{Z}^n$.

Proposition 3 ([Bou13, Corollary 6]). *Let $v_1, \ldots, v_n$ be rational vectors, of denominators $q_1, \ldots, q_n$, such that $q_1 v_1, \ldots, q_n v_n$ form a $\mathbb{Z}$-basis of $\mathbb{Z}^n$, and H a function defined on $\mathbb{T}^n \times D_{r,h}$. Then*

$$[H]_{v_1,\ldots,v_n} = [H].$$

6.2 KAM step

Now we describe an elementary step of our iterative procedure. Such a step consists in pulling back the Hamiltonian H by a transformation of the form

$$\mathcal{F} = (\Phi, \varphi) : (\theta, I, \omega) \mapsto (\Phi(\theta, I, \omega), \varphi(\omega));$$

Φ is a parameter-depending change of coordinates and φ a change of parameters. Moreover, our change of coordinates will be of the form

$$\Phi(\theta, I, \omega) = \Phi_\omega(\theta, I) = (\theta + E(\theta, \omega), I + F(\theta, \omega) \cdot I + G(\theta, \omega))$$

with

$$E : \mathbb{T}^n \times D_h^{\omega_0} \to \mathbb{R}^n, \quad F : \mathbb{T}^n \times D_h^{\omega_0} \to M_n(\mathbb{R}), \quad G : \mathbb{T}^n \times D_h^{\omega_0} \to \mathbb{R}^n$$

and for each fixed parameter ω, Φ_ω will be symplectic. For simplicity, we shall write $\Phi = (E, F, G)$; the composition of such transformations $\mathcal{F} = (\Phi, \varphi) = (E, F, G, \varphi)$ is again a transformation of the same form, and we shall denote by $\mathcal{G}$ the groupoid of such transformations.

Proposition 4. *Let H be as in $(**)$, with $\omega_0 = (1, \bar{\omega}_0) \in \mathbb{R}^n$ non-resonant, consider $0 < \sigma < s$, $0 < \delta < r$, $Q \geq n + 2$, and assume that*

$$\sqrt{\varepsilon} \leq \mu \leq h/2, \quad \sqrt{\varepsilon} \leq r, \quad h \lessdot (Q\Psi(Q))^{-1}, \quad r\mu \lessdot \delta(Q\Psi(Q))^{-1}, \quad (1+\eta) \lessdot Q\sigma^\alpha. \quad (13)$$

Then there exists an $(\alpha, s - \sigma)$-Gevrey symplectic transformation

$$\mathcal{F} = (\Phi, \varphi) = (E, F, G, \varphi) : \mathbb{T}^n \times D_{r-\delta, h/2} \to \mathbb{T}^n \times D_{r,h} \in \mathcal{G},$$

with the estimates

$$\begin{cases} |E|_{\alpha, s-\sigma} \lessdot \Psi(Q)\mu, & |\nabla E|_{\alpha, s-\sigma} \lessdot \sigma^{-\alpha}\Psi(Q)\mu, \\ |F|_{\alpha, s-\sigma} \lessdot \sigma^{-\alpha}\Psi(Q)\mu, & |\nabla F|_{\alpha, s-\sigma} \lessdot \sigma^{-2\alpha}\Psi(Q)\mu, \\ |G|_{\alpha, s-\sigma} \lessdot \sigma^{-\alpha}\Psi(Q)\varepsilon, & |\nabla G|_{\alpha, s-\sigma} \lessdot \sigma^{-2\alpha}\Psi(Q)\varepsilon, \\ |\varphi - \mathrm{Id}|_{\alpha, s-\sigma} \leq \mu, & |\nabla\varphi - \mathrm{Id}|_{\alpha, s-\sigma} \lessdot \sigma^{-\alpha}\mu \end{cases} \quad (14)$$

such that

$$H \circ \mathcal{F}(\theta, I, \omega) = \underbrace{e^+(\omega) + \omega \cdot I}_{N^+(I,\omega)} + \underbrace{A^+(\theta) + B^+(\theta) \cdot I}_{P^+(\theta, I, \omega)} + \underbrace{M^+(\theta, I, \omega) \cdot I^2}_{R^+(\theta, I, \omega)},$$

with the estimates

$$\begin{cases} |A^+|_{\alpha, s-\sigma} \leq \varepsilon/16, & |B^+|_{\alpha, s-\sigma} \leq \mu/4, \\ |e^+ - e \circ \varphi|_{\alpha, s-\sigma} \leq |A|_{\alpha, s}, & |\nabla_I^2 R^+ - \nabla_I^2 R \circ \mathcal{F}|_{\alpha, s-\sigma} \lessdot \eta|F|_{\alpha, s-\sigma}. \end{cases} \quad (15)$$

Proof. We divide the proof of the KAM step into five small steps. Except for the last one, the parameter $\omega \in D_h^{\omega_0}$ will be fixed, so for simplicity, in the first four steps we will drop the dependence on the parameter $\omega \in D_h^{\omega_0}$. Let us first notice that (13) clearly implies the following seven inequalities:

$$h \lessdot (Q\Psi(Q))^{-1} \qquad (16)$$

$$\Psi(Q)\mu \lessdot \sigma^\alpha \qquad (17)$$

$$\varepsilon \leq r\mu \qquad (18)$$

$$r\mu\sigma^{-\alpha}\Psi(Q) \lessdot \delta \qquad (19)$$

$$\mu \lessdot (Q\Psi(Q))^{-1} \qquad (20)$$

$$(1+\eta) \lessdot Q\sigma^\alpha \qquad (21)$$

$$\mu \leq h/2. \qquad (22)$$

It is also important to notice that the implicit constant appearing in (21) is independent of the other implicit constants; we may choose it as large as we want without affecting the other implicit constants. In the first three steps, the term R which contains terms of order at least 2 in I will be ignored, that is we will only consider $\hat{H} = H - R = N + P$.

19

1. Rational approximations of ω_0 and $\omega \in D_h^{\omega_0}$

Since ω_0 is non-resonant, given $Q \geq n+2$, we can apply Proposition 2: there exist n rational vectors $v_1, \ldots, v_n$, of denominators $q_1, \ldots, q_n$, such that $q_1 v_1, \ldots, q_n v_n$ form a $\mathbb{Z}$-basis of $\mathbb{Z}^n$ and for $j \in \{1, \ldots, n\}$,

$$|\omega_0 - v_j| \prec (q_j Q)^{-1}, \quad 1 \leq q_j \prec \Psi(Q).$$

For any $\omega \in D_h^{\omega_0}$, using (16) and $q_j \prec \Psi(Q)$, we have

$$|\omega - v_j| \leq |\omega - \omega_0| + |\omega_0 - v_j| \prec h + (q_j Q)^{-1} \prec (Q\Psi(Q))^{-1} + (q_j Q)^{-1} \prec (q_j Q)^{-1}. \quad (23)$$

2. Successive rational averagings

Let us set $A_1 := A$, $B_1 := B$ so that $P_1(\theta, I) := A_1(\theta) + B_1(\theta) \cdot I$ satisfies $P_1 = P$. Recalling that $[\,.\,]_v$ denotes the averaging along the periodic flow associated to a periodic vector $v \in \mathbb{R}^n$ (see (11)), we define inductively, for $1 \leq j \leq n$,

$$A_{j+1} := [A_j]_{v_j}, \quad B_{j+1} := [B_j]_{v_j}, \quad P_{j+1} := [P_j]_{v_j}$$

so in particular $P_j(\theta, I) = A_j(\theta) + B_j(\theta) \cdot I$ for $1 \leq j \leq n$. Let us also define X_j, for $1 \leq j \leq n$, by

$$X_j(\theta, I) := C_j(\theta) + D_j(\theta) \cdot I$$

where

$$C_j(\theta) = q_j \int_0^1 (A_j - A_{j+1})(\theta + t q_j v_j) t \, dt, \quad D_j(\theta) = q_j \int_0^1 (B_j - B_{j+1})(\theta + t q_j v_j) t \, dt.$$

If we further define N_j by $N_j(I) = e(\omega) + v_j \cdot I$, it is then easy to check, by a simple integration by parts, that the equations

$$\{C_j, N_j\} = A_j - A_{j+1}, \quad \{D_j, N_j\} = B_j - B_{j+1}, \quad 1 \leq j \leq n, \quad (24)$$

and then

$$\{X_j, N_j\} = P_j - P_{j+1}, \quad 1 \leq j \leq n, \quad (25)$$

are satisfied, where $\{\,.\,,\,.\,\}$ denotes the usual Poisson bracket. Moreover, we have the estimates

$$|A_j|_{\alpha,s} \leq |A|_{\alpha,s} \leq \varepsilon, \quad |B_j|_{\alpha,s} \leq |B|_{\alpha,s} \leq \mu, \quad (26)$$

and then

$$|C_j|_{\alpha,s} \leq q_j |A_j|_{\alpha,s} \leq q_j \varepsilon \prec \Psi(Q)\varepsilon, \quad |D_j|_{\alpha,s} \leq q_j |B_j|_{\alpha,s} \leq q_j \mu \prec \Psi(Q)\mu. \quad (27)$$

Next, for any $0 \leq j \leq n$, define $r_j = r - n^{-1}j\delta$ and $s_j = s - (2n)^{-1}j\sigma$. We have $r_n = r - \delta \leq r_j \leq r_0 = r$ while $s_n = s - \sigma/2 \leq s_j \leq s_0 = s$. Let X_j^t be the time-t map of the Hamiltonian flow of X_j. Using (27), together with inequalities (17), (18) and

20

(19), the condition (80) and (82) of Proposition 22, Appendix B, are satisfied, so the latter proposition can be applied: for $1 \leq j \leq n$, X_j^t maps $\mathbb{T}^n \times B_{r_j}$ into $\mathbb{T}^n \times B_{r_{j-1}}$ for all $t \in [0,1]$ and it is of the form

$$X_j^t(\theta, I) = (\theta + E_j^t(\theta), I + F_j^t(\theta) \cdot I + G_j^t(\theta))$$

with

$$\begin{cases} |E_j^t|_{\alpha,s_j} \leq |D_j|_{\alpha,s_{j-1}} \lessdot \Psi(Q)\mu \\ |F_j^t|_{\alpha,s_j} \lessdot \sigma^{-\alpha}|D_j|_{\alpha,s_{j-1}} \lessdot \sigma^{-\alpha}\Psi(Q)\mu \\ |G_j^t|_{\alpha,s_j} \lessdot \sigma^{-\alpha}|C_j|_{\alpha,s_{j-1}} \lessdot \sigma^{-\alpha}\Psi(Q)\varepsilon. \end{cases} \tag{28}$$

Now we define $\Phi^0 := \mathrm{Id}$ to be the identity and inductively $\Phi^j := \Phi^{j-1} \circ X_j^1$ for $1 \leq j \leq n$. Then Φ^j maps $\mathbb{T}^n \times B_{r_j}$ into $\mathbb{T}^n \times B_r$ and one easily checks, by induction using the estimates (28), that Φ^j is still of the form

$$\Phi^j(\theta, I) = (\theta + E^j(\theta), I + F^j(\theta) \cdot I + G^j(\theta))$$

with the estimates, for $j = 1, ..., n$,

$$|E^j|_{\alpha,s_j} \lessdot \Psi(Q)\mu, \quad |F^j|_{\alpha,s_j} \lessdot \sigma^{-\alpha}\Psi(Q)\mu, \quad |G^j|_{\alpha,s_j} \lessdot \sigma^{-\alpha}\Psi(Q)\varepsilon. \tag{29}$$

3. New Hamiltonian

Let us come back to the Hamiltonian $\hat{H} = H - R = N + P = N + P_1$. We claim that for all $0 \leq j \leq n$, we have

$$\hat{H} \circ \Phi^j = N + P_{j+1} + P_{j+1}^+, \quad P_{j+1}^+(\theta, I) = A_{j+1}^+(\theta) + B_{j+1}^+(\theta) \cdot I$$

with the estimates

$$|A_{j+1}^+|_{\alpha,s_j} \lessdot (Q\sigma^\alpha)^{-1}\varepsilon, \quad |B_{j+1}^+|_{\alpha,s_j} \lessdot (Q\sigma^\alpha)^{-1}\mu. \tag{30}$$

Let us prove the claim by induction on $0 \leq j \leq n$. For $j = 0$, we may set $P_1^+ := 0$ and there is nothing to prove. So let us assume that the claim is true for some $j - 1 \geq 0$, and we need to show it is still true for $j \geq 1$. By this inductive assumption, we have

$$\hat{H} \circ \Phi^j = \hat{H} \circ \Phi^{j-1} \circ X_j^1 = (N + P_j + P_j^+) \circ X_j^1$$

with

$$|A_j^+|_{\alpha,s_{j-1}} \lessdot (Q\sigma^\alpha)^{-1}\varepsilon, \quad |B_j^+|_{\alpha,s_{j-1}} \lessdot (Q\sigma^\alpha)^{-1}\mu. \tag{31}$$

Let $S_j = \omega \cdot I - v_j \cdot I$ so that $N = N_j + S_j$ and thus

$$\hat{H} \circ \Phi^j = (N_j + S_j + P_j + P_j^+) \circ X_j^1 = (N_j + S_j + P_j) \circ X_j^1 + P_j^+ \circ X_j^1.$$

Let us consider the first summand of the last sum. Using the equality (25), a standard computation based on Taylor's formula with integral remainder gives

$$(N_j + S_j + P_j) \circ X_j^1 = N + [P_j]_{v_j} + \tilde{P}_{j+1} = N + P_{j+1} + \tilde{P}_{j+1}$$

21

with

$$\tilde{P}_{j+1} = \int_0^1 U_{j+1}^t \circ X_j^t dt, \quad U_{j+1}^t := \{(1-t)P_{j+1} + tP_j + S_j, X_j\}.$$

Clearly, U_{j+1}^t is still of the form

$$U_{j+1}^t(\theta, I) = U_{j+1}^t(\theta, 0) + \nabla_I U_{j+1}^t(\theta, 0) \cdot I$$

as this is true for P_j, S_j, X_j and that this form is preserved under Poisson bracket. Using the estimates for $P_j(\theta, 0)$, $\nabla_I P_j(\theta, 0)$, $X_j(\theta, 0)$, $\nabla_I X_j(\theta, 0)$ (given respectively in (26) and in (27)), the fact that

$$S_j(\theta, 0) = 0, \quad \nabla_I S_j(\theta, 0) = \omega - v_j$$

with the inequality (23), and the estimates for the derivatives and the product of Gevrey functions (given respectively in Proposition 14, Corollary 15 and Proposition 16, Corollary 18, Appendix B), one finds, for all $t \in [0, 1]$

$$|U_{j+1}^t(\theta, 0)|_{\alpha, s_j} \lessdot (\sigma^{-\alpha} q_j \varepsilon \mu + \sigma^{-\alpha} q_j \varepsilon \mu + \sigma^{-\alpha} q_j \varepsilon (q_j Q)^{-1}) \lessdot \sigma^{-\alpha} q_j \varepsilon \mu + (Q\sigma^\alpha)^{-1} \varepsilon.$$

Since $q_j \lessdot \Psi(Q)$, using (20) the latter estimate reduces to

$$|U_{j+1}^t(\theta, 0)|_{\alpha, s_j} \lessdot (Q\sigma^\alpha)^{-1} \varepsilon.$$

Similarly, one obtains

$$|\nabla_I U_{j+1}^t(\theta, 0)|_{\alpha, s_j} \lessdot (Q\sigma^\alpha)^{-1} \mu.$$

Then, using the expression of X_j^t and the associated estimates (28), a direct computation, still using (20), gives

$$|\tilde{P}_{j+1}(\theta, 0)|_{\alpha, s_j} \lessdot (Q\sigma^\alpha)^{-1} \varepsilon$$

and

$$|\nabla_I \tilde{P}_{j+1}(\theta, 0)|_{\alpha, s_j} \lessdot (Q\sigma^\alpha)^{-1} \mu.$$

Using again the estimates of X_j^t given by (28), and the inductive assumption (31), we also find

$$|P_j^+ \circ X_j^1(\theta, 0)|_{\alpha, s_j} \lessdot (Q\sigma^\alpha)^{-1} \varepsilon$$

and

$$|\nabla_I (P_j^+ \circ X_j^1)(\theta, 0)|_{\alpha, s_j} \lessdot (Q\sigma^\alpha)^{-1} \mu.$$

Eventually, we may define

$$P_{j+1}^+ := \tilde{P}_{j+1} + P_j^+ \circ X_j^1$$

so that

$$\hat{H} \circ \Phi^j = N + P_{j+1} + P_{j+1}^+, \quad P_{j+1}^+(\theta, I) = A_{j+1}^+(\theta) + B_{j+1}^+(\theta) \cdot I$$

and these last estimates imply that

$$|A_{j+1}^+|_{\alpha, s_j} \lessdot (Q\sigma^\alpha)^{-1} \varepsilon, \quad |B_{j+1}^+|_{\alpha, s_j} \lessdot (Q\sigma^\alpha)^{-1} \mu.$$

22

The claim is proved. So we may set

$$\Phi := \Phi^n, \quad (E, F, G) := (E^n, F^n, G^n),$$

with, as (29) tells us with $j = n$ and $s_n = s - \sigma/2$,

$$|E|_{\alpha, s-\sigma/2} \lessdot \Psi(Q)\mu, \quad |F|_{\alpha, s\sigma/2} \lessdot \sigma^{-\alpha}\Psi(Q)\mu, \quad |G|_{\alpha, s-\sigma/2} \lessdot \sigma^{-\alpha}\Psi(Q)\varepsilon. \quad (32)$$

Observe that $P_{n+1} = [\cdots [P]_{v_1} \cdots]_{v_n} = [P]_{v_1, \ldots, v_n}$, and thus by Proposition 3, $P_{n+1} = [P]$, and as a consequence

$$\hat{H} \circ \Phi(\theta, I) = e + \omega \cdot I + [A] + [B] \cdot I + A_{n+1}^+(\theta) + B_{n+1}^+(\theta) \cdot I$$

with the estimates

$$|A_{n+1}^+|_{\alpha, s-\sigma/2} \lessdot (Q\sigma^\alpha)^{-1}\varepsilon, \quad |B_{n+1}^+|_{\alpha, s-\sigma/2} \lessdot (Q\sigma^\alpha)^{-1}\mu. \quad (33)$$

4. Estimate of the remainder

Now we take into account the remainder term R that we previously ignored: we have $H = \hat{H} + R$, and therefore

$$H \circ \Phi(\theta, I) = e + \omega \cdot I + [A] + [B] \cdot I + A_{n+1}^+(\theta) + B_{n+1}^+(\theta) \cdot I + R \circ \Phi(\theta, I).$$

Let us decompose

$$R \circ \Phi(\theta, I) = \underbrace{R \circ \Phi(\theta, 0)}_{R_0(\theta)} + \underbrace{\nabla_I(R \circ \Phi)(\theta, 0)}_{R_1(\theta)} \cdot I + \tilde{R}(\theta, I)$$

and let us define

$$\tilde{A} := A_{n+1}^+ + R_0, \quad \tilde{B} := B_{n+1}^+ + R_1.$$

We have $R(\theta, I) = M(\theta, I) \cdot I^2$ and as H and R differ only by terms of order at most one in I, $\nabla_I^2 H = \nabla_I^2 R$ so

$$M(\theta, I) = \int_0^1 (1-t)\nabla_I^2 H(\theta, tI)dt = \int_0^1 (1-t)\nabla_I^2 R(\theta, tI)dt$$

and therefore $|M|_{\alpha, s} \leq \eta$. Then, as $\Phi(\theta, 0) = (\theta + E(\theta), G(\theta))$, we have the expression

$$R_0(\theta) = R(\Phi(\theta, 0)) = M(\theta + E(\theta), G(\theta)) \cdot G(\theta)^2$$

and so using the above estimate on M, together with the estimates on E, G and the estimates for the product and compostion of Gevrey functions (given respectively in Proposition 16 and Proposition 20, Appendix B), we find

$$|R_0|_{\alpha, s-\sigma/2} \lessdot \eta|G|_{\alpha, s-\sigma/2}^2 \lessdot \eta(\sigma^{-\alpha}\Psi(Q)\mu)^2\varepsilon \lessdot \eta(Q\sigma^\alpha)^{-2}\varepsilon \lessdot \eta(Q\sigma^\alpha)^{-1}\varepsilon.$$

Then, we have $\nabla_I R(\theta, I) = \hat{M}(\theta, I) \cdot I^2$ with

$$\hat{M}(\theta, I) = \int_0^1 \nabla_I^2 H(\theta, tI) dt = \int_0^1 \nabla_I^2 R(\theta, tI) dt$$

and hence $|\hat{M}|_{\alpha,s} \leq \eta$. Since

$$|\nabla_I \Phi - \mathrm{Id}|_{\alpha,s-\sigma} = |F|_{\alpha,s-\sigma} \lessdot \sigma^{-\alpha} \Psi(Q) \mu \lessdot 1 \tag{34}$$

we obtain, using the fact that $\varepsilon \leq \mu^2$ and proceeding as before,

$$|R_1|_{\alpha,s-\sigma/2} \lessdot \eta |G|_{\alpha,s-\sigma/2} \lessdot \eta \sigma^{-\alpha} \Psi(Q) \varepsilon \lessdot \eta(\sigma^{-\alpha} \Psi(Q) \mu)\mu \lessdot \eta(Q\sigma^\alpha)^{-1}\mu.$$

These last estimates on R_0 and R_1, together with (33), imply

$$|\tilde{A}|_{\alpha,s-\sigma/2} \lessdot (1+\eta)(Q\sigma^\alpha)^{-1}\varepsilon, \quad |\tilde{B}|_{\alpha,s-\sigma/2} \lessdot (1+\eta)(Q\sigma^\alpha)^{-1}\mu.$$

We can finally now use (21) to ensure that

$$|\tilde{A}|_{\alpha,s-\sigma/2} \leq \varepsilon/16, \quad |\tilde{B}|_{\alpha,s-\sigma/2} \leq \mu/4. \tag{35}$$

It is important to recall here that we may choose the implicit constant in (21) as large as we want (in order to achieve (35)) without affecting any of the other implicit constants. Then observe also that $H \circ \Phi$ and $\tilde{R}$ differ only by terms of order at most one in I, so

$$\nabla_I^2(H \circ \Phi) = \nabla_I^2 \tilde{R}, \quad \nabla_I^2 H = \nabla_I^2 R$$

and therefore using the formula for the Hessian of a composition, (34) and the fact that $\nabla_I^2 \Phi$ is identically zero, one finds

$$|\nabla_I^2 \tilde{R} - \nabla_I^2 R \circ \Phi|_{\alpha,s-\sigma/2} \lessdot \eta |F|_{\alpha,\sigma}. \tag{36}$$

We also set $\tilde{e} := e + [A]$ and observe that

$$|\tilde{e} - e|_{\alpha,s-\sigma/2} \leq |[A]|_{\alpha,s} \leq |A|_{\alpha,s}. \tag{37}$$

5. Change of frequencies and final estimates

Let us now write explicitly the dependence on the parameter $\omega \in D_h^{\omega_0}$: we have

$$H \circ \Phi(\theta, I, \omega) = \tilde{e}(\omega) + (\omega + [B](\omega)) \cdot I + \tilde{A}(\theta, \omega) + \tilde{B}(\theta, \omega) \cdot I + \tilde{R}(\theta, I, \omega).$$

Consider the map $\phi(\omega) := \omega + [B(\omega)]$, it satisfies

$$|\phi - \mathrm{Id}|_{\alpha,s} \leq |[B]|_{\alpha,s} \leq |B|_{\alpha,s} \leq \mu$$

and therefore the conditions (83) of Proposition 23 are satisfied: the first condition of (83) follows, from instance, from condition (17) and the fact that $\Psi(Q) \geq Q \geq 1$, whereas the

second condition of (83) is implied by condition (22). Hence Proposition 23 applies and one finds a unique $\varphi \in G_{\alpha,s-\sigma/2}(D^{\omega_0}_{h/2}, D^{\omega_0}_h)$ such that $\phi \circ \varphi = \mathrm{Id}$ and

$$|\varphi - \mathrm{Id}|_{\alpha,s-\sigma/2} \le |\phi - \mathrm{Id}|_{\alpha,s} \le \mu. \tag{38}$$

We do have $\varphi(\omega) + [B(\varphi(\omega))] = \omega$ and thus, setting $\mathcal{F} := (\Phi, \varphi)$, this implies that

$$H \circ \mathcal{F}(\theta, I, \omega) = H \circ \Phi(\theta, I, \varphi(\omega)) = \tilde{e}(\varphi(\omega)) + \omega \cdot I + \tilde{A}(\theta, \varphi(\omega)) + \tilde{B}(\theta, \varphi(\omega)) \cdot I + \tilde{R}(\theta, I, \varphi(\omega))$$

and at the end we set

$$e^+ := \tilde{e} \circ \varphi, \quad A^+ := \tilde{A} \circ \varphi, \quad B^+ := \tilde{B} \circ \varphi, \quad R^+ := \tilde{R} \circ \varphi.$$

Using once again Proposition 20, the inequalities (35), (36) and (37) imply

$$\begin{cases} |A^+|_{\alpha,s-\sigma} = |\tilde{A} \circ \varphi|_{\alpha,s-\sigma} \le |\tilde{A}|_{\alpha,s-\sigma/2} \le \varepsilon/16, \\ |B^+|_{\alpha,s-\sigma} = |\tilde{B} \circ \varphi|_{\alpha,s-\sigma} \le |\tilde{B}|_{\alpha,s-\sigma/2} \le \mu/4, \\ |e^+ - e \circ \varphi|_{\alpha,s-\sigma} = |(\tilde{e} - e) \circ \varphi|_{\alpha,s-\sigma} \le |\tilde{e} - e|_{\alpha,s-\sigma/2} \le |A|_{\alpha,s}, \\ |\nabla_I^2 R^+ - \nabla_I^2 R \circ \mathcal{F}| = |(\nabla_I^2 \tilde{R} - \nabla_I^2 R \circ \Phi) \circ \varphi|_{\alpha,s-\sigma} \le |\nabla_I^2 \tilde{R} - \nabla_I^2 R \circ \Phi|_{\alpha,s-\sigma/2} \lessdot \eta|F|_{\alpha,s-\sigma}, \end{cases}$$

which were the estimates (15) we needed to prove. The transformation $\mathcal{F} = (\Phi, \varphi) = (E, F, G, \varphi) \in \mathcal{G}$ maps $\mathbb{T}^n \times D_{r-\delta,h/2}$ into $\mathbb{T}^n \times D_{r,h}$ and it follows from (32) and (38) that

$$\begin{cases} |E|_{\alpha,s-\sigma} \lessdot \Psi(Q)\mu, \quad |\nabla E|_{\alpha,s-\sigma} \lessdot \sigma^{-\alpha}\Psi(Q)\mu, \\ |F|_{\alpha,s-\sigma} \lessdot \sigma^{-\alpha}\Psi(Q)\mu, \quad |\nabla F|_{\alpha,s-\sigma} \lessdot \sigma^{-2\alpha}\Psi(Q)\mu, \\ |G|_{\alpha,s-\sigma} \lessdot \sigma^{-\alpha}\Psi(Q)\varepsilon, \quad |\nabla G|_{\alpha,s-\sigma} \lessdot \sigma^{-2\alpha}\Psi(Q)\varepsilon, \\ |\varphi - \mathrm{Id}|_{\alpha,s-\sigma} \le \mu, \quad |\nabla\varphi - \mathrm{Id}|_{\alpha,s-\sigma} \lessdot \sigma^{-\alpha}\mu \end{cases}$$

which were the wanted estimates (14). This concludes the proof. $\qquad\square$

6.3   Iterations and convergence

We now define, for $i \in \mathbb{N}$, the following decreasing geometric sequences:

$$\varepsilon_i := 16^{-i}\varepsilon, \quad \mu_i := 4^{-i}\mu, \quad \delta_i := 2^{-i-2}r, \quad h_i = 2^{-i}h. \tag{39}$$

Next, for a constant $Q_0 \ge n + 2$ to be chosen below, we define Δ_i and Q_i, $i \in \mathbb{N}$, by

$$\Delta_i = 2^i \Delta(Q_0), \quad Q_i = \Delta^{-1}(\Delta_i) = \Delta^{-1}(2^i \Delta(Q_0)), \tag{40}$$

and then we define σ_i, $i \in \mathbb{N}$, by

$$\sigma_i = C Q_i^{-\frac{1}{\alpha}} \tag{41}$$

where $C \geq 1$ is a sufficiently large constant so that the last condition of (13) is satisfied for $\sigma = \sigma_0$ and $Q = Q_0$ (and thus for $\sigma = \sigma_i$ and $Q = Q_i$, for any $i \in \mathbb{N}$); clearly, this constant is of the form $C = \cdot (1 + \eta)^{1/\alpha}$. Finally, we define s_i and r_i, $i \in \mathbb{N}$, by

$$s_0 = s, \quad s_{i+1} = s_i - \sigma_i, \quad r_0 = r, \quad r_{i+1} = r_i - \delta_i. \tag{42}$$

Obviously, we have

$$\lim_{i \to +\infty} r_i = r - \sum_{i \in \mathbb{N}} \delta_i = r/2.$$

We claim that, assuming Δ^{-1} satisfies (A_α), which is equivalent to $\omega_0 \in \mathrm{BR}_\alpha$, we can choose Q_0 sufficiently large so that

$$\lim_{i \to +\infty} s_i \geq s/2 \iff \sum_{i \in \mathbb{N}} \sigma_i \leq s/2.$$

Indeed, since $Q_i = \Delta^{-1}(\Delta_i) = \Delta^{-1}(2^i \Delta(Q_0))$, we have

$$\sum_{i \geq 1} Q_i^{-\frac{1}{\alpha}} = \sum_{i \geq 1} \frac{1}{(\Delta^{-1}(2^i \Delta(Q_0)))^{\frac{1}{\alpha}}} \leq \int_0^{+\infty} \frac{dy}{(\Delta^{-1}(2^y \Delta(Q_0)))^{\frac{1}{\alpha}}} = \int_{\Delta(Q_0)}^{+\infty} \frac{(\ln 2)^{-1} dx}{x(\Delta^{-1}(x))^{\frac{1}{\alpha}}} < +\infty$$

where we made a change of variables $x := 2^y \Delta(Q_0)$, and the last integral converges since Δ^{-1} satisfies (A_α). Now as $\sigma_i = C Q_i^{-\frac{1}{\alpha}}$, we have

$$\sum_{i \in \mathbb{N}} \sigma_i = C \sum_{i \geq 0} Q_i^{-\frac{1}{\alpha}} = C Q_0^{-\frac{1}{\alpha}} + C \sum_{i \geq 1} Q_i^{-\frac{1}{\alpha}} \leq C Q_0^{-\frac{1}{\alpha}} + C \int_{\Delta(Q_0)}^{+\infty} \frac{(\ln 2)^{-1} dx}{x(\Delta^{-1}(x))^{\frac{1}{\alpha}}} \leq s/2$$

provided we choose Q_0 sufficiently large in order to have

$$Q_0^{-\frac{1}{\alpha}} + \int_{\Delta(Q_0)}^{+\infty} \frac{(\ln 2)^{-1} dx}{x(\Delta^{-1}(x))^{\frac{1}{\alpha}}} \leq (2C)^{-1} s. \tag{43}$$

Applying inductively Proposition 4 we will easily obtain the following proposition.

Proposition 5. *Let H be as in $(**)$, with $\omega_0 \in \mathrm{BR}_\alpha$, and fix $Q_0 \geq n+2$ sufficiently large so that (43) is satisfied. Assume that*

$$\sqrt{\varepsilon} \leq \mu \leq h/2, \quad \sqrt{\varepsilon} \leq r, \quad h \cdot < \Delta(Q_0)^{-1}. \tag{44}$$

Then, for each $i \in \mathbb{N}$, there exists an (α, s_i)-Gevrey smooth transformation

$$\mathcal{F}^i = (\Phi^i, \varphi^i) = (E^i, F^i, G^i, \varphi^i) : \mathbb{T}^n \times D_{r_i, h_i} \to \mathbb{T}^n \times D_{r,h} \in \mathcal{G},$$

such that $\mathcal{F}^{i+1} = \mathcal{F}^i \circ \mathcal{F}_{i+1}$, with

$$\mathcal{F}_{i+1} = (\Phi_{i+1}, \varphi_{i+1}) = (E_{i+1}, F_{i+1}, G_{i+1}, \varphi_{i+1}) : \mathbb{T}^n \times D_{r_{i+1}, h_{i+1}} \to \mathbb{T}^n \times D_{r_i, h_i} \in \mathcal{G},$$

26

satisfying the following estimates

$$\begin{cases} |E_{i+1}|_{\alpha,s_{i+1}} \lessdot \Psi(Q_i)\mu_i, & |\nabla E_{i+1}|_{\alpha,s_{i+1}} \lessdot \sigma_i^{-\alpha}\Psi(Q_i)\mu_i, \\ |F_{i+1}|_{\alpha,s_{i+1}} \lessdot \sigma_i^{-\alpha}\Psi(Q_i)\mu_i, & |\nabla F_{i+1}|_{\alpha,s_{i+1}} \lessdot \sigma_i^{-2\alpha}\Psi(Q_i)\mu_i, \\ |G_{i+1}|_{\alpha,s_{i+1}} \lessdot \sigma_i^{-\alpha}\Psi(Q_i)\varepsilon_i, & |\nabla G_{i+1}|_{\alpha,s_{i+1}} \lessdot \sigma_i^{-2\alpha}\Psi(Q_i)\varepsilon_i, \\ |\varphi_{i+1} - \mathrm{Id}|_{\alpha,s_{i+1}} \leq \mu_i, & |\nabla\varphi_{i+1} - \mathrm{Id}|_{\alpha,s_{i+1}} \lessdot \sigma_i^{-\alpha}\mu_i \end{cases} \tag{45}$$

and such that

$$H \circ \mathcal{F}^i(\theta, I, \omega) = \underbrace{e^i(\omega) + \omega \cdot I}_{N^i(I,\omega)} + \underbrace{A^i(\theta) + B^i(\theta) \cdot I}_{P^i(\theta,I,\omega)} + \underbrace{M^i(\theta, I, \omega) \cdot I^2}_{R^i(\theta,I,\omega)}$$

with the estimates

$$\begin{cases} |A^i|_{\alpha,s_{i+1}} \leq \varepsilon_i, & |B^i|_{\alpha,s_{i+1}} \leq \mu_i, \\ |e^{i+1} - e^i \circ \varphi_{i+1}|_{\alpha,s_{i+1}} \leq |A^i|_{\alpha,s_i}, \\ |\nabla_I^2 R^{i+1} - \nabla_I^2 R^i \circ \mathcal{F}_{i+1}|_{\alpha,s_{i+1}} \lessdot \eta|F_{i+1}|_{\alpha,s_{i+1}}. \end{cases} \tag{46}$$

Let us emphasize that the implicit constants in the above proposition depend only on n and α and are thus independent of $i \in \mathbb{N}$.

Proof. For $i = 0$, we let $\mathcal{F}^0$ be the identity, $e^0 := e$, $A^0 := A$, $B^0 := B$, $R^0 := R$, $M^0 := M$ and there is nothing to prove. The general case follows by an easy induction. Indeed, assume that the statement holds true for some $i \in \mathbb{N}$ so that $H \circ \mathcal{F}^i$ is (α, s_i)-Gevrey on the domain $\mathbb{T}^n \times D_{r_i,s_i}$. We want to apply Proposition 4 to this Hamiltonian, with $\varepsilon = \varepsilon_i$, $\mu = \mu_i$, $r = r_i$, $s = s_i$, $h = h_i$, $\sigma = \sigma_i$ and $Q = Q_i$. First, by our choice of Q_0 and δ_0 it is clear that $0 < \sigma_i < s_i$, $0 < \delta_i < r_i$, and $Q_i \geq n + 2$. Then we need to check that the conditions

$$\sqrt{\varepsilon_i} \leq \mu_i \leq h_i/2, \quad \sqrt{\varepsilon_i} \leq r_i, \quad h_i \lessdot \Delta(Q_i)^{-1}, \quad r_i\mu_i \lessdot \delta_i\Delta(Q_i)^{-1}, \quad 1 \lessdot Q_i\sigma_i^\alpha$$

are satisfied. Since

$$\Delta(Q_i) = \Delta(\Delta^{-1}(\Delta_i)) = \Delta_i, \quad 2^{-i}r \leq r_i \leq r,$$

it is sufficient to check the conditions

$$\sqrt{\varepsilon_i} \leq \mu_i \leq h_i/2, \quad \sqrt{\varepsilon_i} \leq 2^{-i}r, \quad h_i \lessdot \Delta_i^{-1}, \quad r\mu_i \lessdot \delta_i\Delta_i^{-1}, \quad 1 \lessdot Q_i\sigma_i^\alpha. \tag{47}$$

The last condition of (47) is satisfied, for all $i \in \mathbb{N}$, simply by the choice of the constant C in the definition of σ_i. As for the other four conditions of (47), using the fact that the sequences ε_i, μ_i, h_i, Δ_i^{-1} and δ_i decrease at a geometric rate with respective ratio $1/16$, $1/4$, $1/2$, $1/2$ and $1/2$, it is clear that they are satisfied for any $i \in \mathbb{N}$ if and only if they are satisfied for $i = 0$. The first three conditions of (47) for $i = 0$ are nothing but (44). Moreover, using our choice of $\delta_0 = r/4$, the fourth condition of (47) for $i = 0$ reads $\mu \lessdot \Delta_0^{-1}$ and this also follows from (44).

Hence Proposition 4 can be applied, and all the conclusions of the statement follow at once from the conclusions of Proposition 4. $\square$

We can finally conclude the proof of Theorem H, by showing that one can pass to the limit $i \to +\infty$ in Proposition 5.

Proof of Theorem H. Recall that we are given $\varepsilon > 0$, $\mu > 0$, $r > 0$, $s > 0$, $h > 0$ and that we define the sequences $\varepsilon_i, \mu_i, \delta_i, h_i$ in (39), and then we chose $Q_0 \geq n + 2$ satisfying (43) to define the sequences Δ_i, Q_i in (40) and σ_i in (41) and finally, s_i and r_i were defined in (42). Moreover, we have

$$
\begin{cases}
\lim_{i \to +\infty} \varepsilon_i = \lim_{i \to +\infty} \mu_i = \lim_{i \to +\infty} h_i = 0, \\
\lim_{i \to +\infty} r_i = r - \sum_{i=0}^{+\infty} \delta_i = r/2, \quad \lim_{i \to +\infty} s_i = s - \sum_{i=0}^{+\infty} \sigma_i \geq s/2
\end{cases}
\tag{48}
$$

and for later use, let us observe that the following series are convergent and can be made as small as one wishes thanks to condition (5) of Theorem H:

$$
\sum_{i=0}^{+\infty} \sigma_i^{-\alpha} \mu_i \leq \sum_{i=0}^{+\infty} Q_i \mu_i = \sum_{i=0}^{+\infty} (\Psi(Q_i))^{-1} \Delta_i \mu_i \leq 2(\Psi(Q_0))^{-1} \Delta_0 \mu = 2Q_0 \mu
\tag{49}
$$

$$
\sum_{i=0}^{+\infty} \mu_i \leq 2\mu
\tag{50}
$$

$$
\sum_{i=0}^{+\infty} \sigma_i^{-\alpha} \Psi(Q_i) \mu_i \leq \sum_{i=0}^{+\infty} \Delta_i \mu_i \leq 2\Delta_0 \mu = 2Q_0 \Psi(Q_0) \mu
\tag{51}
$$

$$
\sum_{i=0}^{+\infty} \Psi(Q_i) \mu_i \leq \sum_{i=0}^{+\infty} Q_i^{-1} \Delta_i \mu_i \leq 2Q_0^{-1} \Delta_0 \mu = 2\Psi(Q_0) \mu
\tag{52}
$$

$$
\sum_{i=0}^{+\infty} \sigma_i^{-\alpha} \Psi(Q_i) \varepsilon_i \leq \sum_{i=0}^{+\infty} \Delta_i \varepsilon_i \leq 2\Delta_0 \varepsilon = 2Q_0 \Psi(Q_0) \varepsilon.
\tag{53}
$$

Now the condition (5) of Theorem H implies that the condition (44) of Proposition 5 is satisfied; what we need to prove is that the sequences given by this Proposition 5 do converge in the Banach space of $(\alpha, s/2)$-Gevrey functions. Recall that $\mathcal{F}^0 = (E^0, F^0, G^0, \varphi^0)$ is the identity, while for $i \geq 0$,

$$
(E^{i+1}, F^{i+1}, G^{i+1}, \varphi^{i+1}) = \mathcal{F}^{i+1} = \mathcal{F}^i \circ \mathcal{F}_{i+1} = (E^i, F^i, G^i, \varphi^i) \circ (E_{i+1}, F_{i+1}, G_{i+1}, \varphi_{i+1})
$$

from which one easily obtains the following inductive expressions:

$$
\begin{cases}
E^{i+1}(\theta, \omega) = E_{i+1}(\theta, \omega) + E^i(\theta + E_{i+1}(\theta, \omega), \varphi_{i+1}(\omega)) \\
F^{i+1}(\theta, \omega) = F_{i+1}(\theta, \omega) + F^i(\theta + E_{i+1}(\theta, \omega), \varphi_{i+1}(\omega)) \cdot (\mathrm{Id} + F_{i+1}(\theta, \omega)) \\
G^{i+1}(\theta, \omega) = (F^i(\theta + E_{i+1}(\theta, \omega), \varphi_{i+1}(\omega)) + \mathrm{Id}) \cdot G_{i+1}(\theta, \omega) + G^i(\theta + E_{i+1}(\theta, \omega), \varphi_{i+1}(\omega)) \\
\varphi^{i+1} = \varphi^i \circ \varphi_{i+1}.
\end{cases}
\tag{54}
$$

Let us first prove that the sequence φ^i converges. We claim that for all $i \in \mathbb{N}$, we have

$$|\nabla \varphi^i|_{\alpha, s_i} \lessdot \prod_{l=0}^{i} (1 + \sigma_l^{-\alpha} \mu_l) \lessdot 1$$

where the fact that the last product is bounded uniformly in $i \in \mathbb{N}$ follows from (49). For $i = 0$, $\varphi^0 = \text{Id}$ and there is nothing to prove; for $i \in \mathbb{N}$ since $\varphi^{i+1} = \varphi^\circ + \varphi_{i+1}$ we have

$$\nabla \varphi^{i+1} = \left(\nabla \varphi^i \circ \varphi_{i+1} \right) \cdot \nabla \varphi_{i+1}$$

so that using the estimate for φ_{i+1} and $\nabla \varphi_{i+1}$ given in (45), Proposition 5, the claim follows by induction. Using this claim, and writing

$$\varphi^{i+1} - \varphi^i = \varphi^i \circ \varphi_{i+1} - \varphi^i = \left(\int_0^1 \nabla \varphi^i \circ (t\varphi_{i+1} + (1-t)\text{Id}) dt \right) \cdot (\varphi_{i+1} - \text{Id})$$

one finds

$$|\varphi^{i+1} - \varphi^i|_{\alpha, s_{i+1}} \lessdot |\varphi_{i+1} - \text{Id}|_{\alpha, s_{i+1}},$$

and therefore

$$|\varphi^{i+1} - \varphi^i|_{\alpha, s_{i+1}} \lessdot \mu_i.$$

Using the convergence of (48) and (50), one finds that the domain of definition of φ^i shrinks to a point and the sequence φ^i converges to a trivial map

$$\varphi^* : \{\omega_0\} \to D_h^{\omega_0}, \quad \varphi^*(\omega_0) := \omega_0^*$$

such that

$$|\omega_0^* - \omega_0| \lessdot \mu.$$

Now let us define

$$V_{i+1}(\theta, \omega) := (\theta + E_{i+1}(\theta, \omega), \varphi_{i+1}(\omega)), \quad V_{i+1} = (\text{Id} + E_{i+1}, \varphi_{i+1})$$

and observe that since $\Psi(Q_i) \geq 1$, then the estimates for E_{i+1}, ∇E_{i+1}, φ_{i+1} and $\nabla \varphi_{i+1}$ given in Proposition 5 implies that

$$|V_{i+1} - \text{Id}|_{\alpha, s_{i+1}} \lessdot \Psi(Q_i)\mu_i, \quad |\nabla V_{i+1} - \text{Id}|_{\alpha, s_{i+1}} \lessdot \sigma_i^{-\alpha} \Psi(Q_i)\mu_i.$$

Using these estimates, and the fact that E^{i+1} can be written as

$$E^{i+1} = E_{i+1} + E^i \circ V_{i+1}$$

we can proceed as before, using the convergence of (51) to show first that

$$|\nabla E^i|_{\alpha, s_i} \lessdot \sum_{l=0}^{i} \sigma_l^{-\alpha} \Psi(Q_l)\mu_l \lessdot 1$$

and then
$$|E^{i+1} - E^i|_{\alpha,s_{i+1}} \lessdot |E_{i+1}|_{\alpha,s_{i+1}} \lessdot \Psi(Q_i)\mu_i.$$

Using the convergence of (48) and (52), this shows that E^i converges to a map
$$E^* : \mathbb{T}^n \times \{\omega_0\} \to \mathbb{T}^n \times D_h^{\omega_0}$$

such that
$$|E^*|_{\alpha,s/2} \lessdot \Psi(Q_0)\mu.$$

For the F^i, we do have the expression
$$F^{i+1} = F_{i+1} + (F^i \circ V_{i+1}) \cdot (\mathrm{Id} + F_{i+1})$$

or alternatively
$$F^{i+1} = (\mathrm{Id} + F^i \circ V_{i+1}) \cdot F_{i+1} + F^i \circ V_{i+1}$$

and thus
$$F^{i+1} - F^i = (\mathrm{Id} + F^i \circ V_{i+1}) \cdot F_{i+1} + F^i \circ V_{i+1} - F^i.$$

As before, using the estimates on F_{i+1} and ∇F_{i+1} given in Proposition 5, one shows that
$$|\nabla F^i|_{\alpha,s_i} \lessdot \sum_{l=0}^{i} \sigma_l^{-2\alpha} \Psi(Q_l)\mu_l$$

but however, here, the sum above is not convergent. Yet we do have
$$\sigma_i^\alpha |\nabla F^i|_{\alpha,s_i} \lessdot \sigma_i^\alpha \sum_{l=0}^{i} \sigma_l^{-2\alpha} \Psi(Q_l)\mu_l \lessdot \sum_{l=0}^{i} \sigma_l^{-\alpha} \Psi(Q_l)\mu_l \lessdot 1$$

from (52) and using the fact that the estimate for V_{i+1} can be written as
$$|V_{i+1} - \mathrm{Id}|_{\alpha,s_{i+1}} \lessdot \sigma_i^\alpha \sigma_i^{-\alpha} \Psi(Q_i)\mu_i$$

one obtains
$$|F^i \circ V_{i+1} - F^i|_{\alpha,s_{i+1}} \lessdot \sigma_i^{-\alpha} \Psi(Q_i)\mu_i.$$

By induction, one shows that
$$|F^i|_{\alpha,s_i} \lessdot \sum_{l=0}^{i} \sigma_l^{-\alpha} \Psi(Q_l)\mu_l \lessdot 1$$

from which one obtains
$$|\mathrm{Id} + F^i \circ V_{i+1}|_{\alpha,s_i} \lessdot 1$$

and as a consequence,
$$|F^{i+1} - F^i|_{\alpha,s_{i+1}} \lessdot \sigma_i^{-\alpha} \Psi(Q_i)\mu_i.$$

Using the convergence of (48) and (51), this shows that F^i converges to a map

$$F^* : \mathbb{T}^n \times \{\omega_0\} \to \mathbb{T}^n \times D_h^{\omega_0}$$

such that

$$|F^*|_{\alpha, s/2} \lessdot Q_0 \Psi(Q_0) \mu.$$

For G^i, we have the expression

$$G^{i+1} = (F^i \circ V_{i+1} + \text{Id}) \cdot G_{i+1} + G^i \circ V_{i+1}$$

and thus

$$G^{i+1} - G^i = (F^i \circ V_{i+1} + \text{Id}) \cdot G_{i+1} + G^i \circ V_{i+1} - G^i.$$

Proceeding exactly as we did for E^i and F^i, using the convergence of (48), (51) and (53), one finds that G^i converges to a map

$$G^* : \mathbb{T}^n \times \{\omega_0\} \to \mathbb{T}^n \times D_h^{\omega_0}$$

such that

$$|G^*|_{\alpha, s/2} \lessdot Q_0 \Psi(Q_0) \varepsilon.$$

In summary, the map $\mathcal{F}^i$ converges to a map

$$\mathcal{F}^* : \mathbb{T}^n \times D_{r/2} \times \{\omega_0\} \to \mathbb{T}^n \times D_{r,h}$$

which belongs to $\mathcal{G}$, of the form

$$\begin{cases} \mathcal{F}^*(\theta, I, \omega_0) = (\Phi_{\omega_0}^*(\theta, I), \omega_0^*), \\ \Phi_{\omega_0}^*(\theta, I) = (\theta + E^*(\theta), I + F^*(\theta) \cdot I + G^*(\theta)) \end{cases}$$

with the estimates

$$|E^*|_{\alpha, s/2} \lessdot \Psi(Q_0) \mu, \quad |F^*|_{\alpha, s/2} \lessdot Q_0 \Psi(Q_0) \mu, \quad |G^*|_{\alpha, s/2} \lessdot Q_0 \Psi(Q_0) \varepsilon, \quad |\omega_0^* - \omega_0| \lessdot \mu. \quad (55)$$

Then from the estimates

$$|A^i|_{\alpha, s_i} \leq \varepsilon_i, \quad |B^i|_{\alpha, s_i} \leq \mu_i,$$

given in (46), Proposition 5, and the convergence (48), it follows that both A^i and B^i converge to zero. Next from the estimates

$$\begin{cases} |e^{i+1} - e^i \circ \varphi_{i+1}|_{\alpha, s_{i+1}} \leq |A^i|_{\alpha, s_i}, \\ |\nabla_I^2 R^{i+1} - \nabla_I^2 R^i \circ \mathcal{F}_{i+1}|_{\alpha, s_{i+1}} \lessdot \eta |F_{i+1}|_{\alpha, s_{i+1}} \end{cases}$$

still given in (46), Proposition 5, one can prove in the same way as we did before, that e^i converges to a trivial map

$$e^* : \{\omega_0\} \to D_h^{\omega_0}, \quad e^*(\omega_0) := e_0^*$$

31

such that

$$|e_0^* - e_{\omega_0^*}| \lessdot \varepsilon \qquad (56)$$

whereas M^i converges to a map

$$M^* : \mathbb{T}^n \times D_{r/2} \times \{\omega_0\} \to \mathbb{T}^n \times D_{r,h}$$

such that, setting $R^*(\theta, I) = M^*(\theta, I)I \cdot I$,

$$|\nabla_I^2 R^* - \nabla_I^2 R_{\omega_0^*}|_{\alpha, s/2} \lessdot \eta Q_0 \Psi(Q_0)\mu. \qquad (57)$$

Therefore we have

$$H \circ \mathcal{F}^*(\theta, I, \omega_0) = H_{\omega_0^*} \circ \Phi_{\omega_0}^*(\theta, I) = e_0^* + \omega_0 \cdot I + R^*(\theta, I),$$

which, together with the previous estimates (55), (56) and (57), is what we wanted to prove. $\qquad\square$

7 Proof of Theorem D, following Bessi

The goal of this short section is to show how Theorem D follows directly from the work of Bessi in [Bes00].

In Bessi, one starts with a non-resonant vector $\omega \in \mathbb{R}^n$ which is assumed to be "exponentially Liouville" in the following sense: there exists $s_0 > 0$ and a sequence $k_j \in \mathbb{Z}^n$ with $|k_j| \to +\infty$ as $j \to +\infty$ for which

$$0 < |k_j \cdot \omega| \le e^{-s_0|k_j|}. \qquad (C_{1,s_0})$$

Given this sequence of $k_j \in \mathbb{Z}^n$, one can find another sequence $\tilde{k}_j \in \mathbb{Z}^n$ such that for all $j \in \mathbb{N}$, $|\tilde{k}_j| \le |k_j|$, $\tilde{k}_j \cdot k_j = 0$ and $|\tilde{k}_j \cdot \omega| \ge c|\tilde{k}_j|$ for some constant $c > 0$ independent of j.

Then one defines the following sequence of Hamiltonians on $\mathbb{R}^n/(2\pi\mathbb{Z}^n) \times \mathbb{R}^n$ (which are similar to the Hamiltonian considered by Arnold in [Arn64]):

$$\begin{cases} H_{\varepsilon,\mu}^{1,j}(\theta, I) := \frac{1}{2}I \cdot I + F_{\varepsilon,\mu}^{1,j}(\theta), \ (\theta, I) \in \mathbb{R}^n/(2\pi\mathbb{Z}^n) \times \mathbb{R}^n \\ F_{\varepsilon,\mu}^{1,j}(\theta) := \varepsilon\nu_{1,j,s}(1 - \cos(k_j \cdot \theta))(1 + \mu\tilde{\nu}_{1,j,s}\cos(\tilde{k}_j \cdot \theta)) \\ 0 < \varepsilon \le 1, \ 0 < \mu \le 1, \ \nu_{1,j,s} := e^{-s|k_j|}, \ \tilde{\nu}_{1,j,s} := e^{-s|\tilde{k}_j|}. \end{cases} \qquad (H_{1,j,s})$$

Observe that the only role of the sequences $\nu_{1,j,s}$ and $\tilde{\nu}_{1,j,s}$ is to ensure that the sequence of perturbations $F_{\varepsilon,\mu}^{1,j}$ satisfy, for all $j \in \mathbb{N}$ and all $0 \le \mu \le 1$:

$$|F_{\varepsilon,\mu}^{1,j}|_s := \sup_{\theta\in\mathbb{C}^n/(2\pi\mathbb{Z}^n), \ |\mathrm{Im}(\theta)\le s|} |F_{\varepsilon,\mu}^{1,j}(\theta)| \le 4\varepsilon.$$

In [Bes00], Bessi proved the following theorem.

Theorem 6 (Bessi)**.** *Assume that $\omega \in \mathbb{R}^n$ satisfy (C_{1,s_0}). Then, if $s_0 > s$, for any $0 \leq \varepsilon \leq 1$, there exists $\mu_\varepsilon > 0$ and $j_\varepsilon \in \mathbb{N}$ such that for any $0 < \mu \leq \mu_\varepsilon$ and any $j \geq j_\varepsilon$, the Hamiltonian system defined in $(H_{1,j,s})$ does not have any invariant torus $\mathcal{T}$ satisfying*

(i) $\mathcal{T}$ projects diffeomorphically to $\mathbb{T}^n$;

(i) There is a C^1 diffeomorphism between $\mathbb{T}^n$ and $\mathcal{T}$ which conjugates the flow on $\mathcal{T}$ to the linear flow on $\mathbb{T}^n$ of frequency ω.

It is clear that it is the regularity of the perturbation, here the analyticity which causes the exponential decay of the Fourier coefficients, that forces the condition (C_{1,s_0}). If the perturbation is assumed to be only of class C^r for some $r \in \mathbb{N}$, then (C_{1,s_0}) can be weakened to cover frequencies ω which are Diophantine with an exponent τ which is related to r (this can be obtained from Bessi's work; one can find a better quantitative result in [CW13], which also uses ideas of [Bes00]).

Here we would like to consider the case where the perturbation is α-Gevrey; we will consider a slight modification of the family of Hamiltonians $(H_{1,j,s})$ to a family of Hamiltonians $(H_{\alpha,j,s})$ depending on $\alpha \geq 1$, which are still analytic but for which the perturbation are bounded and small only in a α-Gevrey norm.

First we need to compute the Gevrey of the function $P_k(\theta) := \cos(k \cdot \theta)$ for an arbitrary $k \in \mathbb{Z}^n$. Using the fact that $(l + 1)^2 \leq 4^l$ for any $l \in \mathbb{N}$ we have

$$|P_k|_{\alpha,s} = c \sup_{l \in \mathbb{N}} \frac{(l+1)^2 s^{\alpha l} |k|^l}{l!^\alpha} \leq c \sup_{l \in \mathbb{N}} \frac{(4|k|s^\alpha)^l}{l!^\alpha} \leq c \left(\sup_{l \in \mathbb{N}} \frac{((4|k|)^{1/\alpha} s)^l}{l!} \right)^\alpha \leq c e^{s\alpha(4|k|)^{\frac{1}{\alpha}}}.$$

Now, given $\alpha \geq 1$, we introduce the following condition on the non-resonant vector $\omega \in \mathbb{R}^n$: there exists $s_0 > 0$ and a sequence $k_j \in \mathbb{Z}^n$ with $|k_j| \to +\infty$ as $j \to +\infty$ for which

$$0 < |k_j \cdot \omega| \leq e^{-s_0\alpha(4|k_j|)^{\frac{1}{\alpha}}}. \qquad (C_{\alpha,s_0})$$

For $\alpha = 1$, this condition reduces to (C_{1,s_0}). Then we consider the following modified sequence of Hamiltonians, which once again corresponds exactly to $H_{1,j,s}$ for $\alpha = 1$:

$$\begin{cases} H^{\alpha,j}_{\varepsilon,\mu}(\theta,I) := \frac{1}{2} I \cdot I + F^{\alpha,j}_{\varepsilon,\mu}(\theta), \ (\theta,I) \in \mathbb{R}^n/(2\pi\mathbb{Z}^n) \times \mathbb{R}^n \\ F^{\alpha,j}_{\varepsilon,\mu}(\theta) := \varepsilon\nu_{\alpha,j,s}(1 - \cos(k_j \cdot \theta))(1 + \mu\tilde{\nu}_{\alpha,j,s}\cos(\tilde{k}_j \cdot \theta)) & (H_{\alpha,j,s}) \\ 0 < \varepsilon \leq 1, \ 0 < \mu \leq 1, \ \nu_{\alpha,j,s} := e^{-s\alpha(4|k_j|)^{\frac{1}{\alpha}}}, \ \tilde{\nu}_{\alpha,j,s} := e^{-s\alpha(4|\tilde{k}_j|)^{\frac{1}{\alpha}}}. \end{cases}$$

With these choices of $\nu_{\alpha,j,s}$ and $\tilde{\nu}_{\alpha,j,s}$ we have that, for all $j \in \mathbb{N}$ and all $0 \leq \mu \leq 1$:

$$|F^{\alpha,j}_{\varepsilon,\mu}|_{s,\alpha} \leq C\varepsilon$$

for some constant $C > 1$ independent of ε and μ. The argument of Bessi goes exactly the same of way for this family of Hamiltonians $(H_{\alpha,j,s})$ under the condition (C_{α,s_0}), and thus we have the following statement.

Theorem 7. *Assume that $\omega \in \mathbb{R}^n$ satisfy (C_{α,s_0}). Then, if $s_0 > s$, for any $0 \leq \varepsilon \leq 1$, there exists $\mu_\varepsilon > 0$ and $j_\varepsilon \in \mathbb{N}$ such that for any $0 < \mu \leq \mu_\varepsilon$ and any $j \geq j_\varepsilon$, the Hamiltonian system defined in $(H_{\alpha,j,s})$ does not have any invariant torus $\mathcal{T}$ satisfying*

- *(i) $\mathcal{T}$ projects diffeomorphically to $\mathbb{T}^n$;*

- *(i) There is a C^1 diffeomorphism between $\mathbb{T}^n$ and $\mathcal{T}$ which conjugates the flow on $\mathcal{T}$ to the linear flow on $\mathbb{T}^n$ of frequency ω.*

Now Theorem 7 implies Theorem D, as if ω satisfies (B_α), then it satisfies (C_{α,s_0}) for some $s_0 > 0$ and it is sufficient to consider a Hamiltonian system as in $(H_{\alpha,j,s})$ with $s < s_0$.

A   On the α-Bruno-Rüssmann condition

Recall from Section 2.2 that the arithmetic function on $[1, +\infty)$ associated with a vector $\omega \in \mathbb{R}^n$ is given by

$$\Psi_\omega(Q) = \max \left\{ |k \cdot \omega|^{-1},\ k \in \mathbb{Z}^n,\ 0 < |k| \leq Q \right\}$$

and that we have introduced a continuous version Ψ which satisfies (4). We have further defined Δ by $\Delta(Q) = Q\Psi(Q)$ and its functional inverse Δ^{-1}.

Also recall that, by definition, if $\alpha \geq 1$, A_α consists of those vectors ω for which

$$\int_{\Delta(1)}^{+\infty} \frac{dx}{x(\Delta^{-1}(x))^{\frac{1}{\alpha}}} < \infty, \tag{A_α}$$

whereas BR_α consists of vectors ω satisfying the α-Bruno-Rüssmann condition:

$$\int_1^{+\infty} \frac{\ln(\Psi_\omega(Q))}{Q^{1+\frac{1}{\alpha}}} dQ < \infty \Leftrightarrow \int_1^{+\infty} \frac{\ln(\Psi(Q))}{Q^{1+\frac{1}{\alpha}}} dQ < \infty. \tag{BR_α}$$

In the proof of Theorem H, we use the following fact.

Proposition 8. *For any $\alpha \geq 1$, $A_\alpha = BR_\alpha$.*

Proof. We aim at showing that the two integrals

$$\int^{+\infty} \frac{dx}{x(\Delta^{-1}(x))^{\frac{1}{\alpha}}} \quad \text{and} \quad \int^{+\infty} \frac{\ln(\Psi(Q))}{Q^{1+\frac{1}{\alpha}}} dQ$$

converge or diverge simultaneously. Equivalently, choosing for example $t = \Delta^{-1}(1)$ and, in the first integral, making the change of variable $x = \Delta(Q)$, we may compare the following two quantities:

$$a_\alpha = \int_t^\infty \frac{d\Delta(Q)}{Q^{\frac{1}{\alpha}}\Delta(Q)} \quad \text{and} \quad b_\alpha = \int_t^\infty \frac{\ln \Delta(Q)}{Q^{1+\frac{1}{\alpha}}} dQ.$$

A (Riemann-Stieltjes) integration by part shows that, if $T > 1$,

$$\int_t^T \frac{d\Delta(Q)}{Q^{\frac{1}{\alpha}}\Delta(Q)} = \frac{\ln \Delta(T)}{T^{\frac{1}{\alpha}}} + \frac{1}{\alpha}\int_t^T \frac{\ln \Delta(Q)}{Q^{1+\frac{1}{\alpha}}}\,dQ. \tag{58}$$

On the one hand, the two integrals in this equality have a (possibly infinite) limit as T tends to $+\infty$, and $\frac{\ln \Delta(T)}{T^{\frac{1}{\alpha}}} \geq 0$, thus (58) yields, as T tends to infinity,

$$a_\alpha \geq \frac{b_\alpha}{\alpha}.$$

On the other hand, since Δ is increasing,

$$\frac{\ln \Delta(T)}{T^{\frac{1}{\alpha}}} = \frac{\ln \Delta(T)}{\alpha}\int_T^{+\infty} \frac{dQ}{Q^{1+\frac{1}{\alpha}}} \leq \frac{1}{\alpha}\int_T^{+\infty} \frac{\ln \Delta(Q)}{Q^{1+\frac{1}{\alpha}}}\,dQ$$

so, letting T tend to $+\infty$, (58) entails

$$a_\alpha \leq \frac{b_\alpha}{\alpha}.$$

So, $a_\alpha = \frac{b_\alpha}{\alpha}$, whence the conclusion. $\qquad\square$

Remark 1. From the proof above, one easily see that if $\omega \in \mathrm{BR}_\alpha$, then

$$\lim_{Q \to +\infty} \frac{\ln(\Delta(Q))}{Q^{1/\alpha}} = 0.$$

But $\ln(\Delta(Q)) = \ln(Q\Psi(Q)) = \ln Q + \ln(\Psi(Q))$ and therefore

$$\lim_{Q \to +\infty} \frac{\ln(\Psi(Q))}{Q^{1/\alpha}} = 0$$

which means that ω satisfies R_α.

We refer to [BF13, Proposition 2.2] for related, more precise results. In the next lemma, we give alternative characterizations of the α-BR condition, so as to facilitate comparisons with other arithmetic conditions.

Lemma 9. *Let $\alpha \geq 1$, $\beta = 1 + \frac{1}{\alpha}$ and $\omega \in \mathbb{R}^n$ non-resonant. The following conditions are equivalent to each other:*

1. $\int_1^{+\infty} \frac{\ln \Psi_\omega(Q)}{Q^\beta}\,dQ < \infty$

2. $\sum_{Q \geq 1} \frac{\ln \Psi_\omega(Q)}{Q^\beta} < \infty$

3. $\sum_{l \geq 0} \frac{\ln \Psi_\omega(2^l)}{2^{l/\alpha}} < \infty$

4. *if $n = 2$ and $\omega = (1, \nu)$, $\sum_{k \geq 1} \frac{\ln q_{k+1}}{q_k^{1/\alpha}} < \infty$, where the q_k's are the main denominators of the continued fraction of ν.*

In the case $\alpha = 1$, the equivalence $(1 \Leftrightarrow 2)$ is proved in [Rüs01], whereas $(2 \Leftrightarrow 3 \Leftrightarrow 4)$ is proved in in [GM10].

Proof. $(1 \Leftrightarrow 2)$ As already noticed in Section 2.2, Ψ_ω is constant on intervals of the form $[Q, Q + 1)$, $Q \in \mathbb{N}_*$. So,

$$\int_1^\infty \frac{\ln \Psi_\omega(Q)}{Q^\beta} \, dQ = \sum_{Q=1}^{+\infty} \ln \Psi_\omega(Q) \int_Q^{Q+1} q^{-\beta} \, dq = \alpha \sum_{Q \geq 1} \ln \Psi_\omega(Q) \left(Q^{-1/\alpha} - (Q+1)^{-1/\alpha} \right).$$

Since the general term of this series is positive and $Q^{-1/\alpha} - (Q+1)^{-1/\alpha} \sim_{Q \to +\infty} Q^{-\beta}$, the first two conditions are equivalent to one another.

$(2 \Leftrightarrow 3)$ We want to show that $f_2 = \sum_{Q \geq 1} \frac{\ln \Psi_\omega(Q)}{Q^\beta}$ and $f_3 = \sum_{l \geq 0} \frac{\ln \Psi_\omega(2^l)}{2^{l/\alpha}}$ are simultaneously finite or infinite. Since Ψ_ω is non-decreasing,

$$f_2 \geq \sum_{l \geq 0} \ln \Psi_\omega(2^l) \sum_{2^l \leq Q \leq 2^{l+1}-1} \frac{1}{Q^\beta},$$

where

$$\sum_{2^l \leq Q \leq 2^{l+1}-1} \frac{1}{Q^\beta} \geq \int_{2^l}^{2^{l+1}} \frac{dQ}{Q^\beta} \geq \frac{\alpha}{2} \frac{1}{2^{l/\alpha}},$$

so that

$$f_2 \geq \frac{\alpha}{2} f_3.$$

Conversely, again using the fact that Ψ_ω is non-decreasing,

$$f_2 \leq \sum_{l \geq 0} \ln \Psi_\omega(2^{l+1}) \sum_{2^l \leq Q \leq 2^{l+1}-1} \frac{1}{Q^\beta},$$

where

$$\sum_{2^l \leq Q \leq 2^{l+1}-1} \frac{1}{Q^\beta} \leq \frac{1}{2^{l\beta}} + \int_{2^l}^{2^{l+1}-1} \frac{dQ}{Q^\beta} \leq \frac{1}{2^{l/\alpha}} \left(\frac{1}{2^l} + \alpha \right) \leq \frac{1+\alpha}{2^{l/\alpha}} = \frac{2^{1/\alpha}(1+\alpha)}{2^{(l+1)/\alpha}}$$

so that

$$f_2 \leq 2^{1/\alpha}(1+\alpha) f_3.$$

$(2 \Leftrightarrow 4)$ Now assume $n = 1$ and $\omega = (1, \nu)$, $\nu \in \mathbb{R}$. Let (p_k/q_k) be the sequence of convergents of ν: one has then the following explicit expression

$$\Psi_\omega(Q) = |q_k \nu - p_k|^{-1}, \quad q_k \leq Q < q_{k+1}$$

36

as was proved in [BF13], Proposition 2.7. Hence,

$$S := \sum_{Q \geq 1} \frac{\ln(\Psi_\omega(Q))}{Q^\beta} = \sum_{k \geq 1} \left(\ln |q_k \nu - p_k|^{-1} \sum_{Q=q_k}^{q_{k+1}-1} \frac{1}{Q^\beta} \right).$$

Bounding the first factor by

$$\ln q_{k+1} \leq \ln |q_k \nu - p_k|^{-1} \leq \ln(2q_{k+1})$$

and the second one by

$$\frac{\alpha}{q_k^{1/\alpha}} - \frac{\alpha}{q_{k+1}^{1/\alpha}} \leq \sum_{Q=q_k}^{q_{k+1}-1} \frac{1}{Q^\beta} \leq \frac{1+\alpha}{q_k^{1/\alpha}}$$

(obtained by comparing the series with an integral) yields

$$\alpha \sum_{k \geq 1} \left(\frac{\ln q_{k+1}}{q_k^{1/\alpha}} - \frac{\ln q_{k+1}}{q_{k+1}^{1/\alpha}} \right) \leq S \leq (1+\alpha) \sum_{k \geq 1} \frac{\ln(2q_{k+1})}{q_k^{1/\alpha}} = (1+\alpha) \sum_{k \geq 1} \left(\frac{\ln 2}{q_k^{1/\alpha}} + \frac{\ln(q_{k+1})}{q_k^{1/\alpha}} \right).$$

Since the sequence $(q_k)_{k \geq 0}$ increases at least geometrically, namely $q_k \geq 2^{k/2}$ for any $k \geq 0$, we have

$$\sum \frac{\ln q_{k+1}}{q_{k+1}^{1/\alpha}} < \infty, \quad \sum \frac{1}{q_k^{1/\alpha}} < \infty$$

and therefore

$$S < \infty \Leftrightarrow \sum \frac{\ln q_{k+1}}{q_k^{1/\alpha}} < \infty.$$

$\square$

Example 10. Let $\nu = \sum_{n \geq 1} 10^{-n!}$. Recall from [Khi63, Theorem 30] that if $p/q \in \mathbb{Q}$ satisfies

$$|p - q\nu| < \frac{1}{2q},$$

then p/q is a best approximation of ν. So, we see that $\sum_{1 \leq n \leq Q} 10^{-n!}$ are best approximations, and thus $q_n = 2^{n!}$. The 4th criterion of Lemma 9 thus tells us that $\omega = (1, \nu) \in \mathrm{BR}_\alpha$ for all $\alpha \geq 1$, while ω is not Diophantine.

B   Gevrey estimates

Let us start by recalling some notations and definitions. Given an integer $m \geq 1$ and $k = (k_1, \ldots, k_m) \in \mathbb{N}^m$, we define

$$|k| = \sum_{i=1}^m k_i, \quad k! = \prod_{i=1}^m k_i!.$$

Given $x \in \mathbb{R}^m$, we set

$$x^k = \prod_{i=1}^{m} x_i^{k_i}.$$

Let K be a compact set of the form

$$K = \mathbb{T}^{m_1} \times \bar{B}^{m_2}, \quad m_1 + m_2 = m,$$

where $\bar{B}^{m_2}$ is the closure of an open subset B^{m_2} of $\mathbb{R}^{m_2}$. Let $f : K \to \mathbb{R}$ be a smooth function, meaning that f extends smoothly to an open neighborhood of K. Such an extension is by no means unique, but note that, by continuity, the partial derivatives of f over K, at any order, do not depend on the extension. For $a \in K$ and $k \in \mathbb{N}^m$ we set

$$\partial^k f(a) = \partial_{x_1}^{k_1} \cdots \partial_{x_m}^{k_m} f(a). \tag{59}$$

Given real numbers $\alpha \geq 1$ and $s > 0$, the function f is said to be (α, s)-Gevrey if

$$|f|_{\alpha,s} := c \sup_{a \in K} \left(\sup_{k \in \mathbb{N}^m} \frac{(|k|+1)^2 s^{\alpha|k|} |\partial^k f(a)|}{|k|!^\alpha} \right) < \infty, \quad c := 4\pi^2/3. \tag{60}$$

The space of such functions will be denoted by $G_{\alpha,s}(K)$, and equipped with the above norm, it is a Banach space. Our definition of Gevrey norm is not quite standard, but up to decreasing or increasing the parameter s, it is comparable to the Gevrey norms that have been used in Hamiltonian perturbation theory (as for instance in [MS02] or in [Pop04]). On the one hand, the role of the factor $(|k|+1)^2$ is to simplify the estimates for the product and composition of Gevrey functions (see respectively Lemmas 17 and 19). On the other hand, the role of the normalizing constant $c > 0$ in the definition is to ensure that $G_{\alpha,s}(K)$ is a Banach algebra (see Lemma 17).

The above definition can be extended to vector-valued functions $f = (f_i)_{1 \leq i \leq p} : K \to \mathbb{R}^p$ for $p \geq 1$ by setting

$$|\partial^k f(a)| := \sum_{1 \leq i \leq p} |\partial^k f_i(a)|, \quad a \in K$$

in (60). The space of such vector-valued functions is still a Banach space with the above norm, and it will be denoted by $G_{\alpha,s}(K, \mathbb{R}^p)$. The case of matrix-valued functions, say with values in the space $M_{m,p}(\mathbb{R})$ of matrix with m rows and p columns, is reduced to the case of vector-valued functions by simply identifying $M_{m,p}(\mathbb{R})$ to $\mathbb{R}^{mp}$.

B.1 Majorant series and Gevrey functions

The definition of Gevrey functions can be conveniently reformulated in terms of majorant series with one variable (see [Kom79], [Kom80] and also [SCK03]).

But first let us consider a formal power series in m variables $X = (X_1, \ldots, X_m)$ with coefficients in a normed real vector space $(E, |\,.\,|_E)$, which is a formal sum of the form

$$A(X) = \sum_{k \in \mathbb{N}^m} A_k X^k, \quad A_k = A_{k_1, \ldots, k_m} \in E.$$

Such a formal series is said to be majorized by another formal power series with real non-negative coefficients

$$B(X) = \sum_{k \in \mathbb{N}^m} B_k X^k, \quad B_k = B_{k_1, \ldots, k_m} \in \mathbb{R},$$

and we write $A \ll B$, if

$$|A_k|_E \leq B_k, \quad \forall\, k \in \mathbb{Z}^m. \tag{61}$$

Next, following [SCK03], we introduce a notion of a smooth function being majorized by a formal power series in one variable. So let $f : K \to \mathbb{R}^p$ be a smooth function, and F be a formal power series in one variable with non-negative coefficients, that we shall write as

$$F(X) = \sum_{l=0}^{+\infty} \frac{F_l}{l!} X^l.$$

We will say that f is majorized by F on K, and we will write $f \ll_K F$ (or $f(x) \ll_K F(X)$), if for all $a \in K$ and all $k \in \mathbb{N}^m$, we have

$$|\partial^k f(a)| \leq F_{|k|}. \tag{62}$$

To better understand this definition, recall that given $f : K \to \mathbb{R}^p$ and $a \in K$, we can define its formal Taylor series at a by

$$T_a f(X) := \sum_{k \in \mathbb{N}^m} \frac{\partial^k f(a)}{k!} X^k,$$

which is a formal power series in m variables that takes values in $\mathbb{R}^p$. To a formal series F in one variable, one can associate a formal series $\hat{F}$ in m variables simply by setting

$$\hat{F}(X_1, \ldots, X_m) := F(X_1 + \cdots + X_m).$$

If is then easy to check that $f \ll_K F$, in the sense of (62), if and only if for all $a \in K$, $T_a f \ll \hat{F}$, in the sense of (61) (with $E = \mathbb{R}^p$ and $|.|_E$ the norm given by the sum of the absolute values of the components).

Now, given $\alpha \geq 1$ and $s > 0$, let us define the following formal power series in one variable

$$M_{\alpha,s}(X) := c^{-1} \sum_{l=0}^{+\infty} \frac{l!^{\alpha-1}}{(l+1)^2} \left(\frac{X}{s^\alpha}\right)^l = c^{-1} \sum_{l=0}^{+\infty} \frac{M_l}{l!} X^l, \quad M_l = \frac{l!^\alpha}{(l+1)^2 s^{\alpha l}}, \quad c = 4\pi^2/3. \tag{63}$$

The following characterization of Gevrey functions is evident from the definitions (60) and (62).

Proposition 11. *If $f : K \to \mathbb{R}^p$ is a smooth function,*

$$|f|_{\alpha,s} = \inf \left\{ C \in [0, +\infty] \,\middle|\, f \ll_K C M_{\alpha,s} \right\}.$$

Henceforth $\alpha \geq 1$ will be fixed, so in the sequel we will simply write $M_{\alpha,s} = M_s$.

B.2 Properties of majorant series

We collect here some properties of majorant series that will be used later on. It is clear how to define the derivatives of a formal power series in one variable, and also a linear combination and the product of two such formal power series. We then have the following lemma.

Lemma 12. *Let* $f, g : K \to \mathbb{R}^p$ *be smooth functions,* F, G *be two formal power series in one variable, and assume that*

$$f \ll_K F, \quad g \ll_K G.$$

Then

$$\partial^k f \ll_K \partial^{|k|} F, \quad k \in \mathbb{N}^m, \tag{64}$$

$$\lambda f + \mu g \ll_K |\lambda| F + |\mu| G, \quad \lambda \in \mathbb{R}, \ \mu \in \mathbb{R}. \tag{65}$$

Moreover, if we define $f \cdot g : K \to \mathbb{R}$ *by*

$$f \cdot g = \sum_{i=1}^{p} f_i g_i,$$

then

$$f \cdot g \ll_K FG. \tag{66}$$

For a proof, we refer to [SCK03], Lemma 2.2, in which the case $p = 1$ is considered; but the general case $p \geq 1$ is entirely similar.

Given two formal power series in one variable F and G, we define the composition $F \odot G$ of F and G by

$$F \odot G(X) := \sum_{l=0}^{+\infty} \frac{F_l}{l!} \left(G(X) - G(0) \right)^l.$$

Lemma 13. *Let* $f : K \to \mathbb{R}^p$ *be a smooth function,* $g : L \to \mathbb{T}^{m_1} \times \mathbb{R}^{m_2}$ *another smooth function such that* $g(L) \subseteq K$, *and assume that*

$$f \ll_K F, \quad g \ll_L G.$$

Then

$$f \circ g \ll_L F \odot G.$$

Once again, for a proof we refer to [SCK03], Lemma 2.3.

B.3 Derivatives

In this section, we will show that the derivatives of a Gevrey function are still Gevrey, at the expense of reducing the parameter $s > 0$; these are analogues of Cauchy estimates for analytic functions.

Proposition 14. *Let $f \in G_{\alpha,s}(K, \mathbb{R}^p)$, and $0 < \sigma < s$. Then for any $k \in \mathbb{N}^m$, $\partial^k f \in G_{\alpha,s-\sigma}(K, \mathbb{R}^p)$ and we have*

$$|\partial^k f|_{\alpha,s-\sigma} \leq \left(\frac{|k|^\alpha}{\sigma^\alpha} \right)^{|k|} |f|_{\alpha,s}.$$

Proof. It is enough to prove the case $|k| = 1$, as the general case follows by an easy induction. From Proposition 11 and (64) of Lemma 12, to prove the case $|k| = 1$ it is sufficient to prove that

$$\partial^1 M_s \ll \sigma^{-\alpha} M_{s-\sigma} \tag{67}$$

where M_s is the formal power series defined in (63). We have

$$\partial^1 M_s(X) = c^{-1} \sum_{l=1}^{+\infty} \frac{l!^\alpha}{(l+1)^2 s^{l\alpha}(l-1)!} X^{l-1} = c^{-1} \sum_{l=0}^{+\infty} \frac{(l+1)!^\alpha}{(l+2)^2 s^{(l+1)\alpha} l!} X^l$$

and hence (67) is true if, for all $l \in \mathbb{N}$,

$$\left(\sigma s^{-1} \right)^\alpha \left(1 - \sigma s^{-1} \right)^{l\alpha} (l+1)^\alpha \leq 1. \tag{68}$$

Since $0 < \sigma s^{-1} < 1$, we have $\ln(1 - \sigma s^{-1}) \leq -\sigma s^{-1}$ and thus

$$\left(1 - \sigma s^{-1} \right)^{l\alpha} (l+1)^\alpha = e^{l\alpha \ln(1-\sigma s^{-1})}(l+1)^\alpha \leq e^{-l\alpha \sigma s^{-1}}(l+1)^\alpha.$$

Let $\lambda = \sigma s^{-1}$, and consider the function $u(x) = (x+1)^\alpha e^{-\alpha\lambda x}$ for $x \geq 0$. This function reaches its maximum at $x = \lambda^{-1}(1-\lambda)$, the value of which is

$$u\left(\lambda^{-1}(1-\lambda) \right) = \lambda^{-\alpha} e^{(\lambda-1)\alpha} \leq \lambda^{-\alpha}.$$

Therefore, for all $l \in \mathbb{N}$, we have

$$\left(1 - \sigma s^{-1} \right)^{l\alpha} (l+1)^\alpha \leq \lambda^{-\alpha} = \left(\sigma s^{-1} \right)^{-\alpha}$$

which is exactly the inequality (68) we wanted to prove. $\qquad \square$

For $f : K \to \mathbb{R}$, let $\nabla f : K \to \mathbb{R}^m$ be the vector-valued function formed by the partial derivatives of f of order one, and more generally, for $f : K \to \mathbb{R}^p$, we let $\nabla f : K \to M_{m,p}(R) \simeq \mathbb{R}^{mp}$ be the matrix-valued function whose columns are given by ∇f_i where $f = (f_i)_{1 \leq i \leq p}$. Then we have the following obvious corollary of Proposition 14.

Corollary 15. *Let $0 < \sigma < s$. If $f \in G_{\alpha,s}(K, \mathbb{R})$, then $\nabla f \in G_{\alpha,s-\sigma}(K, \mathbb{R}^m)$ and*

$$|\nabla f|_{\alpha,s-\sigma} \leq m\sigma^{-a} |f|_{\alpha,s}$$

and if $f \in G_{\alpha,s}(K, \mathbb{R}^p)$, then $\nabla f \in G_{\alpha,s-\sigma}(K, \mathbb{R}^{mp})$ and

$$|\nabla f|_{\alpha,s-\sigma} \leq mp\sigma^{-a} |f|_{\alpha,s}.$$

41

B.4   Products

In this section, we shall prove that the product of Gevrey functions is still a Gevrey function.

Proposition 16. *Let* $f, g \in G_{\alpha,s}(K, \mathbb{R}^p)$. *Then* $f \cdot g \in G_{\alpha,s}(K, \mathbb{R})$ *and we have*

$$|f \cdot g|_{\alpha,s} \leq |f|_{\alpha,s}|g|_{\alpha,s}.$$

Once again, in view of Proposition 11 and (66) of Proposition 12, Proposition 16 is a direct consequence of the following lemma.

Lemma 17. *We have*

$$M_s^2 \ll M_s.$$

The proof given below follows [Lax53]. It is this lemma that motivates the introduction of the normalizing constant in M_s (and thus in the Gevrey norm); without this constant one would have $M_s^2 \ll c^2 M_s$. Let us point out that the proof given below is elementary thanks to the factor $(|k| + 1)^2$ in the definition of M_s; without this factor, the statement is true (with a different normalizing constant) but the proof is more involved (see Lemma 2.7 of [SCK03]).

Proof. Recall that

$$M_s(X) = c^{-1} \sum_{l=0}^{+\infty} \frac{M_l}{l!} X^l, \quad M_l = \frac{l!^{\alpha}}{(l+1)^2 s^{\alpha l}}$$

and so the assertion of the lemma amounts to prove that for all $l \in \mathbb{N}$,

$$\sum_{j=0}^{l} \frac{(j!)^{\alpha-1}((l-j)!)^{\alpha-1}}{(j+1)^2(l-j+1)^2} \leq c\frac{(l!)^{\alpha-1}}{(l+1)^2}, \quad c = \frac{4\pi^2}{3}. \tag{69}$$

Observe that the sum in the left-hand side of (69) is symmetric with respect to $j \mapsto l - j$, and that since $\alpha - 1 \geq 0$, $(j!)^{\alpha-1}((l-j)!)^{\alpha-1} \leq (l!)^{\alpha-1}$ for all $l \in \mathbb{N}$. Hence,

$$\sum_{j=0}^{l} \frac{(j!)^{\alpha-1}((l-j)!)^{\alpha-1}}{(j+1)^2(l-j+1)^2} \leq 2\sum_{j=0}^{l/2} \frac{(j!)^{\alpha-1}((l-j)!)^{\alpha-1}}{(j+1)^2(l-j+1)^2} \leq 2(l!)^{\alpha-1}\sum_{j=0}^{l/2} \frac{1}{(j+1)^2(l-j+1)^2}.$$

Then for any $0 \leq j \leq l/2$, $(l-j+1)^2 \geq (l/2+1)^2 \geq (l+1)^2/4$, and therefore

$$\sum_{j=0}^{l} \frac{(j!)^{\alpha-1}((l-j)!)^{\alpha-1}}{(j+1)^2(l-j+1)^2} \leq \frac{8(l!)^{\alpha-1}}{(l+1)^2}\sum_{j=0}^{l/2} \frac{1}{(j+1)^2} \leq \frac{8(l!)^{\alpha-1}}{(l+1)^2}\sum_{j=0}^{+\infty} \frac{1}{(j+1)^2} = \frac{4\pi^2}{3}\frac{(l!)^{\alpha-1}}{(l+1)^2}$$

which is the inequality we wanted to prove.   □

Proposition 16 can be extended to matrix-valued functions. More precisely, given $f : K \to M_{m,p}(\mathbb{R})$ and $g : K \to M_{p,q}(\mathbb{R})$ where $f = (f_{i,j})_{1 \le i \le m, \, 1 \le j \le p}$ and $g = (g_{j,k})_{1 \le j \le m, \, 1 \le k \le p}$, we define $f \cdot g : K \to M_{m,q}(\mathbb{R})$ by $f \cdot g := ((f \cdot g)_{i,k})_{1 \le i \le m, \, 1 \le k \le q}$ where

$$(f \cdot g)_{i,k} := \sum_{j=1}^{p} f_{i,j} g_{j,k}.$$

Then the following statement is an obvious corollary of Proposition 16.

Corollary 18. *Let $f \in G_{\alpha,s}(K, M_{m,p}(\mathbb{R}))$, $g \in G_{\alpha,s}(K, M_{p,q}(\mathbb{R}))$. Then $f \cdot g \in G_{\alpha,s}(K, M_{m,q}(\mathbb{R}))$ and we have*

$$|f \cdot g|_{\alpha,s} \le |f|_{\alpha,s} |g|_{\alpha,s}.$$

B.5 Compositions

Our goal here is to prove that the composition of Gevrey functions are still Gevrey. But first we need to define two additional formal power series associated to M_s, the latter being defined in (63). So we define $\bar{M}$ by

$$\bar{M}_s(X) := M_s(X) - M_s(0) = c^{-1} \sum_{l=1}^{+\infty} \frac{l!^{\alpha-1}}{(l+1)^2} \left(\frac{X}{s^\alpha}\right)^l \tag{70}$$

and $\tilde{M}$ by

$$\tilde{M}_s(X) := c^{-1} \sum_{l=0}^{+\infty} \frac{(l+1)!^{\alpha-1}}{(l+2)^2} \left(\frac{X}{s^\alpha}\right)^l. \tag{71}$$

It is clear that

$$s^{-\alpha} X \tilde{M}_s(X) = \bar{M}_s(X). \tag{72}$$

Lemma 19. *We have*

$$\tilde{M}_s^2 \ll \tilde{M}_s, \quad \tilde{M}_s \bar{M}_s \ll \bar{M}_s.$$

As for Lemma 17, the factor $(|k|+1)^2$ in the definition of M_s makes the proof simple, but the statement is still true with this factor (see Lemma 2.4 of [SCK03]).

Proof. It is enough to prove the first part of the statement, as the second part of the statement follows from it; indeed, if $\tilde{M}_s^2 \ll \tilde{M}_s$, then using (72) we have

$$\tilde{M}_s(X) \bar{M}_s(X) = s^{-\alpha} X \tilde{M}_s(X) \tilde{M}_s(X) \ll s^{-\alpha} X \tilde{M}_s(X) = \bar{M}_s(X).$$

As in Lemma 17, to prove that $\tilde{M}_s^2 \ll \tilde{M}_s$ one needs to show

$$\sum_{j=0}^{l} \frac{((j+1)!)^{\alpha-1}((l-j+1)!)^{\alpha-1}}{(j+2)^2(l-j+2)^2} \le c \frac{((l+1)!)^{\alpha-1}}{(l+2)^2}, \quad c = \frac{4\pi^2}{3}. \tag{73}$$

The sum in the left-hand side of (73) is still symmetric with respect to $j \mapsto l - j$, and therefore

$$\sum_{j=0}^{l} \frac{((j+1)!)^{\alpha-1}((l-j+1)!)^{\alpha-1}}{(j+2)^2(l-j+2)^2} \leq 2 \sum_{j=0}^{l/2} \frac{((j+1)!)^{\alpha-1}((l-j+1)!)^{\alpha-1}}{(j+2)^2(l-j+2)^2}.$$

Then, for any $l \in \mathbb{N}$ and any $0 \leq j \leq l/2$, since $\alpha - 1 \geq 0$ we have

$$((j+1)!)^{\alpha-1}((l-j+1)!)^{\alpha-1} \leq ((l+1)!)^{\alpha-1}$$

as one may easily check. Then, as in Lemma 17, for any $0 \leq j \leq l/2$, $(l-j+2)^2 \geq (l/2+2)^2 \geq (l+2)^2/4$, and therefore

$$\sum_{j=0}^{l} \frac{((j+1)!)^{\alpha-1}((l-j+1)!)^{\alpha-1}}{(j+2)^2(l-j+2)^2} \leq \frac{8((l+1)!)^{\alpha-1}}{(l+1)^2} \sum_{j=0}^{l/2} \frac{1}{(j+2)^2} \leq \frac{4\pi^2}{3} \frac{((l+1)!)^{\alpha-1}}{(l+2)^2}$$

and this concludes the proof. □

Proposition 20. *Let $f \in G_{\alpha,s}(K, \mathbb{R}^p)$, $0 < \sigma < s$, and $g \in G_{\alpha,s-\sigma}(L, \mathbb{T}^{m_1} \times \mathbb{R}^{m_2})$ such that $g(L) \subseteq K$. Assume that $g = \mathrm{Id} + u$ with*

$$|u|_{\alpha,s-\sigma} \leq \sigma^\alpha. \tag{74}$$

Then $f \circ g \in G_{\alpha,s-\sigma}(L, \mathbb{R}^p)$ and

$$|f \circ g|_{\alpha,s-\sigma} \leq |f|_{\alpha,s}. \tag{75}$$

When f et g are analytic, the analogue of estimate (75) with the supremum norm is obvious, for the obvious reason that the supremum of a function is a non-increasing function of the domain.

As it will be clear in the proof, the conclusions of Proposition 20 holds true under the slightly weaker assumption that

$$|u|_{\alpha,s-\sigma} - \sup_{a \in K} |u(a)| \leq s^\alpha - (s-\sigma)^\alpha$$

but this will not be needed.

Proof. Let

$$a := |f|_{\alpha,s}, \quad b := |u|_{\alpha,s-\sigma}$$

so that, from Proposition 11,

$$f(x) \ll_K a M_s(X), \quad u(x) \ll_L b M_{s-\sigma}(X)$$

and consequently

$$f(x) \ll_K a M_s(X), \quad g(x) \ll_L X + b M_{s-\sigma}(X).$$

We now apply Lemma 13 and, recalling the definition of $\bar{M}_s$ and $\tilde{M}_s$ given respectively in (70) and (71), we obtain

$$
\begin{aligned}
f(g(x)) \;&\ll_L\; aM_s\left(X + b\bar{M}_{s-\sigma}(X)\right) \\
&=\; aM_s(0) + a\bar{M}_s\left(X + b\bar{M}_{s-\sigma}(X)\right) \\
&=\; aM_s(0) + a\bar{M}_s\left(X + b(s-\sigma)^{-\alpha}X\tilde{M}_{s-\sigma}(X)\right) \\
&=\; aM_s(0) + a\sum_{l=1}^{+\infty}\frac{l!^{\alpha-1}}{(l+1)^2}\left(\frac{X + b(s-\sigma)^{-\alpha}X\tilde{M}_{s-\sigma}(X)}{s^\alpha}\right)^l \\
&=\; aM_s(0) + a\sum_{l=1}^{+\infty}\frac{l!^{\alpha-1}}{(l+1)^2}\left(\frac{X}{s^\alpha}\right)^l\left(1 + b(s-\sigma)^{-\alpha}\tilde{M}_{s-\sigma}(X)\right)^l. \quad (76)
\end{aligned}
$$

From the first part of Lemma 19, for any $j \in \mathbb{N}$, we have

$$
\tilde{M}_{s-\sigma}^j \ll \tilde{M}_{s-\sigma}
$$

and therefore

$$
\begin{aligned}
\left(1 + b(s-\sigma)^{-\alpha}\tilde{M}_{s-\sigma}(X)\right)^l \;&=\; \sum_{j=0}^{l}\binom{l}{j}b^j(s-\sigma)^{-j\alpha}\tilde{M}_{s-\sigma}(X)^j \\
&\ll\; \tilde{M}_{s-\sigma}(X)\sum_{j=0}^{l}\binom{l}{j}b^j(s-\sigma)^{-j\alpha} \\
&=\; \tilde{M}_{s-\sigma}(X)\left(1 + b(s-\sigma)^{-\alpha}\right)^l.
\end{aligned}
$$

Now, from (74) we get

$$
b = |u|_{\alpha,s-\sigma} \le \sigma^\alpha \le s^\alpha - (s-\sigma)^\alpha
$$

and thus

$$
1 + b(s-\sigma)^{-\alpha} \le s^\alpha(s-\sigma)^{-\alpha}.
$$

This gives

$$
\left(1 + b(s-\sigma)^{-\alpha}\tilde{M}_{s-\sigma}(X)\right)^l \ll \tilde{M}_{s-\sigma}(X)(s^\alpha(s-\sigma)^{-\alpha})^l
$$

which, together with (76), yields

$$
f(g(x)) \ll_L aM_s(0) + a\tilde{M}_{s-\sigma}(X)\sum_{l=1}^{+\infty}\frac{l!^{\alpha-1}}{(l+1)^2}\left(\frac{X}{(s-\sigma)^\alpha}\right)^l = aM_s(0) + a\tilde{M}_{s-\sigma}(X)\bar{M}_{s-\sigma}(X).
$$

Using the second part of Lemma 19, this gives

$$
f(g(x)) \ll_L aM_s(0) + a\bar{M}_{s-\sigma}(X)
$$

45

and since $M_s(0) = M_{s-\sigma}(0)$, we arrive at

$$f(g(x)) \ll_L a(M_{s-\sigma}(0) + \bar{M}_{s-\sigma}(X)) = aM_{s-\sigma}(X).$$

Using Proposition 11, we eventually obtain

$$|f \circ g|_{\alpha,s-\sigma} \le a = |f|_{\alpha,s}$$

and this concludes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

B.6   Flows

In this section and the next one, we shall state and prove some estimates adapted to the situation considered in §6: that is we consider functions $H = H(\theta, I, \omega)$ which are defined and Gevrey smooth on a domain of the form

$$\mathbb{T}^n \times D_{r,h} = \mathbb{T}^n \times D_r \times D_h^{\omega_0} \subseteq \mathbb{T}^n \times \mathbb{R}^n \times \mathbb{R}^n$$

where D_r is the ball of radius $r > 0$ centered at the origin and $D_h^{\omega_0}$ is an arbitrary ball of radius $h > 0$. In the lemma and proposition below, the variables $\omega \in D_h^{\omega_0}$ play the role of a fixed parameter, hence to simplify the notations we will explicitly suppress the dependence on $\omega \in D_h^{\omega_0}$.

Moreover, throughout this section and the next one, for simplicity we shall write $u \lessdot v$ (respectively $u \mathrel{\cdot\!\!<} v$), if, for some constant $C \ge 1$ which depends only on n and α and could be made explicit, we have $u \le Cv$ (respectively $Cu \le v$).

Let us first start with a vector-valued function $D : \mathbb{T}^n \to \mathbb{R}^n$ which depends only on $\theta \in \mathbb{T}^n$, and that we shall considered as a vector field on $\mathbb{T}^n$.

Lemma 21. *Given $D \in G_{\alpha,s}(\mathbb{T}^n, \mathbb{R}^n)$, let $0 < \sigma < s$ and assume that*

$$|D|_{\alpha,s} \mathrel{\cdot\!\!<} \sigma^\alpha. \tag{77}$$

Then for any $t \in [0,1]$, the time-t map D^t of the flow of D belongs to $G_{\alpha,s-\sigma}(\mathbb{T}^n, \mathbb{T}^n)$ and we have the estimate

$$|D^t - \mathrm{Id}|_{\alpha,s-\sigma} \le |D|_{\alpha,s}. \tag{78}$$

The proof of the above lemma is a variant of the proof of Lemma B.3 in [LMS16].

Proof. The fact that D^t is smooth and defined for all $t \in [0,1]$ (in fact, for all $t \in \mathbb{R}$) follows from the compactness of $\mathbb{T}^n$ and the classical result on the existence and uniqueness of solutions of differential equations (even though this will essentially be re-proved below); the only thing we need to prove is the estimate (78). So let us consider the space $V := C([0,1], G_{\alpha,s-\sigma}(\mathbb{T}^n, \mathbb{T}^n))$ of continuous map from $[0,1]$ to $G_{\alpha,s-\sigma}(\mathbb{T}^n, \mathbb{T}^n)$: given an element $\Phi \in V$ and $t \in [0,1]$, we shall write $\Phi^t := \Phi(t)$ and consequently $\Phi = (\Phi^t)_{t \in [0,1]}$. We equip V with the following norm:

$$||\Phi|| := \max_{t \in [0,1]} |\Phi^t|_{\alpha,s-\sigma}$$

which makes it a Banach space, and if we set $\rho := |D|_{\alpha,s}$, we define

$$B_\rho V := \{\Phi \in V \mid ||\Phi - \mathrm{Id}|| \leq \rho\}.$$

We can eventually define a Picard operator P associated to D by

$$P : B_\rho V \to B_\rho V, \quad \Phi \mapsto P(\Phi)$$

where $P(\Phi) = (P(\Phi)^t)_{t \in [0,1]}$ is defined by

$$P(\Phi)^t := \mathrm{Id} + \int_0^t D \circ \Phi^\tau d\tau.$$

To prove the lemma, it is sufficient to prove that P has a unique fixed point $\Phi_* \in B_\rho V$, as necessarily $(\Phi_*^t)_{t \in [0,1]} = (D^t)_{t \in [0,1]}$. Therefore it is sufficient to prove that P induces a well-defined contraction on $B_\rho V$, as the latter is a complete subset of the Banach space V.

First we need to show that P maps $B_\rho V$ into itself. So assume $\Phi \in B_\rho V$, using (77) this implies that for all $t \in [0,1]$,

$$|\Phi^t - \mathrm{Id}|_{\alpha,s-\sigma} \leq \rho \leq \sigma^\alpha$$

so that (74) of Proposition 20 is satisfied (with $f = D$ and $g = \Phi^t$ for any $t \in [0,1]$) and the latter proposition applies: this gives

$$|D \circ \Phi^t|_{\alpha,s-\sigma} \leq |D|_{\alpha,s} = \rho, \quad t \in [0,1]$$

hence

$$\left| P(\Phi)^t - \mathrm{Id} \right|_{\alpha,s-\sigma} = \left| \int_0^t D \circ \Phi^\tau d\tau \right|_{\alpha,s-\sigma} \leq t\rho \leq \rho, \quad t \in [0,1]$$

and therefore

$$||P(\Phi) - \mathrm{Id}|| \leq \rho.$$

This proves that P maps $B_\rho V$ into itself.

It remains to show that P is a contraction. So let $\Phi_1, \Phi_2 \in B_\rho V$, then for any $t \in [0,1]$,

$$\begin{aligned} P(\Phi_1)^t - P(\Phi_2)^t &= \int_0^t \left(D \circ \Phi_1^\tau - D \circ \Phi_2^\tau \right) d\tau \\ &= \int_0^t \left(\int_0^1 \nabla D \circ (s\Phi_1^\tau + (1-s)\Phi_2^\tau) ds \right) \cdot (\Phi_1^\tau - \Phi_2^\tau) d\tau. \end{aligned}$$

Using Proposition 20, Corollary 15 and Corollary 18, we obtain, for any $t \in [0,1]$,

$$|P(\Phi_1)^t - P(\Phi_2)^t|_{\alpha,s-\sigma} \lessdot \sigma^{-\alpha} |D|_{\alpha,s} \max_{0 \leq \tau \leq t} |\Phi_1^\tau - \Phi_2^\tau|_{\alpha,s-\sigma} \lessdot \sigma^{-\alpha} |D|_{\alpha,s} ||\Phi_1 - \Phi_2||$$

and hence

$$||P(\Phi_1) - P(\Phi_2)|| \lessdot \sigma^{-\alpha} |D|_{\alpha,s} ||\Phi_1 - \Phi_2||.$$

Using (77), we can then insure that P is a contradiction, which concludes the proof. $\qquad \square$

Now let us consider a Hamiltonian function X on $\mathbb{T}^n \times D_r$, of the form

$$X(\theta, I) := C(\theta) + D(\theta) \cdot I, \quad C : \mathbb{T}^n \to \mathbb{R}, \quad D : \mathbb{T}^n \to \mathbb{R}^n. \tag{79}$$

The Hamiltonian equations associated to X are given by:

$$\begin{cases} \dot{\theta}(t) = \nabla_I X(\theta(t), I(t)) = D(\theta(t)), \\ \dot{I}(t) = -\nabla_\theta X(\theta(t), I(t)) = -\nabla C(\theta(t)) - \nabla D(\theta(t)) \cdot I. \end{cases}$$

The equations for θ are uncoupled from the equations of I (and hence can be integrated independently), while the equations for I are affine in I; it is well-known that these facts lead to a simple form of the Hamiltonian flow associated to X (see, for instance, [Vil08]).

Proposition 22. *Let X be as in (79) with $C \in G_{\alpha,s}(\mathbb{T}^n, \mathbb{R})$ and $D \in G_{\alpha,s}(\mathbb{T}^n, \mathbb{R}^n)$. Let $0 < \sigma < s$ and assume that*

$$|D|_{\alpha,s} \cdot< \sigma^\alpha. \tag{80}$$

Then for any $t \in [0,1]$, the time-t map X^t of the Hamiltonian flow of X is of the form

$$X^t(\theta, I) = (\theta + E^t(\theta), I + F^t(\theta) \cdot I + G^t(\theta))$$

where $E^t \in G_{\alpha,s-\sigma}(\mathbb{T}^n, \mathbb{R}^n)$, $F^t \in G_{\alpha,s-\sigma}(\mathbb{T}^n, \mathbb{R}^{n^2})$ and $G^t \in G_{\alpha,s-\sigma}(\mathbb{T}^n, \mathbb{R}^n)$ with the estimates

$$|E^t|_{\alpha,s-\sigma} \leq |D|_{\alpha,s}, \quad |F^t|_{\alpha,s-\sigma} \cdot< \sigma^{-\alpha}|D|_{\alpha,s}, \quad |G^t|_{\alpha,s-\sigma} \cdot< \sigma^{-\alpha}|C|_{\alpha,s}. \tag{81}$$

As a consequence, given $0 < \delta < r$, if we further assume that

$$r\sigma^{-\alpha}|D|_{\alpha,s} + \sigma^{-\alpha}|C|_{\alpha,s} \cdot< \delta \tag{82}$$

then X^t maps $\mathbb{T}^n \times D_{r-\delta}$ into $\mathbb{T}^n \times D_r$.

Proof. The second part of the statement clearly follows from the first part, so let us prove the latter. From the specific form of the Hamiltonian equations associated to X, one has, for any $t \in [0,1]$,

$$X^t(\theta, I) = (\theta + E^t(\theta), I + F^t(\theta) \cdot I + G^t(\theta))$$

with

$$\begin{cases} E^t(\theta) = \int_0^t D(\theta + E^\tau(\theta)) d\tau, \\ F^t(\theta) = -\int_0^t \nabla D(\theta + E^\tau(\theta)) d\tau - \int_0^t \nabla D(\theta + E^\tau(\theta)) \cdot F^\tau(\theta) d\tau \\ G^t(\theta) = -\int_0^t \nabla C(\theta + E^\tau(\theta)) d\tau - \int_0^t \nabla D(\theta + E^\tau(\theta)) \cdot G^\tau(\theta) d\tau. \end{cases}$$

Because of (80), Lemma 21 applies and the flow $D^t(\theta) = \theta + E^t(\theta)$ satisfies (78), and therefore

$$|E^t|_{\alpha,s-\sigma} = |D^t - \text{Id}|_{\alpha,s-\sigma} \leq |D|_{\alpha,s}$$

which gives the first estimate of (22). Using this estimate and (80), we can apply Proposition 20 and Corollary 15 (both with $\sigma/2$ instead of σ) to obtain, for any $0 \leq \tau \leq t \leq 1$,

$$|\nabla D \circ D^\tau|_{\alpha,s-\sigma} \leq |\nabla D|_{\alpha,s-\sigma/2} \cdot< \sigma^{-\alpha}|D|_{\alpha,s}.$$

48

Looking at the expression of F^t, this gives

$$|F^t|_{\alpha, s-\sigma} \lessdot \sigma^{-\alpha} |D|_{\alpha, s} \left(1 + \int_0^t |F^\tau|_{\alpha, s-\sigma} d\tau\right)$$

which, by Gronwall's inequality and (80), implies that for all $t \in [0, 1]$,

$$|F^t|_{\alpha, s-\sigma} \lessdot \sigma^{-\alpha} |D|_{\alpha, s}$$

which is the second estimate of (22). For the third estimate of (22), observe that the same argument yields

$$|G^t|_{\alpha, s-\sigma} \lessdot \sigma^{-\alpha} |C|_{\alpha, s} + \sigma^{-\alpha} |D|_{\alpha, s} \int_0^t |G^\tau|_{\alpha, s-\sigma} d\tau$$

and again, by Gronwall's inequality and (80), for all $t \in [0, 1]$ we have

$$|G^t|_{\alpha, s-\sigma} \lessdot \sigma^{-\alpha} |C|_{\alpha, s}.$$

This concludes the proof. $\qquad\square$

B.7 Inverse functions

In this last section, we shall prove that if a Gevrey map is sufficiently close to the identity, then its local inverse is still Gevrey. To prove this in a setting adapted to §6, let us consider a map ϕ which depends only on $\omega \in D_h$, that is $\phi : D_h \to \mathbb{R}^n$.

Proposition 23. *Given $\phi \in G_{\alpha, s}(D_h, \mathbb{R}^n)$, let $0 < \sigma < s$ and assume that*

$$|\phi - \mathrm{Id}|_{\alpha, s} \lessdot \sigma^\alpha, \quad |\phi - \mathrm{Id}|_{\alpha, s} \leq h/2 \qquad (83)$$

Then there exists a unique $\varphi \in G_{\alpha, s-\sigma}(D_{h/2}, D_h)$ such that $\phi \circ \varphi = \mathrm{Id}$ and

$$|\varphi - \mathrm{Id}|_{\alpha, s-\sigma} \leq |\phi - \mathrm{Id}|_{\alpha, s}. \qquad (84)$$

Proof. Let us define $V := G_{\alpha, s-\sigma}(D_{h/2}, \mathbb{R}^n)$, which is a Banach space with the norm $||\,.\,|| = |\,.\,|_{\alpha, s-\sigma}$, and for $\rho := |\phi - \mathrm{Id}|_{\alpha, s}$, we set

$$B_\rho V := \{\psi \in V \mid ||\psi - \mathrm{Id}|| \leq \rho\}.$$

Let us define the following Picard operator P associated to ϕ:

$$P : B_\rho V \to B_\rho V, \quad \psi \mapsto P(\psi) = \mathrm{Id} - (\phi - \mathrm{Id}) \circ \psi.$$

It is clear that $\phi \circ \varphi = \mathrm{Id}$ if and only if φ is a fixed point of P, and therefore the proposition will be proved once we have shown that P has a unique fixed point in $B_\rho V$, and to do this it is enough to prove that P is a well-defined contraction of $B_\rho V$.

49

First let us prove that P maps $B_\rho V$ into itself. So let $\psi \in B_\rho V$, and using the second part of (83), observe that since

$$\sup_{\omega \in D_{h/2}} |\psi(\omega) - \omega| \leq ||\psi - \mathrm{Id}|| \leq \rho \leq h/2$$

then ψ maps $D_{h/2}$ into D_h. This, together with the first part of (83) allows us to apply Proposition 20 to get

$$||(\phi - \mathrm{Id}) \circ \psi|| = |(\phi - \mathrm{Id}) \circ \psi|_{\alpha, s-\sigma} \leq |\phi - \mathrm{Id}|_{\alpha, s} = \rho$$

and thus

$$||P(\psi) - \mathrm{Id}|| = ||(\phi - \mathrm{Id}) \circ \psi|| \leq \rho,$$

that is, P maps $B_\rho V$ into itself. To show that P is a contraction, using Corollary 15, Corollary 18 and Proposition 20 one gets, for any $\psi_1, \psi_2 \in B_\rho V$:

$$||(\phi - \mathrm{Id}) \circ \psi_1 - (\phi - \mathrm{Id}) \circ \psi_2|| \lessdot \sigma^{-\alpha} |\phi - \mathrm{Id}|_{\alpha, s} ||\psi_1 - \psi_2||$$

and from the first part of (83), one can make sure that P is a contraction. This ends the proof. $\square$

Comment. After this work was made public on Arxiv, an independent and interesting proof of a special case of Theorem E appeared in the preprint [LDP17].

References

[Arn64] V.I. Arnold, *Instability of dynamical systems with several degrees of freedom*, Sov. Math. Doklady **5** (1964), 581–585.

[Bes00] U. Bessi, *An analytic counterexample to the KAM theorem*, Ergodic Theory Dynam. Systems **20** (2000), no. 2, 317–333.

[BF13] A. Bounemoura and S. Fischler, *A diophantine duality applied to the KAM and Nekhoroshev theorems*, Math. Z. **275** (2013), no. 3, 1135–1167.

[BF14] _____, *The classical KAM theorem for Hamiltonian systems via rational approximations*, Regul. Chaotic Dyn. **19** (2014), no. 2, 251–265.

[BF17] A. Bounemoura and J. Féjoz, *Hamiltonian perturbation theory for ultra-differentiable functions*, preprint, 2017.

[Bou11]   A. Bounemoura, *Effective stability for Gevrey and finitely differentiable prevalent Hamiltonians*, Communication in Mathematical Physics **307** (2011), no. 1, 157–183.

[Bou13]   _____, *Normal forms, stability and splitting of invariant manifolds I. Gevrey Hamiltonians*, Regul. Chaotic Dyn. **18** (2013), no. 3, 237–260.

[Bru71]   A. D. Bruno, *Analytical form of differential equations I*, Trans. Moscow Math. Soc. **25** (1971), 131–288.

[Bru72]   A.D. Bruno, *Analytical form of differential equations II*, Trans. Moscow Math. Soc. **26** (1972), 199–239.

[Car03]   T. Carletti, *The Lagrange inversion formula on non-Archimedean fields. Non-analytical form of differential and finite difference equations*, Discrete Contin. Dyn. Syst. **9** (2003), no. 4, 835–858.

[CG03]   L. Cadeddu and T. Gramchev, *Nonlinear estimates in anisotropic Gevrey spaces*, Pliska Stud. Math. Bulgar. **15** (2003), 149–160. MR 2071689

[CM00]   T. Carletti and S. Marmi, *Linearization of analytic and non-analytic germs of diffeomorphisms of* $(\mathbf{C}, 0)$, Bull. Soc. Math. France **128** (2000), no. 1, 69–85.

[CM12]   C. Chavaudret and S. Marmi, *Reducibility of quasiperiodic cocycles under a Brjuno-Rüssmann arithmetical condition*, J. Mod. Dyn. **6** (2012), no. 1, 59–78.

[CW13]   C.-Q. Cheng and L. Wang, *Destruction of Lagrangian torus for positive definite Hamiltonian systems*, Geom. Funct. Anal. **23** (2013), no. 3, 848–866.

[Fay02]   B. R. Fayad, *Weak mixing for reparameterized linear flows on the torus*, Ergodic Theory Dynam. Systems **22** (2002), no. 1, 187–201.

[For94]   G. Forni, *Analytic destruction of invariant circles*, Ergodic Theory Dynam. Systems **14** (1994), no. 2, 267–298.

[Gev18]   M. Gevrey, *Sur la nature analytique des solutions des équations aux dérivées partielles. Premier mémoire.*, Ann. Sci. Ec. Norm. Sup. **35** (1918), 129–190.

[GM10]   A. Giorgilli and S. Marmi, *Convergence radius in the Poincaré-Siegel problem*, Discrete Contin. Dyn. Syst. Ser. S **3** (2010), no. 4, 601–621. MR 2684066

[GY99]   T. Gramchev and M. Yoshino, *Rapidly convergent iteration method for simultaneous normal forms of commuting maps*, Math. Z. **231** (1999), no. 4, 745–770. MR 1709494

[Kar16]   N. Karaliolios, *Local rigidity of Diophantine translations in higher dimensional tori*, preprint, 2016.

[Khi63]   A. Ya. Khintchine, *Continued fractions*, Translated by Peter Wynn, P. Noordhoff, Ltd., Groningen, 1963. MR 0161834

[Kol54]   A. N. Kolmogorov, *On the preservation of conditionally periodic motions for a small change in Hamilton's function*, Dokl. Akad. Nauk. SSSR **98** (1954), 527–530.

[Kom79]   H. Komatsu, *The implicit function theorem for ultradifferentiable mappings*, Proc. Japan Acad. Ser. A Math. Sci. **55** (1979), no. 3, 69–72.

[Kom80]   ———, *Ultradifferentiability of solutions of ordinary differential equations*, Proc. Japan Acad. Ser. A Math. Sci. **56** (1980), no. 4, 137–142.

[KP94]   S. Kuksin and J. Pöschel, *On the inclusion of analytic symplectic maps in analytic Hamiltonian flows and its applications*, Seminar on dynamical systems (1994), 96–116, Birkhäuser, Basel.

[Lax53]   P. D. Lax, *Nonlinear hyperbolic equations*, Comm. Pure Appl. Math. **6** (1953), 231–258.

[LDP17]   J. Lopes Dias and Gaivão J. P., *Renormalization of Gevrey vector fields with a Brjuno type arithmetical condition*, preprint, arXiv:1706.04510, 2017.

[LMS16]   L. Lazzarini, J.-P. Marco, and D. Sauzin, *Measure and capacity of wandering domains in gevrey near-integrable exact symplectic systems*, to appear in Memoirs of the AMS, 2016.

[Mos62]   J. Moser, *On Invariant curves of Area-Preserving Mappings of an Annulus*, Nachr. Akad. Wiss. Göttingen **II** (1962), 1–20.

[MS02]   J.-P. Marco and D. Sauzin, *Stability and instability for Gevrey quasi-convex near-integrable Hamiltonian systems*, Publ. Math. Inst. Hautes Études Sci. **96** (2002), 199–275.

[PM97]   R. Pérez-Marco, *Fixed points and circle maps*, Acta Math. **179** (1997), no. 2, 243–294.

[Pop04]   G. Popov, *KAM theorem for Gevrey Hamiltonians*, Erg. Th. Dyn. Sys. **24** (2004), no. 5, 1753–1786.

[Pös01]   J. Pöschel, *A lecture on the classical KAM theory*, Katok, Anatole (ed.) et al., Smooth ergodic theory and its applications (Seattle, WA, 1999). Providence, RI: Amer. Math. Soc. (AMS). Proc. Symp. Pure Math. 69, 707-732, 2001.

[Pös17]   ———, *On the Siegel-Sternberg linearization theorem*, preprint, 2017.

[Rüs75]   H. Rüssmann, *On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus*, Dynamical systems, theory and applications (Rencontres, Battelle Res. Inst., Seattle, Wash., 1974), Springer, Berlin, 1975, pp. 598–624.Lecture Notes in Phys., Vol. 38.

[Rüs01]   _____, *Invariant tori in non-degenerate nearly integrable Hamiltonian systems*, Regul. Chaotic Dyn. **6** (2001), no. 2, 119–204.

[Sal04]   D.A. Salamon, *The Kolmogorov-Arnold-Moser theorem*, Mathematical Physics Electronic Journal **10** (2004), 1–37.

[SCK03]   C.E. Shin, S.-Y. Chung, and D. Kim, *Gevrey and analytic convergence of Picard's successive approximations*, Integral Transforms Spec. Funct. **14** (2003), no. 1, 19–30.

[Sto13]   L. Stolovitch, *Smooth Gevrey normal forms of vector fields near a fixed point*, Ann. Inst. Fourier (Grenoble) **63** (2013), no. 1, 241–267. MR 3097947

[TZ10]   D. Treschev and O. Zubelevich, *Introduction to the perturbation theory of Hamiltonian systems*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2010.

[Vil08]   J. Villanueva, *Kolmogorov theorem revisited*, J. Differential Equations **244** (2008), no. 9, 2251–2276.

[Yam89]   T. Yamanaka, *A new higher order chain rule and Gevrey class*, Ann. Global Anal. Geom. **7** (1989), no. 3, 179–203.

[Yoc88]   J.-C. Yoccoz, *Linéarisation des germes de difféomorphismes holomorphes de* ($\mathbf{C}, 0$), C. R. Acad. Sci. Paris Sér. I Math. **306** (1988), no. 1, 55–58. MR 929279

[Yoc95]   _____, *Small divisors in dimension one (Petits diviseurs en dimension 1)*, Astérisque. 231. Paris: Société Math. de France, 242 p., 1995.