STABILITY AND INSTABILITY FOR
GEVREY QUASI-CONVEX NEAR-INTEGRABLE
HAMILTONIAN SYSTEMS

by Jean-Pierre MARCO *and* David SAUZIN

In memoriam Michael R. Herman

ABSTRACT

We prove a theorem about the stability of action variables for Gevrey quasi-convex near-integrable Hamiltonian systems and construct in that context a system with an unstable orbit whose mean speed of drift allows us to check the optimality of the stability theorem.

Our stability result generalizes those by Lochak-Neishtadt and Pöschel, which give precise exponents of stability in the Nekhoroshev Theorem for the quasi-convex case, to the situation in which the Hamiltonian function is only assumed to belong to some Gevrey class instead of being real-analytic. For n degrees of freedom and Gevrey-α Hamiltonians, $\alpha \geq 1$, we prove that one can choose $a = 1/2n\alpha$ as an exponent for the time of stability and $b = 1/2n$ as an exponent for the radius of confinement of the action variables, with refinements for the orbits which start close to a resonant surface (we thus recover the result for the real-analytic case by setting $\alpha = 1$).

On the other hand, for $\alpha > 1$, the existence of compact-supported Gevrey functions allows us to exhibit for each $n \geq 3$ a sequence of Hamiltonian systems with wandering points, whose limit is a quasi-convex integrable system, and where the speed of drift is characterized by the exponent $1/2(n-2)\alpha$. This exponent is optimal for the kind of wandering points we consider, inasmuch as the initial condition is located close to a doubly-resonant surface and the stability result holds with precisely that exponent for such an initial condition. We also discuss the relationship between our example of instability, which relies on a specific construction of a perturbation of a discrete integrable system, and Arnold's mechanism of instability, whose main features (partially hyperbolic tori, heteroclinic connections) are indeed present in our system.

CONTENTS

1. Introduction

The present work is devoted to the study of the stability exponents for Gevrey quasi-convex near-integrable Hamiltonian systems, and to the search for an example of an unstable orbit with the highest possible speed of drift. We begin by a short reminder on the Nekhoroshev theory in the analytic category and the question of optimality for the stability exponents. We shall then state our main results of stability and instability in the Gevrey category.

1.1. *Reminder on the analytic case*

1.1.1. Let $\mathbf{T} = \mathbf{R}/\mathbf{Z}$ and $n \geq 2$. According to Poincaré, the "general problem of dynamics" is the study of Hamiltonian systems close to an integrable one. Such a system is generated by a Hamiltonian function on $\mathbf{T}^n \times \mathbf{R}^n$, of the form

$$\mathrm{H}(\theta, r) = h(r) + \varepsilon f(\theta, r),$$

which gives rise to the following vector field

$$\mathrm{X_H} \begin{vmatrix} \dot{\theta}_i = \partial_{r_i} h(r) + \varepsilon \partial_{r_i} f(\theta, r), \\ \dot{r}_i = -\varepsilon \partial_{\theta_i} f(\theta, r), \end{vmatrix} \qquad i = 1, ..., n.$$

The canonical coordinates $(\theta, r) \in \mathbf{T}^n \times \mathbf{R}^n$ are angle-action coordinates for the integrable part h. When $\varepsilon = 0$, the actions r_i are first integrals of the system and the motion takes place on the corresponding invariant tori $\mathbf{T}^n \times \{r\}$, all the solutions being quasiperiodic. What remains of this stability for small $\varepsilon > 0$?

In the 70s, a remarkable achievement of Hamiltonian perturbation theory was the Nekhoroshev Theorem [Nekh77], which asserts that for a generic real-analytic function h and for any real-analytic perturbation f, all solutions are stable at least over exponentially long time intervals: there exist positive numbers a and b, depending only on h, such that for each small enough $\varepsilon > 0$ any initial condition (θ_0, r_0) gives rise to a solution $(\theta(t), r(t))$ which is defined at least for $|t| \leq \exp(\text{const } (\frac{1}{\varepsilon})^a)$ and satisfies $\|r(t) - r(0)\| \leq \text{const } \varepsilon^b$ in that range.

This is a statement of *effective* stability, on finite (but long) time intervals and for all solutions, to be compared with the *perpetual* stability that KAM theory, which started two decades earlier, offers only for a part (but a large part) of the phase space. In fact, if $n = 2$ and h is non-degenerate (or isoenergetically non-degenerate), the KAM Theorem gives more than the Nekhoroshev Theorem, since on each energy level the trajectories are confined on or between invariant tori. It is only for $n \geq 3$ that KAM tori do not *a priori* prevent the projection in action space of a solution from drifting arbitrarily far from its initial location; but such a drift should be exponentially slow, according to the Nekhoroshev Theorem.

An interesting question is to know how large the exponents a and b, but especially a, can be taken in Nekhoroshev's statement: the larger a, the longer the time of stability guaranteed by the theorem; and the larger b, the stronger the confinement of the actions close to their initial values.

1.1.2. The generic condition imposed by Nekhoroshev upon h is a transversality property called steepness. *Quasi-convex* functions provide an important particular case of this property. These are functions h for which there exists $m > 0$ such that, at any point r of the domain of definition of f, the inequality $< \nabla^2 h(r)v, v > \geq m\|v\|^2$ holds for all vectors v orthogonal to $\nabla h(r)$. We require moreover ∇h not to vanish. One can check that such a function is isoenergetically non-degenerate: the mapping $(\lambda, r) \mapsto (h(r), \lambda \nabla h(r))$ is a local diffeomorphism for $\lambda > 0$ (see Section 3.4.4).

The property of quasi-convexity amounts to the convexity of the energy levels of h. It is weaker than strict convexity. On the other hand, if $\hat{h}$ is a strictly convex function of $\hat{r} = (r_1, ..., r_{n-1})$, *i.e.* if there exists $\hat{m} > 0$ such that $< \nabla^2 \hat{h}(\hat{r})\hat{v}, \hat{v} > \geq \hat{m}\|\hat{v}\|^2$ for all $\hat{v} \in \mathbf{R}^{n-1}$, one can check that the function $r = (\hat{r}, r_n) \mapsto h(r) = \hat{h}(\hat{r}) + r_n$ is quasi-convex with $m = \hat{m}(1 + \hat{\Omega})^{-2}$ in any domain in which $\|\nabla \hat{h}(\hat{r})\| \leq \hat{\Omega}$. A non-autonomous periodic perturbation of $\hat{h}$ with $n - 1$ degrees of freedom, say $\hat{h}(\hat{r}) + \varepsilon f(\hat{\theta}, \hat{r}, t)$ where $\hat{\theta} = (\theta_1, ..., \theta_{n-1}) \in \mathbf{T}^{n-1}$ and f is also 1-periodic in t, can thus be viewed as a perturbation of a quasi-convex integrable system with n degrees of freedom[1] and the Nekhoroshev Theorem will apply.

---

[1] Use $H = \hat{h}(\hat{r}) + r_n + \varepsilon f(\hat{\theta}, \hat{r}, \theta_n)$: all the energy-levels of H are identical up to a translation in the r_n-direction and on each one, when using the coordinates $(\hat{\theta}, \hat{r}, \theta_n)$, the corresponding autonomous flow amounts to the flow of $\hat{h} + \varepsilon f(\hat{\theta}, \hat{r}, t)$.

As noticed by the Italian school ([BGG85], [Ga86], [BG86]), it turns out that a shrewd use of convexity leads one to a refined result. The use of convexity in conjunction with energy conservation was even more radical in Lochak's novel method [Lo92] which was designed to obtain the best possible stability exponents a and b, and which works in the quasi-convex case as well. Finally, *the Nekhoroshev Theorem holds with $a = b = 1/2n$ if h is assumed to be quasi-convex*, as proved independently by Lochak and Neishtadt [LN92,LNN93] and by Pöschel [Pö93]. Moreover, beside this global result, one can state local results for neighbourhoods of resonant surfaces: if $m \in \{1, ..., n-1\}$, a set of m independent linear relations with integer coefficients to be satisfied by the $\partial_{r_i} h(r)$ determines a *resonant surface of multiplicity m* in the action space; for the trajectories starting at a distance of order $\varepsilon^{1/2}$ of such a surface, one can take larger exponents, namely $a = b = 1/2(n-m)$.

In fact, the case where $m = n - 1$ (the tori $\{r = r^*\}$ associated with such completely resonant actions r^* are foliated into periodic orbits of the unperturbed system) is the cornerstone of Lochak's periodic orbit method introduced in [Lo92], of which [LN92] is only a slight improvement (see [Lo93] for a non-technical account).

We shall not deal with the general steep case, but we wish to mention a recent work by Niederman [Ni00] according to which one can take $a = b = 1/(2np_1...p_{n-1})$ as global exponents, where the p_i's are the *steepness indices* of h (they are not smaller than 1, and all equal to 1 if h is quasi-convex).

1.1.3. The question of the *optimality* of the exponents obtained for the quasi-convex case is still open for $n \geq 5$, whereas a partial answer is available for $n = 3$ or 4. The optimality question amounts to the search for systems arbitrarily close to integrable which admit unstable orbits, *i.e.* orbits experiencing a noticeable drift in action (say, of order 1 if we leave aside the exponent b), and for an asymptotic upper bound of the time of drift, as close as possible to the lower bound $\exp(\text{const}\,(\frac{1}{\varepsilon})^a)$ provided by the stability result.

The phenomenon of instability in near-integrable systems is usually called (somewhat improperly) Arnold diffusion in reference to Arnold's famous note [Arn64] (see also [AA67]), in which an example of a three-degree-of-freedom system was proposed in view of exploring the complement of KAM tori in the phase space and instability was obtained from heteroclinic connections between whiskered tori. There, Arnold was not concerned about the time of drift of his unstable orbits; on the other hand, he raised the difficult question of the genericity of this phenomenon. Arnold's mechanism of instability has motivated numerous studies about the so-called chains of transition, in more or less general frameworks, and particularly about the possibility of finding orbits shadowing such chains, the computations of transition times, and the exponential smallness of the splitting associated with each torus.

Concerning explicit times of drift for systems close to an integrable Hamiltonian written in action-angle variables, we can quote two results by Bessi [Be96,Be97], who

worked on Arnold's model and on a variant of it with four degrees of freedom. Using Arnold's mechanism of instability and variational methods, Bessi obtains orbits drifting in a time $\exp(\text{const } (\frac{1}{\varepsilon})^{1/2})$ for $n = 3$, and $\exp(\text{const } (\frac{1}{\varepsilon})^{1/4})$ for $n = 4$. These orbits pass close enough to a double resonance, thus the exponents cannot be improved for that kind of trajectories; this shows that the exponent $1/2(n-2)$ for doubly-resonant surfaces is optimal when $n = 3$ or 4.

1.2. *Gevrey stability*

1.2.1. It is the aim of the present paper to enlarge the framework by considering Gevrey functions instead of real-analytic ones, and to tackle the question of optimality in this broader context.

Let $\alpha \geq 1$ a real number and $n \geq 2$ the number of degrees of freedom. For $R > 0$ we denote by $\overline{B}_R$ the closed ball of radius R in $\mathbf{R}^n$ with centre at the origin: we shall consider real-valued functions which are C^∞ in $K = \mathbf{T}^n \times \overline{B}_R$. If $L > 0$, such a function $\varphi = \varphi(\theta, r)$ is said to be Gevrey-(α, L) on K, and we write $\varphi \in G^{\alpha, L}(K)$, if

$$(\mathbf{1.1}) \qquad \|\varphi\|_{\alpha, L} := \sum_{k \in \mathbf{N}^{2n}} \frac{L^{|k|\alpha}}{k!^\alpha} \|\partial^k \varphi\|_{C^0(K)} < \infty.$$

We have used the following notations for multi-indices of derivation:

$$|k| = k_1 + \cdots + k_{2n}, \quad k! = k_1! \dots k_{2n}!, \quad \partial^k = \partial_{x_1}^{k_1} \dots \partial_{x_{2n}}^{k_{2n}},$$

and $(x_1, ..., x_{2n}) = (\theta_1, ..., \theta_n, r_1, ..., r_n)$.

Appendix A is devoted to some useful facts and bibliographical notes concerning these functions. Gevrey-α functions are usually defined by the requirement that $\|\partial^k \varphi\|_{C^0(K)} \leq CM^{|k|} k!^\alpha$ for some $C, M > 0$; we recover this space $G^\alpha(K)$ by taking the union over all positive L of the spaces $G^{\alpha, L}(K)$. The advantage of the definition (1.1) is that each $G^{\alpha, L}(K)$ is a Banach algebra: $\|\varphi\psi\|_{\alpha, L} \leq \|\varphi\|_{\alpha, L} \|\psi\|_{\alpha, L}$ (Lemma A.1). The composition of Gevrey-α functions is also Gevrey-α under appropriate assumptions (Section A.2). We recover real-analytic functions in the special case where $\alpha = 1$; the number L then indicates the size of a complex domain of analytic extension. The Cauchy inequalities admit the following generalization (Lemma A.2): if $0 < \lambda < L$ and $\varphi \in G^{\alpha, L}(K)$,

$$\sum_{k \in \mathbf{N}^{2n}; \, |k|=j} \|\partial^k \varphi\|_{\alpha, L-\lambda} \leq j!^\alpha \lambda^{-j\alpha} \|\varphi\|_{\alpha, L}, \qquad j \geq 0.$$

1.2.2. We shall adapt Lochak's periodic orbit method to the Gevrey case and prove

Theorem **A** *(Stability for the Gevrey quasi-convex case).* — *Let* $\alpha \geq 1$, $n \geq 2$,

(**1.2**) $a = \dfrac{1}{2n\alpha}$ *and* $b = \dfrac{1}{2n}$.

Let R, L, E, ϖ, $m > 0$ *and* $h \in G^{\alpha,L}(\overline{B}_R)$ *such that* $\|h\|_{\alpha,L} \leq E$ *and, for all* $r \in \overline{B}_R$,

(**1.3**) $\varpi \leq \max\limits_{1 \leq i \leq n} |\partial_{r_i} h(r)|$ *and* $\forall v \in (\nabla h(r))^{\perp}$, $m\|v\|^2 \leq <\nabla^2 h(r)v, v>$.

For each $R_0 \in]0, R[$, *there exist positive numbers* ε_*, ε_{**}, C_1, C_2, *which depend only on* n, α, R, R_0, L, E, ϖ, m, *such that, for every Hamiltonian function* $H \in G^{\alpha,L}(\mathbf{T}^n \times \overline{B}_R)$ *satisfying*

$$\varepsilon := \frac{\|H - h\|_{\alpha,L}}{E} \leq \min\{\varepsilon_*, \varepsilon_{**}\},$$

any initial condition $(\theta_0, r_0) \in \mathbf{T}^n \times \overline{B}_{R_0}$ *gives rise to a solution* $(\theta(t), r(t))$ *of* X_H *which is defined at least for* $|t| \leq C_1 \exp(\alpha \left(\frac{\varepsilon_*}{\varepsilon}\right)^a)$ *and satisfies* $\|r(t) - r_0\| \leq C_2\varepsilon^b$ *in that range.*

Like in the analytic case, one can state refined stability results associated with resonances. Let $\mathcal{M}$ be a submodule of codimension $d \in \{1, ..., n-1\}$ of $\mathbf{Z}^n$ and

$$S_{\mathcal{M}} = \{\, r \in \overline{B}_{R_0} \mid \forall k \in \mathcal{M},\ < k, \nabla h(r) > = 0 \,\}$$

be the corresponding resonant surface in action space. One can attach to it positive numbers $c_{\mathcal{M}}$, $c'_{\mathcal{M}}$ and $c''_{\mathcal{M}}$, which essentially depend on the *order* of $\mathcal{M}$ (the size of the multi-integers belonging to a system of generators of $\mathcal{M}$)—see Section 3.4.3 for the details. For a "standard" submodule, *i.e.* when $\mathcal{M}^{\perp}$ is generated by some vectors of the canonical basis of $\mathbf{R}^n$, one can take $c_{\mathcal{M}} = \sqrt{n-1}$ and $c'_{\mathcal{M}} = c''_{\mathcal{M}} = 1$, but in general these numbers are larger. By convention, $\mathcal{M} = \{0\}$ can be treated as a standard submodule ($S_{\{0\}} = \overline{B}_{R_0}$, $d = n$, no particular resonance relation); the following statement can thus be considered as containing the previous one:

Addendum to Theorem **A** *(Solutions starting near a resonant surface).* — *Let*

$$a = \frac{1}{2d\alpha}   \text{and}   b = \frac{1}{2d}.$$

For any $\sigma > 0$, *there exist positive numbers* ε_*, ε_{**}, C_1, C_2, *which depend only on* $c_{\mathcal{M}}$, $c'_{\mathcal{M}}$, $c''_{\mathcal{M}}$, σ *and on* n, α, R, R_0, L, E, ϖ, m, *such that, if* $\varepsilon \leq \min\{\varepsilon_*, \varepsilon_{**}\}$ *and if we assume moreover*

$$\text{dist}(r_0, S_{\mathcal{M}}) \leq \sigma\varepsilon^{1/2},$$

the solution $(\theta(t), r(t))$ *is defined at least for* $|t| \leq C_1 \exp(\alpha \left(\frac{\varepsilon_*}{\varepsilon}\right)^a)$ *and satisfies* $\|r(t) - r_0\| \leq C_2\varepsilon^b$ *in that range.*

The dependence of the involved numbers on the various parameters can be explicited as follows: there exists a positive number c which depends only on α such that,

if one introduces

$$\Omega = L^{-\alpha}E, \quad M = 2^\alpha L^{-2\alpha}E, \quad M' = 6^\alpha L^{-3\alpha}E$$

(majorants of the first-, second- and third-order derivatives of h) and

$$K = \frac{1}{2}(L^\alpha E^{-1} + 2^\alpha L^{-1} + m^{-1}(1 + 2^{\alpha+1}L^{-\alpha})^2)^{-1},$$

$$\rho_0 = \frac{1}{4}\min\left\{1, R - R_0, \frac{K}{M}, \frac{K}{M'}\right\},$$

one can take

$$C_1 = \frac{2L^\alpha}{(2n)^{\alpha-1}\Omega}, \quad C_2 = \frac{2^9 M\Omega c'_{\mathscr{M}}\Delta}{m},$$

$$\text{where} \quad \Delta = \left(\frac{2^{d+1}c^{d-1}_{\mathscr{M}}c''_{\mathscr{M}}\sigma}{K^{d-1}\varpi}\right)^{\frac{1}{d}}\max\left\{1, \frac{1}{2\sigma}\left(\frac{c_{\mathscr{M}}E}{m}\right)^{1/2}\right\},$$

and

$$\varepsilon_*^{1/2d} = \frac{L^\alpha}{c(2n)^{\alpha-1}\Delta}\left(\frac{m}{c_{\mathscr{M}}} + \frac{M^2E}{m^2L^{2\alpha}}\right)^{-1}, \quad \varepsilon_{**} = \min\{\varepsilon_1, \varepsilon_2\},$$

$$\begin{cases} \varepsilon_1^{1/2d} = \dfrac{1}{C_2}\min\left\{L^\alpha, R - R_0, \dfrac{\varpi}{M}\right\}, \\ \varepsilon_2^{1/2d} = \left(\dfrac{c_{\mathscr{M}}\varpi}{2c''_{\mathscr{M}}K\sigma}\right)^{\frac{1}{d}}\min\left\{\dfrac{1}{2}, \dfrac{\rho_0 K}{2c_{\mathscr{M}}c'_{\mathscr{M}}\Omega}\right\} \text{ if } d \geq 2, \quad \varepsilon_2 = \infty \text{ if } d = 1. \end{cases}$$

Sections 3.1–3.4 are devoted to the proof of Theorem A. The first three deal with one-phase averaging and contain a generalization of Neishadt's Theorem [Nei84] to Gevrey classes (Proposition 3.2); their use of Gevrey asymptotic expansions is directly inspired by [RS96]. The fourth one repeats essentially the arguments of [Lo92], [LN92] and [LNN93], explaining how to use quasi-convexity to deduce confinement near a periodic torus, and in fact the whole statement with the help of Dirichlet's Theorem.

In view of the success of approximation techniques in KAM theory, it might look more natural to seek a proof of Theorem A relying on the approximation of Gevrey functions by analytic ones and using the classical Nekhoroshev Theorem or the normal forms which are available for analytic Hamiltonians. In fact, M. Herman had obtained by such a method a first version of the stability result, but without being able to reach the exponents (1.2) of Theorem A; this is why we'll have to make do with such a long proof, which contains the proof of the Nekhoroshev Theorem as a particular case instead of using the analytic result to reach the Gevrey one.

1.2.3. We think that generalizing the Nekhoroshev Theorem to Gevrey Hamiltonian functions is interesting in itself. From a mathematical point of view, this is very natural. Spaces of Gevrey functions have been considered for a long time in other branches of Analysis for themselves; and even if one starts with an analytic Hamiltonian function, it can be desirable to restrict it to some invariant manifold which is not analytic but only Gevrey—this may be the case of a centre manifold—, as C. Simó pointed out to us.

1.3. *Example of instability*

1.3.1. Our main motivation was to compare the exponents of stability that we could obtain in the Gevrey quasi-convex case with the speed of instability we could produce on specific examples. For a given h, as simple as possible, we are interested in particular examples of perturbation in which an unstable orbit can be detected, with a possibility of relating the Gevrey size of the perturbation and the time over which instability develops.

If $\alpha > 1$, the Gevrey class G^α is actually larger than the space of real-analytic functions: it contains in particular compact-supported functions. This gives much more flexibility to constuct examples and explains why the question of optimality is easier in this setting.

Compact-supported functions were already used in [Do86], where Arnold's method was adapted to the C^∞ framework to yield topological instability, but without any concern about the corresponding time of instability.

1.3.2. As for us, we followed a rather different approach. The following examples of near-integrable Hamiltonian systems will be obtained by suspending very explicit discrete systems defined on $\mathbf{T}^{n-1} \times \mathbf{R}^{n-1}$, for which a detailed description of the dynamics is available.

Theorem **B** *(Instability result). — Let* $\alpha > 1$, $n \geq 3$ *and*

$$a^* = \frac{1}{2(n-2)\alpha}.$$

Let $L > 0$ *and* $R > 1$. *There exist a sequence of functions* $(f_j)_{j \geq 0}$ *converging to 0 in the space* $G^{\alpha,L}(\mathbf{T}^n \times \overline{\mathrm{B}}_R)$ *and an increasing sequence of integers* $(\tau_j)_{j \geq 0}$ *such that, for each* $j \geq 0$, *the Hamiltonian system generated by*

$$(\mathbf{1.4}) \qquad \mathscr{H}_j(\theta, r) = \frac{1}{2}(r_1^2 + \cdots + r_{n-1}^2) + r_n + f_j(\theta, r)$$

admits a solution $(\theta(t), r(t))$ *defined at least for* $t \in [0, \tau_j]$ *and for which* $r_1(0) = 0$ *and* $r_1(\tau_j) = 1$. *Moreover, there exist an integer* J *which depends only on* n, *and positive constants*

$C_1 < C_2$ *which depend only on* n, α, L *and* R, *such that the time of drift* τ_j *and the norm* $\varepsilon_j = \|f_j\|_{\alpha,L}$ *are related by*

$$\frac{C_1}{\varepsilon_j^2} \exp(C_1 \left(\frac{1}{\varepsilon_j}\right)^{a^*}) \le \tau_j \le \frac{C_2}{\varepsilon_j^2} \exp(C_2 \left(\frac{1}{\varepsilon_j}\right)^{a^*}), \qquad j \ge J.$$

The proof is given in Sections 2.1–2.4. Note that all the hypotheses of Theorem A are fulfilled; in particular the integrable part $h(r) = \frac{1}{2}(r_1^2 + \cdots + r_{n-1}^2) + r_n$ is the most elementary example of a quasi-convex function.

The functions f_j are in fact independent of the last variable r_n, which means that we could forget the variable r_n without losing information and consider that we are dealing with non-autonomous Hamiltonian functions $\frac{1}{2}(r_1^2 + \cdots + r_{n-1}^2) + f_j(\theta_1, ..., \theta_{n-1}, t, r_1, ..., r_{n-1})$ depending periodically on the time (*cf.* footnote 1).

Moreover, we have restricted these functions to $\mathbf{T}^n \times \overline{B}_R$ only in order to control their Gevrey norms, but their dependence on $r_1, ..., r_{n-1}$ is indeed very mild. The Hamiltonian vector field associated with (1.4) is in fact defined and C^∞ in $\mathbf{T}^n \times \mathbf{R}^n$, and *complete*. In particular, our unstable solution $(\theta(t), r(t))$ can be described for all $t \in \mathbf{R}$: we shall see for instance that $r_1(k\sqrt{\tau_j}) = k/\sqrt{\tau_j}$ for all $k \in \mathbf{Z}$, hence this solution is *biasymptotic to infinity*. This situation is reminiscent of the system studied by Mather [Mat93] and its variants [BT99,DLS00].

1.3.3. Let us now confront our two theorems (with $\alpha > 1$). As already mentioned, the hypotheses of Theorem A are fulfilled so that it applies to the Hamiltonian functions (1.4). We shall see that, provided j is large enough, the second action-component $r_2(t)$ of our unstable solution satisfies

$$\forall t \in \mathbf{R}, \quad 0 < r_2(t) \le 3\sqrt{\varepsilon_j}$$

(see Section 2.4.3). In view of the frequency map $\nabla h(r) = (r_1, ..., r_{n-1}, 1)$, we can thus say that this solution starts $O(\sqrt{\varepsilon_j})$-close to the fixed doubly-resonant surface $\{r_1 = r_2 = 0\}$. Therefore, the value $1/2(n-2)\alpha$ attained by the instability exponent a^* exactly coincides with that of the local stability exponent $a = 1/2d\alpha$ for a resonance submodule of rank 2 and codimension $d = n - 2$. In that respect our result can be considered as optimal.

This establishes the optimality of all the exponents $a = 1/2d\alpha$ for resonances of multiplicity not less than 2 as well, by simply adding $n' \ge 0$ degrees of freedom and considering the solution $(\theta(t), r(t), 0, 0)$ of the Hamiltonian system generated by

$$\mathscr{H}_j(\theta_1, ..., \theta_n, r_1, ..., r_n) + \frac{1}{2}(r_{n+1}^2 + \cdots + r_{n+n'}^2).$$

This yields indeed, for $N = n + n'$ degrees of freedom, a solution which starts $O(\sqrt{\varepsilon_j})$-close to the resonant surface associated with a fixed[2] submodule of rank $n' + 2$ and codimension $d = n - 2$.

One can perform further computations in our example, in order to see why the unstable orbit escapes the influence of resonances of higher multiplicity—see Section 2.4.3 (and Section 3.4.3 for the underlying arithmetics).

It is a general phenomenon that resonances cannot be avoided for a solution which drifts between two distant regions of the action space, because of the existence of at least one fixed simply-resonant hypersurface which separates the given regions and which the solution has to cross. For instance, for the particular integrable part h that we chose, the hypersurfaces $\{r_1 = p/q\}$ are resonant for all rationals p/q and intercept the solutions drifting in the r_1-direction. But one can still wonder about the optimality of the stability exponent $a = 1/2d\alpha$ when $d = n - 1$, *i.e.* about the existence of unstable orbits whose time of drift is characterized by the instability exponent $1/2(n-1)\alpha$ and which remain far enough from all doubly-resonant surfaces (at least those whose order is not too high with respect to ε).

As for the "universal" exponent $a = 1/2n\alpha$, the question of its optimality cannot be raised independently of that of $b = 1/2n$.

1.3.4. We emphasize the difference between our goals, to construct specific perturbations giving rise to instability, and those of Arnold's note [Arn64] which illustrated on a quite simple perturbation a powerful mechanism of instability relying on transition tori. In particular, in Arnold's method, the drifting points are not *a priori* known, but still have to be detected once a transition chain is given. The description of their orbits is necessarily complicated, and quantitative estimates for the speed of drift are even more difficult to obtain. Here, on the contrary, we shall construct examples where we can choose the drifting points from the beginning, and our main task will be the computation of the Gevrey norms of the various functions involved in the construction; in a sense our approach pertains more to functional analysis than to dynamical systems.

As a byproduct, we get new insight in Arnold's mechanism itself, which we do not use but can however illustrate. As a matter of fact, we can choose our example so as to possess a transition chain, of infinite length, shadowed by the solution described in Theorem B. Moreover, the splitting of separatrices (or at least its most significant part), whose determination in analytic examples is notoriously a difficult matter, is *trivially* computed in the same example.

---

[2] We stress that the resonance conditions are *fixed*, *i.e.* independent of j with our notations, because the Addendum to Theorem A also applies to resonances which depend on the perturbed Hamiltonian and on the considered solution. But then the "order" of the resonance and the corresponding numbers $c_{\mathscr{M}}, c'_{\mathscr{M}}, c''_{\mathscr{M}}$ may grow as ε tends to zero and thus need to be taken into account together with the exponent a; this is illustrated in Section 2.4.3.

To be precise we need a refined definition of the splitting, adapted to the particular case we consider here. Let an energy e be given (of course our considerations will not depend of the value of e since the perturbative term f_j does not depend on r_n). We first introduce the manifold $\mathscr{S}_j(e) = \{\theta_n = 0\} \cap \mathscr{H}_j^{(-1)}(\{e\})$: this is a $2(n-1)$-dimensional transverse section for the Hamiltonian flow, that we can identify with the annulus $\mathbf{A}^{n-1}$ through the coordinates $(\bar{\theta}, \bar{r})$, $\bar{\theta} = (\theta_1, ..., \theta_{n-1})$, $\bar{r} = (r_1, ..., r_{n-1})$. The transition tori that we shall consider will be contained in $\mathscr{H}_j^{(-1)}(\{e\})$, $(n-1)$-dimensional, hyperbolic, with n-dimensional invariant manifolds $\mathrm{W}^{\pm}$, and they will possess homoclinic points ω located in $\mathscr{S}_j(e)$. In the coordinates $(\bar{\theta}, \bar{r})$, the intersections of $\mathrm{W}^{\pm}$ with $\mathscr{S}_j(e)$ will be the graphs of suitable functions $\bar{r} = w^{\pm}(\bar{\theta})$; the coordinates $(\bar{\theta}^0, \bar{r}^0)$ of ω will satisfy of course $\bar{r}^0 = w^+(\bar{\theta}^0) = w^-(\bar{\theta}^0)$, but moreover $\theta_2^0 = \cdots = \theta_{n-1}^0 = 0$. One usually defines the splitting of W^+ and W^- at ω as the difference $d(w^+ - w^-)(\bar{\theta}^0)$, which is a symmetric operator (see [LMS03]). But in our example the manifolds $\mathrm{W}^{\pm}$ have an additional property, which will allow us to restrict our attention to a scalar quantity: for $i = 2, ..., n-1$, $(w_i^+ - w_i^-)(\theta_1, 0, ..., 0) \equiv 0$, thus the splitting matrix admits a block decomposition into a "longitudinal" part

$$\mathscr{A}_{|1}(\mathrm{W}^{\pm}, \omega) = \frac{\partial}{\partial \theta_1}(w_1^+ - w_1^-)(\theta_1^0)$$

and a "transversal" part which is the symmetric matrix of order $n-2$ with coefficients $\frac{\partial}{\partial \theta_j}(w_i^+ - w_i^-)(\theta_1^0)$, $i, j \geq 2$. In our situation, the interesting part of the splitting is the longitudinal one, *i.e.* the splitting *in the r_1-direction*, which would be needed to detect heteroclinic connections between consecutive tori of the chain if one would follow Arnold's method.

Addendum to Theorem **B** *(Transition chains and splitting).* — *Let $j \geq 0$ and $e \in \mathbf{R}$. For each $r_1^0 \in \mathbf{R}$, there exists in the energy level $\mathscr{H}_j^{(-1)}(\{e\})$ a $(n-1)$-dimensional invariant torus $\mathrm{T}_{r_1^0}$, whose projection on the r_1-axis is reduced to $\{r_1^0\}$, and which is partially hyperbolic for the Hamiltonian flow, with n-dimensional invariant manifolds.*

Let $q_j = \sqrt{\tau_j}$. The family $(\mathrm{T}_{k/q_j})_{k \in \mathbf{Z}}$ is a transition chain: two consecutive tori $\mathrm{T}_{(k-1)/q_j}$ and T_{k/q_j} are connected by a heteroclinic orbit. The solution $(\theta(t), r(t))$ of Theorem B shadows that chain: for $t = (k + \frac{1}{2})q_j$ with $k \in \mathbf{Z}$, its distance to the torus T_{k/q_j} is less than $1/q_j$ (it is thus exponentially small), and for $(k - \frac{1}{2})q_j < t < (k + \frac{1}{2})q_j$, it remains close to the aforementioned heteroclinic solution.

For each $r_1^0 \in \mathbf{R}$, the torus $\mathrm{T}_{r_1^0}$ has two homoclinic orbits, which intersect the section $\mathscr{S}_j(e)$ at points ω_1, ω_2 for which $\theta_2 = \cdots = \theta_{n-1} = 0$. At these points, the splitting in the r_1-direction of the invariant manifolds $\mathrm{W}^{\pm}(\mathrm{T}_{r_1^0})$ satisfies $|\mathscr{A}_{|1}(\mathrm{W}^{\pm}(\mathrm{T}_{r_1^0}), \omega_i)| = 2\pi/q_j$, $i = 1, 2$.

The proof is given in Section 2.5. Note that the splitting $\mathscr{A}_{|1}$ is exactly of the same order as the gap $1/q_j$ between two tori in the chain; thus by the same token we

get lower and upper bounds for $\mathscr{A}_{|1}$: there exist constants $C_1' > C_2'$ such that

$$\frac{\varepsilon_j}{C_1'} \exp\left(-C_1'\left(\frac{1}{\varepsilon_j}\right)^{a^*}\right) \leq |\mathscr{A}_{|1}(W^\pm(T_{r_1^0}), \omega_i)| \leq \frac{\varepsilon_j}{C_2'} \exp\left(-C_2'\left(\frac{1}{\varepsilon_j}\right)^{a^*}\right).$$

Actually one can say even more: we shall see that the stable and unstable manifolds of two consecutive tori $T_{(k-1)/q_j}$ and T_{k/q_j} are tangent along the heteroclinic orbit we consider, so our example is a limit case of Arnold's mechanism, where the tori are maximally distant. Of course, to construct our drifting points, we do not make use of Arnold's semi-local analysis (lambda-lemma and obstruction property), which would require at least the topological transversality of the intersections. Our drifting points are obtained by a completely different method, but we check *a posteriori* that they shadow the chain we consider.

1.4. *Comments, prospects*

1.4.1. The stability theorem can be extended to ultradifferentiable classes more general than the Gevrey class (defined by such inequalities as (A.2)), and also to the case of C^r functions with $r < \infty$. M. Herman had already devised a method for that, but we could not include it in the present article which was already long enough. In finite differentiability, one can produce easily instability examples, even with a wandering ball. A further article should be devoted to this.

1.4.2. Niederman's recent stability result [Ni00] about the steep case can probably be extended to the Gevrey class. We did not explore that systematically, but we can at least indicate a steep variant of the above example where the instability exponent is increased at the price of relaxing quasi-convexity: for any $p \geq 3$, we can construct a Hamiltonian

$$\mathscr{H}_j(\theta, r) = \frac{1}{2}r_1^2 + \frac{1}{p}r_2^p + \frac{1}{2}(r_3^2 + \cdots + r_{n-1}^2) + r_n + f_j(\theta, r),$$

with a sequence f_j converging to 0 in $G^{\alpha,L}(\mathbf{T}^n \times \overline{B}_R)$, for which there is an unstable orbit of exponent

$$a^* = \frac{1}{p(n-2)\alpha}$$

(we shall not give the details of the construction which is a simple adaptation of the proof of Theorem B). On the one hand, this shows that the requirement of quasi-convexity in Theorem A is necessary for the exponent $a = 1/2n\alpha$ to be valid. On the other hand, this provides an example of an unstable perturbation of a steep integrable system whose exponents are $p_1 = \cdots = p_{n-1} = p - 1$.

1.4.3. Returning to perturbations of a convex or quasi-convex integrable Hamiltonians, it remains to be seen if the time exponent $1/2(n-1)\alpha$ related to simple resonances is optimal; we hope to be able to achieve this by working directly on flows, rather than on mappings like in the proof of Theorem B. The question of knowing whether the time exponent $1/2n\alpha$ can be optimal is more delicate.

But of course the most important and natural sequel to the present work, which we are currently examining, is to try to adapt the new ideas introduced in the case of Gevrey systems to the analytic framework, in order to achieve the fastest possible drift for an analytic near-integrable Hamiltonian system.

2. An example of instability with estimate on the speed of drift

In this part it will be more convenient to work with mappings rather than with flows (the correspondence between Nekhoroshev estimates for flows and for mappings in a rather general perturbative framework was studied in [KP94]). Before proving Theorem B, we need to introduce the setting and give a refined formulation of the result in terms of a discrete dynamical system to be constructed (Proposition 2.1 below). Theorem B will be proved in Section 2.4 as a corollary, by a standard procedure of suspension, and its Addendum in Section 2.5 by a slight modification of the construction.

2.1. *A more precise statement*

2.1.1. If H is a Hamiltonian function generating a complete vector field, we shall denote by Φ^H the corresponding time-1 map; this is a special case of an exact-symplectic transformation. The maps that we shall consider will be compositions of two time-1 maps of Hamiltonian functions on $\mathbf{A}^n = T^*\mathbf{T}^n = \mathbf{T}^n \times \mathbf{R}^n$, where $n \geq 2$ and where $\mathbf{A} = \mathbf{T} \times \mathbf{R}$ is the annulus. These maps are still exact-symplectic, but one cannot expect them to be time-1 maps of Hamiltonian systems on $\mathbf{A}^n$; however, a simple suspension result will prove that they are indeed time-1 maps of Hamiltonian systems on $\mathbf{A}^{n+1}$—beware of the shift of dimension with respect to Sections 1 and 3.

Here we call *integrable* a map of the form Φ^H, where H depends only on the action variables; in that case, $\Phi^H(\theta, r) = (\theta + \nabla H(r), r)$ (notice that Φ^H is well-defined also in the case where H depends only on the angle variables).

Our example of unstable map will be of the form $\Phi^u \circ \Phi^{h+v}$, where $h = \frac{1}{2}(r_1^2 + \cdots + r_n^2)$, and u and v are small Gevrey functions; the unperturbed map Φ^h is the twist map $(\theta, r) \mapsto (\theta + r, r)$ of $\mathbf{A}^n$. The special form of our unstable map, defined as a composition of two time-1 maps, and the fact that we are working in a non-quasianalytic Gevrey class will facilitate the procedure of suspension of Section 2.4.

2.1.2. The statement and construction below contain technical details which may obscure the ideas from which they originate, so we begin with a heuristic description of the system. The main parameter will be an integer q, supposed to be large. We shall split the annulus $\mathbf{A}^n$ and the unperturbed map into two factors: $\Phi^h = \Phi^{\frac{1}{2}r_1^2} \times \Phi^{\frac{1}{2}r_2^2 + \cdots + \frac{1}{2}r_n^2} : \mathbf{A} \times \mathbf{A}^{n-1} \to \mathbf{A} \times \mathbf{A}^{n-1}$, and consider a q-periodic point $a \in \mathbf{A}^{n-1}$ on the second factor. We first want to define a "coupling diffeomorphism" Φ^{u_q} on the product $\mathbf{A}^n$ such that $\Phi^{u_q} \circ \Phi^h$ has a wandering point which drifts along the r_1-axis, and we furthermore require that $u_q \to 0$ when $q \to \infty$, in a suitable Gevrey function space.

On the first factor $\mathbf{A}$, the interesting part of the dynamics will be localized on the union $\mathbf{C}_q$ of the circles $\mathrm{C}_{k/q} = \{(\theta_1, r_1) \in \mathbf{A}, \ r_1 = k/q\}, \ k \in \mathbf{Z}$. Each of these circles is invariant under $\Phi^{\frac{1}{2}r_1^2}$ and supports a q-periodic dynamics, even if q is not the minimal period. On the second factor we shall only consider the orbit $\mathrm{O}(a) = \{a_{(k)}, \ 0 \le k \le q - 1\}$ of $a = a_{(0)}$ under $\Phi^{\frac{1}{2}r_2^2 + \cdots + \frac{1}{2}r_n^2}$.

The coupling diffeomorphism Φ^{u_q} will be chosen so as to satisfy the "synchronization conditions"

$$\Phi^{u_q}((0, r_1), a) = ((0, r_1 + 1/q), a), \quad \text{and}$$

$$\Phi^{u_q}((\theta_1, r_1), a_{(k)}) = ((\theta_1, r_1), a_{(k)}) \ \text{if } k \in \{1, ..., q - 1\},$$

for all $(\theta_1, r_1) \in \mathbf{A}$. Due to the q-periodicity of Φ^h on $\mathbf{C}_q \times \mathrm{O}(a)$, one sees that the point $((0, 0), a)$ is wandering for $\Phi^{u_q} \circ \Phi^h$, and satisfies in particular

$$[\Phi^{u_q} \circ \Phi^h]^{kq}((0, 0), a) = ((0, k/q), a), \qquad k \in \mathbf{Z},$$

while the first components of the other iterates move around the circles of the family $\mathbf{C}_q$. So q^2 iterations of the coupled diffeomorphism $\Phi^{u_q} \circ \Phi^h$ make the point $((0, 0), a)$ drift along the r_1-axis over an interval of length 1. This is all we need in order to estimate our instability time.

As is easily checked, a simple way to get the synchronization is to choose u_q of the form $u_q((\theta_1, r_1), x') = \frac{1}{q} \mathrm{U}(\theta_1) g^{(q)}(x')$ for $((\theta_1, r_1), x') \in \mathbf{A} \times \mathbf{A}^{n-1}$, where $\mathrm{U}'(0) = -1$ and

$$g^{(q)}(a) = 1, \ dg^{(q)}(a) = 0, \ g^{(q)}(a_{(k)}) = 0, \ dg^{(q)}(a_{(k)}) = 0, \ 1 \le k \le q - 1$$

(see Lemma 2.1 below). The size of the function u_q is seen to be of the order of $\|g^{(q)}\|/q$. The main difficulty of the construction will be to ensure the condition $\|g^{(q)}\|/q \to 0$ as $q \to \infty$: by compactness the distance between the initial point a and its nearest iterate tends to 0 as $q \to \infty$, and the values of $g^{(q)}$ on a and this iterate differ by 1; so any Gevrey norm of $g^{(q)}$ will tend to ∞ when $q \to \infty$. One can convince oneself that the construction is not possible with the second factor kept equal to $\Phi^{\frac{1}{2}(r_2^2 + \cdots + r_n^2)}$: the periodic points are equidistributed on periodic tori, and the distance between two of them is just too short.

For this reason we shall add a perturbation to the initial Hamiltonian h, splitting now the dynamics on the second factor $\mathbf{A}^{n-1} = \mathbf{A} \times \mathbf{A}^{n-2}$ into two parts. The first one will be the time-1 map of a pendulum suitably rescaled: $\Phi^{\frac{1}{2}r_2^2 - \frac{1}{N^2}\cos\theta_2}$, with a new large parameter N; and it will be possible to keep the second part on $\mathbf{A}^{n-2}$ equal to the integrable twist map $\Phi^{\frac{1}{2}(r_3^2 + \cdots + r_n^2)}$. The main property of this system is that, due to the presence of the pendulum and its separatrix, one can find q-periodic points $a = a^{(q)}$, with arbitrarily large q, whose distance to the rest of their orbit is of the order of $1/N$.

With a parameter q large enough, the $G^{\alpha,L}$-norms of the functions u_q and $v_N = -\frac{1}{N^2}\cos\theta_2$ will be of the same order $\varepsilon = 1/N^2 \simeq \|g^{(q)}\|/q$: we shall have to take $q = O(\exp(cN^{1/2\alpha(n-1)}))$ (where c depends on α and L). This shows the connection between the instability time τ (for a drift of order one) and the Gevrey size ε of the perturbations u_q and v_N:

$$\tau = q^2 = O\left(\exp\left(2c\left(\frac{1}{\varepsilon}\right)^{\frac{1}{2\alpha(n-1)}}\right)\right),$$

and corresponds to our primary aim.

The hyperbolic fixed point of the pendulum in the second factor induces hyperbolic tori which will enable us to construct transition chains as well. Finally, thanks to a more global formulation of our arguments (the coupling lemma below) and a particular choice of the function U, we shall be able to give a complete description of the intersection of the invariants manifolds attached to the tori with a suitable section; this will give us the estimates on the splitting.

Notations. — When necessary we shall split the set of variables into

$$x_1 = (\theta_1, r_1) \in \mathbf{A}, \quad x_2 = (\theta_2, r_2) \in \mathbf{A}, \quad \hat{x} = (\theta_3, ..., \theta_n, r_3, ..., r_n) \in \mathbf{A}^{n-2}.$$

The reader may consider first the case $n = 2$, which is simpler and of which the general case is an easy generalization.

When a point or a function with several components depends on a parameter, to avoid confusion we use superscripts between brackets for the parameter and subscripts for the components, *e.g.* $x^{(q)} = (x_1^{(q)}, x_2^{(q)}, \hat{x}^{(q)})$.

In order to produce periodic points of the integrable twist map on the annulus, we shall simply use inverses of prime numbers as action variables (because we need mutually prime numbers). We denote by $(p_j)_{j \geq 0}$ the sequence of prime numbers and, for each $j \geq n-1$, we define

(**2.1**) $\qquad N_j = p_j N_j'$ where $N_j' = 1$ if $n = 2$, and $N_j' = p_{j-(n-3)}p_{j-(n-4)}...p_j$ if $n \geq 3$.

2.1.3. The next two sections will be devoted to the proof of

Proposition **2.1** *(Discrete version of the instability example).* — *Let* $n \geq 2$, $\alpha > 1$, $L > 0$ *and*

$$(\textbf{2.2}) \qquad U(\theta_1) = -\frac{1}{2\pi} \sin 2\pi\theta_1, \quad V(\theta_2) = -\cos 2\pi\theta_2.$$

There exist a sequence of functions $(g^{(j)})_{j \geq n-1}$ *in* $G^{\alpha,L}(\mathbf{T}^{n-1})$, *a sequence of real numbers* $(r_2^{(j)})_{j \geq n-1}$ *and positive constants* c *and* Λ *(depending only on* α *and* L*) such that, if for each* $j \geq n-1$ *we define*

$$(\textbf{2.3}) \qquad M_j = 2[cN_j \, e^{2\alpha(\Lambda+(n-2)L)(2\pi p_j)^{1/\alpha}} + 1], \qquad q_j = N_j M_j,$$

we have

$$(\textbf{2.4}) \qquad \frac{1}{q_j} \|g^{(j)}\|_{\alpha,L} \leq \frac{1}{N_j^2}$$

and the transformation

$$(\textbf{2.5}) \qquad \Psi_j = \Phi^{\frac{1}{q_j} U \otimes g^{(j)}} \circ \Phi^{\frac{1}{2}(r_1^2 + r_2^2 + \cdots + r_n^2) + \frac{1}{N_j^2} V(\theta_2)}$$

produces a drift from $r_1 = 0$ *to* $r_1 = 1$ *in* q_j^2 *iterations for the point* $x^{(j)} = (x_1^{(j)}, x_2^{(j)}, \hat{x}^{(j)})$,

$$(\textbf{2.6}) \qquad x_1^{(j)} = (0,0), \quad x_2^{(j)} = (0, r_2^{(j)}), \quad \hat{x}^{(j)} = (0, ..., 0, 1/p_{j-(n-3)}, ..., 1/p_j).$$

We have denoted by $\otimes$ the product of functions depending on separate sets of variables:

$$U \otimes g^{(j)}(\theta) = U(\theta_1) g^{(j)}(\theta_2, ..., \theta_n).$$

Note that the diffeomorphism Ψ_j is of the form $\Phi^{u_j} \circ \Phi^{h+v_j}$, as announced at the beginning of Section 2.1, with

$$(\textbf{2.7}) \qquad \|u_j\|_{\alpha,L} = \left\| \frac{1}{q_j} U \otimes g^{(j)} \right\|_{\alpha,L} \leq \|v_j\|_{\alpha,L} = \left\| \frac{1}{N_j^2} V \right\|_{\alpha,L} = c'/N_j^2$$

(with $c' = \|V\|_{\alpha,L} = 2\pi \|U\|_{\alpha,L}$). So the size of the perturbation, *i.e.* the distance between Ψ_j and its integrable limit Φ^h, tends to 0 as $j \to \infty$ and is easily compared to the time of drift q_j^2; Theorem B will be a simple consequence of Proposition 2.1 (see Section 2.4).
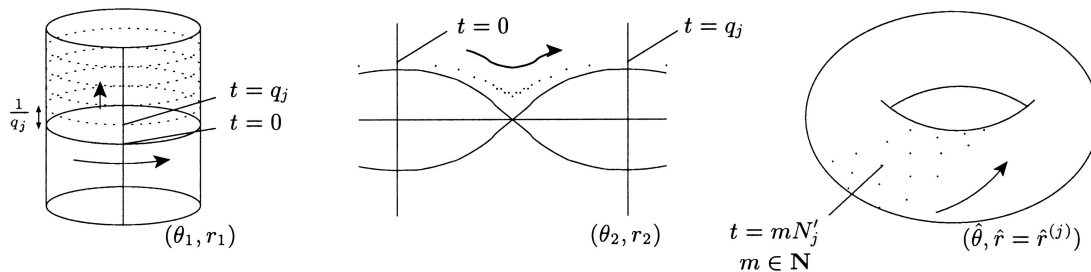
FIG. 1. — Projections of the orbit of $x^{(j)}$.

2.1.4. It will be shown in Section 2.3 how to choose $g^{(j)}$ and $r_2^{(j)}$. The point $x^{(j)}$ is wandering for Ψ_j and its orbit can be described completely. For instance, regarding times which are multiples of q_j,

$$\forall k \in \mathbf{Z}, \quad \Psi_j^{kq_j}(x^{(j)}) = \left(\left(0, \frac{k}{q_j} \right), x_2^{(j)}, \hat{x}^{(j)} \right).$$

More generally, if we set $\hat{r}^{(j)} = (1/p_{j-(n-3)}, ..., 1/p_j)$ and if we denote by $\pi_1, \pi_2, \hat{\pi}$ the projections associated with the decomposition $\mathbf{A}^n = \mathbf{A} \times \mathbf{A} \times \mathbf{A}^{n-2}$, we shall have, for all $k \in \mathbf{Z}$,

$$\hat{\pi} \circ \Psi_j^k(x^{(j)}) = (k\hat{r}^{(j)}, \hat{r}^{(j)}), \quad \pi_1 \circ \Psi_j^k(x^{(j)}) = \left(\frac{k'k''}{q_j}, \frac{k'}{q_j} \right),$$

where $k = k'q_j + k''$ and $0 \leq k'' \leq q_j - 1$, while the x_2-component $\pi_2 \circ \Psi_j^k(x^{(j)})$ follows a q_j-periodic orbit of the pendulum $\frac{1}{2}r_2^2 + \frac{1}{N_j^2}V(\theta_2)$. One can even see the orbit slow down as r_2 comes closer to 0: this corresponds to a stronger influence of the double resonance when r passes closer to $(0, 0, \hat{r}^{(j)})$ (see Figure 1).

We mention here that $x^{(j)}$ is not the only drifting point that we can exhibit for the system (2.5): replacing $(\frac{1}{p_{j-(n-3)}}, ..., \frac{1}{p_j})$ in the definition of $\hat{x}^{(j)}$ in (2.6) by $(\frac{\ell_3^{(j)}}{p_{j-(n-3)}}, ...$ $..., \frac{\ell_n^{(j)}}{p_j})$, where each $\ell_i^{(j)}$ is any integer prime to $p_{j-(n-i)}$, one obtains a new point $x^{(j)}$ which is also drifting from $r_1 = 0$ to $r_1 = 1$ in q_j^2 iterations. This makes quite a large number of unstable orbits if $n \geq 3$. But even if $n = 2$, one can also modify slightly the definition of U and q_j, still preserving the inequality (2.7), in order to make all the points $x = (x_1, x_2, \hat{x})$,

$$(\mathbf{2.8}) \qquad x_1 = \left(\frac{\ell_1^{(j)}}{p_j}, 0 \right), \quad x_2 = (0, r_2^{(j)}), \quad \hat{x} = \left(0, ..., 0, \frac{\ell_3^{(j)}}{p_{j-(n-3)}}, ..., \frac{\ell_n^{(j)}}{p_j} \right)$$

(with $0 \leq \ell_1^{(j)} \leq p_j - 1$ and $\ell_3^{(j)}, ..., \ell_n^{(j)}$ like previously) drift from $r_1 = 0$ to $r_1 = 1$ (see Remark 2.2).

2.2. *Embedding of the standard map and its drifting orbit*

2.2.1. Let us first show where our unstable orbits will come from. At the heart of our instability mechanism lies the q-parametrised family of "standard maps" of the annulus, defined for $q \geq 1$ by

$$\psi_q(\theta_1, r_1) = \left(\theta_1 + qr_1, r_1 - \frac{1}{q}U'(\theta_1 + qr_1)\right),$$

where, in view of (2.2), $U'(\theta_1) = -\cos 2\pi\theta_1$. Since $U'(0) = -1$, one readily computes the orbit of the origin:

$$\forall k \in \mathbf{Z}, \quad \psi_q^k(0, 0) = \left(0, \frac{k}{q}\right).$$

In particular, this orbit drifts from $r_1 = 0$ to $r_1 = 1$ in q iterations.

But ψ_q is not close to integrable, by any manner of means. Yet it will be possible to take advantage of the wandering point that we have for ψ_q: our strategy will be to embed ψ_q into the qth iterate of a near-integrable mapping of $\mathbf{A}^n$, using the fact that ψ_q can be written as

$$(\mathbf{2.9}) \qquad \psi_q = \Phi^{\frac{1}{q}U} \circ F^q, \qquad F = \Phi^{\frac{1}{2}r_1^2} : (\theta_1, r_1) \mapsto (\theta_1 + r_1, r_1).$$

2.2.2. Here comes the only slightly technical tool of our construction, a "coupling lemma" which is in fact a very simple lemma on local skew-products. It will only remain for us to imagine to what situation we must apply it.

Lemma **2.1.** — *Let $m, m' \geq 1$. Suppose we are given two maps, $F : \mathbf{A}^m \to \mathbf{A}^m$ and $G : \mathbf{A}^{m'} \to \mathbf{A}^{m'}$, and two Hamiltonian functions $f : \mathbf{A}^m \to \mathbf{R}$ and $g : \mathbf{A}^{m'} \to \mathbf{R}$ which generate complete vector fields and define time-1 maps Φ^f and Φ^g. Suppose moreover that $a \in \mathbf{A}^{m'}$ is G-periodic, of period q, and that*

$$(\mathbf{2.10}) \qquad g(a) = 1, \quad dg(a) = 0, \quad g(G^k(a)) = 0, \quad dg(G^k(a)) = 0, \quad 1 \leq k \leq q - 1.$$

Then $f \otimes g$ generates a complete Hamiltonian vector field and the mapping

$$\Psi = \Phi^{f \otimes g} \circ (F \times G) : \mathbf{A}^{m+m'} \to \mathbf{A}^{m+m'}$$

satisfies

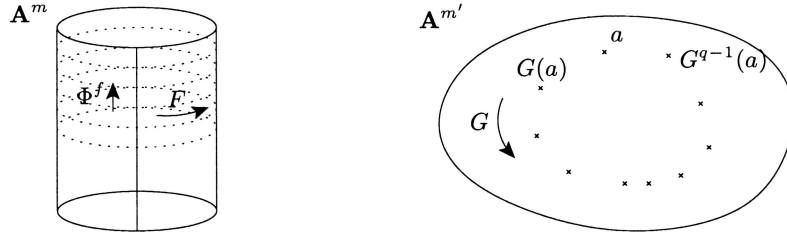$$\forall x \in \mathbf{A}^m, \quad \Psi^q(x, a) = (\Phi^f \circ F^q(x), a).$$

FIG. 2. — Orbits in the coupling lemma: starting with $(x, a) \in \mathbf{A}^m \times \mathbf{A}^{m'}$, the $q - 1$ first iterations of Ψ coincide with those of $F \times G$, the qth iteration with that of $(\Phi^f \circ F) \times G$ thanks to the synchronization conditions (2.10).

We have denoted by $f \otimes g$ the function $(x, x') \mapsto f(x)g(x')$, and by $F \times G$ the mapping $(x, x') \mapsto (F(x), G(x'))$.

Proof. — The Hamiltonian vector field generated by $f \otimes g$ is

$$X_{f \otimes g} \begin{vmatrix} \dot{x} = g(x')X_f(x), \\ \dot{x}' = f(x)X_g(x'). \end{vmatrix}$$

It admits $f(x)$ and $g(x')$ as first integrals. It is complete, since it can be integrated by

$$(2.11) \qquad \Phi^{f \otimes g}(x, x') = (\Phi^{g(x')f}(x), \Phi^{f(x)g}(x')), \qquad (x, x') \in \mathbf{A}^{m+m'}.$$

Let $x \in \mathbf{A}^m$. The points $(F^k(x), G^k(a))_{1 \leq k \leq q-1}$ are fixed points of $\Phi^{f \otimes g}$ because of (2.10) and (2.11). Thus

$$\Psi^k(x, a) = (F^k(x), G^k(a)), \qquad 0 \leq k \leq q - 1.$$

But for the qth iteration, (2.10) and (2.11) yield $\Psi^q(x, a) = \Phi^{f \otimes g}(F^q(x), a) = (\Phi^f(F^q(x)), a)$. $\qquad \square$

Notice that in fact, when F and G are diffeomorphisms, for all $k \in \mathbf{Z}$,

$$\Psi^k(x, a) = (F^{k''} \circ (\Phi^f \circ F^q)^{k'}(x), G^k(a)), \ k = k'q_j + k'', \ 0 \leq k'' \leq q_j - 1.$$

One can summarize Lemma 2.1 by saying that the submanifold $\mathbf{A}^m \times \{a\}$ is invariant under the qth iterate of Ψ, and that (with the natural identification of $\mathbf{A}^m \times \{a\}$ with $\mathbf{A}^m$), the restriction of Ψ^q to $\mathbf{A}^m \times \{a\}$ coincide with $\Phi^f \circ F^q$.

Clearly, for our purposes, we can choose $m = 1$, $F = \Phi^{\frac{1}{2}r_1^2}$ and $f = \frac{1}{q}U$ for the first factor, while the choice of the second factor and of the periodic point is a more delicate matter and will be the aim of Section 2.3. The property $\Phi^{\frac{1}{q}U}(0, r_1) = (0, r_1 + \frac{1}{q})$ will ensure the drift in the r_1-direction of the orbit of $((0, 0), a)$ (see Figure 2).

2.2.3. So, in our construction, the two-dimensional annulus $\mathbf{A} \times \{a\}$ will be invariant under Ψ^q, and the dynamics on it will coincide with that of the standard map ψ_q defined in (2.9).

More precisely, the problem of finding near-integrable unstable systems is reduced to the search for sequences $(g^{(j)})$, $(\mathrm{G}^{(j)})$, $(a^{(j)})$, (q_j) such that:

i) $\mathrm{G}^{(j)} : \mathbf{A}^{n-1} \to \mathbf{A}^{n-1}$ tends to some integrable map G_∞;
ii) $a^{(j)} \in \mathbf{A}^{n-1}$ is $\mathrm{G}^{(j)}$-periodic, of period q_j;
iii) $(g^{(j)}, \mathrm{G}^{(j)}, a^{(j)}, q_j)$ satisfies the synchronization conditions (2.10);
iv) $g^{(j)} : \mathbf{A}^{n-1} \to \mathbf{R}$ is $o(q_j)$ in the Gevrey category.

In that situation $\Psi_j = \Phi^{\frac{1}{q_j} \mathrm{U} \otimes g^{(j)}} \circ (\Phi^{\frac{1}{2}r_1^2} \times \mathrm{G}^{(j)})$ is indeed guaranteed to make $x^{(j)} = ((0,0), a^{(j)}) \in \mathbf{A} \times \mathbf{A}^{n-1}$ drift in q_j^2 iterations, and this number q_j^2 will have to be compared with the distance between Ψ_j and its limit $\Phi^{\frac{1}{2}r_1^2} \times \mathrm{G}_\infty$.

2.3. *Choice of the second factor and proof of Proposition 2.1*

As suggested by (2.5), we set

$$(\mathbf{2.12}) \qquad \mathrm{G}^{(j)} = \mathrm{G}_2^{(j)} \times \hat{\mathrm{G}}, \qquad \mathrm{G}_2^{(j)} = \Phi^{\frac{1}{2}r_2^2 + \frac{1}{N_j^2}\mathrm{V}}, \qquad \hat{\mathrm{G}} = \Phi^{\frac{1}{2}(r_3^2 + \cdots + r_n^2)}$$

(with a product of successive prime numbers for N_j, according to (2.1)). Clearly the sequence $\mathrm{G}^{(j)}$ satisfy the requirement *i)*, and the limit $\Phi^{\frac{1}{2}r_1^2} \times \mathrm{G}_\infty$ is exactly our initial twist map Φ^h. Because of the factor $\mathrm{G}_2^{(j)}$, we shall need some classical results on the flow of the pendulum, which we now recall.

2.3.1. *Some notations and results on the simple pendulum.* — Let

$$\mathrm{P}(\theta_*, r_*) = \frac{1}{2}r_*^2 - \cos 2\pi\theta_*, \qquad (\theta_*, r_*) \in \mathbf{A}.$$

The pendulum described by P has a hyperbolic fixed point at $(\frac{1}{2}, 0)$, whose separatrices are given by $r_* = \pm 2|\cos \pi\theta_*|$; all the solutions of $\mathrm{X_P}$ that avoid the separatrices are periodic, and for each integer $\mathrm{M} \geq 1$ there exists a unique real number $r_*^{(\mathrm{M})} > 2$ such that $(0, r_*^{(\mathrm{M})})$ is Φ^P-periodic of period M. The sequence $(r_*^{(\mathrm{M})})_{\mathrm{M}\geq 1}$ decreases and tends to $r_*^{(\infty)} = 2$ as $\mathrm{M} \to \infty$; the trajectory of $(0, r_*^{(\mathrm{M})})$ lies on an essential invariant circle of the pendulum.

Lemma **2.2.** — *Let* $\sigma = -\frac{1}{2} + \frac{2}{\pi}\arctan e^\pi$. *For all* $\mathrm{M} \in \mathbf{N}^* \cup \{\infty\}$, *the solution*

$$(\theta_*^{(\mathrm{M})}(t), r_*^{(\mathrm{M})}(t)) = \Phi^{t\mathrm{P}}(0, r_*^{(\mathrm{M})})$$

of $\mathrm{X_P}$ *satisfies*

$$\forall t \in \left[\tfrac{1}{2}, \mathrm{M} - \tfrac{1}{2}\right], \quad \sigma \leq \theta_*^{(\mathrm{M})}(t) \leq 1 - \sigma.$$

Proof. — This is an elementary property of the flow of the pendulum: one can use the differential equation $\dot{\theta}_* = \sqrt{2(E^{(M)} + 2\cos^2 \pi\theta_*)}$, with $E^{(M)} = -1 + P(0, r_*^{(M)})$ > 0, to compare $\theta_*^{(M)}(t)$ with the separatrix solution $\theta_*^{(\infty)}(t) = -\frac{1}{2} + \frac{2}{\pi}\arctan e^{2\pi t}$ when $t \in]-\frac{1}{2}, \frac{1}{2}[$. $\qquad\square$

For each M, the function $\theta_*^{(M)}$ is analytic and increasing. Lemma 2.2 shows in particular that we can define its inverse

(**2.13**) $\qquad \tau_*^{(M)} : [-\sigma, \sigma] \to I_M,$

where I_M is an interval contained in $]-\frac{1}{2}, \frac{1}{2}[$ if $M < \infty$ and $I_\infty = [-\frac{1}{2}, \frac{1}{2}]$. In fact, we have the explicit formula

(**2.14**) $\qquad \tau_*^{(M)}(\theta) = \int_0^\theta \dfrac{d\varphi}{\sqrt{(r_*^{(M)})^2 - 4\sin^2 \pi\varphi}}.$

Being an analytic function, $\tau_*^{(M)}$ belongs to each Gevrey class $G^{\alpha, L}([-\sigma, \sigma])$, $\alpha > 1, L > 0$.

Lemma **2.3.** — *For each $\alpha > 1$ and $L > 0$, one can define a finite positive number by*

(**2.15**) $\qquad \Lambda(\alpha, L) = \sup_{1 \le M \le \infty} \|\tau_*^{(M)}\|_{\alpha, L, [-\sigma, \sigma]}.$

Proof. — Choose $\ell > 0$ and $A \in]0, 1[$ such that

$$\forall \theta \in \mathbf{C}, \quad \mathrm{dist}(\theta, [-\sigma, \sigma]) \le \ell \;\Rightarrow\; |\sin \pi\theta| \le A.$$

In view of (2.14), for such values of θ, $\frac{d}{d\theta}\tau_*^{(M)}$ has a modulus $\le \frac{1}{2\sqrt{1-A^2}}$ (because $r_*^{(M)} \ge 2$), and the Cauchy inequalities yield

$$\left\|\left(\tfrac{d}{d\theta}\right)^{k+1}\tau_*^{(M)}\right\|_{C^0([-\sigma, \sigma])} \le \dfrac{k!}{2\ell^k\sqrt{1-A^2}}$$

for all $k \in \mathbf{N}$. The conclusion follows easily. $\qquad\square$

2.3.2. *Choice of the periodic points $a^{(j)}$.* — We now define M_j and $q_j = N_j M_j$ according to (2.3), using $c = \|\mathbf{F}_{\alpha, (4\sigma)^{-1/\alpha}L}\|_{\alpha, (4\sigma)^{-1/\alpha}L}$ and $\Lambda = \Lambda(\alpha, L)$ (with the notations of Lemmas A.3 and 2.3).

The transformation $G_2^{(j)}$ is obtained from the flow of the fixed pendulum system P by rescaling both time and action:

(**2.16**) $\qquad G_2^{(j)} = S^{-1} \circ \Phi^{\frac{1}{N_j}P} \circ S, \quad \text{where} \quad S(\theta_2, r_2) = (\theta_*, r_*) = (\theta_2, N_j r_2).$

There is thus a one-to-one correspondence between the solutions $(\theta_*(t), r_*(t))$ of X_P and those of $X_{\frac{1}{2}r_2^2 + \frac{1}{N_j}V}$:

$$\theta_2(t) = \theta_*(t/N_j), \quad r_2(t) = \tfrac{1}{N_j} r_*(t/N_j),$$

and if we set

$$(\mathbf{2.17}) \qquad x_2^{(j)} = (0, r_2^{(j)}) \in \mathbf{A}, \quad r_2^{(j)} = \frac{r_*^{(M_j)}}{N_j},$$

we obtain a $G_2^{(j)}$-periodic point, of period $N_j M_j = q_j$. Moreover its orbit is well-ordered on the annulus.

As for the remaining components, we set

$$\hat{r}^{(j)} = \left(\frac{1}{p_{j-(n-3)}}, ..., \frac{1}{p_j} \right), \quad \hat{x}^{(j)} = (0, \hat{r}^{(j)}) \in \mathbf{A}^{n-2} = \mathbf{T}^{n-2} \times \mathbf{R}^{n-2}.$$

Since its action components are inverses of mutually prime integers, $\hat{x}^{(j)}$ is a $\hat{G}$-periodic point, of period $N_j' = p_{j-(n-3)}...p_j$. Since N_j' divides q_j, the requirement $ii)$ is satisfied by $G^{(j)}$, $a^{(j)} = (x_2^{(j)}, \hat{x}^{(j)})$ and q_j.

2.3.3. *Choice of $g^{(j)}$ and Gevrey estimates.* — It only remains for us to define $g^{(j)}$ so as to fulfill $iii)$, together with the refined condition (2.4), which in turn gives $iv)$. The idea is that, even taking q_j very large as we did by our choice of M_j, we still have most of the points of $O(a^{(j)})$ outside the region $\{-\sigma < \theta_2 < \sigma\}$ and these can be handled by a compact-supported function; therefore, among the q_j points of $O(a^{(j)})$, only a relatively small number ($\leq N_j$) will need a special treatment.

Indeed, we can provide the following description of our q_j-periodic orbit (see also Figure 3):

- Section 2.3.2 shows that the θ_2-component of $(G^{(j)})^k(a^{(j)})$ coincides with $\theta_*^{(j)}(\frac{k}{N_j})$ for each $k \in \mathbf{Z}$. As a consequence, if we number the iterates by indices k running between $-\frac{N_j-1}{2}$ and $N_j M_j - \frac{N_j+1}{2}$ (notice that N_j is always odd) and decompose the orbit of $a^{(j)}$ into M_j sets of N_j points, Lemma 2.2 ensures that the only points having possibly $\theta_2 \in]-\sigma, \sigma[$ belong to the first of these sets—*i.e.* they correspond to $k \in \{-\frac{N_j-1}{2}, ..., \frac{N_j-1}{2}\}$.
- The first set itself can be decomposed into p_j subsets of N_j' points, in each of which only one point has $\hat{\theta} = 0 \bmod \mathbf{Z}^{n-3}$ (because N_j' divides k) whereas the $N_j' - 1$ remaining ones have $\hat{\theta} = k\hat{r}^{(j)} \neq 0 \bmod \mathbf{Z}^{n-3}$ (because N_j' does not divide k).

We shall construct $g^{(j)}$ as a product of factors taking care of these different kinds of iterates successively (formula (2.18) below).
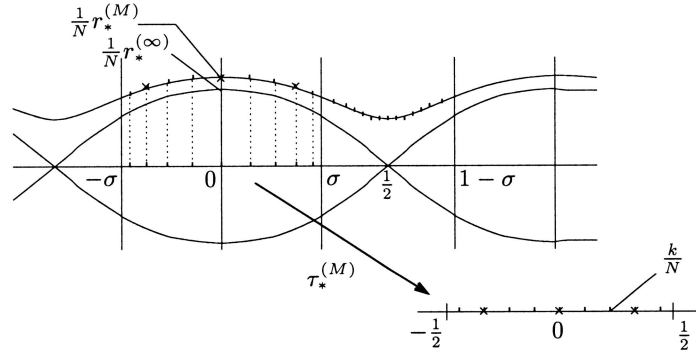
FIG. 3. — View of the orbit of $a^{(j)}$ in the (θ_2, r_2) coordinates. The iterates inside $\{-\sigma < \theta_2 < \sigma\}$ necessarily correspond to $-\frac{N_j-1}{2} \le k \le \frac{N_j-1}{2}$ (among them, those with $N'_j \mid k$ are marked by a cross); under the action of $\tau_*^{(M_j)}$, we get equidistributed points inside $]-\frac{1}{2}, \frac{1}{2}[$.

The following lemma introduces some auxiliary functions and shows where the exponential estimates ultimately come from.

Lemma **2.4.** — *For $p \ge 1$, the function $\eta_p : \mathbf{T} \to \mathbf{R}$ defined by*

$$\eta_p(\theta) = \Big(\frac{1}{p}\sum_{\ell=0}^{p-1} \cos 2\pi\ell\theta\Big)^2, \qquad \theta \in \mathbf{T},$$

satisfies

$$\eta_p(0) = 1, \quad \eta_p'(0) = 0, \quad \eta_p(k/p) = \eta_p'(k/p) = 0, \qquad 1 \le k \le p-1.$$

For all $\alpha > 1$ and $L > 0$,

$$\|\eta_p\|_{\alpha,L} \le e^{2\alpha L(2\pi p)^{1/\alpha}}.$$

Proof. — The proof of the first statement is straightforward. As for the Gevrey norm, using Lemma A.1, we can content ourselves with estimating the norm of the functions

$$\xi_p(\theta) = \frac{1}{p}\sum_{\ell=0}^{p-1} \cos 2\pi\ell\theta, \qquad \theta \in \mathbf{T},$$

for all $p \ge 1$. But for all $k \ge 0$, $\|\xi_p^{(k)}\|_{C^0(\mathbf{T})} < (2\pi p)^k$, thus

$$\forall \alpha \ge 1, \ \forall L > 0, \quad \|\xi_p\|_{\alpha,L} < \sum_{k \ge 0}\frac{1}{k!^\alpha}(2\pi L^\alpha p)^k < e^{\alpha L(2\pi p)^{1/\alpha}}.$$

The result, follows since $\|\eta_p\|_{\alpha,L} \le \|\xi_p\|_{\alpha,L}^2$. $\qquad\square$

In view of the above description of $O(a^{(j)})$, we thus choose the function

$$(\textbf{2.18}) \qquad g^{(j)} = g_2^{(j)} \otimes g_3^{(j)} \otimes \cdots \otimes g_n^{(j)},$$

with

$$g_2^{(j)}(\theta_2) = \eta_{p_j}\big(\tau_*^{(M_j)}(\theta_2)\big) \cdot \mathbf{F}_{\alpha,(4\sigma)^{-1/\alpha}L}(\tfrac{\theta_2}{4\sigma}) \quad \text{for } -\tfrac{1}{2} \le \theta_2 \le \tfrac{1}{2} \,,$$

$$g_i^{(j)}(\theta_i) = \eta_{p_{j-(n-i)}}(\theta_i) \quad \text{for } 3 \le i \le n \,.$$

Thanks to $\mathbf{F}_{\alpha,(4\sigma)^{-1/\alpha}L}$, which belongs to the class of compact-supported functions introduced in Lemma A.3 (extended here by 1-periodicity), our function $g^{(j)}$ vanishes identically outside $\{-\sigma < \theta_2 < \sigma\}$. Its behaviour at the points $(G^{(j)})^k(a^{(j)}) = (\theta^{(k)}, r^{(k)})$ satisfies the condition $iii)$: we have obviously $g^{(j)}(a^{(j)}) = 1$ and $dg^{(j)}(a^{(j)}) = 0$, and $g^{(j)}$ and its differential vanish on the other iterates because this is the case for at least one of the functions $g_i^{(j)}$. Indeed, for any non-zero index k between $-\frac{N_j-1}{2}$ and $N_jM_j - \frac{N_j+1}{2}$, we have

- either $\sigma \le \theta_2^{(k)} \le 1 - \sigma$ and $g_2^{(j)}$ vanishes with its differential at $\theta_2^{(k)}$ because this is the case for $\mathbf{F}_{\alpha,(4\sigma)^{-1/\alpha}L}$,
- or $-\sigma < \theta_2^{(k)} < \sigma$, and then necessarily $-\frac{N_j-1}{2} \le k \le \frac{N_j-1}{2}$, thus $\tau_*^{(M_j)}(\theta_2^{(k)}) = k/N_j \in\,] -\tfrac{1}{2}, \tfrac{1}{2}[$, whereas $\hat{\theta}^{(k)} = k\hat{r}^{(j)}$, and
  - either N_j' does not divide k and at least one of the functions $g_i^{(j)}$ with $i \ge 3$ vanishes with its differential,
  - or $k = k'N_j'$ for some integer k', and $g_2^{(j)}$ vanishes with its differential at $\theta_2^{(k)}$ because this is the case for η_{p_j} at $k/N_j = k'/p_j$.

To end the proof of Proposition 2.1, we just need to check the inequality (2.4). Recalling the definitions of c and Λ given at the beginning of Section 2.3.2 and using Lemma A.1, we get

$$\|g^{(j)}\|_{\alpha,L} \le c \,\|\eta_{p_j} \circ \tau_*^{(M_j)}\|_{\alpha,L}\|\eta_{p_{j-(n-3)}}\|_{\alpha,L}...\|\eta_{p_j}\|_{\alpha,L}$$

with $\|\tau_*^{(M_j)}\|_{\alpha,L} \le \Lambda$. Proposition A.1 allows us to bound $\|\eta_{p_j} \circ \tau_*^{(M_j)}\|_{\alpha,L}$ by $\|\eta_{p_j}\|_{\alpha,\Lambda}$, and Lemma 2.4 then yields $\|g^{(j)}\|_{\alpha,L} \le c \exp\big(2\alpha(\Lambda + (n-2)L)(2\pi p_j)^{1/\alpha}\big) \le M_j/N_j$. Dividing by $q_j = N_jM_j$, we obtain the desired conclusion.

Remark **2.1.** — In fact, it is only at the end that the advantages of the time-1 map $G_2^{(j)}$ of the pendulum and of the twist map $\hat{G}$ appear simultaneously. On the one hand, the equidistribution of the points on the orbit of $\hat{a}^{(j)}$ under $\hat{G}$ allowed us for a very precise control, through the functions $g_i^{(j)}$; their Gevrey norm (or even their analytic norm, since one could set $\alpha = 1$ as far as they alone are concerned) grows exponentially with respect to p_j, but at a rate which decreases with the dimension. On

the other hand, since we needed to divide $g^{(j)}$ by an exponentially large factor so as to make the perturbation $O(1/N_j^2)$, we had to admit larger and larger periods without changing $\|g^{(j)}\|_{\alpha,L}$, and this was the reason for using the pendulum: as M_j goes to infinity, the points different from $(0, r_*^{(j)})$ on its orbit under Φ^P concentrate near the hyperbolic fixed point, still remaining at a finite distance of $(0, r_*^{(j)})$ itself (but here we needed a function vanishing on an interval, hence we could not afford for an analytic function).

Remark **2.2.** — Our claims of § 2.1.4 can now be justified. The reader will convince himself of the accuracy of the description of the orbit of $x^{(j)}$ which was given there. As for the other unstable points obtained by modifying the definition of $\hat{r}^{(j)}$ when $n \geq 3$, this simply amounts to the fact that the new $a^{(j)}$ still satisfies the requirement *ii)* (because the corresponging $\hat{x}^{(j)}$ is still N_j'-periodic under $\hat{G}$) and that $g^{(j)}$ does not depend on the action-variables, thus the requirement *iii)* is still satisfied.

In order to make all the points (2.8) drift, it suffices to perform the following modifications: replace U by

$$U^{(j)}(\theta_1) = -\frac{1}{2\pi p_j} \sin(2\pi p_j \theta_1)$$

and, since this yields an exponentially large function, compensate this by enlarging M_j and thus q_j according to the formula

$$M_j = 2[cN_j \, e^{2\alpha(\Lambda+(n-1)L)(2\pi p_j)^{1/\alpha}} + 1], \qquad q_j = N_j M_j$$

(this maintains the inequality (2.7) because $\|U^{(j)}\|_{\alpha,L} \leq \exp(\alpha L(2\pi p_j)^{1/\alpha})$).

2.4. *Suspension*

We have just described, for arbitrary $\alpha > 1$ and $L > 0$, a discrete dynamical system on $\mathbf{A}^n$ which is of the form $\Phi^u \circ \Phi^{h+v}$, for $h(r) = \frac{1}{2} < r, r >$ and some Gevrey-(α, L) functions $u = u(x)$ and $v = v(x)$, where $x = (\theta, r) \in \mathbf{A}^n = \mathbf{T}^n \times \mathbf{R}^n$ (in fact, u and v do not depend on r in our example). Our aim is now to construct a near-integrable Hamiltonian system $\mathscr{H}$ on $\mathbf{A}^{n+1}$ such that $\Phi^u \circ \Phi^{h+v}$ coincides with the return map of $\Phi^{\mathscr{H}}$, relatively to a suitable $2n$-dimensional section.

We fix $n \geq 1$, $\alpha > 1$, $R > 1$. To begin with, given the function $h(r) = \frac{1}{2} < r, r >$ on $\mathbf{A}^n$ and two arbitrary functions u and v of $G^{\alpha,L}(\mathbf{T}^n \times \overline{B}_R)$, we want to find a non-autonomous Hamiltonian function

$$H = H(x, t), \qquad (x, t) \in \mathbf{A}^n \times \mathbf{T},$$

such that

(**2.19**) $\qquad \Phi^H = \Phi^u \circ \Phi^{h+v}$.

In the left-hand side of (2.19), by Φ^{H} we mean the flow mapping between $t = 0$ and $t = 1$ for the non-autonomous Hamiltonian vector field $\mathrm{X_H}$.

Moreover, we require the function H to be Gevrey-$(\alpha, \mathrm{L_1})$ when restricted to $\mathbf{T}^n \times \overline{\mathrm{B}}_{\mathrm{R}} \times \mathbf{T}$ and to tend to h (considered as a function on $\mathbf{A}^n \times \mathbf{T}$) as $u, v \to 0$. For technical reasons, we shall choose L and $\mathrm{L_1}$ related by

$$(\mathbf{2.20}) \qquad \mathrm{L}^\alpha = \mathrm{L}_1^\alpha \Big(1 + \Big(\mathrm{L}_1^\alpha + \mathrm{R} + \frac{1}{2} \Big) \| \varphi \|_{\alpha, \mathrm{L_1}} \Big),$$

where φ is defined by $\varphi(t) = \dfrac{\mathbf{F}_{\alpha, \mathrm{L_1}}(t)}{\int_{\mathbf{T}} \mathbf{F}_{\alpha, \mathrm{L_1}}}$ for $-\frac{1}{2} \le t \le \frac{1}{2}$ and extended by 1-periodicity. Here $\mathrm{L_1}$ will be an arbitrary positive number, and we shall use the functions u and v associated with the number L defined by (2.20).

Notice that we make use of the Gevrey-$(\alpha, \mathrm{L_1})$ compact-supported function $\mathbf{F}_{\alpha, \mathrm{L_1}}$ introduced in Lemma A.3, and thus need the hypothesis $\alpha > 1$; see [Ku93] and [KP94] (or [PT97]) for a more elaborate approach in the analytic case, which we could have tried to adapt to our case.

2.4.1. *The suspension lemma.* — A simple way of fulfilling (2.19) is to set

$$\mathrm{H}^* = (h + v) \otimes \varphi_0 + u \otimes \varphi_1,$$

where the functions φ_0 and φ_1 depend only on t, φ_0 has compact support $\subset [0, \frac{1}{2}]$ and total mass 1, and φ_1 has compact support $\subset [\frac{1}{2}, 1]$ and total mass 1. For instance, one can choose $\varphi_0(t) = \varphi(t - \frac{1}{4})$ and $\varphi_1(t) = \varphi(t - \frac{3}{4})$, with φ as above.

Observe that we have $\Phi^{\mathrm{H}^*} = \Phi^u \circ \Phi^{h+v}$, because the flow between $t = 0$ and $t = \frac{1}{2}$ along $\mathrm{X_{H^*}}$ yields Φ^{h+v} and its flow between $t = \frac{1}{2}$ and $t = 1$ yields Φ^u. But the limit of H^* as $u, v \to 0$ is $h \otimes \varphi_0$ rather than h.

Other Hamiltonian functions satisfying (2.19) can be deduced from H^*. For instance, if $x^* = \mathscr{F}_t(x)$ is an exact-symplectic transformation of $\mathbf{A}^n$ which depends periodically on t and which is reduced to identity for $t = 0$, the flows between $t = 0$ and $t = 1$ will be the same for $\mathrm{X_{H^*}}$ in coordinates $x^* = (\theta^*, r^*)$ and for the conjugate vector field in coordinates $x = (\theta, r)$. If $\mathscr{F}_t$ admits a generating function $\mathrm{S}_t(\theta, r^*)$, the conjugate vector field is $\mathrm{X_H}$, with

$$\mathrm{H}(\theta, r, t) = \mathrm{H}^*(\theta^*, r^*, t) - \frac{\partial \mathrm{S}_t}{\partial t}(\theta, r^*), \qquad \mathscr{F}_t(\theta, r) = (\theta^*, r^*).$$

Applying this to

$$\tilde{\varphi}_0(t) = \int_0^t (\varphi_0(t_1) - 1) \, dt_1, \quad \mathrm{S}_t(\theta, r^*) = \langle \theta, r^* \rangle + \tilde{\varphi}_0(t) h(r^*),$$

$$\mathscr{F}_t(\theta, r) = (\theta + \tilde{\varphi}_0(t) \nabla h(r), r),$$

we obtain

$$\Phi^{h+f} = \Phi^u \circ \Phi^{h+v}, \qquad f(x,t) = \varphi_0(t)\, u \circ \mathscr{F}_t(x) + \varphi_1(t)\, v \circ \mathscr{F}_t(x).$$

We must furthermore bound the Gevrey-(α, L_1) norm of f in terms of the Gevrey-(α, L) norms of u and v. This can be done using Proposition A.1, but it is better to make use of the peculiar form of $\mathscr{F}_t$, with $\nabla h(r) = r$, as indicated in Remark A.1. The inequality (A.8) is satisfied since $\|\tilde{\varphi}_0\|_{\mathrm{C}^0} \le \frac{1}{2}\|\varphi_0 - 1\|_{\mathrm{C}^0} \le \frac{1}{2}\|\varphi_0\|_{\mathrm{C}^0}$ (because $\varphi_0 \ge 0$) and $\|\tilde{\varphi}_0^{(m+1)}\|_{\mathrm{C}^0} = \|\varphi_0^{(m)}\|_{\mathrm{C}^0}$, thus $(\mathrm{L}_1^\alpha + \mathrm{R})\|\tilde{\varphi}_0\|_{\alpha,\mathrm{L}_1} - \mathrm{R}\|\tilde{\varphi}_0\|_{\mathrm{C}^0} \le (\mathrm{L}_1^\alpha + \mathrm{R})\mathrm{L}_1^\alpha\|\varphi_0\|_{\alpha,\mathrm{L}_1} + \frac{1}{2}\mathrm{L}_1^\alpha\|\varphi_0\|_{\mathrm{C}^0} \le \mathrm{L}^\alpha - \mathrm{L}_1^\alpha$ by our choice of L. We obtain

$$\|f\|_{\alpha,\mathrm{L}_1} \le \|\varphi_0\|_{\alpha,\mathrm{L}_1}\|u\|_{\alpha,\mathrm{L}} + \|\varphi_1\|_{\alpha,\mathrm{L}_1}\|v\|_{\alpha,\mathrm{L}}.$$

Therefore $f \in \mathrm{G}^{\alpha,\mathrm{L}_1}(\mathbf{T}^n \times \overline{\mathrm{B}}_\mathrm{R} \times \mathbf{T})$ and, setting $\mathrm{C} = \|\varphi\|_{\alpha,\mathrm{L}_1}$,

$$\|f\|_{\alpha,\mathrm{L}_1} \le \mathrm{C}(\|u\|_{\alpha,\mathrm{L}} + \|v\|_{\alpha,\mathrm{L}}).$$

As for lower bounds, we can simply remark that, since the functions φ_0 and φ_1 have disjoint supports and satisfy $\varphi_0(1/4) = \varphi_1(3/4) = 1$,

$$\|f\|_{\alpha,\mathrm{L}_1} \ge \max(\|u\|_{\mathrm{C}^0}, \|v\|_{\mathrm{C}^0}).$$

We can now state our definitive result, for which we prefer to introduce an *autonomous* Hamiltonian system on $\mathbf{A}^{n+1}$:

Lemma 2.5. — *Let* $\alpha > 1$, $\mathrm{L}_1 > 0$, $n \ge 1$, $\mathrm{R} > 1$. *Given two functions* $u, v \in \mathrm{G}^{\alpha,\mathrm{L}}(\mathbf{T}^n)$, *with* L *defined by (2.20), there exists* $f \in \mathrm{G}^{\alpha,\mathrm{L}_1}(\mathbf{T}^{n+1} \times \overline{\mathrm{B}}_\mathrm{R}^{n+1})$, *independent of the last action variable* r_{n+1}, *such that if*

$$\mathscr{H}(\theta, r) = \frac{1}{2}(r_1^2 + \cdots + r_n^2) + r_{n+1} + f(\theta, r), \qquad (\theta, r) \in \mathbf{A}^{n+1},$$

the diffeomorphism $\Phi^u \circ \Phi^{\frac{1}{2}<r,r>+v}$ *coincides with the Poincaré map induced by the Hamiltonian flow of* $\mathscr{H}$ *on the section* $\{\theta_n = 0\} \cap \mathscr{H}^{(-1)}(\{e\})$, *for each energy* $e \in \mathbf{R}$.

Moreover, there exists a positive number C, *depending only on* α *and* L_1, *such that*

$$\max(\|u\|_{\mathrm{C}^0}, \|v\|_{\mathrm{C}^0}) \le \|f\|_{\alpha,\mathrm{L}_1} \le \mathrm{C}(\|u\|_{\alpha,\mathrm{L}} + \|v\|_{\alpha,\mathrm{L}}).$$

2.4.2. *Proof of Theorem B.* — Here we have to shift the number of degrees of freedom from n to $n-1$ in Proposition 2.1. Let $n \ge 3$, $\alpha > 1$, $\mathrm{L}_1 > 0$ and $\mathrm{R} > 1$. Proposition 2.1, with data $n-1$, α, and L defined by (2.20), yields a diffeomorphism $\Psi_j = \Phi^{u_j} \circ \Phi^{h+v_j}$ of the annulus $\mathbf{A}^{n-1}$, with a point $x^{(j)}$ drifting between $r_1 = 0$ and $r_1 = 1$ in q_j^2 iterates. Recall that $u_j = \frac{1}{q_j}\mathrm{U} \otimes g^{(j)}$ and $v_j = \frac{1}{\mathrm{N}_j^2}\mathrm{V}$. By Lemma 2.5, one

can find a function $f_j \in G^{\alpha,L_1}(\mathbf{T}^n \times \overline{B}_R)$, independent of r_n, such that Ψ_j is the return map of the flow generated by

$$\mathscr{H}_j(\theta, r) = \frac{1}{2}(r_1^2 + \cdots + r_{n-1}^2) + r_n + f_j(\theta, r), \qquad (\theta, r) \in \mathbf{A}^n.$$

Moreover, if we define $\varepsilon_j = \|f_j\|_{\alpha,L_1}$, we have

$$(\mathbf{2.21}) \qquad \frac{1}{N_j^2} \leq \varepsilon_j \leq C(\|u_j\|_{\alpha,L} + \|v_j\|_{\alpha,L}) \leq C'\frac{1}{N_j^2}$$

with $C' = 2C\|V\|_{\alpha,L}$ for instance (thanks to (2.7) and because $\|v_j\|_{C^0} = \frac{1}{N_j^2}$).

The initial condition $(x^{(j)}, (0,0)) \in \mathbf{A}^{n-1} \times \mathbf{A}$ gives rise to a solution $(x(t),$ $(\theta_n(t), r_n(t)))$ of $X_{\mathscr{H}_j}$ (with $\theta_n(t) = t$) such that, for all $k \in \mathbf{Z}$, $x(k) = \Psi_j^k(x^{(j)})$. Let us now estimate the instability time $\tau_j = q_j^2$ in terms of ε_j.

If $n \geq 4$, thanks to the Prime Number Theorem we can ensure that the $n - 3$ numbers $p_{j-(n-4)}, ..., p_j$ lie in the interval $[\frac{1}{2}p_j, p_j]$ for $j \geq J$, where the integer J depends only on n; thus their product N_j' belongs to the interval $[2^{-(n-3)}p_j^{n-3}, p_j^{n-3}]$ and $N_j = p_j N_j'$ satisfies

$$N_j^{\frac{1}{n-2}} \leq p_j \leq 2N_j^{\frac{1}{n-2}}.$$

If $n = 3$, the previous inequality remains true. In view of (2.3), this implies in all cases

$$2cN_j^2 \exp(\Delta N_j^{\frac{1}{(n-2)\alpha}}) < q_j = N_j M_j < 3cN_j^2 \exp(2\Delta N_j^{\frac{1}{(n-2)\alpha}})$$

with $\Delta = 2\alpha(2\pi)^{1/\alpha}(\Lambda + (n-3)L)$ (the numbers c and Λ depend only on α and L defined by (2.20), thus c and Δ depend only on n, α, L_1, R). Finally, using (2.21), we get

$$\frac{4c^2}{\varepsilon_j^2} \exp(2\Delta\left(\frac{1}{\varepsilon_j}\right)^{\frac{1}{2(n-2)\alpha}}) \leq q_j^2 \leq \frac{9c^2C'^2}{\varepsilon_j^2} \exp(4\Delta C'^{\frac{1}{2(n-2)\alpha}}\left(\frac{1}{\varepsilon_j}\right)^{\frac{1}{2(n-2)\alpha}})$$

and Theorem B is proved.

2.4.3. *Influence zone of some resonances.* — We keep in this paragraph the notation of the previous one concerning the number of degrees of freedom, in view of comparing more precisely the instability time with what could be expected from the Addendum to Theorem A, as announced in § 1.3.3.

We recall that $h(r) = \frac{1}{2}(r_1^2 + \cdots + r_{n-1}^2) + r_n$, thus

$$\nabla h(r) = (r_1, ..., r_{n-1}, 1).$$

We fix n, α, R, R_0, L, and then E, ϖ, m so that the hypotheses of Theorem A are fulfilled.

a) *The resonant surface $\{r_1 = r_2 = 0\}$*

The number $r_*^{(\mathrm{M})}$ was defined in § 2.3.1 by

$$\int_0^1 \frac{d\varphi}{\sqrt{(r_*^{(\mathrm{M})})^2 - 4\sin^2 \pi\varphi}} = \mathrm{M}.$$

It is easily checked that, for $1 \leq \mathrm{M} < \infty$, $2 < r_*^{(\mathrm{M})} \leq 3$, and

$$2\sqrt{\varepsilon_j} < r_2^{(j)} \leq 3\sqrt{\varepsilon_j}, \qquad r_2^{(j)} \sim 2\sqrt{\varepsilon_j} \text{ as } j \to \infty.$$

Let $\mathrm{S} = \{r \in \overline{\mathrm{B}}_\mathrm{R} \mid r_1 = r_2 = 0\}$: this is the resonant surface associated with the standard submodule $\mathbf{Z}^2 \times \{0\}$ of $\mathbf{Z}^n$, and our unstable orbit starts in the domain of stability related to this double resonance. One can thus take $\sigma = 3$ and apply the Addendum to Theorem A. This shows that the value $1/2(n-2)\alpha$ of the instability exponent a^* in Theorem B is optimal for that kind of orbit, as announced in § 1.3.3.

b) *The resonant surfaces $\{r_1 = r_2 = 0$ and $r_i = \hat{\omega}_i,\ i \in \mathrm{I}\}$, $\mathrm{I} \subset \{3, ..., n-1\}$*

From now on we suppose $n \geq 4$ and we test the Addendum to Theorem A by considering resonances of codimension $d < n-2$.

For instance we observe that $\hat{r}_i = 1/p_{j-(n-1-i)} \to 0$ as $j \to \infty$, hence our unstable orbit is close to the resonant surface $\mathrm{S} = \{r_1 = r_2 = 0$ and $r_i = 0,\ 3 \leq i \leq n-1\}$. We can also use the remark of § 2.1.4 and choose an arbitrary vector $\hat{\omega} \in [0, 1[^{n-3}$: using the drifting point (2.8) with $\ell_i^{(j)} = [\hat{\omega}_i\, p_{j-(n-1-i)}] + 1$, we notice that $\frac{1}{p_{j-(n-1-i)}} \leq \frac{\ell_i^{(j)}}{p_{j-(n-1-i)}} - \hat{\omega}_i < \frac{2}{p_{j-(n-1-i)}}$.

In any case, we see that our unstable initial condition can be chosen $\mathrm{O}(1/p_j)$-close to any resonant surface of the form

$$\mathrm{S} = \{r_1 = r_2 = 0 \text{ and } r_i = \hat{\omega}_i,\ i \in \mathrm{I}\},$$

where $\mathrm{I} = \{i_1, ..., i_m\}$ is a non-empty subset of $\{3, ..., n-1\}$ and $(\hat{\omega}_{i_1}, ..., \hat{\omega}_{i_m}, 1)$ is totally resonant. The corresponding resonant submodule is isomorphic to $\mathbf{Z}^2 \times \mathbf{Z}^\mathrm{I}$, its codimension d satisfies $1 \leq d = n-2-\mathrm{card\,I} \leq n-3$, and the distance to the resonant surface $\mathrm{dist}(r^{(j)}(0), \mathrm{S})$ tends to 0 as $j \to \infty$, but this distance is of the order of $1/p_j$, which is not $\mathrm{O}(\sqrt{\varepsilon_j})$, thus one cannot apply the Addendum to Theorem A and it is not surprising that the time of instability q_j^2 be much shorter than $\exp(\mathrm{const}\, (\frac{1}{\varepsilon_j})^{\frac{1}{2d\alpha}})$.

c) *The resonant surfaces $\{r_1 = r_2 = 0$ and $r_i = \hat{r}_i^{(j)},\ i \in \mathrm{I}\}$, $\mathrm{I} \subset \{3, ..., n-1\}$*

One can finally examine the case of variable resonant surfaces, *i.e.* which depend on j. Let I be a non-empty subset of $\{3, ..., n-1\}$ and consider

$$\mathrm{S}_{\mathscr{M}^{(j)}} = \{r_1 = r_2 = 0 \text{ and } r_i = \hat{r}_i^{(j)},\ i \in \mathrm{I}\},$$

where $\mathcal{M}^{(j)}$ is the submodule of $\mathbf{Z}^n$ generated by $(1, 0, 0, ..., 0)$, $(0, 1, 0, ..., 0)$ and the vectors $k_i^{(j)} \in \mathbf{Z}^n$ defined by

$$k_i^{(j)} = (0, ..., 0, \underset{(i^{th} \text{ place})}{p_{j-(n-1-i)}}, 0, ..., 0, -1), \qquad i \in \mathrm{I}.$$

This submodule has codimension is $d = n - 2 - \mathrm{card}\,\mathrm{I}$, $1 \leq d \leq n - 3$, and we must attach to it three numbers c_j, c_j', c_j'' according to § 3.4.3. Computations of the same kind as in the proof of Lemma 3.16 show that $c_j \simeq c_j'' \simeq p_j^{\mathrm{card}\mathrm{I}}$ and $c_j' \simeq p_j$. Since $\mathrm{dist}(r^{(j)}(0), \mathrm{S}_{\mathcal{M}^{(j)}}) = 0$, we may wish to apply the Addendum to Theorem A (with any positive σ), but the involved thresholds $\varepsilon_*^{(j)}$ and $\varepsilon_{**}^{(j)}$ will now depend on j.

If $\mathrm{I} \neq \{3, ..., n-1\}$, $d \geq 2$, then $\varepsilon_{**}^{(j)} = o(\varepsilon_j)$ (because $\varepsilon_2^{(j)}$ is of the order of $(p_j^{d-1}/\mathrm{N}_j)^{2d}$): we cannot apply the Addendum because ε_j is too large.

If $\mathrm{I} = \{3, ..., n-1\}$, $d = 1$, then the argument of the exponential for the stability time is

$$\alpha \left(\frac{\varepsilon_*^{(j)}}{\varepsilon_j} \right)^{1/\alpha} \simeq \mathrm{const}\, p_j^{1/\alpha},$$

which is consistent with the observed instability time q_j^2, but one can check that in that case too ε_j is too large: $\varepsilon_*^{(j)} \leq \mathrm{const}\, \varepsilon_j$ with $0 < \mathrm{const} < 1$.

2.5. *Relations with Arnold's mechanism*

2.5.1. We come back to the notations of Sections 2.1–2.3 concerning the number of degrees of freedom.

The system that we have constructed has many similarities with the classical Arnold's example [Arn64]: the unperturbed system $\Phi^h = \Phi^{\frac{1}{2}r_1^2 + \cdots + \frac{1}{2}r_n^2}$ depends only on the action variables; the perturbation $\frac{1}{\mathrm{N}^2}\mathrm{V}(\theta_2)$, added to h, still preserves the complete integrability (in the usual sense of symplectic geometry) and introduces hyperbolic objects; finally the perturbation $\Phi^{\frac{1}{q}\mathrm{U}\otimes g}$ breaks the complete integrability and produces the drift in action space. One can recognize the effects of the parameters ε and μ in Arnold's example. Arnold's idea was to deduce the existence of drifting points from that of a transition chain formed by heteroclinically connected hyperbolic tori. This led to the problem of finding lower bounds for the homoclinic splitting of separatrices, a question which cannot yet be considered as solved for analytic systems (see [LMS03] for instance); and even in the case where a transition chain is given, some non trivial work remains to be done in order to prove the existence of drifting points and evaluate their speed, using either purely dynamical or variational methods.

It turns out that these questions can be easily investigated in the case of our example of unstable near-integrable Gevrey maps, or at least in a slight modification

of this example: the drifting points we have exhibited in the previous sections shadow chains of transition tori for which the homoclinic splitting can be exactly computed. Our systems will in fact appear as optimal from the point of view of Arnold's mechanism, since the chains we consider are formed by *maximally distant* hyperbolic tori: the stable and unstable manifolds of two consecutive tori are *tangent* along the heteroclinic orbits connecting them, and the intersection breaks down as soon as the distance between them is increased.

Finally, the results of Section 2.4 allow us to recover the Hamiltonian formalism, and the Addendum to Theorem B.

2.5.2. We could have worked directly with the previous system Ψ_j, but in order to simplify the computation of the heteroclinic orbits for our chains it is more convenient to replace it by a slightly different system $\tilde{\Psi}_j$: for each $j \geq n - 1$, we modify the definition of M_j and q_j in (2.3) by replacing Λ by 2Λ inside the exponential, and we replace the function $g^{(j)}$ of (2.18) by $\tilde{g}^{(j)} = \tilde{g}_2^{(j)} \otimes g_3^{(j)} \otimes \cdots \otimes g_n^{(j)}$, with

$$\tilde{g}_2^{(j)}(\theta_2) = g_2^{(j)}(\theta_2)\, \eta_{p_j}\big(\tau_*^{(\infty)}(\theta_2)\big),$$

where η_{p_j} and $\tau_*^{(\infty)}$ were defined in Lemma 2.4 and formula (2.13).

The requirements *ii)* and *iii)* of § 2.2.3 are still satisfied with $a^{(j)} = (x_2^{(j)}, \hat{x}^{(j)})$ defined in the same way as previously, and like at the end of Section 2.3 we obtain $\|\tilde{g}^{(j)}\|_{\alpha,\mathrm{L}} \leq M_j/N_j$, which yields (2.4), because $\|\eta_{p_j} \circ \tau_*^{(\infty)}\|_{\alpha,\mathrm{L}} \leq \|\eta_{p_j}\|_{\alpha,\Lambda}$.

Therefore the conlusions of Proposition 2.1 remain valid for the modified transformation

$$(\textbf{2.22}) \qquad \tilde{\Psi}_j = \Phi^{\frac{1}{q_j}\mathrm{U}\otimes\tilde{g}^{(j)}} \circ \Phi^{\frac{1}{2}(r_1^2+r_2^2+\cdots+r_n^2)+\frac{1}{\mathrm{N}_j^2}\mathrm{V}(\theta_2)} = \Phi^{\frac{1}{q_j}\mathrm{U}\otimes\tilde{g}^{(j)}} \circ \big(\Phi^{\frac{1}{2}r_1^2} \times \mathrm{G}_2^{(j)} \times \hat{\mathrm{G}}\big).$$

As above, the q_j^{th}-power $(\tilde{\Psi}_j)^{q_j}$ leaves the annulus $\mathbf{A} \times \{a^{(j)}\}$ invariant and induces on it the standard map ψ_{q_j}; in particular the point $((0,0), a^{(j)})$ drifts along the r_1-axis. But now the functions $\tilde{g}_2^{(j)}$ and $\tilde{g}^{(j)}$ have new properties.

Lemma **2.6.** — *Let* $w_2^{(j)} = (0, 2/\mathrm{N}_j)$ *be the upper point of the separatrix of the pendulum whose time-1 map is* $\mathrm{G}_2^{(j)}$. *We set as above* $\mathrm{N}_j' = 1$ *if* $n = 2$ *and* $\mathrm{N}_j' = p_{j-(n-3)}\cdots p_j$ *if* $n \geq 3$. *Then* $\tilde{g}_2^{(j)}(w_2^{(j)}) = 1$, $d\tilde{g}_2^{(j)}(w_2^{(j)}) = 0$, *and both* $\tilde{g}_2$ *and its derivative vanish at each point* $(\mathrm{G}_2^{(j)})^{k\mathrm{N}_j'}(w_2^{(j)})$ *with* $k \neq 0$.

The proof is straightforward. The properties of $g_2^{(j)}$ are exactly what we need in order to describe easily the invariant manifolds of the various hyperbolic objects associated with the fixed point $(1/2, 0)$ of the pendulum. We shall limit ourselves to the upper part $(r_2 > 0)$ of these invariant manifolds.

Proposition **2.2** *(Hyperbolic objects and invariant manifolds).* — *Consider the transform-ation* $\tilde{\Psi}_j$ *defined in (2.22), and let* $\mathscr{S}$ *be the surface of equation* $\theta_2 = 0$ *in* $\mathbf{A}^n$.

1. *The* $(2n-2)$-*dimensional annulus* $\mathbf{A} \times \{(1/2, 0)\} \times \mathbf{A}^{n-2}$ *is invariant under* $\tilde{\Psi}_j$ *and normally hyperbolic in* $\mathbf{A}^n$. *For each* $r_1^0 \in \mathbf{R}$ *and* $\hat{r}^0 \in \mathbf{R}^{n-2}$, *the* $(n-1)$-*dimensional torus* $\mathscr{T}_{r_1^0, \hat{r}^0} = \mathrm{C}_{r_1^0} \times \{(1/2, 0)\} \times \mathrm{T}_{\hat{r}^0}$, *where* $\mathrm{C}_{r_1^0} = \mathbf{T} \times \{\hat{r}_1^0\} \subset \mathbf{A}$ *and* $\mathrm{T}_{\hat{r}^0} = \mathbf{T}^{n-2} \times \{\hat{r}^0\} \subset \mathbf{A}^{n-2}$, *is invariant and partially hyperbolic, with n-dimensional invariant manifolds.*

2. *The 2-dimensional annulus* $\mathrm{V}^{(j)} = \mathbf{A} \times \{(1/2, 0)\} \times \{\hat{x}^{(j)}\}$, *is invariant under* $\tilde{\Psi}_j^{\mathrm{N}'_j}$ *and partially hyperbolic in* $\mathbf{A}^n$. *Its 3-dimensional stable and unstable manifolds* $\mathrm{W}^{\pm}(\mathrm{V}^{(j)}, \tilde{\Psi}_j^{\mathrm{N}'_j})$ *admit a 2-dimensional homoclinic annulus inside* $\mathscr{S}$:

$$\mathbf{A} \times \{w_2^{(j)}\} \times \{\hat{x}^{(j)}\} \subset \mathrm{W}^+(\mathrm{V}^{(j)}, \tilde{\Psi}_j^{\mathrm{N}'_j}) \cap \mathrm{W}^-(\mathrm{V}^{(j)}, \tilde{\Psi}_j^{\mathrm{N}'_j}) \cap \mathscr{S}.$$

3. *For each* $r_1^0 \in \mathbf{R}$, *the circle* $\mathscr{C}_{r_1^0}^{(j)} = \mathrm{C}_{r_1^0} \times \{(1/2, 0)\} \times \{\hat{x}^{(j)}\}$ *is invariant under* $\tilde{\Psi}_j^{\mathrm{N}'_j}$ *and partially hyperbolic. Its 2-dimensional stable and unstable manifolds* $\mathrm{W}^{\pm}(\mathscr{C}_{r_1^0}^{(j)}, \tilde{\Psi}_j^{\mathrm{N}'_j})$ *satisfy*

$$\mathrm{C}_{r_1^0} \times \{w_2^{(j)}\} \times \{\hat{x}^{(j)}\} \subset \mathrm{W}^+\left(\mathscr{C}_{r_1^0}^{(j)}, \tilde{\Psi}_j^{\mathrm{N}'_j}\right) \cap \mathscr{S},$$
$$\Phi^{\frac{1}{q_j}\mathrm{U}}(\mathrm{C}_{r_1^0}) \times \{w_2^{(j)}\} \times \{\hat{x}^{(j)}\} \subset \mathrm{W}^-\left(\mathscr{C}_{r_1^0}^{(j)}, \tilde{\Psi}_j^{\mathrm{N}'_j}\right) \cap \mathscr{S}.$$

Proof. — Let $\mathscr{O}_\sigma = \{(\theta, r) \in \mathbf{A}^n \mid \sigma < \theta_2 < 1 - \sigma\}$: we observe that $\Phi^{\frac{1}{q_j}\mathrm{U}\otimes\tilde{g}^{(j)}}$ is reduced to identity in $\mathscr{O}_\sigma$ because of the compact support function that we have used in the definition of $g_2^{(j)}$. Thus, in $\mathscr{O}_\sigma$, $\tilde{\Psi}_j$ coincides with the integrable map $\Phi^{\frac{1}{2}r_1^2} \times \mathrm{G}_2^{(j)} \times \hat{\mathrm{G}}$, which is easily seen to possess all the hyperbolic invariant sets mentioned in Proposition 2.2. Moreover, the local stable and unstable manifolds of each of these sets coincide for both maps.

Therefore, only the statements concerning the global invariant manifolds of $\mathrm{V}^{(j)}$ and $\mathscr{C}_{r_1^0}^{(j)}$ for each r_1^0 remain to be checked. They will be obtained by a variant of Lemma 2.1 in which the point

$$w^{(j)} = (w_2^{(j)}, \hat{x}^{(j)}) \in \mathbf{A} \times \mathbf{A}^{n-2}$$

will play the role of a periodic point of infinite period for $\mathrm{G}^{(j)} = \mathrm{G}_2^{(j)} \times \hat{\mathrm{G}}$.

Let $\mathrm{F} = \Phi^{\frac{1}{2}r_1^2}$ and consider an arbitrary point $x_1 = (\theta_1, r_1)$ in $\mathbf{A}$. Thanks to the formula (2.11), we see that for each $k \in \mathbf{Z}^*$, $(\mathrm{F}^k(x_1), (\mathrm{G}^{(j)})^k(w^{(j)}))$ is a fixed point of $\Phi^{\frac{1}{q_j}\mathrm{U}\otimes\tilde{g}^{(j)}}$, because $g^{(j)}$ and its differential vanish at $(\mathrm{G}^{(j)})^k(w^{(j)})$ (use Lemma 2.6 if N'_j divides k; if not, at least one of the functions $g_i^{(j)}$ with $i \geq 3$ vanishes with its differential). Hence

$$\tilde{\Psi}_j^k(x_1, w^{(j)}) = \left(\mathrm{F}^k(x_1), (\mathrm{G}_2^{(j)})^k(w_2^{(j)}), \hat{\mathrm{G}}^k(\hat{x}^{(j)})\right) \qquad \text{for } k \geq 1,$$

and $(x_1, w^{(j)}) \in \mathrm{W}^+(\mathscr{C}_{r_1^0}^{(j)}, \tilde{\Psi}_j^{\mathrm{N}'_j}) \subset \mathrm{W}^+(\mathrm{V}^{(j)}, \tilde{\Psi}_j^{\mathrm{N}'_j})$.
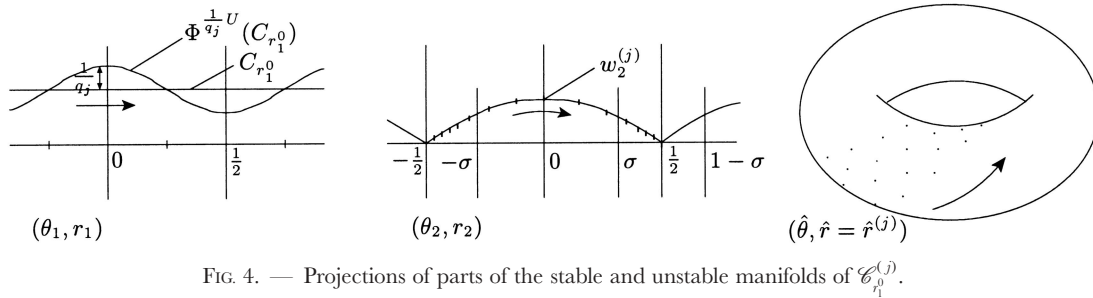
FIG. 4. — Projections of parts of the stable and unstable manifolds of $\mathscr{C}_{r_1^0}^{(j)}$.

As for backward iterates, since $\tilde{g}^{(j)}$ takes the value 1 at $w^{(j)}$ (still with vanishing differential), the formula (2.11) now yields $\Phi^{\frac{1}{q_j}\mathrm{U}\otimes\tilde{g}^{(j)}}(x_1, w^{(j)}) = (\Phi^{\frac{1}{q_j}\mathrm{U}}(x_1), w^{(j)})$, hence

$$\tilde{\Psi}_j^{-1}(\Phi^{\frac{1}{q_j}\mathrm{U}}(x_1), w^{(j)}) = (\mathrm{F}^{-1}(x_1), (\mathrm{G}^{(j)})^{-1}(w^{(j)})),$$

and $\tilde{\Psi}_j^{-k}(\Phi^{\frac{1}{q_j}\mathrm{U}}(x_1), w^{(j)}) = (\mathrm{F}^{-k}(x_1), (\mathrm{G}^{(j)})^{-k}(w^{(j)}))$ for $k \geq 1$, therefore $(\Phi^{\frac{1}{q_j}\mathrm{U}}(x_1), w^{(j)})$ $\in \mathrm{W}^-(\mathscr{C}_{r_1^0}^{(j)}, \tilde{\Psi}_j^{\mathrm{N}_j'}) \subset \mathrm{W}^-(\mathrm{V}^{(j)}, \tilde{\Psi}_j^{\mathrm{N}_j'})$ (see Figure 4). □

2.5.3. Proposition 2.2 describes in particular the intersections of the stable and unstable manifolds of the one-parameter familly of circles $\mathscr{C}_{r_1^0}^{(j)}$ with a suitable section, from which one easily deduces the existence of heteroclinic connections. Remark that

$$\Phi^{\frac{1}{q_j}\mathrm{U}}(\mathrm{C}_{r_1^0}) = \left\{(\theta_1, r_1) \in \mathbf{A}, \ r_1 = r_1^0 + \frac{1}{q_j}\cos 2\pi\theta_1\right\}.$$

As a consequence, the invariant circles $\mathscr{C}_{r_1}^{(j)}$ and $\mathscr{C}_{r_1'}^{(j)}$ have a heteroclinic connection as soon as $|r_1 - r_1'| \leq \frac{1}{q_j}$; the invariants manifolds intersect transversely when $|r_1 - r_1'| < \frac{1}{q_j}$, and tangentially in the limit case $|r_1 - r_1'| = \frac{1}{q_j}$. In particular, we have

Corollary **2.1** (*The transition chain*). — *The familly* $(\mathscr{C}_{k/q_j}^{(j)})_{k\in\mathbf{Z}}$ *is a chain in Arnold's sense: two consecutive circles* $\mathscr{C}_{k/q_j}^{(j)}$ *and* $\mathscr{C}_{(k+1)/q_j}^{(j)}$ *have a heteroclinic connection, which intersects* $\mathscr{S}$ *at the point* $((0, (k+1)/q_j), w_2^{(j)}, \hat{x}^{(j)})$. *The stable and unstable manifolds* $\mathrm{W}^+(\mathscr{C}_{(k+1)/q_j}^{(j)}, \tilde{\Psi}_j^{\mathrm{N}_j'})$ *and* $\mathrm{W}^-(\mathscr{C}_{k/q_j}^{(j)}, \tilde{\Psi}_j^{\mathrm{N}_j'})$ *are tangent along that heteroclinic orbit.*

One can see that the orbit of the wandering point $x^{(j)}$ shadows that chain, passing close to each of the circles in turn. For $k \in \mathbf{Z}$, the iterate $\tilde{\Psi}_j^{(k+\frac{1}{2})q_j}(x^{(j)})$ is close to the circle $\mathscr{C}_{k/q_j}^{(j)}$, since this point has $r_1 = \frac{k}{q_j}$, $\theta_2 = \frac{1}{2}$ and $\hat{x} = \hat{x}^{(j)}$, while the remaining

coordinates are $\theta_1 = \frac{k}{2}$ and $r_2 = \frac{1}{N_j} r_*^{(M_j)}(\frac{M_j}{2})$; a simple computation shows indeed that

$$\mathrm{dist}\left(\tilde{\Psi}_j^{(k+\frac{1}{2})q_j}(x^{(j)}), \mathscr{C}_{k/q_j}^{(j)} \right) < \frac{1}{q_j}$$

(because $\delta = r_*^{(M_j)}(\frac{M_j}{2})$ is the minimal value of r_* on a trajectory $(\theta_*(t), r_*(t))$ of the pendulum which is periodic of period M_j: integrating $\frac{d\theta_*}{dt} = r_* \geq \delta$, we get $1 > M_j \delta$).

The remaining points of the orbit drift along the heteroclinic connections between two consecutive circles. But the existence of this shadowing orbit could hardly be proved by the usual methods, since the intersections of the invariant manifolds are only tangential.

2.5.4. The tori $\mathscr{T}_{r_1^0, \hat{r}^{(j)}}$ will give rise to the tori $\mathrm{T}_{r_1^0}$ which are mentioned in the Addendum to Theorem B after the suspension procedure described in Section 2.4. We restricted ourselves to $\hat{\theta} = 0$ by considering the circles $\mathscr{C}_{r_1^0}^{(j)}$ in order to have a better control of the dynamics, but this is sufficient for our purpose, as shows the following lemma:

Lemma **2.7.** — *For each $r_1^0 \in \mathbf{R}$,*

$$\mathrm{W}^{\pm}(\mathscr{C}_{r_1^0}^{(j)}, \tilde{\Psi}_j^{N_j'}) = \mathrm{W}^{\pm}(\mathscr{T}_{r_1^0, \hat{r}^{(j)}}, \tilde{\Psi}_j) \cap (\mathbf{A} \times \mathbf{A} \times \{\hat{x}^{(j)}\}).$$

Proof. — Let us denote by $\hat{\pi}$ the third projection associated with the decomposition $\mathbf{A}^n = \mathbf{A} \times \mathbf{A} \times \mathbf{A}^{n-2}$ and by $\hat{g}$ the function $g_3^{(j)} \otimes \cdots \otimes g_n^{(j)}$. We shall use the following consequence of the identity (2.11): any point x of $\mathbf{A}^n$ whose projection $\hat{\pi}(x)$ is a critical point of $\hat{g}$ satisfies $\hat{\pi} \circ \Phi^{\frac{1}{q} U \otimes \tilde{g}_2 \otimes \hat{g}}(x) = \hat{\pi}(x)$ (we omit most of the indices j to simplify the notations). Therefore, for all $(x_1, x_2, \hat{x}) \in \mathbf{A}^n$,

$$d\hat{g}(\hat{\mathrm{G}}(\hat{x})) = 0 \implies \hat{\pi} \circ \tilde{\Psi}_j(x_1, x_2, \hat{x}) = \hat{\mathrm{G}}(\hat{x}),$$

$$d\hat{g}(\hat{x}) = 0 \implies \hat{\pi} \circ \tilde{\Psi}_j^{-1}(x_1, x_2, \hat{x}) = \hat{\mathrm{G}}^{-1}(\hat{x}).$$

In particular, we obtain:

$$(\textbf{2.23}) \qquad \tilde{\Psi}_j^{N_j'}(\mathbf{A} \times \mathbf{A} \times \{\hat{x}^{(j)}\}) = \mathbf{A} \times \mathbf{A} \times \{\hat{x}^{(j)}\}.$$

Let us now consider the unstable manifolds of $\mathscr{C}_{r_1^0}^{(j)}$ and $\mathscr{T}_{r_1^0, \hat{r}^{(j)}}$ (the corresponding property for the stable manifold would be proved along the same lines). They can be defined as the union of the *local* unstable manifolds and their positive iterates under $\tilde{\Psi}_j^{N_j'}$. But

$$\mathrm{W}_{loc}^{-}(\mathscr{T}_{r_1^0, \hat{r}^{(j)}}, \tilde{\Psi}_j) = \left\{ (x_1, x_2, \hat{x}) \in \mathscr{O}_\sigma \mid r_1 = r_1^0, \ r_2 = -\frac{2}{N_j} \cos \pi \theta_2, \ \hat{r} = \hat{r}^{(j)} \right\},$$

because $\tilde{\Psi}_j$ and $\Phi^{\frac{1}{2}r_1^2} \times G_2^{(j)} \times \hat{G}$ coincide when restricted to $\mathcal{O}_\sigma$; and for the same reason,

$$W_{loc}^-(\mathscr{C}_{r_1^0}^{(j)}, \tilde{\Psi}_j^{N_j'}) = \left\{(x_1, x_2, \hat{x}) \in \mathcal{O}_\sigma \mid r_1 = r_1^0, \ r_2 = -\frac{2}{N_j} \cos \pi\theta_2, \ \hat{x} = \hat{x}^{(j)}\right\}.$$

Thus $W_{loc}^-(\mathscr{C}_{r_1^0}^{(j)}, \tilde{\Psi}_j^{N_j'}) = W_{loc}^-(\mathscr{T}_{r_1^0, \hat{r}^{(j)}}, \tilde{\Psi}_j) \cap (\mathbf{A} \times \mathbf{A} \times \{\hat{x}^{(j)}\})$, and by (2.23),

$$\tilde{\Psi}_j^{kN_j'}\left(W_{loc}^-(\mathscr{C}_{r_1^0}^{(j)}, \tilde{\Psi}_j^{N_j'})\right) = \tilde{\Psi}_j^{kN_j'}(W_{loc}^-(\mathscr{T}_{r_1^0, \hat{r}^{(j)}}, \tilde{\Psi}_j)) \cap (\mathbf{A} \times \mathbf{A} \times \{\hat{x}^{(j)}\})$$

for all $k \in \mathbf{N}$.   □

2.5.5. We can finally address the question of the homoclinic splitting of the invariant manifolds, which is the only point of the Addendum to Theorem B still to be explained, in view of Lemma 2.7. The remarkable fact is the separation of the splitting into a *transversal* part, *i.e.* the splitting of the two invariant manifolds $W^\pm(V^{(j)})$ at each point of their intersection along the homoclinic annulus, and a *longitudinal* part, *i.e.* the splitting of the invariant manifolds $W^\pm(\mathscr{C}_{r_1^0}^{(j)})$ within the homoclinic annulus. This phenomenon corresponds to the block decomposition of the splitting matrix mentioned in § 1.3.4 and justified by Lemma 2.7.

We shall be interested only in the longitudinal splitting: the circles $\mathscr{C}_{r_1^0}^{(j)}$ are the only interesting objects in the case $n = 2$, and even when $n \geq 3$, the splitting of $W^\pm(\mathscr{C}_{r_1^0}^{(j)})$ embodies more refined information than that of the bigger invariant manifolds $W^\pm(\mathscr{T}_{r_1^0, \hat{r}^{(j)}})$. Moreover, as far as the drift along the r_1-axis alone is concerned, the longitudinal splitting is the only relevant parameter, and its determination is very simple here, since its amounts to the determination of the angle between the two circles $C_{r_1^0} \times \{w_2^{(j)}\} \times \{\hat{x}^{(j)}\}$ and $(\Phi^{\frac{1}{q_j}U}(C_{r_1^0})) \times \{w_2^{(j)}\} \times \{\hat{x}^{(j)}\}$ in the homoclinic annulus, at their intersection points.

Corollary **2.2** *(The longitudinal homoclinic splitting).* — *For each r_1^0, the circle $\mathscr{C}_{r_1^0}^{(j)}$ has two homoclinic orbits, which intersect $\mathscr{S}$ at the points $((\pm 1/4, 0), w_2^{(j)}, \hat{x}^{(j)})$. The longitudinal splitting angle $\mathscr{A}$ between $W^+(\mathscr{C}_{r_1^0}^{(j)}, \Psi_j^{N_j'})$ and $W^-(\mathscr{C}_{r_1^0}^{(j)}, \Psi_j^{N_j'})$ at each of these points satisfies $|\mathscr{A}| = \frac{2\pi}{q_j}$.*

Proof. — This fact is remarkably elementary: it suffices to compute the derivative of the longitudinal splitting function $\theta_1 \mapsto \frac{1}{q_j} \cos 2\pi\theta_1$ at the intersection points!   □

3. Stability theorem for Gevrey classes in the quasiconvex case

3.1. *Gevrey averaging around a periodic torus*

3.1.1. Periodic tori (*i.e.* totally resonant tori) of the unperturbed system will play a key role in the proof of Theorem A. To obtain stability over exponentially long times in the neighbourhood of a periodic torus of frequency-vector ω, the only "analytic" step is the elimination of the non-resonant part (with respect to ω) of the Hamiltonian up to an exponentially small remainder. In the aforementioned articles [Lo92,LN92,LNN93], the reduction to the normal form is achieved through a large number of successive canonical changes of variables. Once this normal form is obtained, the Hamiltonian character of the vector field and the quasi-convexity of the unperturbed system are used to derive stability near the torus (see Section 3.4); but they are not necessary for the reduction to the normal form itself, which amounts in fact to a one-phase averaging process.

As noticed in [Lo92], the exponential smallness with respect to ε of the remainder in the normal form, obtained after a number of iterations which is of the order of a negative power of ε, should be related to the Gevrey divergence of the formal series of classical perturbation theory. There is indeed a *formal* canonical change of variables which reduces the system to a *formal* resonant normal form, without any non-resonant remainder. But even if the data are analytic, those power series of ε are expected to have zero radius of convergence (generically).

The Gevrey growth of the coefficients of the formal series was proved in a simple problem in [Sa92]. But [RS96] gives a more systematic approach, whose formalism we shall adapt to our framework, and explains how to recover a normal form with exponentially small remainder from the formal solution and thus Neishtadt's Theorem [Nei84] on the elimination of a fast phase.[3] The generalization with respect to [Sa92] or [RS96] will consist in dealing with Gevrey instead of analytic data, and also in staying in the Hamiltonian framework (the Hamiltonian theorem proposed in [RS96] does not fit our purpose). This last point is a slight complication rather than an advantage: the Hamiltonian character is not necessary for obtaining a resonant normal form, but we need to verify that it can be used to find a symplectic normalizing transformation.

3.1.2. The data of our problem will be Gevrey-α functions and we shall begin by looking for *formal* solutions: power series in a small parameter ε whose coefficients are Gevrey-α functions. A formal series $\tilde{a} = \sum_{j \in \mathbf{N}} \varepsilon^j a_j$ with coefficients in a normed vector space V is said to be Gevrey-$(\alpha + 1)$ if there exist $C, M > 0$ such that

$$\forall j \in \mathbf{N}, \quad \|a_j\|_V \le CM^j j!^\alpha.$$

---

[3] See also [CD93] and [Bae95] for the use of Gevrey bounds in dynamic bifurcation problems.

Not all the authors follow this convention for indexing formal Gevrey classes, but we adopt it here to be consistent with Gevrey classes of functions: this way Gevrey-1 functions are analytic whereas Gevrey-1 series would be convergent. The key point will be the arising of Gevrey-$(\alpha + 1)$ formal series associated with Gevrey-α data. This jump of Gevrey index lies at the heart of our work.

3.1.3. Changing somewhat the setting with respect to Theorem A, we shall now state results which apply to a Hamiltonian of the form

$$< \omega, r > + f(\theta, r),$$

where ω is a fixed totally resonant frequency-vector and $f = \varepsilon F$ a small Gevrey-α perturbation. Their proof will be given in Sections 3.2 and 3.3; it is only in Section 3.4 that we shall come back to the original setting.

We thus consider perturbations of the periodic linear flow $\dot{\theta} = \omega$ on the torus $\mathbf{T}^n$ with frequency-vector $\omega \in \mathbf{R}^n \setminus \{0\}$. The common period of the corresponding orbits is denoted by $T > 0$, *i.e.* the components of $T\omega$ are integers without any non-trivial common divisor. Such a vector ω is said to be *rational of period* T. This allows us to fix $\ell \in \mathbf{Z}^n$ such that

$$(\mathbf{3.1}) \qquad < \ell, \omega > = \frac{1}{T}$$

and to define an $(n - 1)$-dimensional torus

$$(\mathbf{3.2}) \qquad \Sigma_0 = \{\theta \in \mathbf{T}^n \mid \ < \ell, \theta > \in \mathbf{Z}\}$$

which is transverse to the linear flow (and each trajectory meets Σ_0 exactly once per period).

A function φ on $\mathbf{T}^n$ is said to be *resonant* (with respect to ω) if it is constant along any trajectory of the linear flow induced by ω, *i.e.* if $< \omega, \partial_\theta \varphi > = 0$. Observe that if f is resonant, due to the Hamiltonian form of the equations of motion, $< \omega, r >$ is a first integral of the vector field associated with $< \omega, r > + f$.

The resonant part of a continuous function φ is defined as

$$(\mathbf{3.3}) \qquad \mathscr{R}_\omega \varphi(\theta) = \frac{1}{T} \int_0^T \varphi(\theta + t\omega) \, dt.$$

In the case of functions of θ and other variables, the definition of the operator $\mathscr{R}_\omega$ is extended by treating these extra variables as parameters.

We shall use the following notations for the Poisson bracket of two functions on the phase space:

$$L_\psi \varphi = \{\psi, \varphi\} = < \partial_r \psi, \partial_\theta \varphi > - < \partial_\theta \psi, \partial_r \varphi >.$$

Proposition **3.1.** *— Consider the Hamiltonian system generated by*

$$H_\varepsilon(\theta, r) = < \omega, r > + \varepsilon F(\theta, r),$$

*where ω is a rational vector of period T, and $F \in G^{\alpha, L_0}(\mathbf{T}^n \times \overline{B}_R)$ for some $\alpha \geq 1$ and $L_0, R > 0$.
There exists a unique pair $(\tilde{A}, \tilde{B})$ of formal series in ε with coefficients in $C^\infty(\mathbf{T}^n \times \overline{B}_R)$,*

$$\tilde{A}(\theta, r; \varepsilon) = \sum_{j \in \mathbf{N}} \varepsilon^j A_j(\theta, r), \qquad \tilde{B}(\theta, r; \varepsilon) = \sum_{j \in \mathbf{N}} \varepsilon^j B_j(\theta, r),$$

such that

　　i) each coefficient A_j of $\tilde{A}$ is resonant,
　　ii) each coefficient B_j of $\tilde{B}$ vanishes on $\Sigma_0 \times \overline{B}_R$,
　　iii)

(**3.4**) $$\exp(\varepsilon L_{\tilde{B}}) H_\varepsilon = < \omega, r > + \varepsilon \tilde{A}(\theta, r; \varepsilon).$$

*Moreover $A_0 = \mathscr{R}_\omega F$ and these formal series are Gevrey-$(\alpha + 1)$ in ε in the following sense:
their coefficients are Gevrey-$(\alpha, L_0/2)$ and there exists a positive number c depending only on α such
that*

$$\|A_j\|_{\alpha, \frac{L_0}{2}} \leq c\bar{N}(c^2 L_0^{-2\alpha} T\bar{F})^j j!^\alpha \quad \text{for } j \geq 1,$$

$$\|B_j\|_{\alpha, \frac{L_0}{2}} \leq cT\bar{N}(c^2 L_0^{-2\alpha} T\bar{F})^j j!^\alpha \quad \text{for } j \geq 0,$$

where $\bar{F} = \|F\|_{\alpha, L_0}$ and $\bar{N} = \|F - \mathscr{R}_\omega F\|_{\alpha, L_0}$ (thus $\bar{N} \leq 2\bar{F}$).

The proof of this proposition is in Section 3.2. We chose $L = L_0/2$ for simplicity,
but the coefficients of $\tilde{A}$ and $\tilde{B}$ are Gevrey-(α, L) and admit at most Gevrey-$(\alpha + 1)$
growth in $G^{\alpha, L}(\mathbf{T}^n \times \overline{B}_R)$ for all $L \in]0, L_0[$.
We shall also use an additional technical information:

Addendum to Proposition **3.1**. *— The coefficients of the formal series $\tilde{C} = - < \omega, \partial_\theta \tilde{B} >$
satisfy*

$$\|C_j\|_{\alpha, \frac{L_0}{2}} \leq c\bar{N}(c^2 L_0^{-2\alpha} T\bar{F})^j j!^\alpha \quad \text{for } j \geq 0,$$

where c is a positive number which depends only on α.

3.1.4. Of course, in the equation (3.4), the operator $\exp(\varepsilon L_{\tilde{B}})$ must be understood as the formal Lie series associated with the auxiliary vector field $X_{\varepsilon\tilde{B}}$ generated by the formal Hamiltonian function $\varepsilon\tilde{B}$. With the notation of the beginning of Section 2.1, the formal symplectic transformation $\Phi^{\varepsilon\tilde{B}}$, which is the flow at time $\tau = 1$ for $\frac{dx}{d\tau} = X_{\varepsilon\tilde{B}}(x; \varepsilon)$, is of the form $\mathrm{Id} + \mathcal{O}(\varepsilon)$, thus the formula

$$\varphi \circ \Phi^{\varepsilon\tilde{B}} = \exp(\varepsilon L_{\tilde{B}})\varphi = \varphi + \sum_{j \geq 1} \frac{\varepsilon^j}{j!} L_{\tilde{B}}^j \varphi$$

makes sense at a formal level for any function φ on the phase space (using the Taylor formula to expand the left-hand side). In particular,

$$H_\varepsilon \circ \Phi^{\varepsilon\tilde{B}} = <\omega, r> + \varepsilon\tilde{A}(\theta, r; \varepsilon).$$

Thus, Proposition 3.1 must be understood as a statement about a "formal resonant normal form". We may consider $\Phi^{\varepsilon\tilde{B}}$ as providing formal coordinates in which $<\omega, r>$ is a first integral of the motion.

There exists a non-Hamiltonian version of this proposition: if $m \in \mathbf{N}$ and $F_1, ..., F_n, G_1, ..., G_m$ are Gevrey-α functions of $\theta \in \mathbf{T}^n$ and $y \in \mathbf{R}^m$ for $\|y\| \leq R$, the vector field

$$\left| \begin{array}{l} \dot{\theta} = \omega + \varepsilon F(\theta, y) \\ \dot{y} = \varepsilon G(\theta, y) \end{array} \right.$$

can be formally conjugated to a formal resonant vector field, and the normal form as well as the normalizing series are Gevrey-$(\alpha + 1)$ in ε. In the case where $\omega = (1, 0, ..., 0)$, *i.e.* when we have a slow-fast system with one rapid phase θ_1, the role of the normalizing transformation is to eliminate the rapid phase. We then recover a generalization to Gevrey classes of the results of [Sa92] and [RS96] on the Gevrey character of the formal series.

3.1.5. The next step is the passage from the formal normal form to an approximate Gevrey normal form.

Proposition **3.2.** — *Let* $\omega \in \mathbf{R}^n \backslash \{0\}$ *a rational vector of period* T, *and* $f \in G^{\alpha, L_0}(\mathbf{T}^n \times \overline{B}_R)$ *with some* $R, L_0 > 0$. *We set*

$$\varepsilon = \|f\|_{\alpha, L_0}, \quad \varepsilon' = \|f - \mathcal{R}_\omega f\|_{\alpha, L_0},$$

and consider the Hamiltonian system generated by

$$H(\theta, r) = <\omega, r> + f(\theta, r).$$

There exists a positive number c which depends only on α such that: if $\delta \in]0, R[$ and $\varepsilon T \leq \nu$, where

$$\nu = c^{-1} L_0^\alpha \min\{\delta, L_1^\alpha\}, \quad L_1 = \frac{L_0}{16(2n)^{\frac{\alpha-1}{\alpha}}},$$

there exist $\mathscr{A}, \mathscr{E}, \Psi^{(1)}, ..., \Psi^{(2n)} \in G^{\alpha, L_1}(\mathbf{T}^n \times \overline{B}_{R-\delta})$ satisfying

    i) $\mathscr{A}$ is resonant with respect to ω,
    ii) $Id + \Psi$ is a symplectic transformation whose image is contained in $\mathbf{T}^n \times \overline{B}_R$,
    iii)

$$(\mathbf{3.5}) \qquad H \circ (Id + \Psi) = < \omega, r > + \mathscr{A} + \mathscr{E},$$

    iv) the following inequalities hold:

$$\|\mathscr{E}\|_{\alpha, L_1} \leq c \, \varepsilon' \exp\left(-\alpha \left(\frac{\nu}{\varepsilon T}\right)^{1/\alpha}\right)$$

$$\|\mathscr{A} - \mathscr{R}_\omega f\|_{\alpha, L_1} \leq c \, \varepsilon' \frac{\varepsilon T}{\nu}, \qquad \sum_{i=1}^{2n} \|\Psi^{(i)}\|_{\alpha, L_1} \leq c L_0^{-\alpha} \varepsilon T.$$

Thus we shall have at our disposal new coordinates in which $\frac{d}{dt} < \omega, r >$ is exponentially small with respect to the size of the perturbation f.

The proof of Proposition 3.2 is in Section 3.3. Observe that $\varepsilon' \leq 2\varepsilon$ and that in this statement εT should be considered as the true small parameter rather than ε.

One can also write a non-Hamiltonian version of Proposition 3.2. In the case of a slow-fast system with one rapid phase, when $\omega = (1, 0, ..., 0)$, the statement amounts to a generalization of Neishtadt's Theorem [Nei84] and Ramis-Schäfke's result [RS96] to Gevrey classes: the rapid phase can be eliminated up to an exponentially small error.

3.2. *Formal normal form associated with a periodic torus*

This section is devoted to the proof of Proposition 3.1. Let $R, L_0, L_1 > 0$ with $L_1 < L_0$; at the last moment we shall choose $L_1 = \frac{L_0}{2}$. We shall use the notations

$$V_0 = G^{\alpha, L_0}(\mathbf{T}^n \times \overline{B}_R), \quad V_1 = G^{\alpha, L_1}(\mathbf{T}^n \times \overline{B}_R), \quad K = \mathbf{T}^n \times \overline{B}_R.$$

Of course $V_0 \subset V_1 \subset C^\infty(K)$. We fix $F \in V_0$ and $H_\varepsilon(\theta, r) = < \omega, r > + \varepsilon F(\theta, r)$. The operators of partial differentiation are denoted by $\partial_{x_1}, ..., \partial_{x_{2n}}$. We introduce the auxiliary number $L = \frac{L_1 + L_0}{2} \in]L_1, L_0[$ and the intermediate space $V = G^{\alpha, L}(K)$.

3.2.1. *Algorithm for the formal solution.* — When expanding the conjugation equation (3.4), we shall meet the linear equation

$$\text{(3.6)} \qquad A + <\omega, \partial_\theta B> = \varphi,$$

where φ is a known function of (θ, r) which belongs to $C^\infty(K)$ and the unknown real-valued functions A and B are required to satisfy

i) A belongs to $C^\infty(K)$ and is resonant with respect to ω,
ii) B belongs to $C^\infty(K)$ and vanishes on $\Sigma_0 \times \overline{B}_R$.

Lemma **3.1.** — *The equation (3.6) admits a unique solution* (A, B) *which satisfies i) and ii). It is given by the following formulas:*

$$A = \mathscr{R}\varphi, \quad B = \mathscr{N}\varphi,$$

with $\mathscr{R} = \mathscr{R}_\omega$ *defined by (3.3) and*

$$\text{(3.7)} \qquad \mathscr{N}\varphi(\theta, r) = \int_0^{T<\ell, \tilde{\theta}>} (\varphi - \mathscr{R}\varphi)(\theta - t\omega, r)\, dt$$

where $\tilde{\theta}$ *is any lift of* θ *in* $\mathbf{R}^n$. *Moreover, for all* $k \in \mathbf{N}^{2n}$, $\|\partial^k \mathscr{R}\varphi\|_{C^0(K)} \leq \|\partial^k \varphi\|_{C^0(K)}$ *and*

$$\text{(3.8)} \qquad \|\partial^k \mathscr{N}\varphi\|_{C^0(K)} \leq \tfrac{T}{2}\|\partial^k(\varphi - \mathscr{R}\varphi)\|_{C^0(K)}, \quad \|\partial^k \mathscr{N}\varphi\|_{C^0(K)} \leq \tfrac{T}{2}\|\partial^k \varphi\|_{C^0(K)}.$$

Remark **3.1.** — Here, $\mathscr{R}$ stands for "Resonant" and $\mathscr{N}$ for "Non-resonant"; we depart here from a notation which is sometimes used in Hamiltonian perturbation theory, according to which $\mathscr{N}$ corresponds to the null space of a certain operator (containing the averaged Hamiltonians), and $\mathscr{R}$ to its range (containing those Hamiltonians that can be transformed away).

Proof. — Since $<\omega, \partial_\theta B>$ has obviously zero resonant part, A necessarily coincides with $\mathscr{R}\varphi$. Let $\psi = \varphi - \mathscr{R}\varphi$. The equation (3.6) is equivalent to

$$-\frac{d}{dt}[B(\theta - t\omega, r)] = \psi(\theta - t\omega, r).$$

The formula (3.7) is then obtained by integrating between $t_0 = 0$ and $t_1 = T < \ell, \tilde{\theta} >$, a value which was chosen to ensure $\theta - t_1\omega \in \Sigma_0$. The resulting formula is independent of the chosen lift of θ because $\mathscr{R}\psi = 0$.

One can prove the inequalities which are indicated at the end of the statement by selecting $\tilde{\theta}$ such that $| < \ell, \tilde{\theta} > | \leq \tfrac{1}{2}$; for the last one, notice that

$$\mathcal{N}\varphi(\theta, r) =$$

$$(1 \pm <\ell, \tilde{\theta}>) \int_0^{T<\ell,\tilde{\theta}>} \varphi(\theta - t\omega, r)\, dt \mp <\ell, \tilde{\theta}> \int_0^{T(\pm 1 + <\ell,\tilde{\theta}>)} \varphi(\theta - t\omega, r)\, dt,$$

hence $|\mathcal{N}\varphi(\theta, r)| \leq 2T| <\ell, \tilde{\theta}> | \cdot \big(1 - | <\ell, \tilde{\theta}> |\big) \cdot \|\varphi\|_{C^0(K)}$. □

The following lemma contains the first claims of Proposition 3.1 and gives an algorithm to find the formal solution $(\tilde{A}, \tilde{B})$. We have defined two operators of $C^\infty(K)$, $\mathcal{R}$ and $\mathcal{N}$; we extend them by linearity to $C^\infty(K)[[\varepsilon]]$.

*Lemma **3.2.** — There exists a unique pair $(\tilde{A}, \tilde{B})$ of formal series of $C^\infty(K)[[\varepsilon]]$ satifying i), ii) and iii) of Proposition 3.1, and it can be obtained as follows: define a functional in $C^\infty(K)[[\varepsilon]]$ by*

$$\forall \varphi \in C^\infty(K)[[\varepsilon]], \quad \mathcal{F}(\varphi) = \exp(\varepsilon L_{\mathcal{N}\varphi})F + \gamma(\varepsilon L_{\mathcal{N}\varphi})(\mathcal{R}\varphi - \varphi),$$

with $\gamma(t) = \sum_{p \geq 1} \frac{1}{(p+1)!} t^p$; then the sequence of iterates $(\mathcal{F}^\nu(0))_{\nu \in \mathbf{N}}$ converges formally towards a series $\varphi^* \in V[[\varepsilon]]$ and the solution $(\tilde{A}, \tilde{B})$ is given by

$$\tilde{A} = \mathcal{R}\varphi^*, \quad \tilde{B} = \mathcal{N}\varphi^*.$$

In particular, the first coefficients of φ^* and $\tilde{A}$ are $\varphi_0^* = F$ and $A_0 = \mathcal{R}F$.

Proof. — Setting $h(\theta, r) = <\omega, r>$ and using the identity $\exp(t) = 1 + t + t\gamma(t)$, we see that the condition *iii)* amounts to

$$\tilde{A} - L_{\tilde{B}}h = \exp(\varepsilon L_{\tilde{B}})F + \gamma(\varepsilon L_{\tilde{B}})L_{\tilde{B}}h.$$

We notice that $L_{\tilde{B}}h = - <\omega, \partial_\theta \tilde{B}>$ and we introduce a new unknown formal series

$$\varphi = \tilde{A} + <\omega, \partial_\theta \tilde{B}>.$$

According to Lemma 3.1 and because of conditions *i)* and *ii)*, there is a one-to-one correspondence between φ and $(\tilde{A}, \tilde{B})$: as soon as φ is known, we recover the solution by the formulas

$$\tilde{A} = \mathcal{R}\varphi, \quad \tilde{B} = \mathcal{N}\varphi.$$

We are thus left with the equation $\varphi = \mathcal{F}(\varphi)$. The relation

$$\mathcal{F}(\varphi + \mathcal{O}(\varepsilon^j)) = \mathcal{F}(\varphi) + \mathcal{O}(\varepsilon^{j+1}), \qquad \varphi \in C^\infty(K)[[\varepsilon]], \; j \in \mathbf{N},$$

(which stems from the fact that $\forall p \geq 1$, $(L^p_{\mathcal{N}\varphi + \mathcal{O}(\varepsilon^j)} - L^p_{\mathcal{N}\varphi})\psi = \mathcal{O}(\varepsilon^j)$) implies the formal convergence of the sequence of iterates $(\mathcal{F}^\nu(0))_{\nu \in \mathbf{N}}$ towards a unique fixed point φ^* for $\mathcal{F}$ in $C^\infty(K)[[\varepsilon]]$. It lies in $V[[\varepsilon]]$, since each of its coefficients is obtained from $F \in V_0$ by a finite number of operations involving $\mathcal{R}$, $\mathcal{N}$ and differentiation. The first coefficient is already visible on $\mathcal{F}(0) = F$. □

3.2.2. *Notations for Gevrey majorant series.* — We follow closely [RS96], but working in Gevrey classes requires some adaptations (see also [Wa00]).

Beside the operators $\mathscr{R}$ and $\mathscr{N}$, which must be considered as bounded operators, the functional $\mathscr{F}$ of Lemma 3.2 involves the unbounded operator $\varepsilon L_{\mathscr{N}\varphi}$. The Gevrey growth of the coefficients of the formal solution will eventually stem from these repeated differentiations.

In order to deal with that phenomenon, a convenient way of taking advantage of the "generalized Cauchy inequalities" of Lemma A.2 is to define, for each $j \in \mathbf{N}$,

$$\|\varphi\|_{\alpha,\mathrm{L},j} = \sup_{\zeta \in]0,1[} \{(1-\zeta)^{j\alpha}\|\varphi\|_{\alpha,\zeta\mathrm{L}}\}$$

$$= \sup_{\zeta \in]0,1[} \{(1-\zeta)^{j\alpha} \sum_{k\in\mathbf{N}^{2n}} \frac{\zeta^{|k|\alpha}\mathrm{L}^{|k|\alpha}}{k!^{\alpha}}\|\partial^k\varphi\|_{\mathrm{C}^0(\mathrm{K})}\}$$

and $\mathrm{E}_{\alpha,\mathrm{L},j}(\mathrm{K}) = \{\varphi \in \mathrm{C}^{\infty}(\mathrm{K}) \mid \|\varphi\|_{\alpha,\mathrm{L},j} < \infty\}$. This is an increasing sequence of spaces, since

$$\|\varphi\|_{\alpha,\mathrm{L}} = \|\varphi\|_{\alpha,\mathrm{L},0} \geq \|\varphi\|_{\alpha,\mathrm{L},1} \geq \ldots \geq \|\varphi\|_{\alpha,\mathrm{L},j} \geq \ldots$$

Moreover, since $\frac{\mathrm{L}_1}{\mathrm{L}} \in]0,1[$,

$$(\mathbf{3.9}) \qquad \|\varphi\|_{\alpha,\mathrm{L}_1} \leq (1-\tfrac{\mathrm{L}_1}{\mathrm{L}})^{-j\alpha}\|\varphi\|_{\alpha,\mathrm{L},j},$$

thus

$$\mathrm{V}_0 = \mathrm{G}^{\alpha,\mathrm{L}_0}(\mathrm{K}) \subset \mathrm{E}_{\alpha,\mathrm{L},j}(\mathrm{K}) \subset \mathrm{V}_1 = \mathrm{G}^{\alpha,\mathrm{L}_1}(\mathrm{K}), \qquad j \in \mathbf{N}.$$

We now come to the main tool of our analysis of Gevrey formal series, an adaptation of a lemma which is due to Nagumo and Malgrange (see the references in [RS96]).

Lemma **3.3.** — *If $j \in \mathbf{N}$ and $\varphi \in \mathrm{E}_{\alpha,\mathrm{L},j}(\mathrm{K})$, then all first-order partial derivatives of φ belong to $\mathrm{E}_{\alpha,\mathrm{L},j+1}(\mathrm{K})$ and*

$$\sum_{i=1}^{2n} \|\partial_{x_i}\varphi\|_{\alpha,\mathrm{L},j+1} \leq \beta(j+1)^{\alpha}\|\varphi\|_{\alpha,\mathrm{L},j},$$

where $\beta = \left(\frac{e}{\mathrm{L}}\right)^{\alpha}$.

Proof. — Let $\zeta \in]0,1[$. For all $\zeta' \in]\zeta,1[$, Lemma A.2 yields

$$\sum_{i=1}^{2n} \|\partial_{x_i}\varphi\|_{\alpha,\zeta\mathrm{L}} \leq \mathrm{L}^{-\alpha}\frac{1}{(\zeta'-\zeta)^{\alpha}}\|\varphi\|_{\alpha,\zeta'\mathrm{L}}.$$

We conclude by choosing $\zeta' \in\ ]\zeta, 1[$ in such a way that

$$\frac{1}{(\zeta' - \zeta)(1 - \zeta')^j} \leq \frac{e(j + 1)}{(1 - \zeta)^{j+1}}$$

(one can take $\zeta' = \zeta + \frac{1-\zeta}{e}$ if $j = 0$, and $\zeta' = \zeta + \frac{1-\zeta}{j+1}$ if $j \geq 1$). $\qquad\qquad \square$

We are now in a position to introduce "Gevrey majorant series":

Definition **3.1.** — *Let* $\varphi = \sum \varepsilon^j \varphi_j \in V_1[[\varepsilon]]$ *and* $\bar{\varphi} = \sum \xi^j \bar{\varphi}_j \in \mathbf{R}[[\xi]]$. *We say that* $\bar{\varphi}$ *is an* (α, L)-*Gevrey majorant series for* φ, *and we use the notation* $\varphi \ll_{\alpha,\mathrm{L}} \bar{\varphi}$, *when*

$$\varphi_j \in \mathrm{E}_{\alpha,\mathrm{L},j}(\mathrm{K}) \quad and \quad \|\varphi_j\|_{\alpha,\mathrm{L},j} \leq j!^\alpha \bar{\varphi}_j$$

for each $j \in \mathbf{N}$.

Remark **3.2.** — Notice that, if moreover $\bar{\varphi}$ is a convergent power series, then φ is a Gevrey-$(\alpha + 1)$ series in $V_1[[\varepsilon]]$: indeed, denoting by C and M positive numbers such that $\bar{\varphi}_j \leq \mathrm{C}\mathrm{M}^j$ for all $j \geq 0$, we derive

$$\|\varphi_j\|_{\alpha,\mathrm{L}_1} \leq \mathrm{C}\mathrm{M}^j(1 - \tfrac{\mathrm{L}_1}{\mathrm{L}})^{-j\alpha} j!^\alpha$$

from the inequality (3.9).

Lemma **3.4.** — *Let* $\beta = \left(\frac{e}{\mathrm{L}}\right)^\alpha$.

  i) If $\varphi \ll_{\alpha,\mathrm{L}} \bar{\varphi}$, $\psi \ll_{\alpha,\mathrm{L}} \bar{\psi}$ *and* $\lambda, \mu \in \mathbf{R}$, *then* $\lambda\varphi + \mu\psi \ll_{\alpha,\mathrm{L}} |\lambda|\bar{\varphi} + |\mu|\bar{\psi}$ *and*
    $\varphi\psi \ll_{\alpha,\mathrm{L}} \bar{\varphi}\bar{\psi}$.

  ii) If $\varphi \ll_{\alpha,\mathrm{L}} \bar{\varphi}$, *then* $\mathscr{R}\varphi \ll_{\alpha,\mathrm{L}} \bar{\varphi}$ *and* $\mathscr{N}\varphi \ll_{\alpha,\mathrm{L}} \frac{\mathrm{T}}{2}\bar{\varphi}$.

  iii) If $\varphi \ll_{\alpha,\mathrm{L}} \bar{\varphi}$, $\varepsilon\partial_{x_i}\varphi \ll_{\alpha,\mathrm{L}} \beta\xi\bar{\varphi}$ *for all* $i = 1, ..., 2n$.

  iv) If $\partial_{x_i}\chi \ll_{\alpha,\mathrm{L}} \bar{\chi}$ *for all* $i = 1, ..., 2n$, *and if* $a \in V_1$ *satisfies* $\sum_{i=1}^{2n} \|\partial_{x_i}a\|_{\alpha,\mathrm{L}} \leq \bar{a}_0$,

$$\{\chi, a\} \ll_{\alpha,\mathrm{L}} \bar{\chi}\bar{a}_0.$$

  v) If $\partial_{x_i}\chi \ll_{\alpha,\mathrm{L}} \bar{\chi}$ *for all* $i = 1, ..., 2n$ *and* $\varphi \ll_{\alpha,\mathrm{L}} \bar{\varphi}$,

$$\{\chi, \varepsilon\varphi\} \ll_{\alpha,\mathrm{L}} \beta\xi\bar{\chi}\bar{\varphi}.$$

Proof. — *i)* Using Lemma A.1 and the definition of $\|\,.\,\|_{\alpha,\mathrm{L},j}$, we get

$$\|\varphi_{j_1}\psi_{j_2}\|_{\alpha,\mathrm{L},j_1+j_2} \leq \|\varphi_{j_1}\|_{\alpha,\mathrm{L},j_1}\|\psi_{j_2}\|_{\alpha,\mathrm{L},j_2} \leq (j_1 + j_2)!^\alpha \bar{\varphi}_{j_1}\bar{\psi}_{j_2}.$$

ii) Use (3.8).

iii) Lemma 3.3 yields, for each $j \geq 1$,

$$(\mathbf{3.10}) \qquad \sum_{i=1}^{2n} \|\partial_{x_i}\varphi_{j-1}\|_{\alpha,\mathrm{L},j} \leq \beta j!^\alpha \bar{\varphi}_{j-1}.$$

iv) We observe that

$$\|\{\chi_j, a\}\|_{\alpha,\mathrm{L},j} \leq \sum_{i=1}^{n} \|\partial_{x_{i+n}}\chi_j \partial_{x_i} a\|_{\alpha,\mathrm{L},j} + \sum_{i=1}^{n} \|\partial_{x_i}\chi_j \partial_{x_{i+n}} a\|_{\alpha,\mathrm{L},j}$$

$$\leq \left(\max_{1 \leq i \leq 2n} \|\partial_{x_i}\chi_j\|_{\alpha,\mathrm{L},j}\right) \sum_{i=1}^{2n} \|\partial_{x_i} a\|_{\alpha,\mathrm{L}} \leq j!^{\alpha} \bar{\chi}_j \bar{a}_0.$$

v) Similarly,

$$\|\{\chi_{j_1}, \varphi_{j_2-1}\}\|_{\alpha,\mathrm{L},j_1+j_2} \leq \left(\max_{1 \leq i \leq 2n} \|\partial_{x_i}\chi_{j_1}\|_{\alpha,\mathrm{L},j_1}\right) \sum_{i=1}^{2n} \|\partial_{x_i}\varphi_{j_2-1}\|_{\alpha,\mathrm{L},j_2}$$

and we can use (3.10): $\|\{\chi_{j_1}, \varphi_{j_2-1}\}\|_{\alpha,\mathrm{L},j_1+j_2} \leq \beta(j_1+j_2)!^{\alpha} \bar{\chi}_{j_1} \bar{\varphi}_{j_2}.$ □

3.2.3. *Gevrey character of the formal solution.* — We have obtained φ^* of Lemma 3.2 as the unique formal solution of the equation

$$\varphi^* = \mathcal{F}(\varphi^*) = \exp(\varepsilon \mathrm{L}_{\mathcal{N}\varphi^*})\mathrm{F} + \gamma(\varepsilon \mathrm{L}_{\mathcal{N}\varphi^*})(\mathcal{R}\varphi^* - \varphi^*).$$

Since $\varphi_0^* = \mathrm{F}$, we can introduce a new unknown formal series ψ^* defined by

(**3.11**) $\qquad \varphi^* = \mathrm{F} + \varepsilon \psi^*.$

The corresponding equation for ψ^* is $\psi^* = \mathcal{G}(\psi^*)$, where the functional $\mathcal{G}$ of $\mathrm{V}_1[[\varepsilon]]$ is easily deduced from $\mathcal{F}$. With the notation $\mathrm{G} = \mathcal{N}\mathrm{F}$, the result of the computation is

$$\mathcal{G}(\psi) = \sum_{p \geq 0} \frac{\varepsilon^p}{(p+1)!} \mathrm{L}_{\mathrm{G}+\varepsilon\mathcal{N}\psi}^{p+1}\left(\mathrm{F} + \frac{1}{p+2}(\mathcal{R}\mathrm{F} - \mathrm{F})\right)$$

$$+ \sum_{p \geq 0} \frac{\varepsilon^{p+1}}{(p+2)!} \mathrm{L}_{\mathrm{G}+\varepsilon\mathcal{N}\psi}^{p+1}(\mathcal{R}\psi - \psi).$$

We have again the property $\mathcal{G}(\psi + \mathcal{O}(\varepsilon^j)) = \mathcal{G}(\psi) + \mathcal{O}(\varepsilon^{j+1})$ which ensures that

(**3.12**) $\qquad \psi^* = \lim_{\nu \to \infty} \mathcal{G}^{\nu}(0) \in \mathrm{V}_1[[\varepsilon]].$

Let $\bar{\mathrm{F}} = \|\mathrm{F}\|_{\alpha,\mathrm{L}_0}$, $\bar{\mathrm{N}} = \|\mathrm{F} - \mathcal{R}\mathrm{F}\|_{\alpha,\mathrm{L}_0}$, and

(**3.13**) $\qquad \bar{\mathrm{F}}_0 = e^{-\alpha}(\frac{\mathrm{L}_0}{\mathrm{L}} - 1)^{-\alpha}\bar{\mathrm{F}}, \quad \bar{\mathrm{N}}_0 = e^{-\alpha}(\frac{\mathrm{L}_0}{\mathrm{L}} - 1)^{-\alpha}\bar{\mathrm{N}}.$

Using Lemma A.2, we see that

(**3.14**) $\qquad \sum_{i=1}^{2n} \|\partial_{x_i}\mathrm{F}\|_{\alpha,\mathrm{L}} \leq \beta\bar{\mathrm{F}}_0, \quad \sum_{i=1}^{2n} \|\partial_{x_i}(\mathcal{R}\mathrm{F} - \mathrm{F})\|_{\alpha,\mathrm{L}} \leq \beta\bar{\mathrm{N}}_0.$

Lemma **3.5.** — *If $a \in V_1$ and $\psi, \chi \in V_1[[\varepsilon]]$ with*

$$\psi \ll_{\alpha,L} \bar{\psi}, \qquad \chi \ll_{\alpha,L} \bar{\chi}, \qquad \sum_{i=1}^{2n} \|\partial_{x_i} a\|_{\alpha,L} \leq \bar{a}_0,$$

the following inequalities hold for all $p \in \mathbf{N}$:

$$\varepsilon^p L_{G+\varepsilon\mathcal{N}\psi}^{p+1} a \ll_{\alpha,L} (\tfrac{T}{2}\beta)^{p+1}(\bar{N}_0 + \xi\bar{\psi})^{p+1}(\beta\xi)^p \bar{a}_0,$$

$$\varepsilon^{p+1} L_{G+\varepsilon\mathcal{N}\psi}^{p+1} \chi \ll_{\alpha,L} (\tfrac{T}{2}\beta)^{p+1}(\bar{N}_0 + \xi\bar{\psi})^{p+1}(\beta\xi)^{p+1} \bar{\chi}.$$

Proof. — We first observe, thanks to (3.8) and (3.14), that

$$\|\partial_{x_i} G\|_{\alpha,L} \leq \tfrac{T}{2}\beta\bar{N}_0, \qquad i = 1, ..., 2n.$$

Equivalenlty, we can treat $\partial_{x_i} G$ as a formal series $V_1[[\varepsilon]]$ which is reduced to its first coefficient and write $\partial_{x_i} G \ll_{\alpha,L} \tfrac{T}{2}\beta\bar{N}_0$. We now prove the lemma by induction on p. Lemma 3.4 implies that

$$\partial_{x_i}(G + \varepsilon\mathcal{N}\psi) \ll_{\alpha,L} \tfrac{T}{2}\beta(\bar{N}_0 + \xi\bar{\psi}), \qquad i = 1, ..., 2n.$$

We thus obtain for $p = 0$, by Lemma 3.4 *iv)* and *v)*:

$$L_{G+\varepsilon\mathcal{N}\psi} a = \{G + \varepsilon\mathcal{N}\psi, a\} \ll_{\alpha,L} \tfrac{T}{2}\beta(\bar{N}_0 + \xi\bar{\psi})\bar{a}_0,$$

$$\varepsilon L_{G+\varepsilon\mathcal{N}\psi} \chi = \{G + \varepsilon\mathcal{N}\psi, \varepsilon\chi\} \ll_{\alpha,L} \tfrac{T}{2}\beta(\bar{N}_0 + \xi\bar{\psi})\beta\xi\bar{\chi}.$$

Assuming that the desired inequalities hold at rank $p \geq 0$, we conclude by applying Lemma 3.4 *v)* to

$$\varepsilon^{p+1} L_{G+\varepsilon\mathcal{N}\psi}^{p+2} a = \left\{G + \varepsilon\mathcal{N}\psi, \varepsilon(\varepsilon^p L_{G+\varepsilon\mathcal{N}\psi}^{p+1} a)\right\},$$

$$\varepsilon^{p+2} L_{G+\varepsilon\mathcal{N}\psi}^{p+2} \chi = \left\{G + \varepsilon\mathcal{N}\psi, \varepsilon(\varepsilon^{p+1} L_{G+\varepsilon\mathcal{N}\psi}^{p+1} \chi)\right\}. \qquad \square$$

Lemma **3.6.** — *The formal series ψ^* defined by (3.11) admits an (α, L)-Gevrey majorant series*

$$\psi^* \ll_{\alpha,L} \frac{2\bar{N}_0}{\xi}\mathcal{M}(\tfrac{T}{2}\beta^2 \bar{F}_0 \xi),$$

where $\mathcal{M}(\zeta) = 2\zeta\left(1 - 6\zeta + \left[(1 - 6\zeta)^2 - 32\zeta^2\right]^{1/2}\right)^{-1}$.

Proof. — Suppose $\psi \in V_1[[\varepsilon]]$, $\psi \ll_{\alpha,L} \bar{\psi}$. Because of Lemma 3.4 and of the inequality (3.14), for each $p \in \mathbf{N}$ the function $a = F + \frac{1}{p+2}(\mathscr{R}F - F)$ and the formal series $\chi = \mathscr{R}\psi - \psi$ satisfy

$$\chi \ll_{\alpha,L} 2\bar{\psi}, \qquad \sum_{i=1}^{2n} \|\partial_{x_i} a\|_{\alpha,L} \leq 2\beta\bar{F}_0.$$

Thus we can apply the previous lemma and obtain an "(α, L)-Gevrey majorant functional" for $\mathscr{G}$:

$$\psi \ll_{\alpha,L} \bar{\psi} \;\Rightarrow\; \mathscr{G}(\psi) \ll_{\alpha,L} \sum_{p \geq 0} (\tfrac{T}{2}\beta^2)^{p+1}(\bar{N}_0 + \xi\bar{\psi})^{p+1}\xi^p(2\bar{F}_0 + \xi\bar{\psi}).$$

Using again $\bar{N}_0 \leq 2\bar{F}_0$, we enlarge slightly this majorant and define

$$\bar{\mathscr{G}}(\bar{\psi}) = \frac{T}{2}\beta^2 \frac{2\bar{N}_0\bar{F}_0 + 4\bar{F}_0\xi\bar{\psi} + \xi^2\bar{\psi}^2}{1 - \frac{T}{2}\beta^2\xi(2\bar{F}_0 + \xi\bar{\psi})}.$$

The unique fixed point $\bar{\psi}^*$ of $\bar{\mathscr{G}}$ in $\mathbf{R}[[\xi]]$ is the limit of the sequence of iterates $(\bar{\mathscr{G}}^\nu(0))$; it is thus an (α, L)-Gevrey majorant series for ψ^*. It can be computed exactly; enlarging slightly its coefficients, we obtain the desired result. $\qquad\square$

The last lemma allows us to conlude the proof of Proposition 3.1 and of the Addendum. We have indeed $\varphi^* = F + \varepsilon\psi^*$ and $\varepsilon \ll_{\alpha,L} \xi$, thus

$$\tilde{A} = \mathscr{R}\varphi^* \ll_{\alpha,L} \bar{F} + 2\bar{N}_0\mathscr{M}(\tfrac{T}{2}\beta^2\bar{F}_0\xi),$$

$$\tilde{B} = \mathscr{N}\varphi^* \ll_{\alpha,L} \tfrac{T}{2}\bar{N} + T\bar{N}_0\mathscr{M}(\tfrac{T}{2}\beta^2\bar{F}_0\xi),$$

and $\tilde{C} = \mathscr{R}\varphi^* - \varphi^* \ll_{\alpha,L} \bar{N} + 4\bar{N}_0\mathscr{M}(\tfrac{T}{2}\beta^2\bar{F}_0\xi)$. With our choice of $L_1 = \frac{L_0}{2}$, the implied numbers are $\beta = (\frac{4e}{3})^\alpha L_0^{-\alpha}$, $\bar{F}_0 = (\frac{3}{e})^\alpha\bar{F}$ and $\bar{N}_0 = (\frac{3}{e})^\alpha\bar{N}$. Writing $\mathscr{M}(\zeta) = \sum_{j \geq 1}\mathscr{M}_j\zeta^j$ with $0 \leq \mathscr{M}_j \leq c_1^j$ for some $c_1 > 0$, we conclude like in the remark following Definition 3.1.

3.3. *Normal form with exponentially small non-resonant remainder*

We shall prove Proposition 3.2 by applying Proposition 3.1 with $F = \frac{f}{\|f\|_{\alpha,L_0}}$ and using some "Gevrey techniques" which we now explain.

3.3.1. *Gevrey-$(\alpha + 1)$ asymptotic expansions.* — **a)** We refer *e.g.* to [Bal94] and [RS96] for a more usual framework (see also [Mal95], [Ra80,Ra84,Ra93], [To90]). We have adapted the definitions to the case of real-valued Gevrey-α functions, with some technicalities to deal more easily with nonlinear problems.

Let $\varepsilon_0 > 0$ and V a Banach algebra whose norm we denote by $\| \, . \, \|$. We shall consider functions defined in $]0, \varepsilon_0]$ with values in V; the role of V will be played later by $G^{\alpha, L}(\mathbf{T}^n \times \overline{B}_R)$ for some L, R > 0.

Definition **3.2.** — *A function* $\varphi :]0, \varepsilon_0] \to$ V *is said to admit a Gevrey-$(\alpha + 1)$ asymptotic expansion if there exist* $C, M_0 > 0$ *and* $\tilde{\varphi} = \sum_{j \geq 0} \varepsilon^j \varphi_j \in$ V$[[\varepsilon]]$ *such that*

$$\forall \varepsilon \in]0, \varepsilon_0], \ \forall j \in \mathbf{N}, \quad \varepsilon^{-j} \| \varphi(\varepsilon) - \sum_{p=0}^{j-1} \varepsilon^p \varphi_p \| \leq C M_0^{-j\alpha} j!^\alpha.$$

In such a situation we use the notation

$$\rho_0 \varphi(\varepsilon) = \varphi(\varepsilon), \qquad \rho_j \varphi(\varepsilon) = \varepsilon^{-j} (\varphi(\varepsilon) - \sum_{p=0}^{j-1} \varepsilon^p \varphi_p), \qquad j \geq 1.$$

The relation $\rho_j \varphi(\varepsilon) = \varphi_j + \varepsilon \rho_{j+1} \varphi(\varepsilon)$ implies

$$(\mathbf{3.15}) \qquad \varphi_j = \lim_{\varepsilon \to 0} \rho_j \varphi(\varepsilon), \qquad j \in \mathbf{N}.$$

In particular the asymptotic expansion $\tilde{\varphi}$ is uniquely determined by φ; we shall denote it by $\tilde{\varphi} = \mathscr{J} \varphi$. It is necessarily a Gevrey-$(\alpha + 1)$ formal series in V$[[\varepsilon]]$: $\| \varphi_j \| \leq C M_0^{-j\alpha} j!^\alpha$ for all $j \in \mathbf{N}$.

Thus the supremum of the numbers $\frac{M_0^{j\alpha}}{j!^\alpha} \sup_{\varepsilon \in]0, \varepsilon_0]} \| \rho_j \varphi(\varepsilon) \|$ is finite. But, even if it means replacing M_0 by a slightly larger number M, we prefer to use a series rather than a supremum.

Definition **3.3.** — *Let* M > 0. *A function* $\varphi :]0, \varepsilon_0] \to$ V *is said to admit a Gevrey-$(\alpha + 1)$ asymptotic expansion of type* M *if there exists* $\tilde{\varphi} = \sum_{j \geq 0} \varepsilon^j \varphi_j \in$ V$[[\varepsilon]]$ *such that*

$$\| \varphi \|_{\alpha, M, V} = \sum_{j \geq 0} \frac{M^{j\alpha}}{j!^\alpha} \sup_{\varepsilon \in]0, \varepsilon_0]} \| \rho_j \varphi(\varepsilon) \| < \infty$$

with the same notations as above. The space of such functions is denoted by $\mathscr{A}_{\alpha, M}([0, \varepsilon_0], V)$ *(and the space of all functions from* $]0, \varepsilon_0]$ *to V which admit a Gevrey-$(\alpha + 1)$ asymptotic expansion is nothing but the union over all positive numbers* M *of the spaces* $\mathscr{A}_{\alpha, M}([0, \varepsilon_0], V)$*).*

Lemma **3.7.** — *The function* $\|\,.\,\|_{\alpha,\mathrm{M},\mathrm{V}}$ *is a norm which endows* $\mathscr{A}_{\alpha,\mathrm{M}}([0,\varepsilon_0],\mathrm{V})$ *with a structure of Banach algebra. In particular, if* $\varphi,\psi \in \mathscr{A}_{\alpha,\mathrm{M}}([0,\varepsilon_0],\mathrm{V})$, $\mathscr{J}(\varphi\psi) = \mathscr{J}(\varphi)\,\mathscr{J}(\psi)$ *and*

$$\|\varphi\psi\|_{\alpha,\mathrm{M},\mathrm{V}} \leq \|\varphi\|_{\alpha,\mathrm{M},\mathrm{V}}\|\psi\|_{\alpha,\mathrm{M},\mathrm{V}}.$$

Proof. — We let the reader check that $\mathscr{A}_{\alpha,\mathrm{M}}([0,\varepsilon_0],\mathrm{V})$ is a Banach space. Let $\varphi,\psi \in \mathscr{A}_{\alpha,\mathrm{M}}([0,\varepsilon_0],\mathrm{V})$, $\mathscr{J}\varphi = \sum \varepsilon^j \varphi_j$, $\mathscr{J}\psi = \sum \varepsilon^j \psi_j$. Let $\varepsilon \in\,]0,\varepsilon_0]$ and $j \in \mathbf{N}$. Using the identity

$$\varphi(\varepsilon) = \sum_{p=0}^{j-1} \varepsilon^p \varphi_p + \varepsilon^j \rho_j \varphi(\varepsilon)$$

and then the analogous identities for ψ with $p' < j - p$, we compute

$$\varphi(\varepsilon)\psi(\varepsilon) = \sum_{p=0}^{j-1}\sum_{p'=0}^{j-p-1} \varepsilon^{p+p'}\varphi_p\psi_{p'} + \varepsilon^j \sum_{p=0}^{j-1}\varphi_p\,\rho_{j-p}\psi(\varepsilon) + \varepsilon^j(\rho_j\varphi(\varepsilon))\psi(\varepsilon).$$

Thus $\mathscr{J}(\varphi\psi) = \mathscr{J}(\varphi)\,\mathscr{J}(\psi)$ and $\rho_j(\varphi\psi) = \sum_{p=0}^{j-1}\varphi_p\,\rho_{j-p}\psi + (\rho_j\varphi)(\rho_0\psi)$, hence by virtue of (3.15):

$$(\mathbf{3.16}) \qquad \sup_{]0,\varepsilon_0]}\|\rho_j(\varphi\psi)\| \leq \sum_{p=0}^{j} \sup_{]0,\varepsilon_0]}\|\rho_p\varphi\|\,\sup_{]0,\varepsilon_0]}\|\rho_{j-p}\psi\|.$$

The inequality for the product follows then easily. $\qquad\square$

Observe that, if $\varphi \in \mathscr{A}_{\alpha,\mathrm{M}}([0,\varepsilon_0],\mathrm{V})$, $\|\varepsilon\varphi\|_{\alpha,\mathrm{M},\mathrm{V}} \leq (\varepsilon_0 + \mathrm{M}^\alpha)\|\varphi\|_{\alpha,\mathrm{M},\mathrm{V}}$ (as a consequence of the previous computations, or directly because $\rho_j(\varepsilon\varphi) = \rho_{j-1}(\varphi)$ for $j \geq 1$).

Lemma **3.8.** — *Suppose* $0 < \mathrm{M}_1 < \mathrm{M}$ *and* $\varphi \in \mathscr{A}_{\alpha,\mathrm{M}}([0,\varepsilon_0],\mathrm{V})$. *Then*

$$\mathscr{J}\varphi = 0 \quad \Rightarrow \quad \forall \varepsilon \in\,]0,\varepsilon_0],$$

$$\|\varphi(\varepsilon)\| \leq \|\varphi\|_{\alpha,\mathrm{M},\mathrm{V}}(1 - \tfrac{\mathrm{M}_1}{\mathrm{M}})^{-\alpha}\,\exp(-\alpha\mathrm{M}_1\varepsilon^{-1/\alpha}).$$

Thus the functions in the kernel of the operator $\mathscr{J}$ *are exponentially small with respect to* ε.

Proof. — Assume $\mathscr{J}\varphi = 0$ and let $\varepsilon \in\,]0,\varepsilon_0]$. For all $j \geq 0$,

$$\varepsilon^{-j}\|\varphi(\varepsilon)\| = \|\rho_j\varphi(\varepsilon)\| \leq \mathrm{M}^{-j\alpha}j!^\alpha\|\varphi\|_{\alpha,\mathrm{M},\mathrm{V}}.$$

Raising these expressions at the power $1/\alpha$, multiplying by the appropriate factor and taking the sum over j, we obtain:

$$\|\varphi(\varepsilon)\|^{1/\alpha}\,\exp(\mathrm{M}_1\varepsilon^{-1/\alpha}) = \sum_{j\geq 0}\frac{\mathrm{M}_1^j}{j!}\|\rho_j\varphi(\varepsilon)\|^{1/\alpha} \leq \|\varphi\|_{\alpha,\mathrm{M},\mathrm{V}}^{1/\alpha}(1 - \tfrac{\mathrm{M}_1}{\mathrm{M}})^{-1},$$

which yields the desired inequality. $\qquad\square$

b) It can be shown that the kernel of $\mathscr{J}$ is non-trivial if $V \neq \{0\}$. We shall now exhibit an inverse to the right for this operator by means of the formal Borel transform and of an incomplete Laplace transform.

Definition **3.4.** — *Let* $M_0 > 0$. *A formal series* $\tilde{\varphi} = \sum_{j \geq 0} \varepsilon^j \varphi_j$ *in* $V[[\varepsilon]]$ *is said to be Gevrey-*$(\alpha + 1)$ *of type* M_0, *and we write* $\tilde{\varphi} \in V[[\varepsilon]]_{\alpha, M_0}$ *in that case, if*

$$\|\tilde{\varphi}\|_{\alpha, M_0, V} = \sum_{j \geq 0} \frac{M_0^{j\alpha}}{\Gamma((j+1)\alpha)} \|\varphi_j\| < \infty.$$

The V-*valued function*

$$\hat{\phi}(\zeta) = \sum_{j \geq 0} \frac{\zeta^{(j+1)\alpha - 1}}{\Gamma((j+1)\alpha)} \varphi_j$$

is then continuous in $[0, M_0]$: *it is the formal Borel transform of* $\varepsilon \tilde{\varphi}$, *we define the truncated sum of* $\tilde{\varphi}$ *by the incomplete Laplace integral*

$$\mathscr{T}_{\alpha, M_0} \tilde{\varphi}(\varepsilon) = \varepsilon^{-1} \int_0^{M_0} \hat{\phi}(\zeta) \, e^{-\zeta \, \varepsilon^{-1/\alpha}} \, d\zeta.$$

Notice that $\zeta \hat{\phi}(\zeta)$ is a holomorphic function of $\xi = \zeta^\alpha$ in the disk $\{|\xi| < M_0^\alpha\}$, which is continuous in its closure, and $\|\zeta \hat{\phi}(\zeta)\| \leq |\xi| \, \|\tilde{\varphi}\|_{\alpha, M_0, V}$. Besides, $\varepsilon \mathscr{T}_{\alpha, M_0} \tilde{\varphi}(\varepsilon)$ is an entire function of $\varepsilon^{-1/\alpha}$.

As an example, the last part of Proposition 3.1 and its Addendum can be rephrased as follows: $\tilde{A}' = \frac{\tilde{A} - \mathscr{R}_\omega F}{\varepsilon}$, $\tilde{B}$ and $\tilde{C}$ belong to $V[[\varepsilon]]_{\alpha, M_0}$ with $V = G^{\alpha, \frac{L_0}{2}}(\mathbf{T}^n \times \overline{B}_R)$ as soon as $M_0^{-\alpha} \geq 2c^2 L_0^{-2\alpha} T \overline{F}$, and

$$(\mathbf{3.17}) \qquad \|\tilde{A}'\|_{\alpha, M_0, V} \leq \frac{4\alpha^2 c}{\Gamma(\alpha)} M_0^{-\alpha} \overline{N}, \quad \|\tilde{B}\|_{\alpha, M_0, V} \leq \frac{2c}{\Gamma(\alpha)} T \overline{N}, \quad \|\tilde{C}\|_{\alpha, M_0, V} \leq \frac{2c}{\Gamma(\alpha)} \overline{N}.$$

(Indeed, one can use Hölder's inequality to check that $\Gamma((j+1)\alpha)/j!^\alpha \geq \Gamma(\alpha)$.)

Lemma **3.9.** — *Suppose* $0 < \alpha M < M_0$ *and* $\tilde{\varphi} \in V[[\varepsilon]]_{\alpha, M_0}$. *Then* $\mathscr{T}_{\alpha, M_0} \tilde{\varphi} \in \mathscr{A}_{\alpha, M}(\mathbf{R}^+, V)$ *and*

$$\mathscr{J} \, \mathscr{T}_{\alpha, M_0} \tilde{\varphi} = \tilde{\varphi}, \quad \|\mathscr{T}_{\alpha, M_0} \tilde{\varphi}\|_{\alpha, M, V} \leq \kappa [1 - (\tfrac{\alpha M}{M_0})^\alpha]^{-1} \|\tilde{\varphi}\|_{\alpha, M_0, V},$$

where κ *is a positive number which depends only on* α.

Proof. — Let $\varphi = \mathscr{T}_{\alpha,\mathrm{M}_0} \tilde{\varphi}$, $\varepsilon > 0$, $z = \varepsilon^{-1/\alpha}$. For $j \in \mathbf{N}$, we set

$$\rho_j \varphi(\varepsilon) = \varepsilon^{-j}(\varphi(\varepsilon) - \sum_{p=0}^{j-1} \varepsilon^p \varphi_p) = \varepsilon^{-(j+1)}\Big(\int_0^{\mathrm{M}_0} \hat{\phi}(\zeta)\, e^{-z\zeta}\, d\zeta - \sum_{p=0}^{j-1} \varepsilon^{p+1} \varphi_p \Big).$$

Using the identity $\varepsilon^{p+1} = \int_0^\infty \frac{\zeta^{(p+1)\alpha-1}}{\Gamma((p+1)\alpha)}\, e^{-z\zeta}\, d\zeta$, we can write

$$\varepsilon^{j+1} \rho_j \varphi(\varepsilon) =$$

$$\int_0^{\mathrm{M}_0} \sum_{p\geq0} \zeta^{(p+1)\alpha-1} \frac{\varphi_p}{\Gamma((p+1)\alpha)}\, e^{-z\zeta}\, d\zeta - \sum_{p=0}^{j-1} \int_0^\infty \zeta^{(p+1)\alpha-1} \frac{\varphi_p}{\Gamma((p+1)\alpha)}\, e^{-z\zeta}\, d\zeta$$

$$= -\sum_{p=0}^{j-1} \Big(\int_{\mathrm{M}_0}^\infty \zeta^{(p+1)\alpha-1}\, e^{-z\zeta}\, d\zeta \Big) \frac{\varphi_p}{\Gamma((p+1)\alpha)} + \int_0^{\mathrm{M}_0} \sum_{p=j}^\infty \zeta^{(p+1)\alpha-1} \frac{\varphi_p}{\Gamma((p+1)\alpha)}\, e^{-z\zeta}\, d\zeta.$$

We observe that in the integrals of last line, either $\mathrm{M}_0^{-1}\zeta \geq 1$ and $p < j$, or $\mathrm{M}_0^{-1}\zeta \leq 1$ and $p \geq j$: in all cases $(\mathrm{M}_0^{-1}\zeta)^{(p+1)\alpha-1} \leq (\mathrm{M}_0^{-1}\zeta)^{(j+1)\alpha-1}$, therefore $\zeta^{(p+1)\alpha-1} \leq \zeta^{(j+1)\alpha-1}\mathrm{M}_0^{p\alpha-j\alpha}$ and

$$\varepsilon^{j+1} \|\rho_j \varphi(\varepsilon)\| \leq \Big(\int_0^\infty \zeta^{(j+1)\alpha-1}\, e^{-z\zeta}\, d\zeta \Big) \Big(\sum_{p=0}^\infty \mathrm{M}_0^{p\alpha-j\alpha} \frac{\|\varphi_p\|}{\Gamma((p+1)\alpha)} \Big).$$

Hence

$$\forall j \in \mathbf{N},\ \forall \varepsilon > 0, \quad \|\rho_j \varphi(\varepsilon)\| \leq \Gamma((j+1)\alpha)\mathrm{M}_0^{-j\alpha} \|\tilde{\varphi}\|_{\alpha,\mathrm{M}_0,\mathrm{V}}.$$

Using the Stirling formula we now choose $\kappa = \kappa(\alpha) > 0$ such that

$$\forall j \in \mathbf{N}, \quad \Gamma((j+1)\alpha)^{1/\alpha} \leq \kappa^{1/\alpha}\alpha^j j!$$

and the conclusion follows easily. $\qquad\qquad\square$

c) Our strategy to prove Proposition 3.2 will be to apply the operator $\mathscr{T}_{\alpha,\mathrm{M}_0}$ to the formal series $\tilde{\mathrm{A}}$ and $\tilde{\mathrm{B}}$ of Proposition 3.1 with $\mathrm{V} = \mathrm{G}^{\alpha,\mathrm{L}_0/2}(\mathbf{T}^n \times \overline{\mathrm{B}}_\mathrm{R})$ and a well-chosen number $\mathrm{M}_0 > 0$. Denoting their truncated sums $\mathscr{A}$ and $\mathscr{B}$, we shall then prove that the flow at time $\tau = 1$ for $\mathrm{X}_{\varepsilon\mathscr{B}}$, which is a symplectic map Φ, satisfies $\mathscr{J}(\mathrm{H}_\varepsilon \circ \Phi) = \langle \omega, r \rangle + \varepsilon\tilde{\mathrm{A}}$ and conclude that $\mathrm{H}_\varepsilon \circ \Phi - \langle \omega, r \rangle - \varepsilon\mathscr{A}$ is exponentially small.

3.3.2. *Gevrey-$(\alpha+1)$ asymptotics for composition of functions and for flows.* — In the sequel we suppose

$$0 < R_1 < R, \qquad K_1 = \mathbf{T}^n \times \overline{B}_{R_1} \subset K = \mathbf{T}^n \times \overline{B}_R,$$

$$0 < 4(2n)^{\frac{\alpha-1}{\alpha}} L_1 < L, \qquad V = G^{\alpha,L}(K) \subset V_1 = G^{\alpha,L_1}(K_1)$$

and we deal with functions of ε which take their values in V or V_1. Let $\varepsilon_1, M > 0$. When considering a member φ of a space like $\mathscr{A}_{\alpha,M}([0, \varepsilon_1], V)$, we shall sometimes use the notation $\varphi_\varepsilon(x)$ or $\varphi(x; \varepsilon)$ for the value at x of the function $\varphi(\varepsilon)$.

a) We need to inquire about the composition of functions which admit Gevrey-$(\alpha + 1)$ asymptotic expansions of type M. All subsequent composition products must be understood at fixed ε, with respect to the variable x.

Lemma **3.10** *(Composition of asymptotics).* — *Suppose that* $Y \in \mathscr{A}_{\alpha,M}([0, \varepsilon_1], V)$ *and* $\Psi = (\Psi_{[1]}, ..., \Psi_{[2n]}) \in \mathscr{A}_{\alpha,M}([0, \varepsilon_1], V_1^{2n})$. *If*

$$\varepsilon_1 \max\{ \|\Psi_{[1]}\|_{\alpha,M,V_1}, ..., \|\Psi_{[2n]}\|_{\alpha,M,V_1} \} \leq \min \left\{ R - R_1, \frac{L^\alpha}{2^\alpha (2n)^{\alpha-1}} - L_1^\alpha \right\}$$

then $Y \circ (\mathrm{Id} + \varepsilon \Psi) \in \mathscr{A}_{\alpha,M}([0, \varepsilon_1], V_1)$, $\mathscr{J}(Y \circ (\mathrm{Id} + \varepsilon \Psi)) = \mathscr{J}Y \circ (\mathrm{Id} + \varepsilon \mathscr{J}\Psi)$ *and*

$$\|Y \circ (\mathrm{Id} + \varepsilon \Psi)\|_{\alpha,M,V_1} \leq$$

$$\|Y\|_{\alpha,M,V} \exp\left((\tfrac{2M}{L})^\alpha (\|\Psi_{[1]}\|_{\alpha,M,V_1} + \cdots + \|\Psi_{[2n]}\|_{\alpha,M,V_1}) \right).$$

Proof. — To shorten the notations, we define for each $j \in \mathbf{N}$ a function $\rho_{j,L}^*$ on $\mathscr{A}_{\alpha,M}([0, \varepsilon_1], V)$ by

$$\rho_{j,L}^*(\varphi) = \sup_{\varepsilon \in]0, \varepsilon_1]} \|\rho_j \varphi(\varepsilon)\|_{\alpha,L}.$$

The inequality (3.16) is easily generalized: for $s, j \in \mathbf{N}$ and $\varphi^0, \varphi^1, ..., \varphi^s \in \mathscr{A}_{\alpha,M}([0, \varepsilon_1], V_1)$,

$$(\mathbf{3.18}) \qquad \rho_{j,L_1}^*(\varphi^0 \varphi^1 ... \varphi^s) \leq \sum_{\substack{j_0, j_1, ..., j_s \in \mathbf{N} \\ j_0 + j_1 + \cdots + j_s = j}} \rho_{j_1,L_1}^*(\varphi^0) \rho_{j_1,L_1}^*(\varphi^1) ... \rho_{j_s,L_1}^*(\varphi^s).$$

We thus suppose that we are given $Y \in \mathscr{A}_{\alpha,M}([0, \varepsilon_1], V)$ and $\Psi \in \mathscr{A}_{\alpha,M}([0, \varepsilon_1], V_1^{2n})$, and we study $X = Y \circ (\mathrm{Id} + \varepsilon \Psi)$. By Taylor formula, for all $\varepsilon \in]0, \varepsilon_1]$ and $j \in \mathbf{N}$,

$$(\mathbf{3.19}) \qquad X_\varepsilon = \sum_{\ell \in \mathbf{N}^{2n}; \, |\ell| \leq j-1} \frac{\varepsilon^{|\ell|}}{\ell!} (\partial_x^\ell Y_\varepsilon) \Psi_\varepsilon^\ell + \varepsilon^j \sigma_j(\varepsilon)$$

with

$$(\textbf{3.20}) \qquad \sigma_j(\varepsilon) = \sum_{\ell \in \mathbf{N}^{2n};\ |\ell|=j} \frac{1}{\ell!} \Big[\int_0^1 (\partial_x^\ell Y_\varepsilon) \circ (\mathrm{Id} + \varepsilon t \Psi_\varepsilon) j (1-t)^{j-1} dt \Big] \Psi_\varepsilon^\ell.$$

The assumption on $\varepsilon_1 \|\Psi_{[i]}\|_{\alpha,M,V_1}$ allows us to apply Corollary A.1 with L/2 in place of L; we get

$$\forall \varepsilon \in [0,\varepsilon_1], \quad \|\sigma_j(\varepsilon)\|_{\alpha,L_1} \leq \sum_{\ell \in \mathbf{N}^{2n};\ |\ell|=j} \frac{1}{\ell!} \|\partial_x^\ell Y_\varepsilon\|_{\alpha,\frac{L}{2}} \prod_{i=1}^{2n} \rho_{0,L_1}^*(\Psi_{[i]})^{\ell_i}.$$

We now see from (3.19) that X admits $\mathscr{J}Y \circ (\mathrm{Id} + \varepsilon \mathscr{J}\Psi)$ as asymptotic expansion and that

$$\rho_j(X_\varepsilon) = \sum_{|\ell| \leq j-1} \frac{1}{\ell!} \rho_j(\varepsilon^{|\ell|}(\partial_x^\ell Y_\varepsilon)\Psi_\varepsilon^\ell) + \sigma_j(\varepsilon).$$

For ℓ fixed such that $|\ell| \leq j-1$, we have $\rho_j(\varepsilon^{|\ell|}(\partial_x^\ell Y_\varepsilon)\Psi_\varepsilon^\ell) = \rho_{j-|\ell|}((\partial_x^\ell Y_\varepsilon)\Psi_\varepsilon^\ell)$; the inequality (3.18) thus yields

$$\rho_{j,L_1}^*(\varepsilon^{|\ell|}(\partial_x^\ell Y)\Psi^\ell) \leq$$

$$\sum_{j_0+j_1+\cdots+j_{|\ell|}=j-|\ell|} \rho_{j_0,L_1}^*(\partial_x^\ell Y) \prod_{i=1}^{2n} \rho_{j_{\hat{\ell}_{i-1}+1},L_1}^*(\Psi_{[i]})\cdots\rho_{j_{\hat{\ell}_i},L_1}^*(\Psi_{[i]})$$

with the notations $\hat{\ell}_0 = 0$ and $\hat{\ell}_i = \ell_1 + \cdots + \ell_i$ for $1 \leq i \leq 2n$. Using the inequality $L_1 \leq L/2$, we end up with

$$(\textbf{3.21}) \qquad \rho_{j,L_1}^*(X) \leq \sum_{\substack{|\ell| \leq j \\ j_0+j_1+\cdots+j_{|\ell|}=j-|\ell|}} \frac{1}{\ell!} \rho_{j_0,\frac{L}{2}}^*(\partial_x^\ell Y) \prod_{i=1}^{2n} \rho_{j_{\hat{\ell}_{i-1}+1},L_1}^*(\Psi_{[i]})\cdots\rho_{j_{\hat{\ell}_i},L_1}^*(\Psi_{[i]}).$$

Applying Lemma A.2 to the functions $\rho_{j_0}Y(\varepsilon) \in G^{\alpha,L}(K)$, we obtain the inequalities

$$\forall \ell \in \mathbf{N}^{2n}, \quad \frac{1}{\ell!^\alpha} \Big(\frac{L}{2}\Big)^{|\ell|\alpha} \rho_{j_0,\frac{L}{2}}^*(\partial_x^\ell Y) \leq \rho_{j_0,L}^*(Y).$$

We can now obtain a convergent sum over j from the inequalities (3.21):

$$\|X\|_{\alpha,M,V_1} \leq \sum_{\ell \in \mathbf{N}^{2n}} \sum_{j_0,j_1,\ldots,j_{|\ell|} \in \mathbf{N}} \frac{M^{j\alpha}}{j!^\alpha \ell!} \rho_{j_0,\frac{L}{2}}^*(\partial_x^\ell Y) \prod_{i=1}^{2n} \rho_{j_{\hat{\ell}_{i-1}+1},L_1}^*(\Psi_{[i]})\cdots\rho_{j_{\hat{\ell}_i},L_1}^*(\Psi_{[i]}),$$

where $j = j_0 + j_1 + \cdots + j_{|\ell|} + |\ell|$. Indeed, since $j! \geq \ell! j_0! j_1! \ldots j_{|\ell|}!$, we have

$$\|X\|_{\alpha,M,V_1} \leq \sum_{\ell \in \mathbf{N}^{2n}} \sum_{j_0 \in \mathbf{N}} \frac{M^{(j_0+|\ell|)\alpha}}{j_0!^\alpha \ell!^{\alpha+1}} \rho^*_{j_0,\frac{L}{2}} (\partial_x^\ell Y) \prod_{i=1}^{2n} \|\Psi_{[i]}\|^{\ell_i}_{\alpha,M,V_1}$$

$$\leq \sum_{j_0 \in \mathbf{N}} \frac{M^{j_0\alpha}}{j_0!^\alpha} \rho^*_{j_0,L}(Y) \sum_{\ell \in \mathbf{N}^{2n}} \frac{1}{\ell!} (\tfrac{2M}{L})^{|\ell|\alpha} \prod_{i=1}^{2n} \|\Psi_{[i]}\|^{\ell_i}_{\alpha,M,V_1},$$

and the last expression is bounded by $\|Y\|_{\alpha,M,V} \exp\big((\tfrac{2M}{L})^\alpha (\|\Psi_{[1]}\|_{\alpha,M,V_1} + \cdots + \|\Psi_{[2n]}\|_{\alpha,M,V_1})\big)$.   □

b) Since we have used the Lie method to produce a formal mapping in Proposition 3.1, we also need to study the flow of a vector field whose components admit Gevrey-$(\alpha+1)$ asymptotic expansions of type M.

Lemma **3.11** *(Asymptotics of flows).* — *Let* $Y = (Y_{[1]}, ..., Y_{[2n]}) \in \mathscr{A}_{\alpha,M}([0,\varepsilon_1], V^{2n})$ *and consider the vector field* $\frac{dx}{d\tau} = \varepsilon Y_\varepsilon(x)$ *for each* $\varepsilon \in [0,\varepsilon_1]$. *Let* $C \leq \min\{\frac{R-R_1}{\varepsilon_1}, \frac{1}{\varepsilon_1}(\frac{L^\alpha}{2^\alpha(2n)^{\alpha-1}} - L_1^\alpha)\}$. *If*

$$(\mathbf{3.22}) \qquad \sum_{1 \leq i \leq 2n} \|Y_{[i]}\|_{\alpha,M,V} \leq C \exp\big(-(\tfrac{2M}{L})^\alpha C\big)$$

and

$$(\mathbf{3.23}) \qquad \sum_{1 \leq i_1 \leq 2n} \max_{1 \leq i_2 \leq 2n} \|\partial_{x_{i_2}} Y_{[i_1]}\|_{\alpha,M,V} \leq \frac{1}{2(\varepsilon_1 + M^\alpha)} \exp\big(-(\tfrac{2M}{L})^\alpha C\big),$$

then the flow at time $\tau = 1$ *of* $\varepsilon Y_\varepsilon$ *is a well defined mapping of the form* $\mathrm{Id} + \varepsilon \Psi_\varepsilon$, *with* $\Psi \in \mathscr{A}_{\alpha,M}([0,\varepsilon_1], V_1^{2n})$. *Moreover* $\mathrm{Id} + \varepsilon \mathscr{J} \Psi(x;\varepsilon)$ *coincides with the flow at time 1 of the formal vector field* $\frac{dx}{d\tau} = \varepsilon \mathscr{J} Y(x;\varepsilon)$, *and*

$$\|\Psi_{[1]}\|_{\alpha,M,V_1} + \cdots + \|\Psi_{[2n]}\|_{\alpha,M,V_1} \leq C.$$

Proof. — Let us consider $\mathscr{A} = \mathscr{A}_{\alpha,M}([0,\varepsilon_1], V_1^{2n})$ as a Banach space for the norm

$$\|\Psi\| = \|\Psi_{[1]}\|_{\alpha,M,V_1} + \cdots + \|\Psi_{[2n]}\|_{\alpha,M,V_1}, \quad \Psi = (\Psi_{[1]}, ..., \Psi_{[2n]}) \in \mathscr{A}.$$

If $\Psi \in \mathscr{A}$ with $\|\Psi\| \leq C$, using (3.22) we can apply Lemma 3.10 and define a functional $\mathscr{G}$:

$$\mathscr{G}(\Psi) = Y \circ (\mathrm{Id} + \varepsilon \Psi) \in \mathscr{A}, \quad \|\mathscr{G}(\Psi)\| \leq C.$$

If moreover $\Psi' \in \mathscr{A}$ with $\|\Psi'\|_{\alpha,M,V_1} \leq C$, we get $\|\mathscr{G}(\Psi') - \mathscr{G}(\Psi)\| \leq \frac{1}{2}\|\Psi' - \Psi\|$ by virtue of (3.23) and of the identity

$$\mathscr{G}(\Psi')_{[i]} - \mathscr{G}(\Psi)_{[i]} =$$

$$< \int_0^1 \partial_x Y_{[i]} \circ (\mathrm{Id} + \varepsilon[(1-t)\Psi + t\Psi']) \, dt, \, \varepsilon(\Psi' - \Psi) >.$$

We now consider the Banach space $C^0([0,1], \mathscr{A})$ of all continuous $\mathscr{A}$-valued functions of $\tau \in [0,1]$, equipped with the norm of the supremum over $[0,1]$. The flow of the vector field $\frac{dx}{d\tau} = \varepsilon Y_\varepsilon(x)$ can be written $\mathrm{Id} + \varepsilon \Psi_{\varepsilon,\tau}$, where Ψ is a fixed point for the functional

$$(\mathscr{F}(\Psi))_{\varepsilon,\tau} = \int_0^\tau \mathscr{G}(\Psi_{\varepsilon,\tau_1}) \, d\tau_1 \quad \text{if} \quad \|\Psi\|_{C^0([0,1],\mathscr{A})} \leq C.$$

The theorem of contractions applies and we find Ψ as the limit of the iterates $\mathscr{F}^\nu(0)$, whereas $\mathscr{J}\Psi$ is the limit of $\mathscr{J}(\mathscr{F}^\nu(0))$ and thus satisfies the equation corresponding to $\mathscr{J}Y$. The Ψ of the lemma is obtained by restriction to $\tau = 1$. $\qquad\square$

c) In fact, the previous lemma will be used only in the special case where Y is the vector field generated by a Hamiltonian function of the form $\mathscr{T}_{\alpha,M_0}\tilde{B}$, according to the notation of Definition 3.4. We shall use the same number κ as in Lemma 3.9.

Lemma **3.12** *(Special case).* — *Suppose* $M_0 > \alpha M$ *and* $\tilde{B} \in V[[\varepsilon]]_{\alpha,M_0}$. *Let* $\mathscr{B} = \mathscr{T}_{\alpha,M_0}\tilde{B}$ *and consider the Hamiltonian vector field* $\frac{dx}{d\tau} = \varepsilon X_{\mathscr{B}_\varepsilon}(x)$ *for each* $\varepsilon \in [0, \varepsilon_1]$. *If*

$$\Delta \leq \min\{ R - R_1, \, \tfrac{L^\alpha}{4^\alpha (2n)^{\alpha-1}} - L_1^\alpha \}$$

and

$$(3.24) \qquad \|\tilde{B}\|_{\alpha,M_0,V} \leq \frac{1}{\varepsilon_1 \kappa}[1 - (\tfrac{\alpha M}{M_0})^\alpha] \exp\!\left(-(\tfrac{4}{L})^\alpha \Delta \tfrac{M^\alpha}{\varepsilon_1}\right) \times$$

$$\min\left\{ (\tfrac{L}{2})^\alpha \Delta, \, \tfrac{1}{2}(\tfrac{L^2}{8})^\alpha (1 + \tfrac{M^\alpha}{\varepsilon_1})^{-1} \right\},$$

then the flow at time $\tau = 1$ *of* $\varepsilon X_{\mathscr{B}_\varepsilon}$ *is a well defined symplectic mapping of the form* $\mathrm{Id} + \varepsilon \Psi_\varepsilon$, *with* $\Psi \in \mathscr{A}_{\alpha,M}([0, \varepsilon_1], V_1^{2n})$. *Moreover* $\mathrm{Id} + \varepsilon \mathscr{J}\Psi(x; \varepsilon)$ *coincides with* $\exp(\varepsilon L_{\tilde{B}_\varepsilon})$, *and*

$$\|\Psi_{[1]}\|_{\alpha,M,V_1} + \cdots + \|\Psi_{[2n]}\|_{\alpha,M,V_1} \leq \frac{\Delta}{\varepsilon_1}.$$

Proof. — The conclusion will follow from the possibility of applying the previous lemma with $X_{\mathscr{B}}$ in place of Y and $L/2$ in place of L. We just to need to check that

$$\sum_{1 \le i \le 2n} \|\partial_{x_i} \mathscr{B}\|_{\alpha, M, V'} \le C \, \exp\!\big(-(\tfrac{4M}{L})^\alpha C\big),$$

$$\sum_{1 \le i_1 \le 2n} \max_{1 \le i_2 \le 2n} \|\partial_{x_{i_2}} \partial_{x_{i_1}} \mathscr{B}\|_{\alpha, M, V'} \le \frac{1}{2(\varepsilon_1 + M^\alpha)} \exp\!\big(-(\tfrac{4M}{L})^\alpha C\big),$$

where $V' = G^{\alpha, L/2}(K)$ and $C = \Delta/\varepsilon_1$.

But Lemma A.2 shows that $\partial_x^m \tilde{B} \in V'[[\varepsilon]]_{\alpha, M_0}$ for each $m \in \mathbf{N}^{2n}$ and that

$$\sum_{m \in \mathbf{N}^{2n}} \frac{1}{m!^\alpha} (\tfrac{L}{2})^{|m|\alpha} \|\partial_x^m \tilde{B}\|_{\alpha, M_0, V'} \le \|\tilde{B}\|_{\alpha, M_0, V}.$$

Applying Lemma 3.9, we get $\partial_x^m \mathscr{B} = \mathscr{T}_{\alpha, M_0}(\partial_x^m \tilde{B}) \in \mathscr{A}_{\alpha, M}(\mathbf{R}^+, V')$ for each $m \in \mathbf{N}^{2n}$ and

$$\sum_{m \in \mathbf{N}^{2n}} \frac{1}{m!^\alpha} (\tfrac{L}{2})^{|m|\alpha} \|\partial_x^m \mathscr{B}\|_{\alpha, M, V'} \le \kappa \big[1 - (\tfrac{\alpha M}{M_0})^\alpha\big]^{-1} \|\tilde{B}\|_{\alpha, M_0, V}.$$

Extracting the terms which correspond to $|m| = 1$ or $|m| = 2$ and using (3.24), we obtain the desired inequalities. □

3.3.3. *Gevrey-$(\alpha+1)$ normal form.* — We now prove Proposition 3.2. Let us suppose that its hypotheses are fulfilled. We use the same notations as in the statement, except that we define

$$\varepsilon_0 = \|f\|_{\alpha, L_0}, \quad F = \frac{1}{\varepsilon_0} f,$$

and we consider ε as a free parameter in $]0, \varepsilon_0]$: the chain of reasoning which follows will deal with the Hamiltonian function

$$H_\varepsilon(\theta, r) = \, <\omega, r> + \varepsilon F(\theta, r),$$

and the identity $\varepsilon = \varepsilon_0$ will be restored at the end only. Thus, $\|F\|_{\alpha, L_0} = 1$ and $\|F - \mathscr{R}_\omega F\|_{\alpha, L_0} = \varepsilon'/\varepsilon_0$, where $\varepsilon' = \|f - \mathscr{R}_\omega f\|_{\alpha, L_0}$.

We define $L = \frac{L_0}{2}$, $L_1 = \frac{L_0}{16(2n)^{\frac{\alpha-1}{\alpha}}}$ and $\Delta = \min\{\delta, L_1^\alpha\}$. Observe that $4(2n)^{\frac{\alpha-1}{\alpha}} L_1 < L < L_0$. We shall use the notations

$$0 < R_1 = R - \delta < R, \qquad K_1 = \mathbf{T}^n \times \overline{B}_{R_1} \subset K = \mathbf{T}^n \times \overline{B}_R,$$

and

$$V_0 = G^{\alpha, L_0}(K) \subset V = G^{\alpha, L}(K) \subset V_1 = G^{\alpha, L_1}(K_1).$$

a) We define an $(n-1)$-dimensional torus Σ_0 associated with ω as in (3.1) and (3.2), and apply Proposition 3.1 to F. We obtain a formal resonant series $\tilde{A}$, and

a formal series $\tilde{B}$ vanishing on $\Sigma_0 \times \overline{B}_R$, such that the formal flow $\mathrm{Id} + \varepsilon\tilde{\Psi}$ at time $\tau = 1$ of $X_{\varepsilon\tilde{B}}$ satisfies

$$H_\varepsilon \circ (\mathrm{Id} + \varepsilon\tilde{\Psi}) = <\omega, r> + \varepsilon\tilde{A}.$$

Of course all these objects are formal series with respect to ε. Setting $\tilde{C} = - <\omega, \partial_\theta\tilde{B}>$, Proposition 3.1 and its Addendum provide also a positive number c_a depending only on α and such that

$$\tilde{A}' = \frac{\tilde{A} - \mathcal{R}_\omega F}{\varepsilon}, \tilde{B}, \tilde{C} \in V[[\varepsilon]]_{\alpha,M_0}, \qquad M_0^{-\alpha} = c_a L_0^{-2\alpha} T,$$

and

$$\|\tilde{A}'\|_{\alpha,M_0,V} \leq c_a M_0^{-\alpha} \frac{\varepsilon'}{\varepsilon_0}, \quad \|\tilde{B}\|_{\alpha,M_0,V} \leq c_a T \frac{\varepsilon'}{2\varepsilon_0} \leq c_a T, \quad \|\tilde{C}\|_{\alpha,M_0,V} \leq c_a \frac{\varepsilon'}{\varepsilon_0}.$$

b) In order to apply Lemma 3.9, we define $M > 0$ by

$$M^{-\alpha} = c_b L_0^{-\alpha} \Delta^{-1} T$$

with $c_b = c_a \max\{\alpha^\alpha, 4^{\alpha+1}e\kappa\}$ (where κ is the positive number which appears in the statement of Lemma 3.9; one can check that $\alpha M \leq M_0/16$ because $L_0^\alpha \Delta^{-1} \geq 16^\alpha$). This way we define functions of (θ, r, ε) which belong to $\mathscr{A}_{\alpha,M}(\mathbf{R}^+, V)$,

$$\mathscr{A}' = \mathscr{T}_{\alpha,M_0}\tilde{A}', \quad \mathscr{B} = \mathscr{T}_{\alpha,M_0}\tilde{B}, \quad \mathscr{C} = \mathscr{T}_{\alpha,M_0}\tilde{C},$$

and which satisfy

$$\|\mathscr{A}'\|_{\alpha,M,V} \leq c_b' M_0^{-\alpha} \frac{\varepsilon'}{\varepsilon_0}, \quad \|\mathscr{B}\|_{\alpha,M,V} \leq c_b' T \frac{\varepsilon'}{2\varepsilon_0},$$

$$\|\mathscr{C}\|_{\alpha,M,V} \leq c_b' \frac{\varepsilon'}{\varepsilon_0}, \qquad c_b' = \frac{16}{15}\kappa c_a.$$

As functions of θ, the first one is resonant and the second one vanishes on Σ_0. Notice that the numbers c_b and c_b' depend only on α.

c) The number c_b was chosen sufficiently large to ensure the fulfillment of the hypotheses of Lemma 3.12 with

$$\varepsilon_1 = M^\alpha = (c_b L_0^{-\alpha} \Delta^{-1} T)^{-1}.$$

The flow $\Phi^{\varepsilon\mathscr{B}_\varepsilon}$ at time $\tau = 1$ of $X_{\varepsilon\mathscr{B}_\varepsilon}$ is thus a symplectic transformation of the form $\mathrm{Id} + \varepsilon\Psi_\varepsilon$, with $\Psi \in \mathscr{A}_{\alpha,M}([0, \varepsilon_1], V_1^{2n})$, $\mathscr{J}\Psi = \tilde{\Psi}$ and

$$\|\Psi^{(1)}\|_{\alpha,M,V_1} + \cdots + \|\Psi^{(2n)}\|_{\alpha,M,V_1} \leq \frac{\Delta}{\varepsilon_1} = c_b L_0^{-\alpha} T.$$

d) Let $h(r) = <\omega, r>$ and $\mathscr{A}_\varepsilon(\theta, r) = \mathscr{R}_\omega F(\theta, r) + \varepsilon \mathscr{A}'_\varepsilon(\theta, r)$. We now compare the conjugate Hamiltonian $H_\varepsilon \circ \Phi^{\varepsilon \mathscr{B}_\varepsilon}$ and the resonant Hamiltonian $h + \varepsilon \mathscr{A}_\varepsilon$:

$$\mathscr{E}_\varepsilon = \frac{1}{\varepsilon}\left(H_\varepsilon \circ \Phi^{\varepsilon \mathscr{B}_\varepsilon} - (h + \varepsilon \mathscr{A}_\varepsilon)\right) = \mathscr{C}^*_\varepsilon + (F - \mathscr{R}_\omega F) \circ \Phi^{\varepsilon \mathscr{B}_\varepsilon} - \varepsilon \mathscr{A}'_\varepsilon,$$

$$\mathscr{C}^*_\varepsilon = \frac{1}{\varepsilon}\left((h + \varepsilon \mathscr{R}_\omega F) \circ \Phi^{\varepsilon \mathscr{B}_\varepsilon} - (h + \varepsilon \mathscr{R}_\omega F)\right).$$

Since $\varepsilon \mathscr{C}_\varepsilon = - <\omega, \varepsilon \partial_\theta \mathscr{B}_\varepsilon>$ is precisely the derivative of h along the vector field $X_{\varepsilon \mathscr{B}_\varepsilon}$, we can write

$$\mathscr{C}^*_\varepsilon = \int_0^1 (\mathscr{C}_\varepsilon + \varepsilon \mathscr{D}_\varepsilon) \circ \Phi^{\tau \varepsilon \mathscr{B}_\varepsilon} \, d\tau, \qquad \mathscr{D}_\varepsilon = \{\mathscr{B}_\varepsilon, \mathscr{R}_\omega F\}.$$

According to Lemma 3.10, we have

$$\mathscr{E} \in \mathscr{A}_{\alpha, M}([0, \varepsilon_1], V_1), \quad \mathscr{J}\mathscr{E} = 0, \quad \|\mathscr{E}\|_{\alpha, M, V_1} \leq c_d \frac{\varepsilon'}{\varepsilon_0},$$

with c_d depending only on α (indeed, we can apply Lemma 3.10 to $(F - \mathscr{R}_\omega F) \circ (\mathrm{Id} + \varepsilon \Psi_\varepsilon)$, but also to $(\mathscr{C}_\varepsilon + \varepsilon \mathscr{D}_\varepsilon) \circ \Phi^{\tau \varepsilon \mathscr{B}_\varepsilon} = (\mathscr{C}_\varepsilon + \varepsilon \mathscr{D}_\varepsilon) \circ (\mathrm{Id} + \varepsilon \Psi_{\tau, \varepsilon})$ using the proof of Lemma 3.11 to bound $\|\Psi\|_{C^0([0,1], \mathscr{A})}$.

We are in a position to apply Lemma 3.8 with $M_1 = M/2 = \frac{1}{2}\varepsilon_1^{1/\alpha}$: for each $\varepsilon \in [0, \varepsilon_1]$,

$$\|\mathscr{E}_\varepsilon\|_{\alpha, L_1} \leq 2^\alpha \exp(-\frac{\alpha}{2}(\varepsilon_1/\varepsilon)^{1/\alpha})\|\mathscr{E}\|_{\alpha, M, V_1}.$$

The choice of $c = c(\alpha)$ is now easy and we conclude by setting $\varepsilon = \varepsilon_0$ (observe that we need to bound $\varepsilon_0 T$ in order to ensure $\varepsilon_0 \leq \varepsilon_1$) and by renaming $\varepsilon \mathscr{A}, \varepsilon \mathscr{E}, \varepsilon \Psi$ as $\mathscr{A}, \mathscr{E}, \Psi$ respectively.

3.4. *Proof of Theorem A*

We now explain how to deduce our Gevrey version of the Nekhoroshev Theorem from Proposition 3.2 by adapting the method presented in [Lo92], [LN92] and [LNN93]. The periodic orbit method consists in finding, close to a given initial condition (θ_0, r_0), a periodic torus $\mathbf{T}^n \times \{r^*\}$ of the unperturbed system, and to use the previous work in conjunction with quasi-convexity to show stability over exponentially long time intervals for all solutions starting near this torus.

For any $r \in \mathbf{R}^n$ and $R > 0$ we shall use the notation $\overline{B}(r, R)$ for the closed ball of centre r and radius R. We shall treat all at once the global result and the local results associated with resonant surfaces $S_{\mathscr{M}}$, using the simple observation that $S_{\{0\}} = \overline{B}_{R_0}$ with the notations of Section 1.2.

From now on we fix a function h satisfying the hypotheses of Theorem A: $h \in G^{\alpha,L}(\overline{B}_R)$, $\|h\|_{\alpha,L} \leq E$ and we are given $\varpi, m > 0$ such that the inequalities (1.3) hold, *i.e. h is quasi-convex.* We also fix $R_0 \in]0, R[$ and define $R^* = \frac{R+R_0}{2}$.

We consider $n \geq 1$ and $R, R_0, L, E, \varpi, m > 0$ as data, with respect to which the dependence of all subsequent constants will be made explicit (but we do not try to track the dependence with respect to α).

3.4.1. *Normal form for the original Hamiltonian.* — We first rephrase Proposition 3.2 in the original setting. If $\omega \in \mathbf{R}^n$, we still denote by $\mathscr{R}_\omega g$ the resonant part of a continuous function g with respect to ω as defined by (3.3).

Lemma 3.13. — *Suppose that we are given* $r^* \in \overline{B}_{R^*}$ *with* $\nabla h(r^*)$ *rational of period* T, *and a Hamiltonian function* $H \in G^{\alpha,L}(\mathbf{T}^n \times \overline{B}_R)$ *satisfying*

$$\|H - h\|_{\alpha,L} \leq \gamma \mu^2 E,$$

for some $\gamma > 0$ *and* $\mu \in]0, 1]$. *Let* $\tilde{R} > 0$ *such that* $5\mu\tilde{R} \leq R - R^*$.

Then there exists a positive number c, which depends only on α, *such that, if*

$$(\mathbf{3.25}) \qquad \mu T \leq \nu, \quad \text{where} \quad \nu = \left[cEL^{-\alpha}\left(\gamma + (1 + 10\tilde{R}L^{-\alpha})^2 \right) \right]^{-1} \min\{\tilde{R}, (2n)^{1-\alpha}L^\alpha\},$$

there exist $g, \Psi^{(1)}, ..., \Psi^{(2n)} \in C^\infty(\mathbf{T}^n \times \overline{B}(r^*, 4\mu\tilde{R}))$ *satisfying*

i) $\mathrm{Id} + \Psi$ *is a symplectic transformation whose image is contained in* $\mathbf{T}^n \times \overline{B}(r^*, 5\mu\tilde{R})$,
ii) $H \circ (\mathrm{Id} + \Psi) = h + g$,
iii) $\|g\|_{C^0(\mathbf{T}^n \times \overline{B}(r^*, 4\mu\tilde{R}))} \leq c\mu^2\gamma E$,

$$\sum_{i=1}^{n} \|\partial_{\theta_i}(g - \mathscr{R}_{\nabla h(r^*)}g)\|_{C^0(\mathbf{T}^n \times \overline{B}(r^*, 4\mu\tilde{R}))} \leq$$

$$c(2n)^{\alpha-1}\mu^2\gamma EL^{-\alpha} \exp\left(-\alpha\left(\frac{\nu}{\mu T}\right)^{1/\alpha}\right),$$

$$\sum_{i=n+1}^{2n} \|\Psi^{(i)}\|_{C^0(\mathbf{T}^n \times \overline{B}(r^*, 4\mu\tilde{R}))} \leq \frac{\mu^2 T}{\nu}\tilde{R}.$$

Proof. — Let $\omega = \nabla h(r^*)$ and $f = H - h$. We shall use the rescaled variable $\tilde{r} = \frac{r-r^*}{\mu}$ instead of r. We thus deal with the new Hamiltonian function

$$\tilde{H}(\theta, \tilde{r}) = \frac{1}{\mu}H(\theta, r^* + \mu\tilde{r}) = \frac{1}{\mu}h(r^*) + <\omega, \tilde{r}> + \tilde{f}(\theta, \tilde{r}),$$

where $\tilde{f}$ is easily deduced from the Taylor formula applied to h. Indeed, writing

$$h(r^* + \mu\tilde{r}) = h(r^*) + \mu < \omega, \tilde{r} > + \mu^2 \tilde{h}(\tilde{r}),$$

$$\tilde{h}(\tilde{r}) = \int_0^1 < \nabla^2 h(r^* + t\mu\tilde{r})\tilde{r}, \tilde{r} > (1 - t)\, dt,$$

we obtain

$$\tilde{f}(\theta, \tilde{r}) = \mu\tilde{h}(\tilde{r}) + \frac{1}{\mu} f(\theta, r^* + \mu\tilde{r}).$$

Since $\mu \leq 1$, we easily check that $\tilde{f} \in G^{\alpha, \tilde{L}}(\mathbf{T}^n \times \overline{B}_{5\tilde{R}})$ for any $\tilde{L} \in]0, L[$ (of course we use Lemma A.2 to bound the second-order derivatives of h). Let us choose $\tilde{L} = 2^{-1/\alpha}L$: we find

$$\tilde{\varepsilon} = \|\tilde{f}\|_{\alpha, \tilde{L}} \leq c_0 \mu \tilde{E}, \quad \tilde{\varepsilon}' = \|\tilde{f} - \mathcal{R}_\omega \tilde{f}\|_{\alpha, \tilde{L}} \leq 2\mu\gamma E,$$

for some $c_0 = c_0(\alpha) > 0$, with $\tilde{E} = E(\gamma + (1 + 10\tilde{R}L^{-\alpha})^2)$.

Applying Proposition 3.2 with $\delta = \tilde{R}$ yields a symplectic transformation $\mathrm{Id} + \tilde{\Psi}$, a resonant function $\tilde{\mathcal{A}} = \mathcal{R}_\omega \tilde{f} + \tilde{\mathcal{A}}'$ and an exponentially small remainder $\tilde{\mathcal{E}}$, such that

$$\tilde{H} \circ (\mathrm{Id} + \tilde{\Psi}) = \frac{1}{\mu} h(r^*) + < \omega, \tilde{r} > + \tilde{\mathcal{A}} + \tilde{\mathcal{E}}$$

and these functions belong to $G^{\alpha, \tilde{L}_1}(\mathbf{T}^n \times \overline{B}_{4\tilde{R}})$, with $\tilde{L}_1 = 2^{-4}(2n)^{-\frac{\alpha-1}{\alpha}}\tilde{L}$. All this provided that $\tilde{\varepsilon}T \leq \tilde{\nu} = c_1^{-1}\tilde{L}^\alpha \min\{\tilde{R}, \tilde{L}_1^\alpha\}$, for some $c_1 = c_1(\alpha)$.

Coming back to the original variables, we define the symplectic transformation $\mathrm{Id} + \Psi$ by

$$\Psi^{(i)}(\theta, r) = \tilde{\Psi}^{(i)}\left(\theta, \frac{r - r^*}{\mu}\right),$$

$$\Psi^{(n+i)}(\theta, r) = \mu\tilde{\Psi}^{(n+i)}\left(\theta, \frac{r - r^*}{\mu}\right), \quad i = 1, ..., n.$$

The result is $H \circ (\mathrm{Id} + \Psi) = h + g$ with

$$g(\theta, r) = \mathcal{R}_\omega f(\theta, r) + \mu(\tilde{\mathcal{A}}' + \tilde{\mathcal{E}})\left(\theta, \frac{r - r^*}{\mu}\right), \quad g - \mathcal{R}_\omega g = \mu(\tilde{\mathcal{E}} - \mathcal{R}_\omega\tilde{\mathcal{E}}).$$

We let the reader check the desired inequalities for $c = c(\alpha)$ large enough.   □

3.4.2. *Confinement by quasi-convexity.* — The interest of making the non-resonant part $g - \mathscr{R}_{\nabla h(r^*)}g$ of the perturbation exponentially small lies in the possibility of applying a geometric argument which makes use of quasi-convexity to produce a confinement of all orbits starting near $\mathbf{T}^n \times \{r^*\}$.

We introduce $\Omega = L^{-\alpha}E$, $M = 2^\alpha L^{-2\alpha}E$ and $M' = 6^\alpha L^{-3\alpha}E$, so that

$$(\mathbf{3.26}) \qquad \sum_{|k|=1} \|\partial^k h\|_{C^0(\overline{B}_R)} \leq \Omega, \quad \sum_{|k|=2} \|\partial^k h\|_{C^0(\overline{B}_R)} \leq M, \quad \sum_{|k|=3} \|\partial^k h\|_{C^0(\overline{B}_R)} \leq M'.$$

(The third number will be used later, in (3.33).)

We first observe that, at each point r of $\overline{B}_R$, the property of quasi-convexity, which was initially imposed in the hyperplane orthogonal to $\nabla h(r)$, can be extended to the complement of a cone around $\nabla h(r)$ like in [Pö93]:

$$(\mathbf{3.27}) \qquad \forall v \in \mathbf{R}^n, \quad |<\nabla h(r), v>| \leq \ell_0 \|v\| \;\Rightarrow\; <\nabla^2 h(r)v, v> \geq m_0 \|v\|^2,$$
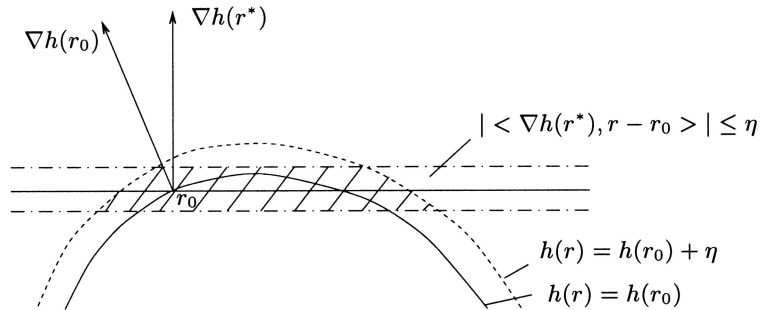
with, for instance,

$$(\mathbf{3.28}) \qquad \ell_0 = \frac{m\varpi}{8M}, \quad m_0 = \frac{m}{2}.$$

Lemma **3.14.** — *Suppose that* $r^*, r_0, r \in \overline{B}_R$ *and* $\rho, \eta > 0$ *satisfy*

$$\|r_0 - r^*\| \leq \rho, \quad h(r) - h(r_0) \leq \eta, \quad |<\nabla h(r^*), r - r_0>| \leq \eta.$$

We assume moreover that $\rho \leq \frac{\ell_0}{4M}$. *Then*

$$\|r - r_0\| \leq 2 \max \left\{ \frac{M\rho + \sqrt{m_0 \eta}}{m_0}, \frac{2\eta}{\ell_0} \right\} \quad \text{or} \quad \|r - r_0\| \geq \frac{2\ell_0}{m_0}.$$

Corollary **3.1.** — *Let* $\lambda = 8M/m_0$ *and* $r^* \in \overline{B}_{R^*}$ *with a rational frequency-vector* $\nabla h(r^*)$. *We suppose* $\tau, \rho > 0$, $g \in C^2(\mathbf{T}^n \times \overline{B}(r^*, \lambda \rho))$ *with*

$$(\mathbf{3.29}) \qquad \|g\|_{C^0(\mathbf{T}^n \times \overline{B}(r^*, \lambda\rho))} \leq \frac{1}{2}m_0 \rho^2, \quad \sum_{i=1}^n \|\partial_{\theta_i}(g - \mathscr{R}_{\nabla h(r^*)}g)\|_{C^0(\mathbf{T}^n \times \overline{B}(r^*, \lambda\rho))} \leq \frac{m_0 \rho^2}{\Omega \tau},$$

and $\rho \leq \min\{\frac{\ell_0}{5M}, \frac{R-R^*}{\lambda}\}$. Then any initial condition $(\theta_0, r_0) \in \mathbf{T}^n \times \overline{B}(r^*, \rho)$ gives rise to a solution $(\theta(t), r(t)) \in \mathbf{T}^n \times \overline{B}(r^*, \lambda\rho)$ of the Hamiltonian vector field generated by $h(r) + g(\theta, r)$ which is defined at least for $|t| \leq \tau$ and satisfies $\|r(t) - r_0\| \leq \lambda\rho/2$ in that range.

Proof of the lemma. — Let $v = r - r_0$. Bounding the components of $\nabla h(r_0) - \nabla h(r^*)$ by (3.26), we obtain

$$(\mathbf{3.30}) \qquad | < \nabla h(r_0), v > | \leq \eta + M\rho\|v\|.$$

Let us assume that $\|v\| \geq \frac{4\eta}{\ell_0}$. We define a function

$$\varphi(t) = < \nabla h(r_0 + tv), v >, \qquad t \in [0, 1].$$

The previous inequalities and the assumption on ρ imply $|\varphi(0)| \leq \frac{\ell_0}{2}\|v\|$. But, according to quasi-convexity as expressed by (3.27),

$$\forall t \in [0, 1], \quad |\varphi(t)| > \ell_0\|v\| \quad \text{or} \quad \varphi'(t) \geq m_0\|v\|^2.$$

From that we easily deduce that

$$\forall t \in [0, 1], \quad \varphi(t) \geq \min\{\ell_0\|v\|, \varphi(0) + m_0\|v\|^2 t\}.$$

If we now assume $\|v\| \leq \frac{\ell_0}{2m_0}$, observing that $\varphi(0) + m_0\|v\|^2 \leq \ell_0\|v\|$, we obtain

$$\eta \geq h(r) - h(r_0) = \int_0^1 \varphi(t)\,dt \geq \varphi(0) + \frac{1}{2}m_0\|v\|^2.$$

Hence, by (3.30),

$$\eta \geq \frac{1}{2}m_0\|v\|^2 - \eta - M\rho\|v\|,$$

and the conclusion follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Proof of the corollary. — Let $t \mapsto x(t) = (\theta(t), r(t)$ be the solution, whose interval of definition we denote by $] - \tau_1, \tau_2[$, and fix $t \in [-\tau, \tau] \cap\,] - \tau_1, \tau_2[$ (thus $r(t) \in \overline{B}(r^*, \lambda\rho) \subset \overline{B}_R$ by hypothesis). On the one hand, energy conservation implies

$$\left| h(r(t)) - h(r_0) \right| = |g(x(t)) - g(x(0))| \leq m_0\rho^2.$$

On the other hand, the assumption on the non-resonant part of g implies

$$\left| \frac{d}{dt} < \nabla h(r^*), r(t) > \right| = | < \nabla h(r^*), \nabla_\theta g(x(t)) > | \leq \frac{m_0\rho^2}{\tau},$$

hence $| < \nabla h(r^*), r(t) - r_0 > | \leq \eta = m_0\rho^2$. We are in a position to apply Lemma 3.14: we obtain $\|r(t) - r_0\| \leq 2\max\{\frac{M\rho + \sqrt{m_0\eta}}{m_0}, \frac{2\eta}{\ell_0}\}$ (the other possibility is excluded by our assumption $\|r(t) - r_0\| \leq \lambda\rho$), from which we deduce $\|r(t) - r_0\| \leq \lambda\rho/2$ (using $\lambda > 8$) and $\|r(t) - r^*\| \leq (1 + \frac{\lambda}{2})\rho < \lambda\rho$.

This shows that $[-\tau, \tau] \subset\,] - \tau_1, \tau_2[$, since taking t close to the boundary of the interval of definition of the solution would yield $x(t)$ close to the boundary of $\mathbf{T}^n \times \overline{B}(r^*, \lambda\rho)$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

3.4.3. *Use of Dirichlet's Theorem.* — Given an initial condition (θ_0, r_0), we need to find a nearby periodic torus $\mathbf{T}^n \times \{r^*\}$ in order to use the previous results. We shall use an elementary fact from the theory of Diophantine approximation, the Dirichlet principle (see *e.g.* [Ca59]), to approach the given frequency-vector $\omega_0 = \nabla h(r_0)$ by a rational vector ω, close enough to it but whose period is not too large. Section 3.4.4 will then show how to find r^* with $\nabla h(r^*)$ proportional to ω.

But first, since we wish to be able to improve our results when the initial condition lies near a resonant surface $S_{\mathcal{M}}$, we need some algebraic preliminaries. We shall follow [Lo92] about the notion of "order" of a resonance submodule, with slight modifications. We denote by $\|\cdot\|_\infty$ the sup norm in $\mathbf{R}^n$, while $\|\cdot\|$ stands as always for the Euclidean norm, and we denote by $|||\cdot|||_\infty$ and $|||\cdot|||$ the corresponding operator norms for square n-rowed matrices. Let $(e_1, ..., e_n)$ be the canonical basis of $\mathbf{R}^n$.

If we are given a submodule $\mathcal{M}$ of $\mathbf{Z}^n$ of rank $m \in \{0, ..., n-1\}$, in order to describe the $\mathcal{M}$-resonant vectors ω of $\mathbf{R}^n$ (*i.e.* the vector space $\mathcal{M}^\perp$), we can choose a basis $(k_1, ..., k_m)$ of $\mathcal{M}$ and build a matrix K with m rows and n columns such that $\omega \in \mathcal{M}^\perp \Leftrightarrow \mathrm{K}\omega = 0$ and the elements of K are integers (we view the members of $\mathcal{M}$ as linear forms). According to a classical result from linear algebra (see *e.g.* [Art91], Chap. 12), we can diagonalize this matrix, and write $\mathrm{K} = \mathrm{B}\Delta\mathrm{A}$ with $\mathrm{A} \in \mathrm{G}\ell_n(\mathbf{Z})$, $\mathrm{B} \in \mathrm{G}\ell_m(\mathbf{Z})$, $\Delta = [\mathrm{D} \,|\, 0_{m \times (n-m)}]$, where $\mathrm{D} = \mathrm{diag}(p_1, ..., p_m)$, the p_i are positive integers and $p_i | p_j$ for $i \leq j$.

The numbers p_i are called the *invariant factors* of K; they are uniquely determined. What is interesting for us is the matrix A, since a vector ω belongs to $\mathcal{M}^\perp$ if and only if the m first components of $\mathrm{A}\omega$ vanish.

Definition **3.5.** — *Consider all such decompositions* $\mathrm{K} = \mathrm{B}\Delta\mathrm{A}$; *we call* $c'_{\mathcal{M}}$ *the minimal value of* $|||\mathrm{A}|||_\infty$, *and* $c''_{\mathcal{M}}$ *the minimal value of* $|||\mathrm{A}^{-1}|||_\infty$ *for the decompositions such that* $c'_{\mathcal{M}} = |||\mathrm{A}|||_\infty$. *Moreover, if the codimension* $d = n - m$ *is greater than 1, we define* $c_{\mathcal{M}}$ *to be the minimal value of* $\sqrt{d-1}|||\mathrm{A}^{-1}|||$ *for the decompositions such that* $c'_{\mathcal{M}} = |||\mathrm{A}|||_\infty$ *and* $c''_{\mathcal{M}} = |||\mathrm{A}^{-1}|||_\infty$. *If* $d = 1$ *we put* $c_{\mathcal{M}} = 1$.

The numbers $c_{\mathcal{M}}, c'_{\mathcal{M}}, c''_{\mathcal{M}}$ depend on $\mathcal{M}$ only, and not on the particular basis $(k_1, ..., k_m)$ that we have chosen at the beginning, since any other basis $(k'_1, ..., k'_m)$ would have given rise to a matrix K' of the form PK with $\mathrm{P} \in \mathrm{G}\ell_m(\mathbf{Z})$, and we would have used the minimal decompositions $(\mathrm{PB})\Delta\mathrm{A}$ instead of $\mathrm{B}\Delta\mathrm{A}$.

Remark **3.3.** — If $m = 0$, $\mathcal{M} = \{0\}$ (and $d = n$), we have $c_{\mathcal{M}} = \sqrt{d-1}$ and $c'_{\mathcal{M}} = c''_{\mathcal{M}} = 1$. This is also the case for a "standard" resonance submodule, *i.e.* when $\mathcal{M}^\perp$ is generated by some vectors of the canonical basis of $\mathbf{R}^n$.

Lemma **3.15.** — *Let* $\mathcal{M}$ *a submodule of* $\mathbf{Z}^n$ *of rank* $m \in \{0, ..., n-1\}$ *and codimension* $d = n - m$. *For any* ω_0 *in* $\mathcal{M}^\perp \setminus \{0\}$ *and for any real number* $\mathrm{Q} > 1$, *there exist a positive*

number T *and a rational vector* $\omega \in \mathscr{M}^\perp$ *of period* T *such that*

$$\frac{1}{c'_{\mathscr{M}}} \leq \mathrm{T}\|\omega_0\|_\infty < c''_{\mathscr{M}}\mathrm{Q}, \quad \mathrm{T}\|\omega - \omega_0\| \leq c_{\mathscr{M}}\,\mathrm{Q}^{-\frac{1}{d-1}}.$$

In the case of maximal rank, i.e. $d = 1$, *one can take* $\mathrm{Q} = 1$ *and the last inequality must be interpreted as the equality* $\omega - \omega_0 = 0$.

Proof. — Let us proceed as indicated above to describe $\mathscr{M}^\perp$: we write $\mathrm{K} = \mathrm{B}\Delta\mathrm{A}$, with minimal $|||\mathrm{A}|||_\infty$, so that

$$\omega \in \mathscr{M}^\perp \Leftrightarrow \mathrm{A}\omega \in \mathrm{Vect}(e_{m+1}, ..., e_n).$$

Let $\omega_0 \in \mathscr{M}^\perp$ and $\mathrm{Q} > 1$. Setting $\omega_0^* = \mathrm{A}\omega_0$ and $\xi = \|\omega_0^*\|_\infty$, we can renumber the indices between $m + 1$ and n in such a way that $\omega_0^* = \xi(0, ..., 0, \pm 1, \omega_0')$ for some $\omega_0' \in \mathbf{R}^{d-1}$. The question is now reduced to approximation in $\mathbf{R}^{d-1}$. Dirichlet's Theorem yields $q \in \mathbf{N}^*$ and $\ell' \in \mathbf{Z}^{d-1}$ such that

$$q < \mathrm{Q}, \quad \|q\omega_0' - \ell'\|_\infty \leq \mathrm{Q}^{-\frac{1}{d-1}}.$$

Let $\omega^* = \xi(0, ..., 0, \pm 1, \frac{\ell'}{q})$ and $\omega = \mathrm{A}^{-1}\omega^* \in \mathscr{M}^\perp$. We have $\|\omega^* - \omega_0^*\| \leq \sqrt{d-1}\frac{\xi}{q}\mathrm{Q}^{-\frac{1}{d-1}}$, thus $\|\omega - \omega_0\| \leq \frac{c_{\mathscr{M}}\xi}{q}\mathrm{Q}^{-\frac{1}{d-1}}$. One checks easily that ω is rational and that its period T satisfies

$$\frac{1}{\xi} \leq \mathrm{T} \leq \frac{q}{\xi}.$$

The conclusion stems from the observation that $\frac{1}{c'_{\mathscr{M}}} \leq \frac{1}{\xi}\|\omega_0\|_\infty \leq c''_{\mathscr{M}}$.   □

Remark **3.4.** — Another possible choice would have been $c_{\mathscr{M}} = c''_{\mathscr{M}}\sqrt{n-1}$ (with $c'_{\mathscr{M}}$ and $c''_{\mathscr{M}}$ defined as before), as easily seen by a slight modification of the end of the proof (observe that $\|\omega^* - \omega_0^*\|_\infty \leq \frac{\xi}{q}\mathrm{Q}^{-\frac{1}{d-1}}$, thus $\mathrm{T}\|\omega - \omega_0\|_\infty \leq c''_{\mathscr{M}}\mathrm{Q}^{-\frac{1}{d-1}}$).

Apart from the case of standard submodules, the computation of the numbers $c_{\mathscr{M}}, c'_{\mathscr{M}}, c''_{\mathscr{M}}$ may be quite tedious. Still, we can give an example which is used in § 2.4.3:

Lemma **3.16.** — *Let* $p_1 < \cdots < p_{n-1}$ *be positive integers such that any two of them are mutually prime. Let* $\mathscr{M}$ *be the submodule of* $\mathbf{Z}^n$ *of rank* $m = n - 1$ *generated by the vectors* $k_1, ..., k_m$, *where*

$$(\mathbf{3.31}) \qquad k_i = (0, ..., 0, p_i, 0, ..., 0, -1), \qquad 1 \leq i \leq m.$$
$$\underset{(i^{th} \text{ place})}{}$$

Then $p_m \leq c'_{\mathscr{M}} < (m-1)p_m$ *and* $p_1...p_m \leq c''_{\mathscr{M}} < m(m-1)^m p_m^m$ *for* $m \geq 2$, *whereas* $c'_{\mathscr{M}} = c''_{\mathscr{M}} = p_1 + 1$ *for* $m = 1$.

Proof. — Let

$$K = \begin{pmatrix} p_1 & & -1 \\ & \ddots & \vdots \\ & & p_m & -1 \end{pmatrix}.$$

We have a decomposition

$$K = \begin{pmatrix} 1 & & 0 \\ & \ddots & \vdots \\ & & 1 & 0 \end{pmatrix} A, \qquad A = \begin{pmatrix} p_1 & & -1 \\ & \ddots & \vdots \\ & & p_m & -1 \\ u_1 & \cdots & u_m & 0 \end{pmatrix},$$

with $A \in G\ell_n(\mathbf{Z})$, as soon as the integers $u_1, ..., u_m$ satisfy

(**3.32**) $$\sum_{i=1}^{m} u_i \check{P}_i = \pm 1, \qquad \text{where} \quad \check{P}_i = p_1 ... p_{i-1} p_{i+1} ... p_m \quad \text{for} \quad i = 1, ..., m.$$

Let us suppose $m \geq 2$. The existence of solutions for (3.32) is equivalent to the g.c.d. of the numbers $\check{P}_i$ being 1 and stems directly from our hypothesis on $p_1, ..., p_m$; thus all the invariant factors are equal to 1.

Let us choose a solution $u = (u_1, ..., u_m)$ of (3.32) with $+1$ in the right-hand side. By adding to it an appropriate linear combination of the vectors of $\mathbf{Z}^n$

$$(-p_1, 0, ..., 0, p_i, 0, ..., 0), \qquad i \geq 2$$
$$\underset{(i^{th} \text{ place})}{}$$

(which are solutions of the corresponding homogeneous equation), we can produce a solution u^* with $|u_i^*| < p_i/2$ for $i \geq 2$. The equation $u_1^* \check{P}_1 = 1 - u_2^* \check{P}_2 - \cdots - u_m^* \check{P}_m$ implies $|u_1^*| \leq (m-1)p_1/2$; hence $|u_1^*| + \cdots + |u_m^*| < (m-1)p_m$. Since $|||A|||_\infty$ is nothing but the maximum of the sums of the absolute values of the elements of a row, we end up with $c'_\mathcal{M} \leq |||A|||_\infty < (m-1)p_m$.

Let us now suppose $K = B\Delta A$ with $B \in G\ell_m(\mathbf{Z})$, $A \in G\ell_{m+1}(\mathbf{Z})$, $|||A|||_\infty = c'_\mathcal{M}$ and $|||A^{-1}|||_\infty = c''_\mathcal{M}$. We retain from the previous inequality that each element of A has absolute value less than $(m-1)p_m$. Each cofactor of A has absolute value less than $(m-1)^m p_m^m$, hence $c''_\mathcal{M} < m(m-1)^m p_m^m$.

To obtain lower bounds, introduce the matrix $C = B^{-1}$ and denote by $C_1, ..., C_m$ its columns. The equation $CK = \Delta A$ yields

$$A = \begin{pmatrix} \boxed{p_1 C_1} & \cdots & \boxed{p_m C_m} & \boxed{-\sum C_i} \\ u_1 & \cdots & u_m & u_{m+1} \end{pmatrix},$$

with some integers $u_1, ..., u_{m+1}$. The determinant of C being ± 1, the column C_m has at least one non-zero element, thus $c'_{\mathcal{M}} \geq p_m$, and the last cofactor of A is precisely $p_1...p_m$, hence $c''_{\mathcal{M}} \geq p_1...p_m$.

The easy case $m = 1$ is left to the reader. $\square$

Notice that we had $c_{\mathcal{M}} = 1$ because $d = n - m = 1$.

3.4.4. *Use of isoenergetic non-degeneracy.* — We now introduce

$$K = \frac{1}{2}\big(L^\alpha E^{-1} + 2^\alpha L^{-\alpha} + m^{-1}(1 + 2^{\alpha+1}L^{-\alpha})^2\big)^{-1},$$

(**3.33**)

$$\rho_0 = \frac{1}{4}\min\left\{\frac{K}{M'}, \frac{K}{M}, 1, R - R_0\right\}.$$

Lemma **3.17.** — *For all* $r_0 \in \overline{B}_{R_0}$ *and* $\omega \in \overline{B}(\nabla h(r_0), K\rho_0)$, *there exist* $r^* \in \overline{B}_{R^*}$ *and* $\xi \in [\frac{1}{2}, 2]$ *such that*

$$h(r^*) = h(r_0), \quad \nabla h(r^*) = \xi\omega, \quad \|r^* - r_0\| \leq K^{-1}\|\omega - \omega_0\|.$$

Proof. — Let $\Phi : (\xi, r) \mapsto (h(r), \xi\nabla h(r))$. We shall prove that Φ induces a diffeomorphism from $\overline{B}((1, r_0), 2\rho_0)$ to a neighbourhood of $(h(r_0), \nabla h(r_0))$ which contains a ball of radius $K\rho_0$.

It is sufficient to check that, for each $(\xi, r) \in [\frac{1}{2}, 2] \times \overline{B}_R$, the linearized map $D\Phi(\xi, r)$ is invertible and

(**3.34**) $\|D\Phi(\xi, r)^{-1}\| \leq K^{-1}.$

Indeed, one can check that each second-order partial derivative of Φ is bounded by $\frac{K}{4\rho_0}$, and we have ensured $2\rho_0 \leq \frac{1}{2}$ and $R_0 + 2\rho_0 \leq R^*$.

Let $(\xi, r) \in [\frac{1}{2}, 2] \times \overline{B}_R$, $\omega = \nabla h(r)$ and $A = \nabla^2 h(r)$. The linearized map $D\Phi(\xi, r)$ is easily computed:

$$D\Phi(\xi, r)(u, v) = (<\omega, v>, u\omega + \xi A v), \qquad (u, v) \in \mathbf{R} \times \mathbf{R}^n.$$

To invert it, we use the projections Π and $\Pi^\perp$ from $\mathbf{R}^n$ onto $\mathbf{R}\omega$ and $\omega^\perp$, and the following consequence of (1.3):

$$A_\perp = (\Pi^\perp \circ A)_{|\omega^\perp} \text{ is invertible and } \|A_\perp^{-1}\| \leq m^{-1}.$$

If $(u_1, v_1) \in \mathbf{R} \times \mathbf{R}^n$, one can determine $D\Phi(\xi, r)^{-1}(u_1, v_1) = (u, v)$ by

$$v = \frac{u_1}{\|\omega\|^2}\omega + A_\perp^{-1}(\xi^{-1}\Pi^\perp v_1 - \frac{u_1}{\|\omega\|^2}\Pi^\perp A\omega), \quad u\omega = v_1 - \xi A v.$$

The inequality $\|(u_1, v_1)\| \leq K^{-1}\|(u, v)\|$ follows easily from $\|\omega\| \leq L^{-\alpha}E$ and $\|A\| \leq 2^\alpha L^{-2\alpha}E$. $\square$

We now suppose that $\mathcal{M}$ is a submodule of $\mathbf{Z}^n$ of codimension $d \in \{1, ..., n\}$ and put together the last two lemmas:

Corollary 3.2. — *For all $r_0 \in \overline{B}_{R_0}$ and $Q > 1$ such that*

$$Q \geq \left(\frac{2c_{\mathcal{M}}c'_{\mathcal{M}}\Omega}{\rho_0 K}\right)^{d-1},$$

there exists $r^* \in \overline{B}_{R^*}$ such that $\nabla h(r^*)$ is rational and its period T satisfies

$$(\mathbf{3.35}) \qquad \frac{1}{2c'_{\mathcal{M}}\Omega} \leq T \leq \frac{2c''_{\mathcal{M}}}{\varpi}Q, \quad \|r^* - r_0\| \leq \operatorname{dist}(r_0, S_{\mathcal{M}}) + c_{\mathcal{M}}K^{-1}T^{-1}Q^{-\frac{1}{d-1}},$$

where $c_{\mathcal{M}}, c'_{\mathcal{M}}, c''_{\mathcal{M}}$ were introduced in Lemma 3.15. In the case of maximal rank, i.e. when $d = 1$, one can take $Q = 1$ and the last inequality must be interpreted as $\|r^* - r_0\| \leq \operatorname{dist}(r_0, S_{\mathcal{M}})$.

Proof. — Let $r'_0 \in S_{\mathcal{M}}$ such that $\|r'_0 - r_0\| = \operatorname{dist}(r_0, S_{\mathcal{M}})$ and apply the last two lemmas with r'_0 in place of r_0. $\qquad\square$

3.4.5. *End of the proof of Theorem A.* — Let $H \in G^{\alpha,L}(\mathbf{T}^n \times \overline{B}_R)$ and $\varepsilon = \|H - h\|_{\alpha,L}/E$. Let $\sigma > 0$ and $(\theta_0, r_0) \in \mathbf{T}^n \times \overline{B}_R$ such that $\operatorname{dist}(r_0, S_{\mathcal{M}}) \leq \sigma\varepsilon^{1/2}$.

a) We set

$$Q = Q_1^{d-1}\varepsilon^{-\frac{d-1}{2d}}, \qquad Q_1 = \left(\frac{\varpi c_{\mathcal{M}}}{2Kc''_{\mathcal{M}}\sigma}\right)^{\frac{1}{d}}.$$

In view of applying Corollary 3.2, we impose a first "validity threshold" if $d \geq 2$:

$$(\mathbf{T1}) \qquad \varepsilon^{\frac{1}{2d}} \leq Q_1 \min\left\{\frac{1}{2}, \frac{\rho_0 K}{2c_{\mathcal{M}}c'_{\mathcal{M}}\Omega}\right\}.$$

This way Q satisfies the hypotheses of the corollary (no such condition is needed if $d = 1$), and we obtain $r^* \in \overline{B}_{R^*}$ such that $\nabla h(r^*)$ is rational of period T and (3.35) is fulfilled. As a consequence (whether $d \geq 2$ or not),

$$(\mathbf{3.36}) \qquad \frac{\varpi}{2c''_{\mathcal{M}}Q_1^{d-1}}\varepsilon^{\frac{1}{2}} \leq \frac{\varepsilon^{\frac{1}{2d}}}{T} \leq 2c'_{\mathcal{M}}\Omega\varepsilon^{\frac{1}{2d}},$$

and $\sigma\varepsilon^{1/2} \leq c_{\mathcal{M}}K^{-1}T^{-1}Q_1^{-1}\varepsilon^{1/2d}$, hence the above choice of Q_1 ensures that

$$\|r^* - r_0\| \leq \frac{2c_{\mathcal{M}}}{KQ_1} \cdot \frac{\varepsilon^{\frac{1}{2d}}}{T}.$$

b) Let $\lambda = 8M/m_0$ and

$$\Gamma = \left(\frac{2c''_{\mathscr{M}}Q_1^{d-1}}{\varpi}\right)^2, \quad \Delta = \max\left\{\frac{2c_{\mathscr{M}}}{KQ_1}, \left(\frac{c_{\mathscr{M}}\Gamma E}{2m_0}\right)^{1/2}\right\}, \quad \mu_1 = \frac{\lambda\Delta}{L^\alpha}.$$

Observe that this definition of Δ agrees with the one given after the statement of Theorem A. In view of applying Lemma 3.13, we set

$$\mu = \mu_1 \frac{\varepsilon^{\frac{1}{2d}}}{T}, \quad \gamma = \mu_1^{-2}\Gamma, \quad \tilde{R} = \lambda\mu_1^{-1}\Delta = L^\alpha,$$

and we notice that the above definitions of Γ and γ together with (3.36) ensure $\varepsilon \leq \gamma\mu^2$.

Three new thresholds appear: to ensure $\mu \leq 1$, we require

$$\textbf{(T2)} \qquad \varepsilon^{\frac{1}{2d}} \leq \frac{1}{2c'_{\mathscr{M}}\Omega\mu_1},$$

and $5\mu\tilde{R} \leq R - R^*$ will be guaranteed by

$$\textbf{(T3)} \qquad \varepsilon^{\frac{1}{2d}} \leq \frac{1}{2\lambda c'_{\mathscr{M}}\Omega\Delta} \cdot \frac{R - R_0}{10};$$

Lemma 3.13 applies provided (3.25) is satisfied, but we reinforce slightly this condition: we require

$$\textbf{(T4)} \qquad \varepsilon^{\frac{1}{2d}} \leq \frac{\nu}{2\lambda\mu_1}, \qquad \nu = \frac{1}{c(2n)^{\alpha-1}}\left(\frac{2m_0}{c_{\mathscr{M}}\lambda^2} + 2^7 EL^{-2\alpha}\right)^{-1},$$

where c is some positive number which depends only on α (the above number ν is not smaller than the one indicated in (3.25) thanks to the definitions of Δ and γ).

We get a close-to-identity symplectic transformation $\Phi = \text{Id} + \Psi$ and a transformed Hamiltonian $H \circ \Phi = h + g$ which is defined in $\mathbf{T}^n \times \overline{B}(r^*, 4\mu\tilde{R})$ and whose non-resonant part is exponentially small.

c) In view of applying Corollary 3.1 to $H \circ \Phi$, we set

$$\rho = \frac{2\mu\tilde{R}}{\lambda} = 2\Delta\frac{\varepsilon^{\frac{1}{2d}}}{T}.$$

Thanks to the definition of Δ, we can check that $\|g\|_{C^0(\mathbf{T}^n \times \overline{B}(r^*, \lambda\rho))} \leq \frac{1}{2}m_0\rho^2$. The other condition in (3.29) derives from the choice

$$\tau = \frac{2L^\alpha}{(2n)^{\alpha-1}\Omega}\exp\left(\alpha\left(\frac{\nu}{\mu_1\varepsilon^{\frac{1}{2d}}}\right)^{\frac{1}{\alpha}}\right),$$

using the information on the non-resonant part of g. But we also need to impose

$$\rho \leq \min\left\{\frac{\ell_0}{5M}, \frac{R - R^*}{\lambda}\right\},$$

and this can be achieved by requiring

(**T5**) $$\varepsilon^{\frac{1}{2d}} \leq \frac{1}{4c'_{\mathscr{M}}\Omega\Delta} \min\left\{\frac{\ell_0}{5M}, \frac{R - R_0}{2\lambda}\right\}.$$

Let $(\theta'_0, r'_0) = \Phi(\theta_0, r_0)$. The definition of Δ implies that $\|r^* - r_0\| \leq \mu\tilde{R}/\lambda$, thus $\|r^* - r'_0\| \leq \frac{\mu\tilde{R}}{\lambda} + 2\frac{\mu^2 T}{\nu}\tilde{R} \leq \rho$. Therefore, according to Corollary 3.1, this initial condition gives rise to a solution $(\theta'(t), r'(t))$ of the Hamiltonian vector field generated by $h + g$ which is defined for $|t| \leq \tau$ and satisfies $\|r'(t) - r'_0\| \leq \lambda\rho/2$; this corresponds to a solution $(\theta(t), r(t)) = \Phi^{-1}(\theta'(t), r'(t))$ of X_H, which satisfies $\|r(t) - r_0\| \leq \lambda\rho$.

To conlude, we define $\varepsilon_* > 0$ by

$$\varepsilon_*^{\frac{1}{2d}} = \frac{L^\alpha}{2^{16}c(2n)^{\alpha-1}\Delta}\left(\frac{m}{c_{\mathscr{M}}} + \frac{M^2 E}{m^2 L^{2\alpha}}\right)^{-1} \leq \frac{\nu}{2\lambda\mu_1}$$

and replace (T4) by $\varepsilon \leq \varepsilon_*$; we retain that $\tau \geq C_1 \exp(\alpha\left(\frac{\varepsilon_*}{\varepsilon}\right)^{1/2d\alpha})$, and we use (3.36) to bound $\lambda\rho$ by $C_2\varepsilon^{1/2d}$. Gathering the definitions of the five thresholds and the various constants that we have encountered, and enlarging c, we end up with the claim of Section 1.2.

A. Appendix on Gevrey classes

A.1. *Elementary properties*

We fix $\alpha \geq 1$, $n_1, n_2 \in \mathbf{N}$ and $K = \mathbf{T}^{n_1} \times \overline{B}_R$ where $\overline{B}_R$ is a closed ball in $\mathbf{R}^{n_2}$. The total number of variables will be $N = n_1 + n_2$. We recall the definitions

(**A.1**) $$\|\varphi\|_{\alpha,L} := \sum_{k \in \mathbf{N}^N} \frac{L^{|k|\alpha}}{k!^\alpha}\|\partial^k\varphi\|_{C^0(K)},$$

$$G^{\alpha,L}(K) = \{\varphi \in C^\infty(K) \mid \|\varphi\|_{\alpha,L} < \infty\}, \quad G^\alpha(K) = \bigcup_{L>0} G^{\alpha,L}(K).$$

The definition (A.1), which we did not find in the literature, is interesting only when dealing with fixed L; but if we let L vary freely, we recover the usual definition of Gevrey-α functions:

(**A.2**) $$\forall\varphi \in C^\infty(K), \quad \varphi \in G^\alpha(K) \Leftrightarrow \exists B, C > 0 \mid \forall k \in \mathbf{N}^N,$$

$$\|\partial^k\varphi\|_{C^0(K)} \leq CB^{|k|}\mathscr{M}_{|k|},$$

where $\mathscr{M}_\ell = \ell!^\alpha$. Other sequences $(\mathscr{M}_\ell)$ give rise to more general *ultradifferentiable classes* (or *Carleman classes*) instead of the Gevrey classes $G^\alpha(K)$. M. Gevrey [Ge18] himself already proved the stability under multiplication and composition of this space of functions. This question was studied in the framework of ultradifferentiable mappings by Roumieu [Ro62], and also by Dyn'kin [Dy80] with quite different methods. We cannot quote all the references on this subject, but the interested reader may consult [Ko79], [Gr99], [Th97] and the references therein. Apart from the theory of partial differential equations where they have been widely used, Gevrey functions were also considered in connection with dynamical systems problems, for instance in [El97], [GP95] and [Po00].

We could have used a norm like $\sup_{k \in \mathbf{N}^N} \frac{L^{|k|\alpha}}{k!^\alpha} \|\partial^k \varphi\|_{C^0(K)}$ instead of (A.1), but the following two lemmas would not have assumed such a simple form:

Lemma **A.1.** — *Let* $L > 0$.

$$(\mathbf{A.3}) \qquad \forall \varphi, \psi \in G^{\alpha,L}(K), \quad \|\varphi\psi\|_{\alpha,L} \leq \|\varphi\|_{\alpha,L} \|\psi\|_{\alpha,L}.$$

Proof. — As $\alpha \geq 1$, Leibniz rule implies

$$\|\partial^k(\varphi\psi)\|_{C^0(K)} \leq \sum_{k=k^{(1)}+k^{(2)}} \left(\frac{k!}{k^{(1)}!k^{(2)}!}\right)^\alpha \|\partial^{k^{(1)}}\varphi\|_{C^0(K)} \|\partial^{k^{(2)}}\psi\|_{C^0(K)}. \qquad \square$$

Lemma **A.2.** — *Suppose* $0 < \lambda < L$ *and* $\varphi \in G^{\alpha,L}(K)$. *Then all partial derivatives of* φ *belong to* $G^{\alpha,L-\lambda}(K)$ *and*

$$\sum_{m \in \mathbf{N}^N} \frac{\lambda^{|m|\alpha}}{m!^\alpha} \|\partial^m \varphi\|_{\alpha,L-\lambda} \leq \|\varphi\|_{\alpha,L}.$$

In particular, for all $j \in \mathbf{N}$,

$$\sum_{m \in \mathbf{N}^N; \, |m|=j} \|\partial^m \varphi\|_{\alpha,L-\lambda} \leq j!^\alpha \lambda^{-j\alpha} \|\varphi\|_{\alpha,L}.$$

Proof. — We observe that, for each $k \in \mathbf{N}^N$,

$$\sum_{\substack{\ell,m \in \mathbf{N}^N \\ \ell+m=k}} \frac{(L-\lambda)^{|\ell|\alpha}\lambda^{|m|\alpha}}{\ell!^\alpha m!^\alpha} = \prod_{i=1}^N \sum_{\substack{\ell_i,m_i \in \mathbf{N} \\ \ell_i+m_i=k_i}} \left(\frac{(L-\lambda)^{\ell_i}\lambda^{m_i}}{\ell_i!m_i!}\right)^\alpha$$

$$\leq \prod_{i=1}^N \left(\sum_{\substack{\ell_i,m_i \in \mathbf{N} \\ \ell_i+m_i=k_i}} \frac{(L-\lambda)^{\ell_i}\lambda^{m_i}}{\ell_i!m_i!}\right)^\alpha \leq \frac{L^{|k|\alpha}}{k!^\alpha}.$$

The following series is thus convergent:

$$\sum_{m \in \mathbf{N}^N} \frac{\lambda^{|m|\alpha}}{m!^\alpha} \|\partial^m \varphi\|_{\alpha, L-\lambda} = \sum_{\ell, m \in \mathbf{N}^N} \frac{(L-\lambda)^{|\ell|\alpha} \lambda^{|m|\alpha}}{\ell!^\alpha m!^\alpha} \|\partial^{\ell+m} \varphi\|_{C^0(K)} \leq \|\varphi\|_{\alpha, L}.$$

Extracting the terms which correspond to $|m| = j$, we obtain

$$\lambda^{j\alpha} \sum_{|m|=j} \frac{1}{m!^\alpha} \|\partial^m \varphi\|_{\alpha, L-\lambda} \leq \|\varphi\|_{\alpha, L}$$

and we can use the inequalities $m!^\alpha \leq j!^\alpha$ to conclude. $\qquad\square$

An important and well-known property of Gevrey-α classes with $\alpha > 1$ is the existence of compact-supported functions:

Lemma **A.3.** — *Let $\alpha > 1$ and $L > 0$. There exists a non-negative function $\mathbf{F}_{\alpha, L}$ in $G^{\alpha, L}([-\frac{1}{2}, \frac{1}{2}])$ whose support is included in $[-\frac{1}{4}, \frac{1}{4}]$ and such that $\mathbf{F}_{\alpha, L}(0) = 1$ and $\mathbf{F}'_{\alpha, L}(0) = 0$.*

Proof. — We define $p > 0$ by

$$\alpha = 1 + \frac{1}{p}$$

and, for each $\lambda > 0$, a function $f_\lambda \in C^\infty(\mathbf{R})$ by

$$f_\lambda(x) = \begin{vmatrix} 0 & \text{if } x \leq 0 \\ \exp(-\frac{\lambda\sqrt{2}}{x^p}) & \text{otherwise.} \end{vmatrix}$$

The conclusion will stem from the fact that $f_\lambda \in G^{\alpha, L}(\mathbf{R})$ for λ large enough with respect to p and L. Indeed it will be sufficient to consider then $x \mapsto f_\lambda(\frac{1}{4} + x) f_\lambda(\frac{1}{4} - x)/f_\lambda(\frac{1}{4})^2$ (we could also easily produce a "bump function" by using integration).

Let us thus bound the derivatives of f_λ on $\mathbf{R}$. Let $k \in \mathbf{N}$ and $x > 0$. We observe that $f_{\lambda|\mathbf{R}^+}$ extends to a holomorphic function on $\mathbf{C} - \mathbf{R}^-$. Let $\delta = \frac{\pi}{4} \min\{1, \frac{1}{p}\}$ and $\Sigma_\delta = \{z \in \mathbf{C} \mid |\arg z| \leq \delta\}$. The closed disk D_x of centre x and radius $(x \sin \delta)$ is the largest disk centred at x contained in Σ_δ, and the Cauchy inequalities yield

$$|f_\lambda^{(k)}(x)| \leq \frac{k!}{(x \sin \delta)^k} \max_{D_x} |f_\lambda|.$$

We see that $\max_{D_x} |f_\lambda| \leq \exp(-\frac{\lambda}{(2x)^p})$ (because if $z = r e^{i\theta} \in D_x$, $\Re e(z^{-p}) = r^{-p} \cos(p\theta) \geq 1/(\sqrt{2}|z|^p)$ and $|z| \leq 2x$), and the maximum of $y \mapsto y^k e^{-\lambda y^p}$ is easily seen to be $(\frac{k}{\lambda p e})^{k/p}$, therefore

$$|f_\lambda^{(k)}(x)| \leq \left(\frac{2}{\sin \delta}\right)^k \left(\frac{k}{\lambda p e}\right)^{k/p} k!.$$

Now, as soon as $\lambda > (2L^\alpha / \sin \delta)^p / p$, we get the convergence of $\sum \frac{L^{k\alpha}}{k!^\alpha} \|f_\lambda^{(k)}\|_{C^0(\mathbf{R})}$ by Stirling formula. □

For ultradifferentiable classes other than Gevrey, the question of the existence of compact-supported functions is solved by Carleman's theorem which gives a necessary and sufficient condition of quasi-analyticity (see *e.g.* [Th96]).

A.2. *Composition of Gevrey functions*

As already mentioned, the composition of Gevrey functions was considered in [Ge18], [Ro62], [Dy80] (see also [El97]). Our definition (A.1) gives rise to a particularly simple statement:

Proposition **A.1.** — *Let* L_1 *and* L *be positive numbers, and let* K_1 *and* K *be like in the previous section, with total numbers of variables* M *and* N *respectively. Let* $Y \in G^{\alpha,L}(K)$ *and let* $u = (u_{[1]}, ..., u_{[N]})$ *be a mapping whose components belong to* $G^{\alpha,L_1}(K_1)$. *If* $u(K_1) \subset K$ *and*

$$(\mathbf{A.4}) \qquad \|u_{[1]}\|_{\alpha,L_1} - \|u_{[1]}\|_{C^0(K_1)}, ..., \|u_{[N]}\|_{\alpha,L_1} - \|u_{[N]}\|_{C^0(K_1)} \leq \frac{L^\alpha}{M^{\alpha-1}},$$

then $Y \circ u \in G^{\alpha,L_1}(K_1)$ *and* $\|Y \circ u\|_{\alpha,L_1} \leq \|Y\|_{\alpha,L}$.

Proof. — We shall use the notation $(\mathbf{N}^M)^* = \{k \in \mathbf{N}^M ; |k| \geq 1\}$. The assumption (A.4) amounts to

$$(\mathbf{A.5}) \qquad A_{[i]} = \sum_{k \in (\mathbf{N}^M)^*} \frac{L_1^{|k|\alpha}}{k!^\alpha} \|\partial_x^k u_{[i]}\|_{C^0(K_1)} \leq A = \frac{L^\alpha}{M^{\alpha-1}}, \qquad i = 1, ..., N.$$

Let $X = Y \circ u$. We have $\|X\|_{C^0(K_1)} \leq \|Y\|_{C^0(K)}$ and, for each $k \in (\mathbf{N}^M)^*$,

$$\frac{1}{k!}\partial_x^k X = \sum_{\ell \in (\mathbf{N}^N)^*} \frac{1}{\ell!}(\partial_y^\ell Y) \circ u \sum_{\substack{k^1,...,k^{|\ell|} \in (\mathbf{N}^M)^* \\ k^1 + \cdots + k^{|\ell|} = k}} \frac{1}{k^1! ... k^{|\ell|}!} \prod_{i=1}^N \partial_x^{k^{\hat{\ell}_{i-1}+1}} u_{[i]} ... \partial_x^{k^{\hat{\ell}_i}} u_{[i]}$$

with the notations $\hat{\ell}_0 = 0$ and $\hat{\ell}_i = \ell_1 + \cdots + \ell_i$ for $1 \leq i \leq N$. Thus, for $k \in (\mathbf{N}^M)^*$,

$$\frac{1}{k!}\|\partial_x^k X\|_{C^0(K_1)} \leq \sum_{\ell \in (\mathbf{N}^N)^*} \frac{1}{\ell!} \|\partial_y^\ell Y\|_{C^0(K)} \sum_{\substack{k^1,...,k^{|\ell|} \in (\mathbf{N}^M)^* \\ k^1 + \cdots + k^{|\ell|} = k}} \frac{1}{k^1! ... k^{|\ell|}!} \cdot$$

$$(\mathbf{A.6})$$

$$\cdot \prod_{i=1}^N \|\partial_x^{k^{\hat{\ell}_{i-1}+1}} u_{[i]}\|_{C^0(K_1)} ... \|\partial_x^{k^{\hat{\ell}_i}} u_{[i]}\|_{C^0(K_1)}.$$

The property

(**A.7**) $\qquad \forall s \geq 1, \ \ \forall k^1, ..., k^s \in (\mathbf{N}^{\mathrm{M}})^*, \quad k = k^1 + \cdots + k^s \quad \Rightarrow \quad \dfrac{k!}{k^1!...k^s!} \geq \mathrm{M}^{-s}s!$

will be checked later. Using it now, we see that for each multi-index in the sum of the right-hand side of (A.6),

$$1 \leq \frac{\mathrm{M}^{|\ell|}k!}{|\ell|!\,k^1!...k^{|\ell|}!} \leq \frac{\mathrm{M}^{|\ell|}k!}{\ell!\,k^1!...k^{|\ell|}!} \leq \frac{\mathrm{M}^{|\ell|\alpha}k!^{\alpha}}{\ell!^{\alpha}k^1!^{\alpha}...k^{|\ell|}!^{\alpha}},$$

hence

$$\frac{1}{k!^{\alpha-1}\ell!\,k^1!...k^{|\ell|}!} \leq \frac{\mathrm{M}^{|\ell|(\alpha-1)}}{\ell!^{\alpha}k^1!^{\alpha}...k^{|\ell|}!^{\alpha}}.$$

Multiplying (A.6) by $\mathrm{L}_1^{|k|\alpha}/k!^{\alpha-1}$ and taking the sum over k, we thus obtain

$$\sum_{k \in \mathbf{N}^{\mathrm{M}}} \frac{\mathrm{L}_1^{|k|\alpha}}{k!^{\alpha}} \|\partial_x^k \mathrm{X}\|_{\mathrm{C}^0(\mathrm{K}_1)} \leq \sum_{\ell \in \mathbf{N}^{\mathrm{N}}} \frac{\mathrm{M}^{|\ell|(\alpha-1)}}{\ell!^{\alpha}} \|\partial_y^\ell \mathrm{Y}\|_{\mathrm{C}^0(\mathrm{K})} \mathrm{B}_\ell$$

with $\mathrm{B}_0 = 1$ and

$$\mathrm{B}_\ell = \sum_{k^1,...,k^{|\ell|} \in (\mathbf{N}^{\mathrm{M}})^*} \frac{\mathrm{L}_1^{|k|\alpha}}{k^1!^{\alpha}...k^{|\ell|}!^{\alpha}} \prod_{i=1}^{\mathrm{N}} \|\partial_x^{k^{\hat{\ell}_{i-1}+1}} u_{[i]}\|_{\mathrm{C}^0(\mathrm{K}_1)}...\|\partial_x^{k^{\hat{\ell}_i}} u_{[i]}\|_{\mathrm{C}^0(\mathrm{K}_1)}.$$

But, in view of (A.5),

$$\mathrm{B}_\ell = \mathrm{A}_{[1]}^{\ell_1}...\mathrm{A}_{[\mathrm{N}]}^{\ell_{\mathrm{N}}} \leq \mathrm{A}^{|\ell|}$$

and $\mathrm{M}^{\alpha-1}\mathrm{A} = \mathrm{L}^{\alpha}$, hence $\mathrm{X} \in \mathrm{G}^{\alpha,\mathrm{L}_1}(\mathrm{K}_1)$ and $\|\mathrm{X}\|_{\alpha,\mathrm{L}_1} \leq \|\mathrm{Y}\|_{\alpha,\mathrm{L}}$.

Only (A.7) remains to be checked. We begin with the case $\mathrm{M} = 1$, by induction on s: the inequality is obviously satisfied if $s = 1$, and the elementary property

$$\forall k', k'' \in \mathbf{N}^*, \quad \frac{(k' + k'')!}{k'!k''!} \geq k' + k''$$

allows us to pass from rank s to rank $s + 1$: if $k^1, ..., k^{s+1} \in \mathbf{N}^*$,

$$\frac{(k^1 + \cdots + k^s + k^{s+1})!}{k^1!...k^s!k^{s+1}!} = \frac{(k^1 + \cdots + k^s)!}{k^1!...k^s!} \frac{(k^1 + \cdots + k^s + k^{s+1})!}{(k^1 + \cdots + k^s)!k^{s+1}!}$$
$$\geq s!(s + 1).$$

Let us now consider the case $\mathrm{M} \geq 2$. We are given an array $(k_i^j)_{1 \leq i \leq \mathrm{M}, 1 \leq j \leq s}$ of non-negative integers with M rows and s columns, and the quantity to be studied is

$$\frac{k!}{k^1!...k^s!} = \pi_1...\pi_{\mathrm{M}},$$

with π_i depending only on the ith row: $\pi_i = \frac{(k_i^1 + \cdots + k_i^s)!}{k_i^1!...k_i^s!} \geq 1$. We rearrange the rows and columns as follows:

– For $j = 1$, we select the smallest i such that $k_i^1 \geq 1$ (such an index i exists because $|k^1| \geq 1$), we put the corresponding row at the top of the array and renumber the rows. Now $k_1^1 \geq 1$; we rearrange the next columns so that $k_1^1, ..., k_1^{p_1} \geq 1$ and $k_1^{p_1+1} = ... = k_1^s = 0$ if $p_1 < s$. If $p_1 = s$ we stop here, if not we continue with the reordering but shall not touch again the first row nor the first p_1 columns.

– For $j = 2$, we select the smallest i such that $k_i^{p_1+1} \geq 1$ (such an index i exists because $|k^{p_1+1}| \geq 1$, and necessarily $i \geq 2$ since $k_1^{p_1+1} = 0$), we put the corresponding row in second position and renumber the next rows. Now $k_2^{p_1+1} \geq 1$, and we rearrange the next columns so that $k_2^{p_1+1}, ..., k_2^{p_1+p_2} \geq 1$, and $k_2^{p_1+p_2+1} = ... = k_2^s = 0$ if $p_1 + p_2 < s$. If $p_1 + p_2 = s$ we stop here, if not we continue with the reordering but shall not touch again the first two rows nor the first $p_1 + p_2$ columns.

– And so on for $j = 2, ..., s$.

With such a renumbering of the integers k_i^j, the value of the product $\pi_1...\pi_M$ is not altered, there are integers $M_1 \geq 1$ and $p_1, ..., p_{M_1} \geq 1$ such that

$$M_1 \leq M, \quad p_1 + \cdots + p_{M_1} = s$$

and the first M_1 rows ($M_1 \leq M$) have been reordered: for $1 \leq i \leq M_1$,

$$k_i^{\hat{p}_{i-1}+1}, ..., k_i^{\hat{p}_i} \geq 1$$

(still with the notations $\hat{p}_0 = 0$ and $\hat{p}_i = p_1 + \cdots + p_i$ for $1 \leq i \leq M_1$). Now

$$\pi_i \geq \frac{(k_i^{\hat{p}_{i-1}+1} + ... + k_i^{\hat{p}_i})!}{k_i^{\hat{p}_{i-1}+1}!...k_i^{\hat{p}_i}!}, \qquad 1 \leq i \leq M_1$$

(because the ratio of these two quantities is again a multinomial coefficient), and the last quantity is $\geq p_i!$ from the $M = 1$ case. Since $\frac{(p_1+\cdots+p_{M_1})!}{p_1!...p_{M_1}!} \leq M_1^s$, we have $p_1!...p_{M_1}! \geq M_1^{-s}s!$, hence

$$\frac{k!}{k^1!...k^s!} \geq \pi_1...\pi_{M_1} \geq M^{-s}s!.$$

$\square$

We use in Section 3.3.2 the following corollary of our composition result:

Corollary **A.1.** — *Suppose* $0 < R_1 < R$ *and* $0 < (2n)^{\frac{\alpha-1}{\alpha}}L_1 < L$ *and let* $K_1 = \mathbf{T}^n \times \overline{B}_{R_1} \subset K = \mathbf{T}^n \times \overline{B}_R$. *If* $Y \in G^{\alpha,L}(K)$, $\varphi = (\varphi_{[1]}, ..., \varphi_{[2n]}) \in (G^{\alpha,L_1}(K_1))^{2n}$ *and*

$$\|\varphi_{[1]}\|_{\alpha,L_1}, ..., \|\varphi_{[2n]}\|_{\alpha,L_1} \leq \min\left\{R - R_1, \frac{L^\alpha}{(2n)^{\alpha-1}} - L_1^\alpha\right\},$$

then $Y \circ (\text{Id} + \varphi) \in G^{\alpha,L_1}(K_1)$ *and* $\|Y \circ (\text{Id} + \varphi)\|_{\alpha,L_1} \leq \|Y\|_{\alpha,L}$.

Proof. — Let $N = 2n$ and $u = (u_{[1]}, ..., u_{[N]}) = \text{Id} + \varphi$. We have $u(K_1) \subset K$, because $\|\varphi_{[i]}\|_{C^0(K_1)} \leq R - R_1$ for each i. On the other hand, $\|\partial_x^k u_{[i]}\|_{C^0(K_1)} \leq 1 + \|\partial_x^k \varphi_{[i]}\|_{C^0(K_1)}$ if $\partial_x^k = \partial_{x_i}$, and $\|\partial_x^k u_{[i]}\|_{C^0(K_1)} = \|\partial_x^k \varphi_{[i]}\|_{C^0(K_1)}$ if $|k| \geq 2$, so

$$\|u_{[i]}\|_{\alpha, L_1} - \|u_{[i]}\|_{C^0(K_1)} = \sum_{k \in (\mathbf{N}^M)^*} \frac{L_1^{|k|\alpha}}{k!^\alpha} \|\partial_x^k u_{[i]}\|_{C^0(K_1)}$$

$$\leq L_1^\alpha + \|\varphi_{[i]}\|_{\alpha, L_1} \leq \frac{L^\alpha}{(2n)^{\alpha-1}}.$$

We can thus apply Proposition A.1. $\qquad\square$

Remark **A.1.** — We did not manage to get rid of the terms $M^{\alpha-1}$ and $(2n)^{\alpha-1}$ which involve the dimension in our composition results. These terms appear in the combinatorics, and we do not know how to replace them by 1, apart from the analytic or one-dimensional cases.

However, it may happen that the dependence of u with respect to the various variables in Proposition A.1 is well separated and that the hypotheses can be weakened. For instance, we use the following result in Section 2.4:

Suppose $R > 0$ and $0 < L_1 < L$, and let $Y \in G^{\alpha, L}(\mathbf{T}^n \times \overline{B}_R)$ and $\varphi \in G^{\alpha, L_1}(\mathbf{T})$. Consider

$$u(\theta, r, t) = (\theta + \varphi(t)r, r), \qquad \theta \in \mathbf{T}^n, \ r \in \overline{B}_R, \ t \in \mathbf{T}.$$

If

(**A.8**) $\qquad (L_1^\alpha + R)\|\varphi\|_{\alpha, L_1} - R\|\varphi\|_{C^0(\mathbf{T})} \leq L^\alpha - L_1^\alpha,$

then $Y \circ u \in G^{\alpha, L_1}(\mathbf{T}^n \times \overline{B}_R)$ and $\|Y \circ u\|_{\alpha, L_1} \leq \|Y\|_{\alpha, L}$.

The proof consists in adapting the proof of Proposition A.1 to the peculiar form of u, but the details are tedious. (Begin with the case where Y does not depend on r, from which the general case is easily deduced, and observe that the left-hand side of (A.5) becomes $A_{[i]} = L_1^\alpha + (L_1^\alpha + R)\|\varphi\|_{\alpha, L_1} - R\|\varphi\|_{C^0(\mathbf{T})}$; it is sufficient to require $A_{[i]} \leq L^\alpha$ because $\partial_\theta^q \partial_r^p \partial_t^\ell u_{[i]} = 0$ as soon as $|q| + |p| \geq 2$, and we are thus essentially dealing with one index $\ell \in \mathbf{N}$ like in a one-dimensional case.)

Acknowledgements

REFERENCES

[Arn64] V. I. ARNOLD, Instability of dynamical systems with several degrees of freedom, *Dokl. Akad. Nauk SSSR* **156** (1964), 9–12; *Soviet Math. Dokl.* **5** (1964), 581–585.

[AA67] V. I. ARNOLD and A. AVEZ, *Problèmes ergodiques de la mécanique classique*, Gauthier-Villars, Paris (1967).

[Art91] M. ARTIN, *Algebra*, Prentice Hall, Englewood Cliffs, New Jersey (1991).

[Bae95] C. BAESENS, Gevrey series and dynamical bifurcations for analytic slow-fast mappings, *Nonlinearity* **8** (1995), 179–201.

[Bal94] W. BALSER, *From divergent power series to analytic functions*, Lecture Notes in Mathematics 1582, Springer-Verlag, Berlin Heidelberg (1994).

[BG86] G. BENETTIN and G. GALLAVOTTI, Stability of motions near resonances in quasi-integrable Hamiltonian systems, *J. Phys. Stat.* **44** (1986), 293–338.

[BGG85] G. BENETTIN, L. GALGANI and A. GIORGILLI, A proof of Nekhoroshev's theorem for the stability times in nearly integrable Hamiltonian systems, *Celestial Mech.* **37** (1985), 1–25.

[Be96] U. BESSI, An approach to Arnold's diffusion through the calculus of variations, *Nonlinear Anal. TMA* **26** (1996), 1115–1135.

[Be97] U. BESSI, Arnold's example with three rotators, *Nonlinearity* **10** (1997), 763–781.

[BT99] S. BOLOTIN and D. TRESCHEV, Unbounded growth of energy in nonautonomous Hamiltonian systems, *Nonlinearity* **12** (1999), 365–388.

[CD93] M. CANALIS-DURAND, Solution formelle Gevrey d'une équation singulièrement perturbée, *Asymptotic Analysis* **7** (1993), 1–32.

[Ca59] J. W. S. CASSELS, *An introduction to the geometry of numbers*, Grundlehren Math. Wiss. 99, Springer-Verlag, Berlin Göttingen Heidelberg (1959).

[DLS00] A. DELSHAMS, R. DE LA LLAVE and T. M. SEARA, A geometric approach to the existence of orbits with unbounded energy in generic periodic perturbations by a potential of generic geodesic flows of T^2, *Comm. Math. Phys.* **209** (2000), 353–392.

[Do86] R. DOUADY, Stabilité ou instabilité des points fixes elliptiques, *Ann. Sc. Éc. Norm. Sup.* **21** (1988), 1–46.

[Dy80] E. M. DYN'KIN, Pseudo-analytic extension of smooth functions. The uniform scale, *Twelve papers in analysis*, Amer. Math. Soc. Transl. **115** (1980), 33–58.

[El97] L. H. ELIASSON, Discrete one-dimensional quasi-periodic Schrödinger operators with pure point spectrum, *Acta Math.* **179** (1997), 153–196.

[Ga86] G. GALLAVOTTI, Quasi-integrable mechanical systems, *Phénomènes critiques, systèmes aléatoires, théories de jauge*, part II, Les Houches 1984, K. Osterwalder and R. Stora eds., North-Holland, Amsterdam New York (1986), 539–624.

[Ge18] M. GEVREY, Sur la nature analytique des solutions des équations aux dérivées partielles, *Ann. Sc. Éc. Norm. Sup.* **35** (1928), 129–190.

[Gr99] T. GRAMCHEV, Gevrey class, *Encyclopaedia of Mathematics*, Suppl. II, Kluwer, Doordrecht (1999).

[GP95] T. GRAMCHEV and G. POPOV, Nekhoroshev type estimates for billiard ball maps, *Ann. Inst. Fourier* **45** (1995), 859–895.

[Ko79] H. KOMATSU, The implicit function theorem for ultradifferentiable mappings, *Proc. Japan Acad. Ser. A* **55** (1979), 69–72.

[Ku93] S. KUKSIN, On the inclusion of an almost integrable analytic symplectomorphism into a Hamiltonian flow, *Russian J. Math. Phys.* **1** (1993), 191–207.

[KP94] S. KUKSIN and J. PÖSCHEL, On the inclusion of analytic symplectic maps in analytic Hamiltonian flows and its applications, *Seminar on Dynamical Systems (St. Petersburg, 1991)*, Progr. Nonlinear Differential Equations Appl. 12, Birkhäuser, Basel (1994), 96–116.

[Lo92] P. LOCHAK, Canonical perturbation theory via simultaneous approximation, *Usp. Mat. Nauk* **47** (1992), 59–140; *Russian Math. Surveys* **47** (1992), 57–133.

[Lo93] P. LOCHAK, Hamiltonian perturbation theory: periodic orbits, resonances and intermittency, *Nonlinearity* **6** (1993), 885–904.

[LN92] P. LOCHAK and A. I. NEISHTADT, Estimates in the theorem of N. N. Nekhorocheff for systems with a quasi-convex Hamiltonian, *Chaos* **2** (1992), 495–499.

[LNN93] P. LOCHAK, A. I. NEISHTADT and L. NIEDERMAN, Stability of nearly integrable convex Hamiltonian systems over exponentially long times, *Proc. 1991 Euler Institute Conf. on Dynamical Systems*, Birkhäuser, Boston (1993).

[LMS03] P. LOCHAK, J.-P. MARCO and D. SAUZIN, On the splitting of the invariant manifolds in multidimensional near-integrable Hamiltonian systems, *Memoirs of the Amer. Math. Soc.* **163** (2003).

[Mat93] J. MATHER, Variational construction of connecting orbits, *Ann. Inst. Fourier* **43** (1993), 1349–1386.

[Mal95] B. MALGRANGE, Resommation des séries divergentes, *Expo. Math.* **13** (1995), 163–222.

[Nei84] A. I. NEISHTADT, The separation of motions in systems with rapidly rotating phase, *Prikl. Mat. Mekh.* **48**, 2 (1984), 197–204; *J. Appl. Math. Mech.* **48**, 2 (1984), 133–139.

[Nekh77] N. N. NEKHOROSHEV, An exponential estimate of the time of stability of nearly integrable Hamiltonian systems, *Usp. Mat. Nauk* **32** (1977), 5–66; *Russian Math. Surveys* **32** (1977), 1–65.

[Ni00] L. NIEDERMAN, Exponential stability for small perturbations of steep integrable Hamiltonian systems, Preprint University of Paris-Sud 2000–73 (2000), to appear in *Ergod. Th. and Dyn. Syst.*

[Po00] G. POPOV, Invariant tori, effective stability, and quasimodes with exponentially small error terms. I. Birkhoff normal forms, *Ann. Henri Poincaré* **1** (2000), 223–248.

[Pö93] J. PÖSCHEL, Nekhoroshev estimates for quasi-convex Hamiltonian systems, *Math. Z.* **213** (1993), 187–216.

[PT97] A. V. PRONIN and D. V. TRESCHEV, On the inclusion of analytic maps into analytic flows, *Reg. Chaot. Dyn.* **2**, 2 (1997), 14–24.

[Ra80] J.-P. RAMIS, Les séries *k*-sommables et leurs applications, *Proc. 1979 'Les Houches' (Analysis, Microlocal Calculus and Relativistic Quantum Theory)*, Lecture Notes in Physics 126, Springer-Verlag, Berlin Heidelberg (1980), 178–199.

[Ra84] J.-P. RAMIS, Théorèmes d'indices Gevrey pour les équations différentielles ordinaires, *Mem. Am. Math. Soc.* **296** (1984).

[Ra93] J.-P. RAMIS, *Séries divergentes et théories asymptotiques*, Panoramas et synthèses 1, Soc. Math. de France (1993).

[Ro62] C. ROUMIEU, Ultradistributions définies sur $\mathbf{R}^n$ et sur certaines classes de variétés différentiables, *J. Analyse Math.* **10** (1962–1963), 153–192.

[RS96] J.-P. RAMIS and R. SCHÄFKE, Gevrey separation of slow and fast variables, *Nonlinearity* **9** (1996), 353–384.

[Sa92] D. SAUZIN, Caractère Gevrey des solutions formelles d'un problème de moyennisation, *C. R. Acad. Sci. Paris* **315** (1992), Série I, 991–995.

[Th96] V. THILLIEZ, Quelques propriétés de quasi-analyticité, *Gazette des Mathématiciens* **70** (1996), 49–68.

[Th97] V. THILLIEZ, Sur les fonctions composées ultradifférentiables, *J. Math. Pures Appl.* **76** (1997), 499–524.

[To90] J.-C. TOUGERON, *An introduction to the theory of Gevrey expansions and to the Borel-Laplace transform with some applications*, Cours à Rennes et Toronto (1989–90).

[Wa00] G. WALLET, La variété des équations surstables, *Bull. Soc. math. France* **128** (2000), 497–528.

J.-P. M.
Université Paris VI, UMR 7586,
Analyse algébrique,
4 place Jussieu,
75252 Paris cedex 05, France,
marco@math.jussieu.fr

D. S.
CNRS – IMCCE, UMR 8028,
Astronomie et systèmes dynamiques,
77 av. Denfert-Rochereau,
75014 Paris, France,
sauzin@imcce.fr