Nekoroshev estimates and instability for Gevrey class Hamiltonians

David Sauzin*

Research trimester on dynamical systems, Centro Ennio de Giorgi
Pisa, February 2002

Abstract

When a completely integrable Hamiltonian system h, which is written in action-angle coordinates, is perturbed, the action variables remain stable over exponentially long time intervals. The hypotheses are the quasi-convexity of h and the Gevrey-α regularity of h and of the perturbation, with $\alpha \geq 1$. This is a generalization of Nekhoroshev's Theorem (which corresponds to the analytic case, $\alpha = 1$). The stability time is governed by an exponent which can be chosen to be $1/2n\alpha$ in general, but which can be improved in $1/2(n-2)\alpha$ for the orbits passing close to a double resonance of h.

For $\alpha > 1$, the existence of Gevrey-α functions with compact support allows us to prove the optimality of the previous result: for three degrees of freedom or more, we construct systems exhibiting unstable orbits for which the speed of drift is optimal.

These results were obtained in a joint work with Jean-Pierre Marco which started as a collaboration with Michael Herman [MS03a].

1 Introduction

1.1 The present text gives an account of a joint work [MS03a] with Jean-Pierre Marco[1], which had begun as a collaboration with Michael R. Herman about one year before his sudden death in November 2000. The subject is the exponential stability for near-integrable Hamiltonian systems and the search for examples of perturbations giving rise to instability, *viz.* wandering points or even wandering domains, with the highest possible speed of drift. This kind of instability pertains to the topic of "Arnold diffusion"; we are not concerned here with the problem of the genericity of this phenomenon, but rather with the explicit construction of suitable perturbations.

One thing we must stress from the beginning is that we work in the Gevrey category rather than in the analytic one; our stability result is thus a generalization of the classical one available for analytic Hamiltonians (see Chapter 5 of [Gio03] in the same volume), while our construction of unstable systems relies on the existence of Gevrey functions with compact support—a fact which makes the designing of examples much simpler than in the analytic category.

We wish here to give all the details of the construction of the instability example in the simplest case, but we shall content ourselves with brief indications on the stability part.

1.2 So, we start with a completely integrable Hamiltonian system $h(r)$ written in action-angle coordinates (θ, r): $\theta \in \mathbb{T}^n$ with $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ and $n \geq 2$, the action variable r belongs to some domain in $\mathbb{R}^n$ and the equations of motions are $\dot{\theta} = \nabla h(r)$, $\dot{r} = 0$. All solutions are quasi-periodic and the phase space is foliated into invariant n-tori on which r is constant (see *e.g.* Section 2.5 of [Gio03]).

*CNRS – IMCCE, UMR 8028, *Astronomie et systèmes dynamiques*, 77 av. Denfert-Rochereau, 75014 Paris. E-mail: sauzin@imcce.fr

[1]Université Paris VI, UMR 7586, *Analyse algébrique*, 4 place Jussieu, 75252 Paris CEDEX 05. E-mail: marco@math.jussieu.fr

In the words of Poincaré, the "general problem of dynamics" is to investigate to what extent the features of this integrable situation persist under small perturbation, *i.e.* when we consider a Hamiltonian $H(\theta, r) = h(r) + f(\theta, r)$ close to h.

When h satisfies a certain property of non-degeneracy, KAM theory offers a first answer, which can be viewed as a result of perpetual stability for a large part of the phase space: a lot of invariant tori do persit under perturbation, they are only slightly deformed and still carry quasi-periodic motions of the perturbed system. In particular, the action variable corresponding to these solutions vary little for all times. If the number n of degrees of freedom is 2, each energy level (which is an invariant manifold of dimension 3) is divided by these tori into regions in which the action variable is confined: each initial condition gives rise either to a solution which lies on one of these tori or to a solution which is trapped between two of them. In none of these cases can r vary much.

For 3 degrees of freedom, the KAM tori do not separate the energy levels any longer, and Arnold [Arn64] proposed an example of perturbation in view of exploring their complement in the phase space. He showed the possibility of obtaining "unstable" orbits (or "drifting" orbits), *i.e.* solutions $(\theta(t), r(t))$ for which $r(t)$ undergoes a variation of at least, say, one unity, in a system which is however arbitrarily close to integrable. This is the phenomenon now called Arnold diffusion. The mechanism used by Arnold to produce instability relied on heteroclinic connections between whiskered tori. In that paper, he was not concerned about the time of drift of his unstable orbits; on the other hand, he raised the difficult question of the genericity (in the space of Hamiltonian perturbations) 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.

On the other hand, Nekhoroshev's Theorem of exponential stability [Nekh77] puts a limitation on the speed to which instability phenomena can develop. Indeed, it asserts that all solutions of the perturbed system have a nearly constant action over exponentially long time intervals. The assumptions for this are the analyticity of the Hamiltonian and a generic property called "steepness" for the integrable part h. Observe that this is a statement of "effective" stability, on finite (long) time intervals, concerning the *whole* phase space.

More precisely, there exist $a, b > 0$ such that, with the notations $H = h(r) + f(\theta, r)$ and $\varepsilon = \|f\|$, all solutions $(\theta(t), r(t))$ of the Hamiltonian vector field generated by H satisfy $\|r(t) - r(0)\| \leq \text{const}\, \varepsilon^b$ for $|t| \leq \exp(\text{const}(\frac{1}{\varepsilon})^a)$. Here the "stability exponents" a and b depend only on h. We shall henceforth restrict ourselves to integrable parts h which are convex, or *quasi-convex*, meaning that their levels are strictly convex hypersurfaces (quasi-convexity is a weaker requirement than the convexity of the function h itself; see (2) below); quasi-convexity is a particular case of steepness[2] and, according to [LN92, LNN93] or [Pö93], one can then take $a = b = \frac{1}{2n}$.

The problem of the optimality of the time exponent a is related to Arnold diffusion. Indeed, to know the largest possible value of a, we should find systems arbitrarily close to integrable which admit unstable orbits, experiencing a noticeable drift in action in a time comparable to the lower bound $\exp(\text{const}(\frac{1}{\varepsilon})^a)$ imposed by the stability result. This optimality question is still open for $n \geq 5$, whereas a partial answer is available for $n = 3$ or 4 [Be96, Be97], based on Arnold's model.

1.3 Michael Herman had the idea of enlarging the framework by considering Gevrey-α functions instead of real-analytic ones; these are smooth functions φ whose successive partial derivatives are bounded by $c(\varphi)M(\varphi)^{|k|}(|k|!)^\alpha$, where k is the multi-index of derivation and $\alpha \geq 1$ is a fixed parameter. One recovers the analytic case by choosing $\alpha = 1$, whereas for $\alpha > 1$ the space of Gevrey-α functions displays a property of non-quasianalyticity: it contains non-zero functions which vanish identically on intervals.

[2] We refer the interested reader to [Ni00] for a discussion of the values of the stability exponents in the general steep case.

M. Herman soon obtained a stability result analogous to Nekhoroshev's in this broader context (still in the quasi-convex case), and began to design a method to construct in the non-analytic case ($\alpha > 1$) examples of unstable systems arbitrarily close to $h(r) = \frac{1}{2}(r_1^2 + \dots r_{n-1}^2) + r_n$. He proposed to Jean-Pierre Marco and me to work together to improve both the stability and instability results, in view of making them match.

The first version of the Gevrey stability result was obtained by approximation techniques, relying on the known results for the analytic case, but led to stability exponents which were too small; we found that, by repeating completely Lochak's proof of the Nekhoroshev Theorem [Lo92], we could reach exponents $a = \frac{1}{2n\alpha}$ and $b = \frac{1}{2n}$ which agree with the existing values when $\alpha = 1$. We shall give an account of that work in Section 2.

As for the instability result, it is only after M. Herman's death that we arrived to an example displaying "instability exponents" as low as permitted by the stability theorem: we obtain optimality for any $n \geq 3$ and $\alpha > 1$ in the case of initial conditions close enough to a double resonance—see Section 3.

2 Gevrey stability

2.1 Gevrey functions

Let $n \geq 2$ and $R > 0$. We recall the notation $\mathbb{T} = \mathbb{R}/\mathbb{Z}$; we denote by B_R the open ball of $\mathbb{R}^n$ of radius R with center at the origin and by $\overline{B}_R$ its closure.

Let $\alpha \geq 1$ and $L > 0$. A real-valued C^∞ function $\varphi = \varphi(\theta, r)$ on $K = \mathbb{T}^n \times \overline{B}_R$ is said to be Gevrey-(α, L), and we write $\varphi \in G^{\alpha,L}(K)$, if

$$\|\varphi\|_{\alpha,L} := \sum_{k \in \mathbb{N}^{2n}} \frac{L^{|k|\alpha}}{k!^\alpha} \|\partial^k \varphi\|_{C^0(K)} < \infty, \tag{1}$$

with the standard notations $|k| = k_1 + \dots + k_{2n}$, $k! = k_1! \cdots k_{2n}!$, $\partial^k = \partial_{x_1}^{k_1} \cdots \partial_{x_{2n}}^{k_{2n}}$, and with $(x_1, \dots, x_{2n}) = (\theta_1, \dots, \theta_n, r_1, \dots, r_n)$. The space $G^\alpha(K)$ of Gevrey-α functions on K is defined as the union over all $L > 0$ of the spaces $G^{\alpha,L}(K)$.

Observe that $G^1(K)$ is exactly the space of real analytic functions on K: if $\varphi \in G^{1,L}(K)$, the number L indicates the size of a complex domain $K_L = \{ z \in (\mathbb{C}/\mathbb{Z})^n \times \mathbb{C}^n \mid \mathrm{dist}_\infty(z, K) < L \}$ where the analytic extension $\tilde{\varphi}$ of φ is holomorphic (the index ∞ refers to the norm of the supremum of the components), and $\|\tilde{\varphi}\|_{C^0(\overline{K}_L)} \leq \|\varphi\|_{1,L}$; conversely, if $\tilde{\varphi}$ is holomorphic in $\overline{K}_L$ and if $\varphi = \tilde{\varphi}_{|K}$ is real-valued, the Cauchy inequalities yield $\|\varphi\|_{1,\lambda} \leq e^{\lambda/L} \|\tilde{\varphi}\|_{C^0(\overline{K}_L)}$ for each $\lambda < L$. We can thus think of $\| \,.\, \|_{1,L}$ as of the usual analytic norms.

For $\alpha > 1$, it easy to exhibit non-analytic Gevrey-α functions. For instance, setting $\alpha = 1 + \frac{1}{p}$ and using the functions $\varphi_c \in G^\alpha([-1,1])$ defined for $c > 0$ by $\varphi_c(x) = \exp(-c/x^p)$ if $x > 0$ and $\varphi_c(x) = 0$ if $x \leq 0$, one can construct "bump functions" in $G^{\alpha,L}$, which vanish identically outside a given interval I and whose value is 1 at each point of a given subinterval of I.

Apart from the theory of partial differential equations where they have been widely used, Gevrey functions were also already considered in connection with problems of dynamical systems. We did not find in the literature the definition of the norms $\| \,.\, \|_{\alpha,L}$, but we introduced them because they provide us Banach algebras. The stability under multiplication and composition of the space of Gevrey functions is a classical topic, but with our definition we obtain such a simple inequality as

$$\|\varphi\psi\|_{\alpha,L} \leq \|\varphi\|_{\alpha,L} \|\psi\|_{\alpha,L}$$

for the multiplication (thanks to the Leibniz formula), and an inequality of the form $\|\varphi \circ u\|_{\alpha,L_1} \leq \|\varphi\|_{\alpha,L}$ under suitable hypotheses on the vector-valued function u whose components are assumed to be Gevrey-(α, L_1) (see [MS03a, Appendix]).

As for the classical Cauchy inequaties, of constant use when dealing with analytic functions, they admit an immediate generalisation: if φ is Gevrey-(α, L) and if $0 < \lambda < L$,

$$\sum_{|m|=j} \|\partial^m \varphi\|_{\alpha, L-\lambda} \leq j!^\alpha \lambda^{-j\alpha} \|\varphi\|_{\alpha,L}, \qquad j \geq 0.$$

2.2 The stability theorem

By adapting Lochak's periodic orbit method [Lo92] to the Gevrey case, we prove in [MS03a] the following stability theorem for any $\alpha \geq 1$ and $L > 0$:

Theorem 1 *Let $n \geq 2$ and*

$$a = \frac{1}{2n\alpha}, \quad b = \frac{1}{2n}.$$

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

$$\varpi \leq \max_{1 \leq i \leq n} |\partial_{r_i} h(r)| \quad \text{and} \quad \forall v \in (\nabla h(r))^\perp, \quad m\|v\|^2 \leq \langle \nabla^2 h(r) v, v \rangle. \tag{2}$$

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

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

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

Moreover, like in the analytic case ([Lo92],[Lo93]), there are improved estimates for nearly-resonant initial conditions. An action r_0 is said to be non-resonant if $\nabla h(r_0)$ (which is the frequency-vector for the Kronecker flow $\dot{\theta} = \nabla h(r_0)$ induced by the unperturbed system on the invariant torus $\{r = r_0\}$) satisfies

$$\langle k, \nabla h(r_0) \rangle = 0 \Rightarrow k = 0, \qquad k \in \mathbb{Z}^n$$

(the torus $\{r = r_0\}$ is then densely filled by any solution of the corresponding Kronecker flow). For a resonant action r_0, the set $\{\, k \in \mathbb{Z}^n \mid \langle k, \nabla h(r_0) \rangle = 0 \,\}$ is a submodule of $\mathbb{Z}^n$, whose rank m is called the *multiplicity* of the resonance (the torus $\{r = r_0\}$ is then foliated in $(n-m)$-tori which are invariant for the corresponding Kronecker flow). Conversely, if $\mathcal{M}$ is a submodule of rank $m \in \{1, \ldots, n-1\}$ of $\mathbb{Z}^n$, we define the corresponding resonant surface in action space as

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

If $\operatorname{dist}(r_0, S_{\mathcal{M}}) \leq \sigma \varepsilon^{1/2}$ for some positive number σ (i.e. if the action r_0 is nearly $\mathcal{M}$-resonant), the stability exponents which we obtain for the solutions starting in $\mathbb{T}^n \times \{r_0\}$ jump to

$$a = \frac{1}{2(n-m)\alpha} \quad \text{and} \quad b = \frac{1}{2(n-m)}. \tag{3}$$

Of course, the numbers $\varepsilon_*, c, C_1, C_2$ in the corresponding refined statement also depend on $\mathcal{M}$ and σ (possible values are explicitly computed in [MS03a]).

This means that there are improved stability estimates for solutions passing close to a resonance—the higher the multiplicity of the resonance, the longer the stability time. This phenomenon might look surprising, as resonances are often viewed as causes of instability, but it is a very direct by-product of Lochak's periodic orbit method (see [Lo93] for a non-technical account in the analytic case). In fact, the case of maximal resonance, *i.e.* $m = n - 1$ (the tori associated with such completely resonant actions are foliated into periodic orbits of the unperturbed system) is the cornerstone of the method: in a sense, the method begins by obtaining the exponents $a = 1/2\alpha$ and $b = 1/2$ near such a periodic torus.

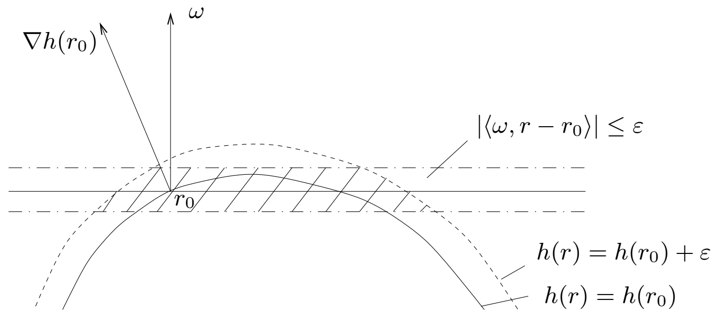

Figure 1: The inequalities $|h(r) - h(r_0)| \leq \varepsilon$ and $|\langle \omega, r - r_0 \rangle| \leq \varepsilon$ force confinement, due to the convexity of Σ_{r_0}.

2.3 The periodic orbit method

We now sketch the proof of Theorem 1 and its refinement for nearly resonant initial conditions. We shall keep the discussion at the level of heuristics and avoid the numerous technical details which make the proof contained in [MS03a, Section 3] quite long. We thus fix R, R_0, h and $H = h + f$ as in the hypotheses of Theorem 1 and put $\varepsilon = \|f\|_{\alpha,L}$. We shall use the auxiliary notations $R^* = \frac{R+R_0}{2}$, and $\Omega = L^{-\alpha}E$, $M = 2^\alpha L^{-2\alpha}E$, so that

$$\sum_{|k|=1} \|\partial^k h\|_{C^0(\overline{B}_R)} \leq \Omega, \quad \sum_{|k|=2} \|\partial_k h\|_{C^0(\overline{B}_R)} \leq M. \tag{4}$$

2.3.1 Confinement by convexity

Any solution $(\theta(t), r(t))$ of the perturbed system has constant energy, thus $|h(r(t)) - h(r(0))| \leq \varepsilon$ for all t. In other words, $r(t)$ remains always close to the hypersurface $\Sigma_{r_0} = h^{-1}(h(r_0))$ determined by the initial action r_0. But this hypersurface is strictly convex, its intersection with any hyperplane in $\mathbb{R}^n$ is thus bounded.

Suppose now $\omega \in \mathbb{R}^n$ satisfies $\langle \omega, \partial_\theta f(\theta, r) \rangle = 0$ for all (θ, r). In that case, the function $\langle \omega, r \rangle$ would be a first integral of the motion, since $\frac{d}{dt}\langle \omega, r(t) \rangle = -\langle \omega, \partial_\theta f \rangle$, and the action $r(t)$ would be trapped in the bounded region $\{|h(r) - h(r_0)| \leq \varepsilon\} \cap \{\langle \omega, r \rangle = \langle \omega, r_0 \rangle\}$ (in fact in the connected component of r_0 in this intersection).

The same argument is still valid if $\langle \omega, r \rangle$ is only an approximate first integral, but it yields then confinement only for finite time intervals. If indeed we assume that $\omega \in \mathbb{R}^n$ satisfies $|\langle \omega, \partial_\theta f(\theta, r) \rangle| \leq \varepsilon/\tau$, we obtain $|\langle \omega, r(t) - r_0 \rangle| \leq \varepsilon$ as long as $|t| \leq \tau$. The action is still trapped in a bounded region of small diameter (see Figure 1), for a long time if τ is large compared to ε. A detailed computation would show that the diameter of confinement region is essentially $\varepsilon^{1/2}$.

The idea will be to use $\omega = \nabla h(r^*)$ with r^* well chosen, close to r_0, and to find a canonical change of coordinates Ψ such that $H \circ \Psi = h + g$ with $\langle \omega, \partial_\theta g \rangle$ exponentially small. The change of coordinates will be close to identity and defined only in a neighbourhood of r^*, *i.e.* in a domain of the form $\mathbb{T}^n \times \overline{B}(r^*, \rho^*)$ with a small $\rho^* > 0$. In the slightly distorted coordinates provided by Ψ, we shall be in a position to apply the above arguments, which lead precisely to

Lemma 2.1 *Let $\lambda = 8M/m$ and $r^* \in \overline{B}_{R^*}$. Suppose $\tau, \rho > 0$ and $g \in C^2(\mathbb{T}^n \times \overline{B}(r^*, \lambda\rho))$ satisfy*

$$|g(\theta, r)| \leq m\rho^2, \quad |\langle \nabla h(r^*), \partial_\theta g(\theta, r) \rangle| \leq \frac{m\rho^2}{\tau}$$

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

We do not give here any more details, for which the reader is once for all referred to [MS03a]. We have indicated here the precise values of several constants just to emphasize the role of quasi-convexity. However, most of the time in the sequel, we shall content ourselves with writing "a constant", or "const", for the various numbers which are involved in the other technical statements and which depend only on the parameters n, α, R, ϖ, m, etc.

2.3.2 Resonant normal form

We say that the Hamiltonian system $H = h + f$ has been put in "normal form" when a local canonical change of coordinates has been performed in such a way that the new Hamiltonian function $H = h + g$ (where $g = f + h\circ\Psi - h$) enjoys some interesting property. In our case, in view of the previous paragraph, the required property is the smallness of $\langle\omega, \partial_\theta g\rangle$, with a certain $\omega\in\mathbb{R}^n$.

Observe that $\langle\omega,\partial_\theta g\rangle = 0$ means that, for each value r of the action, g is constant along the solutions of the Kronecker flow $\dot\theta=\omega$. Therefore, if ω is non-resonant, by continuity this amounts to g being independent of the angles θ_i (and in that case, each component r_i is a first integral of the Hamiltonian system generated by $h+g$). If on the contrary ω is completely resonant, for instance if $\omega = \frac{1}{T}(1,0,\dots,0)$, this condition is much less demanding and we obtain only $\langle\omega, r\rangle$ as first integral (in the given example of completely resonant vector, we are only requiring that g do not depend on the first angle). In such a situation, the Hamiltonian $h+g$ is said to be "resonant" with respect to ω.

Suppose r^* is a given completely resonant action, and $\omega = \nabla h(r^*)$. There exists $T>0$ such that the components of $T\omega$ are integers without any non-trivial common divisor: this number is the common period of the solutions of $\dot\theta = \omega$. We define the *resonant part* of a continuous function φ as

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

Obviously, a function is resonant with respect to ω iff it coincides with its resonant part.

We shall be able to obtain the following *(approximately) resonant normal form*, where μ stands for a small parameter which can be taken proportional to $\varepsilon^{1/2}$ but which we prefer to leave free for the moment (because the true small parameter will rather be μT):

Proposition 2.1 *If $H = h + f$ satisfies*

$$\|f\|_{\alpha,L} \le \mathrm{const}\,\mu^2, \qquad \mu T\le\nu,$$

where ν is a constant, and if $0 < 2\mu\tilde R \le R - R^*$, there exists a μT-close to identity symplectic transformation Ψ defined on $\mathbb{T}^n\times\overline{B}(r^*,\mu\tilde R)$, such that the last n components of $\Psi - \mathrm{Id}$ satisfy

$$\sum_{i=1}^n \|\Psi^{(n+i)} - r_i\|_{C^0} \le \frac{\mu T}{\nu}\mu\tilde R,$$

and such that the new Hamiltonian reads $H\circ\Psi = h + g$ with $\|g\|_{C^0}\le \mathrm{const}\,\mu^2$ and

$$\sum_{i=1}^n \|\partial_{\theta_i}(g - \mathcal{R}_\omega g)\|_{C^0} \le \mathrm{const}\,\mu^2\,\exp\big(-\alpha\Big(\frac{\nu}{\mu T}\Big)^{1/\alpha}\big), \qquad \omega = \nabla h(r^*).$$

We shall comment on the proof of this proposition later (Section 2.4), which is in fact a *one-phase averaging* result.

Observe that Proposition 2.1 and Lemma 2.1 immediately yield the desired confinement property for initial conditions close to the completely resonant torus $\mathbb{T}^n\times\{r = r^*\}$, and therefore the conclusion of Nekhoroshev's Theorem as described in the second part of Section 2.2 when $m = n - 1$.

If indeed we fix a submodule $\mathcal{M}$ of rank $n-1$ in $\mathbb{Z}^n$, we can find $k^* \in \mathbb{Z}^n$ such that $S_{\mathcal{M}} = \{ r \in \overline{B}_{R_0} \mid \exists T > 0, \nabla h(r) = \pm \frac{1}{T} k^* \}$. Thus, if we are given (θ_0, r_0) with $\mathrm{dist}(r_0, S_{\mathcal{M}}) \le \sigma \varepsilon^{1/2}$, we can find a completely resonant action r^* such that $\|r_0 - r^*\| \le \mathrm{const}\, \varepsilon^{1/2}$ and $\nabla h(r^*) = \pm \frac{1}{T} k^*$. In view of (4), the period T satisfies $\frac{1}{\Omega}\|k^*\|_\infty \le T \le \frac{1}{\varpi}\|k^*\|_\infty$. Hence, applying Proposition 2.1 with $\mu = \varepsilon^{1/2}$ (assuming ε small enough), we obtain a resonant normal form valid in a small ball containing r_0, and Lemma 2.1 in the coordinates $(\tilde\theta, \tilde r) = \Psi^{-1}(\theta, r)$ yields confinement at a distance $\lambda\rho/2 = \mathrm{const}\, \varepsilon^{1/2}$ over an exponentially long time $\tau = \exp(\mathrm{const} \big(\frac{1}{\varepsilon}\big)^{1/2\alpha})$.

2.3.3 Use of Dirichlet's principle

For the other initial conditions $(\theta_0, r_0) \in \mathbb{T}^n \times \overline{B}_{R_0}$, we just need to find a completely resonant action r^* close enough to r_0 with a control of the size of its period T. This is easily done with the help of Dirichlet's principle, according to which, given $\omega' \in \mathbb{R}^{n-1}$ and $Q > 1$, there exist $\ell' \in \mathbb{Z}^{n-1}$ and $q \ge 1$ such that $\|q\omega' - \ell'\|_\infty \le Q^{-\frac{1}{n-1}}$ and $q < Q$. Here we start with $\omega_0 = \nabla h(r_0)$, which we can write $\xi(\pm 1, \omega_0')$ with $\xi = \|\omega_0\|_\infty$ (after renumbering the components if necessary), and we obtain a completely resonant vector $\omega = \xi(\pm 1, \frac{\ell_0'}{q})$ whose period T lies between $1/\xi$ and q/ξ; we retain

$$T\|\nabla h(r_0) - \omega\|_\infty \le \|\nabla h(r_0)\|_\infty Q^{-\frac{1}{n-1}}, \quad 1 \le T\|\nabla h(r_0)\|_\infty \le Q.$$

To pass from ω to a completely resonant action r^*, we use the following easy consequence of quasi-convexity:

Lemma 2.2 *For all $r_0 \in \overline{B}_{R_0}$ and ω close enough to $\nabla h(r_0)$, there exist $r^* \in \overline{B}_{R^*}$ and $\xi \in [\frac{1}{2}, 2]$ such that*

$$h(r^*) = h(r_0), \quad \nabla h(r^*) = \xi\omega, \quad \|r^* - r_0\| \le \mathrm{const}\, \|\omega - \omega_0\|.$$

(The Implicit Function Theorem yields the local invertibility of $(\xi, r) \mapsto (h(r), \xi\nabla h(r))$; this last property is called "isoenergetic non-degeneracy".)

Choosing $Q = \big(\frac{1}{\varepsilon}\big)^{\frac{n-1}{2n}}$, we end up with

$$\|r_0 - r^*\| \le \mathrm{const}\, \mu, \qquad \varepsilon^{1/2} \le \mu = \frac{1}{T}\varepsilon^{1/2n} \le \varepsilon^{1/2n}.$$

The conclusion is reached like previously, by applying Proposition 2.1 (with the above value of μ and $\tilde R = \mathrm{const}$) and then Lemma 2.1 (with $\rho = \mathrm{const}\, \mu$ and $\tau = \mathrm{const} \exp(\alpha \big(\frac{\nu}{\mu T}\big)^{1/\alpha})$).

The intermediary cases of nearly m-resonant initial actions are treated similarly, with an improvement stemming from the fact that Dirichlet's principle can be applied essentially in $\mathbb{R}^{n-m-1}$ instead of $\mathbb{R}^{n-1}$ (see [MS03a, Section 3.4.3]).

2.4 Gevrey averaging

We now explain the proof of Proposition 2.1, which is the analytic core of the periodic orbit method. This part corresponds to Section 3 of [MS03a], but we shall not reproduce here the lengthy calculations which are needed; we shall try to emphasize rather the strategy which we followed. The arguments of Section 2.3 were adapted from [Lo92, LN92, LNN93], and the case $\alpha = 1$ of the normal form indicated in Proposition 2.1 (or a variant of it) can be found in these references, however we depart here from them to adopt a different path (already suggested by Lochak [Lo92]).

Indeed, faced with the necessity of repeating the derivation of the normal form to adapt it to the Gevrey case, we seized the opportunity to incorporate new ingredients which make an extensive use of *Gevrey formal series*, which are generically divergent but with a control on the rate of divergence. A formal series $\tilde{a} = \sum_{j\in\mathbb{N}} \varepsilon^j a_j$ with coefficients in a normed vector space V is said to be Gevrey-$(\alpha+1)$ if there exist $C, M > 0$ such that

$$\|a_j\|_V \le CM^j j!^\alpha, \qquad j \in \mathbb{N}.$$

Observe that α should be zero to guarantee that such a formal series has positive radius of convergence. On the contrary, we assume $\alpha \ge 1$ and the key point will be the arising of Gevrey-$(\alpha+1)$ formal *series* in a situation whose data are Gevrey-α *functions*. This jump of Gevrey index lies at the heart of our work on averaging.

We thus suppose that we are given a Hamiltonian $H = h(r)+f(\theta, r)$ and a T-periodic resonant action r^*. Let $\omega = \nabla h(r^*)$; our aim is to eliminate the non-resonant part (with respect to ω) of the Hamiltonian up to an exponentially small remainder. In the aforementioned articles, this is achieved by a large number of successive canonical changes of coordinates, iterating an averaging procedure, and the number of iterations is of the order of a negative power of the small parameter. As noticed in [Lo92], this can be related to the Gevrey divergence of the formal series of classical perturbation theory. There is indeed a formal symplectic transformation which is a formal power series of the small parameter and which puts the Hamiltonian in a formal normal form without any non-resonant remainder. Of course this formal series is expected to be divergent, but Gevrey estimates on the growth of its coefficients are sufficient to construct, *e.g.* by summation to the least term, an approximate resonant normal form like the one we seek.[3]

Here we can take $\mu = \mathrm{const}(\|f\|_{\alpha,L})^{1/2}$ as small parameter, but it will be convenient to perform some scalings. Using $\tilde{r} = \frac{r-r^*}{\mu}$ instead of r and setting $\tilde{\theta} = \theta$, we define a conformal-symplectic change of variables, giving rise to the new Hamiltonian

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

where the first term is an additive constant of no importance, while $\tilde{f}$ is easily computed from the Taylor formula applied to h and from f. Supposing $\mu \le 1$, one can check that $\tilde{f} \in G^{\alpha,\tilde{L}}(\mathbb{T}^n \times \overline{B}_{5\tilde{R}})$ with

$$\tilde{\varepsilon} = \|\tilde{f}\|_{\alpha,\tilde{L}} \le \mathrm{const}\,\mu.$$

We take $\tilde{\varepsilon}$ as new small parameter. Dropping the tildes, we see that the desired conclusion will follow from the study of a Gevrey perturbation of the linear Hamiltonian $\langle\omega, r\rangle$.

Proposition 2.2 *Let $\omega \in \mathbb{R}^n$ be a completely resonant vector of period T and consider a Hamiltonian of the form*

$$H(\theta, r) = \langle \omega, r\rangle + \varepsilon F(\theta, r),$$

where $F \in G^{\alpha,L}(\mathbb{T}^n \times \overline{B}_R)$ has norm 1. Then, if εT is smaller than a certain constant ν and $0 < \delta < R$, there exist a function $\mathcal{A}$ in $G^{\alpha,L/2}(\mathbb{T}^n \times \overline{B}_{R-\delta})$ and an εT-close to identity symplectic transformation Ψ whose components belong to $G^{\alpha,L/2}(\mathbb{T}^n \times \overline{B}_{R-\delta})$ and whose image is contained in $\mathbb{T}^n \times \overline{B}_R$, such that

$$H \circ \Psi = \langle\omega, r\rangle + \mathcal{A}$$

and $\mathcal{A}$ is nearly resonant: $\|\mathcal{A} - \mathcal{R}_\omega \mathcal{A}\|_{\alpha,L/2} \le \mathrm{const}\,\varepsilon\, \exp(-\alpha\big(\frac{\nu}{\varepsilon T}\big)^{1/\alpha})$.

When returning to the original variables, this proposition easily implies Proposition 2.1. Its proof is obtained from the study of formal series

$$\tilde{\Psi} = \sum_{j\ge 0} \varepsilon^j \Psi_j(\theta, r), \quad \tilde{A} = \sum_{j\ge0} \varepsilon^j A_j(\theta, r),$$

[3] Such estimates were already obtained in [Sa92] for a simpler problem. A more systematic approach is indicated in [RS96], whose formalism we have developed and adapted to our Hamiltonian problem of averaging with Gevrey data.

whose coefficients are Gevrey-(α, L) functions, and such that $\tilde{\Psi}$ is a formal symplectic transformation conjugating H with $\langle\omega, r\rangle + \varepsilon\tilde{A}(\theta, r)$ and $\tilde{A}$ is resonant with respect to ω (*i.e.* $\tilde{A}_j = \mathcal{R}_\omega \tilde{A}_j$ for all j).

The existence of $\tilde{A}$ and $\tilde{\Psi}$ is checked by plugging formal expansions in the conjugation equation. It is convenient to look for $\tilde{\Psi}$ as a Lie series, *i.e.* as the time-ε flow of some auxiliary (formal) Hamiltonian $\tilde{B}$—see Section 4 of [Gio03]. One obtains linear equations

$$A_j + \langle\omega, \partial_\theta B_j\rangle = \varphi_j,$$

with $\varphi_0 = F$ and, for $j \geq 1$, φ_j determined by $F, A_0, B_0, \ldots, A_{j-1}, B_{j-1}$. These equations are solved inductively by $A_j = \mathcal{R}_\omega\varphi_j$ and $B_j = \mathcal{N}_\omega\varphi_j$, where the operator $\mathcal{N}_\omega$ takes the form

$$\mathcal{N}_\omega\varphi(\theta, r) = \int_0^{T\theta_1} (\varphi - \mathcal{R}_\omega\varphi)(\theta - t\omega, r)\, dt$$

in the case where $\omega = \frac{1}{T}(1, 0, \ldots, 0)$ (there is an analogous formula in the general case).

The point is that $\mathcal{R}_\omega$ and $\mathcal{N}_\omega$ are bounded operators:

$$\|\partial^k \mathcal{R}_\omega\varphi\|_{C^0(K)} \leq \|\partial^k\varphi\|_{C^0(K)}, \quad \|\partial^k \mathcal{N}_\omega\varphi\|_{C^0(K)} \leq \frac{T}{2}\|\partial^k\varphi\|_{C^0(K)}.$$

But the formula defining φ_j in terms of the previous coefficients involve the unbounded operator $\varphi \mapsto L_{\mathcal{N}\varphi}F$, where L stands for the Lie derivative: it turns out that, to compute the jth coefficient of $\tilde{A}$ or $\tilde{B}$, we need to differentiate j times the given function F and these repeated differentiations, not being compensated by the division by the appropriate factor, tend to create Gevrey divergence. What we can prove is

$$\|A_j\|_{\alpha,L/2}, \|B_j\|_{\alpha,L/2} \leq \mathrm{const}^{j+1}\, T^j j!^\alpha, \qquad j \geq 0.$$

This is a technical work, in which the techniques of *Gevrey majorant series* introduced in [RS96] prove to be efficient.

The above inequalities are sufficient to deduce Proposition 2.2, because any Gevrey-$(\alpha + 1)$ series $\tilde{a}$ can be realized as the asymptotic expansion of some function $a(\varepsilon)$ which satisfies, for small $\varepsilon > 0$, inequalities of the form

$$\|a(\varepsilon) - \sum_{j=0}^{J-1} \varepsilon^j a_j\|_V \leq CM^J J!^\alpha \varepsilon^J, \qquad J \geq 1. \tag{5}$$

Such a function $a(\varepsilon)$ can be viewed as an approximate sum of the possibly divergent series $\tilde{a}(\varepsilon)$ (one can define it *e.g.* by using a truncated Borel-Laplace integral) and is not unique, but any two functions a and a^* satisfying (5) must be exponentially close one to the other:

$$\|a(\varepsilon) - a^*(\varepsilon)\|_V \leq \exp\big(-\mathrm{const}\big(\frac{1}{\mu}\big)^{1/\alpha}\big).$$

Applying this to $\tilde{A}$ and $\tilde{B}$, with $V = G^{\alpha,L/2}$, one obtains V-valued functions of ε, the first of which is a resonant funtion $\mathcal{A}^*$, the second of which generates a close to identity symplectic transformation Ψ, and one can check that $H \circ \Psi$ is exponentially close to $\langle\omega, r\rangle + \mathcal{A}^*$.

3 Instability Theorem

3.1 The statement for $n \geq 3$

Theorem 1 is valid for the perturbations of any quasi-convex integrable part $h(r)$. A class of examples of quasi-convex functions is provided by the suspension of $(n-1)$-degree-of-freedom convex Hamiltonians. Indeed, if $\hat{h}(r_1,\dots,r_{n-1})$ is strictly convex, one can check that $h(r) = \hat{h}(r_1,\dots,r_{n-1}) + r_n$ is quasi-convex, and this is useful when dealing with non-autonomous time-periodic perturbations of $\hat{h}$: with the notations $\hat{\theta}$ for the first $n-1$ angles and $\hat{r}$ for the first $n-1$ components of the action, the non-autonomous vector field generated by $\hat{h}(\hat{r}) + f(\hat{\theta},\hat{r},t)$ is equivalent to the autonomous vector field generated by the n-degree-of-freedom Hamiltonian $H(\theta,r) = \hat{h}(\hat{r}) + r_n + f(\hat{\theta},\hat{r},\theta_n)$ (all energy levels of H are isomorphic and, on each one, θ_n is essentially the time, whereas the variable r_n can be forgotten without losing information).

From now on we suppose $\alpha > 1$ and $n \geq 3$. The next theorem shows that the time exponents given in Section 2.2 are optimal in the case of at least double resonances ($m \geq 2$) for the most elementary example of quasi-convex integrable part, $h(r) = \frac{1}{2}(r_1^2 + \cdots + r_{n-1}^2) + r_n$.

Theorem 2 *Let $a^* = \frac{1}{2(n-2)\alpha}$, $L > 0$ and $R > 1$. There exist a sequence of functions $(f_j)_{j\geq 0}$ converging to 0 in the space $G^{\alpha,L}(\mathbb{T}^n \times \overline{B}_R)$ and an increasing sequence of integers $(\tau_j)_{j\geq 0}$ such that, for each $j \geq 0$, the Hamiltonian system generated by*

$$\mathcal{H}_j(\theta,r) = \frac{1}{2}(r_1^2 + \cdots + r_{n-1}^2) + r_n + f_j(\theta,r) \tag{6}$$

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

$$\frac{C_1}{\varepsilon_j^2}\exp\Big(C_1\Big(\frac{1}{\varepsilon_j}\Big)^{a^*}\Big) \leq \tau_j \leq \frac{C_2}{\varepsilon_j^2}\exp\Big(C_2\Big(\frac{1}{\varepsilon_j}\Big)^{a^*}\Big), \qquad j \geq J.$$

We shall see that, provided j is large enough, the second action-component $r_2(t)$ of our unstable solution satisfies $0 < r_2(t) \leq 3\sqrt{\varepsilon_j}$. In view of the frequency map $\nabla h(r) = (r_1,\dots,r_{n-1},1)$, we can thus say that this solution starts $O(\sqrt{\varepsilon_j})$-close to the doubly-resonant surface $\{r_1 = r_2 = 0\}$. It is immediate to deduce from that, by a slight modification of $\mathcal{H}_j$, examples of nearly m-resonant unstable initial conditions for any $m \geq 2$ with instability exponent $a^* = \frac{1}{2(n-m)\alpha}$.

In other words, the time exponent a of (3) is optimal for $m \geq 2$, $\alpha > 1$ and $n \geq 3$.

The functions f_j will in fact be independent of the last variable r_n, which means that we can consider that we are dealing with non-autonomous Hamiltonian functions in $n-1$ degrees of freedom depending periodically on time.

Our unstable solution $(\theta(t),r(t))$ can be described for all $t \in \mathbb{R}$ (the Hamiltonian vector field is in fact defined in the whole of $\mathbb{T}^n \times \mathbb{R}^n$ and is complete); we shall see for instance that $r_1(k\sqrt{\tau_j}) = k/\sqrt{\tau_j}$ for all $k \in \mathbb{Z}$, hence this solution is *biasymptotic to infinity.*

Since Theorem 2 provides examples of wandering points, it is natural to ask whether one could obtain similarly examples of wandering balls in near-integrable systems. The answer is yes, as shown recently [MS03b] by a suitable modification of the proof:[4] one can construct an arbitrarily small Gevrey perturbation of $h(r) = \frac{1}{2}(r_1^2 + \cdots + r_{n-1}^2) + r_n$ which admits a wandering ball (the smaller the perturbation, the smaller the diameter of this wandering domain); we obtain a small open set which is transported by the Hamiltonian flow from $r_1 = -\infty$ to $r_1 = +\infty$, but the time needed to go from $r_1 = 0$ to $r_1 = 1$ is somewhat longer as in the case of the above unstable point (governed by an instability exponent larger than the above a^*).

3.2 Discrete analogue of the result

In the sequel, we shall explain the proof of Theorem 2 in the case where $n = 3$; see [MS03a] for the details and the modifications which are needed when $n > 3$.

[4]This recent work is partly based on ideas by M. Herman which were already sufficient in the C^k category when $k \leq n-2$.

It will be more convenient to work with mappings rather than with flows. 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.

Let $h(r_1, r_2) = \frac{1}{2}(r_1^2 + r_2^2)$. The maps that we shall consider will be of the form $\Phi^u \circ \Phi^{h+v}$: compositions of two time-1 maps of Hamiltonian functions on $\mathbb{A}^2 = T^*\mathbb{T}^2 = \mathbb{T}^2 \times \mathbb{R}^2$, with small Gevrey-$\alpha$ functions for u and v. Thus they will be close to the integrable twist map $\Phi^h(\theta, r) = (\theta + r, r)$. Moreover, our functions u will depend on θ only, thus $\Phi^u(\theta, r) = (\theta, r - \nabla u(\theta))$.

These maps are still exact-symplectic and will appear, by a simple suspension procedure, as time-1 maps of non-autonomous Hamiltonian systems, or as sections of autonomous Hamiltonian systems of $\mathbb{A}^3 = T^*\mathbb{T}^3$ of the desired form.

Here is a more precise version of our instability result in the discrete setting:

Proposition 3.1 *Let $\alpha > 1$, $L > 0$ and*

$$U(\theta_1) = -\frac{1}{2\pi}\sin 2\pi\theta_1, \quad V(\theta_2) = -\cos 2\pi\theta_2. \tag{7}$$

There exist a sequence of functions $(g^{(j)})_{j\geq 1}$ in $G^{\alpha,L}(\mathbb{T}^2)$, a sequence of real numbers $(r_2^{(j)})_{j\geq 1}$ and a positive constant c (depending only on α and L) such that, if for each $j \geq 1$ we define

$$N_j = 2j + 1, \qquad M_j = 2[cN_j\, e^{cN_j^{1/\alpha}} + 1], \qquad q_j = N_j M_j, \tag{8}$$

we have

$$\tfrac{1}{q_j}\|g^{(j)}\|_{\alpha,L} \leq \frac{1}{N_j^2} \tag{9}$$

and the transformation

$$\Psi_j = \Phi^{\frac{1}{q_j}U\otimes g^{(j)}} \circ \Phi^{h(r_1,r_2)+\frac{1}{N_j^2}V(\theta_2)} \tag{10}$$

produces a drift from $r_1 = 0$ to $r_1 = 1$ in q_j^2 iterations for the point $x^{(j)} = (0, 0, 0, r_2^{(j)})$.

The notation $\otimes$ simply corresponds to the product of functions depending on different variables: $U \otimes g^{(j)}(\theta) = U(\theta_1)g^{(j)}(\theta_2)$.

Observe that $\|\frac{1}{q_j}U \otimes g^{(j)}\|_{\alpha,L} \leq \|\frac{1}{N_j^2}V\|_{\alpha,L} = \frac{\text{const}}{N_j^2}$. Theorem 2 follows from this proposition, because one can write $\Psi_j = \Phi^{h(r)+f_j(\theta,r,t)}$ with a t-periodic Gevrey function f_j whose norm ε_j, lying between $\frac{1}{N_j^2}$ and $\frac{\text{const}}{N_j^2}$, is easily compared with $\tau_j = q_j^2$ (see [MS03a], Sec. 2.4). We shall now indicate the proof of Proposition 3.1.

3.3 Unstable orbit, coupling lemma, exponential estimates

3.3.1 Consider, for a positive integer q, the map defined on $\mathbb{A} = \mathbb{T} \times \mathbb{R}$ by $\psi(\theta_1, r_1) = (\theta_1 + qr_1, r_1 + \frac{1}{q}\cos 2\pi(\theta_1 + qr_1)$. The orbit of the origin under ψ is unstable, drifting between $r_1 = 0$ and $r_1 = 1$ in q iterates, since $\psi^k(0,0) = (0, \frac{k}{q})$ for all $k \in \mathbb{Z}$. This follows from the identity

$$\psi = \Phi^{\frac{1}{q}U} \circ F^q, \quad \text{where } \ F = \phi^{\frac{1}{2}r_1^2},$$

because $U'(0) = -1$.

The map ψ is not close to integrable when q is large, but we shall construct a near-integrable map Ψ such that ψ is a subsystem of Ψ^q: this will yield an orbit of Ψ which undergoes a drift of length 1 in q^2 iterates.

3.3.2 To that end we shall use the following lemma with the F defined above and with the pendulum map $G = \Phi^{\frac{1}{2}r_2^2+\frac{1}{N_j^2}V(\theta_2)}$ (although the lemma can be formulated in a greater generality; for instance in [MS03a] we use the map $G = \Phi^{\frac{1}{2}r_2^2+\frac{1}{N_j^2}V(\theta_2)} \times \Phi^{\frac{1}{2}(r_3^2+\cdots+r_{n-1}^2)}$ of A^{n-2} when the number n of degrees of freedom is larger than 3).

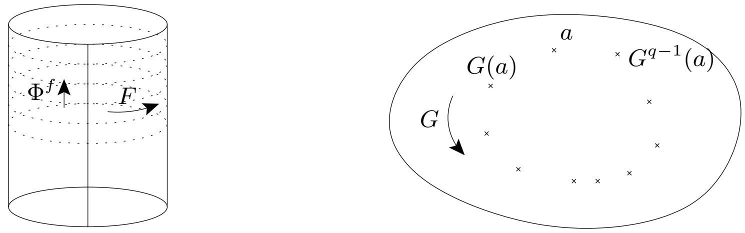

Figure 2: Starting with $(x_1, a) \in \mathbb{A}\times\mathbb{A}$, the $q-1$ first iterations of Ψ coincide with those of $F\times G$, the q^{th} iteration with that of $(\Phi^f \circ F)\times G$ thanks to the condition (11).

Lemma 3.1 *Suppose that $a \in \mathbb{A}$ is G-periodic of period q, and that $g : \mathbb{A} \to \mathbb{R}$ satisfies*

$$g(a) = 1, \qquad dg(a) = 0, \qquad g(G^k(a)) = 0, \quad dg(G^k(a)) = 0, \qquad 1 \le k \le q-1. \tag{11}$$

The mapping

$$\Psi = \Phi^{\frac{1}{q}U\otimes g} \circ (F\times G) \ : \ \mathbb{A}^2 \ \to \ \mathbb{A}^2$$

satisfies

$$\forall x_1 \in \mathbb{A}, \quad \Psi^q(x_1, a) = (\psi(x_1), a),$$

where $\psi = \Phi^{\frac{1}{q}U} \circ F^q$.

Proof. Use the identity $\Phi^{f\otimes g}(x_1,x_2) = (\Phi^{g(x_2)f}(x_1), \Phi^{f(x_1)g}(x_2))$ with $f = \frac{1}{q}U$ and see Figure 2. Notice that the precise form of f, F or G does not matter, but only the "synchronization" condition (11). □

3.3.3 We have already defined $N_j = 2j+1$ and thus G. There remains only to indicate how we choose a and g (depending on j) and why we use $q = q_j$ defined by (8); Lemma 3.1 will then produce an unstable orbit $((0,0),a)$, drifting in q^2 iterates, for the corresponding map Ψ.

Let M be a positive integer and put $q = N_j M$. The q-periodic points of G correspond to the M-periodic points of the time-1 map Φ^P of the standard pendulum, $P(\theta_*, r_*) = \frac{1}{2}r_*^2 + V(\theta_*)$, by the scaling

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

For each M we define $r_*^{(M)} \in]2,3[$ so that $(0, r_*^{(M)})$ is Φ^P-periodic of period M (and we denote by $(\theta_*^{(M)}(t), r_*^{(M)}(t))$ the corresponding solution of the flow of P); this way $a = (0, \frac{1}{N_j} r_*^{(M)})$ is G-periodic of period q.

When M increases, the Φ^P-periodic orbit that we consider approaches the upper separatrix of the pendulum, but we observe that, no matter how large M is,

$$\forall t \in [\tfrac{1}{2}, M - \tfrac{1}{2}], \quad \sigma \le \theta_*^{(M)}(t) \le 1 - \sigma,$$

where $\sigma = -\frac{1}{2} + \frac{2}{\pi}\arctan e^{\pi}$ ($\sigma = \theta_*^{(\infty)}(\frac{1}{2})$ if we use the time-parametrisation of the separatrix such that $(\theta_*^{(\infty)}(0), r_*^{(\infty)}(0)) = (0,2)$). In other words, in this orbit, only the points corresponding to times between $-\frac{1}{2}$ and $\frac{1}{2}$ (modulo M) have a chance to fall in the strip $\{-\sigma < \theta_* < \sigma\}$ (modulo 1); and among the q iterates $G^k(a)$, only those with $-\frac{N_j}{2} < k < \frac{N_j}{2}$ (modulo q) have a chance to fall in the strip $\{-\sigma < \theta_2 < \sigma\}$ (see Figure 3).

3.3.4 Consequently, we can take $g(\theta_2) = \xi(\theta_2)\varphi_\sigma(\theta_2)$ with a *fixed* Gevrey-(α, L) function φ_σ vanishing identically in $[\sigma, 1-\sigma]$ and satisfying $\varphi_\sigma(0) = 1$ and $\varphi'_\sigma(0) = 0$ (such a functions exists because $\alpha > 1$), and a Gevrey-(α, L) function ξ defined in $[-\sigma, \sigma]$ and satisfying the condition (11)

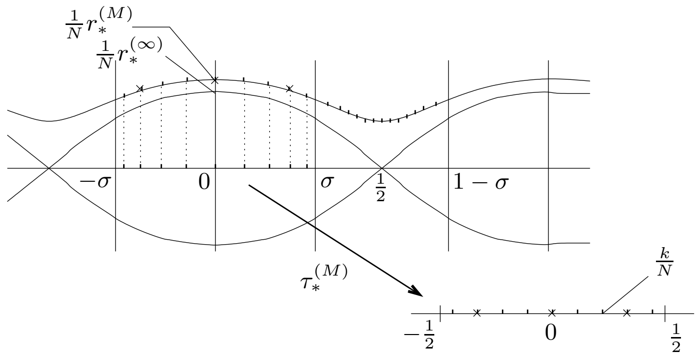

Figure 3: Orbit of a in the (θ_2, r_2) coordinates. The iterates inside $\{-\sigma < \theta_2 < \sigma\}$ necessarily correspond to $-\frac{N_j-1}{2} \le k \le \frac{N_j-1}{2}$ under the action of $\tau_*^{(M_j)}$, we get equidistributed points inside $\left]-\frac{1}{2}, \frac{1}{2}\right[$.

for the N_j points corresponding to $-j \le k \le j$: the function g will then satisfy (11) for all the values of k.

We choose $\xi = \eta_{N_j} \circ \tau_*^{(M)}$, where $\tau_*^{(M)} : [-\sigma, \sigma] \to \left]-\frac{1}{2}, \frac{1}{2}\right[$ is defined by $\tau_*^{(M)} \circ \theta_*^{(M)} = \mathrm{Id}$ and η_{N_j} is defined by

Lemma 3.2 *For $N \ge 1$, the function $\eta_N : \mathbb{T} \to \mathbb{R}$ defined by*

$$\eta_N(\theta) = \Big(\frac{1}{N}\sum_{\ell=0}^{N-1} \cos 2\pi\ell\theta\Big)^2, \qquad \theta \in \mathbb{T},$$

satisfies

$$\eta_N(0) = 1, \qquad \eta_N'(0) = 0, \qquad \eta_N(k/N) = \eta_N'(k/N) = 0, \qquad 1 \le k \le N-1.$$

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

$$\|\eta_N\|_{\alpha,\Lambda} \le e^{2\alpha\Lambda(2\pi N)^{1/\alpha}}.$$

(The proof of this lemma is straightforward.)

The idea is that the flow of the pendulum near the separatrix in the strip $\{-\sigma \le \theta_* \le \sigma\}$ is straightened by $\tau_*^{(M)}$, which leaves us with the case of N_j equidistributed points on the circle. Moreover, one can check that $\tau_*^{(M)}$ is bounded independently of M by some Λ, and therefore that our function g satisfies $\|g\|_{\alpha,L} \le \mathrm{const}\, \|\eta_{N_j}\|_{\alpha,\Lambda}$ (applying the composition lemma devised for the Gevrey-α case alluded to in Section 2.1). Now the reader can see the origin of our choice of $M = M_j$ in (8): it is just intended to guarantee $\frac{1}{N_j M}\|g\|_{\alpha,L} \le \frac{1}{N_j^2}$. We let the reader convince himself that $r_2^{(j)} = \frac{1}{N_j} r_*^{(M_j)}$ and $g^{(j)} = (\eta_{N_j} \circ \tau_*^{(M_j)})\varphi_\sigma$ satisfy all the desired properties.

4 Concluding remarks

4.1 The system that we have constructed has many similarities with the classical Arnold's example [Arn64]: the unperturbed system $\Phi^h = \Phi^{\frac{1}{2}r_1^2 + \frac{1}{2}r_2^2}$ depends only on the action variables; the perturbation $\frac{1}{N^2}V(\theta_2)$, added to h, still preserves the complete integrability (in the usual sense of symplectic geometry) and introduces hyperbolic objects; finally the perturbation $\Phi^{\frac{1}{q}U\otimes g}$ 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 is still the subject of current studies (see [LMS03] for instance, and [Ma03]); and even in the case where a transition chain is given, there remains some non-trivial work to do 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 shadow chains of transition tori for which the homoclinic splitting can be exactly computed. Our systems in fact appears as optimal from the point of view of Arnold's mechanism, since the chains we consider are formed by *maximally distant* hyperbolic tori (in our case, these tori are the circles $\{\theta_1 \in \mathbb{T}, r_1 = \frac{k}{q}, \theta_2 = \frac{1}{2}, r_2 = 0\}$ with $k \in \mathbb{Z}$): 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. (But the non-transversality of these intersections would prevent us from using Arnold's arguments to produce the drifting orbits; our system is rather a limit-case of Arnold's mechanism.)

4.2 To pass from wandering points to wandering domains in [MS03b], Moser's stability theorem for the elliptic fixed points of an area-preserving map of the plane is used twice: first, to obtain a small wandering domain $\mathcal{B}$ for the map $\psi = \Phi^{\frac{1}{q}U} \circ F^q$ introduced at the beginning of the construction in Section 3.3, but with a slightly modified potential U (to create ellipticity and ensure the twist condition after quotient by $r_1 \mapsto r_1 + \frac{1}{q}$); second, to obtain a small globally G-periodic domain $\mathcal{A}$ for a map $G = \Phi^{\frac{1}{2}r_2^2 + \frac{1}{N_j^2}V(\theta_2)} \circ \Phi^w$, with a well chosen function w. An easy modification of the coupling lemma 3.1, with the set $\mathcal{A}$ instead of the G-periodic point a, shows indeed that $\mathcal{B} \times \mathcal{A}$ will be wandering for $\Psi = \Phi^{\frac{1}{q}U \otimes g} \circ (\phi^{\frac{1}{2}r_1^2} \times G)$, for g well chosen.

The modifications needed for $n > 3$ are quite similar in the case of the wandering domain and in the case of the wandering point.

4.3 We also hope to be able to retain some of the ideas introduced in the Gevrey case to tackle the analytic one. Encouraging results were already obtained by J.-P. Marco [Ma03] for the problem of lower estimates of the homoclinic splitting in analytic examples of near-integrable systems.

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.

[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.

[Gio03] A. Giorgilli, "Notes on exponential stability of Hamiltonian systems," same volume.

[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.

[LMS03] P. Lochak, J.-P. Marco and D. Sauzin, "On the splitting of invariant manifolds in multi-dimensional near-integrable Hamiltonian systems", *to appear in Memoirs of the A.M.S.*

[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*, Birkhaüser, Boston (1993).

[Ma03] J.-P. Marco, "Uniform lower bounds of the splitting for analytic symplectic systems," *in preparation.*

[MS03a] J.-P. Marco and D. Sauzin, "Stability and instability for Gevrey quasi-convex near-integrable Hamiltonian systems," *Publ. Math. I.H.E.S.* **96** (2003), 77 pages.

[MS03b] J.-P. Marco and D. Sauzin, "Wandering domains for Gevrey perturbations of integrable systems," *in preparation.*

[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," to appear in *Ergod. Th. and Dyn. Syst.*

[Pö93] J. Pöschel, "Nekhoroshev estimates for quasi-convex Hamiltonian systems," *Math. Z.* **213** (1993), 187–216.

[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.