HAL open science

Nekhoroshev estimates for finitely differentiable quasi-convex Hamiltonians

Abed Bounemoura

▶ To cite this version:

Abed Bounemoura. Nekhoroshev estimates for finitely differentiable quasi-convex Hamiltonians. Journal of Differential Equations, 2010, 249 (11), pp.2905-2920. hal-00454594v2

HAL Id: hal-00454594
https://hal.science/hal-00454594v2

Submitted on 7 Jun 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Nekhoroshev estimates for finitely differentiable quasi-convex Hamiltonians

Abed Bounemoura

June 7, 2010

Abstract

A major result about perturbations of integrable Hamiltonian systems is the Nekhoroshev theorem, which gives exponential stability for all solutions provided the system is analytic and the integrable Hamiltonian is generic. In the particular but important case where the latter is quasi-convex, these exponential estimates have been generalized by Marco and Sauzin if the Hamiltonian is Gevrey regular, using a method introduced by Lochak in the analytic case. In this paper, using the same approach, we investigate the situation where the Hamiltonian is assumed to be only finitely differentiable, for which it is known that exponential stability does not hold but nevertheless we prove estimates of polynomial stability.

1 Introduction

In this paper, we are concerned with the stability properties of near-integrable Hamiltonian systems of the form

$$\begin{cases} H(\theta, I) = h(I) + f(\theta, I) \\ |f| < \varepsilon \ll 1 \end{cases}$$

where $(\theta, I) \in \mathbb{T}^n \times \mathbb{R}^n$ are action-angle coordinates for the integrable part h and f is a small perturbation in some suitable topology defined by a norm $|\,.\,|$. More precisely, we are interested in the evolution of the action variables $I(t)$, which are trivially constant in the absence of perturbation.

1. The first main result in this direction is given by an application of the KAM theory on the persistence of quasi-periodic solutions (see [Pös01] for a recent exposition). Assuming h satisfies some non-degeneracy condition and the system is analytic, if ε is sufficiently small there exists a constant c such that

$$|I(t) - I_0| \leq c\sqrt{\varepsilon}, \quad t \in \mathbb{R},$$

for "most" initial actions I_0, more precisely for a set of large measure but with empty interior. When $n = 2$, this is even true for all solutions provided h is isoenergetically non degenerate, but for $n > 3$, the famous example of Arnold ([Arn64]) shows that there exist "unstable" solutions, along which the variation of the actions can be arbitrarily large no matter how small the perturbation is. From its very beginning, KAM theory was known to hold for non-analytic Hamiltonians (see [Mos62] in the context of twist maps). It is now well established in various regularity classes, including the C^∞ case (essentially by Herman, see [Bos86] and [Féj04]) and the Gevrey case ([Pop04]). Following ideas of Moser, the theorem also holds if H is only of class C^k, with $k > 2n$ (see [Pös82], [Sal04], [SZ89] and also [Alb07] for a refinement), even though the minimal number of derivatives is still an open question, except in a special case for $n = 2$ ([Her86]).

2. Another fundamental result, which complements KAM theory, is given by Nekhoroshev's theorem ([Nek77], [Nek79]). If the integrable part h satisfies some generic condition and the system is analytic, then for ε sufficiently small there exist positive constants c_1, c_2, c_3, a and b such that

$$|I(t) - I_0| \leq c_1 \varepsilon^b, \quad |t| \leq c_2 \exp(c_3 \varepsilon^{-a}),$$

for all initial actions I_0. Hence all solutions are stable, not for all time, but for an exponentially long time. In the special case where h is strictly quasi-convex, a completely new proof of these estimates was given by Lochak ([Loc92]) using periodic averaging and simultaneous Diophantine approximation. The method of Lochak has had many applications, in particular it was used by Marco and Sauzin to extend Nekhoroshev's theorem to the Gevrey regular case under the quasi-convexity assumption ([MS02]).

3. However, no such estimates have been studied when the Hamiltonian is merely finitely differentiable, and this is the content of the present paper. We will prove below (Theorem 2.1) that if H is of class C^k, for $k \geq 2$, and h quasi-convex, then one has the stability estimates

$$|I(t) - I_0| \leq c_1 \varepsilon^{\frac{1}{2n}}, \quad |t| \leq c_2 \varepsilon^{-\frac{k-2}{2n}},$$

for some positive constants c_1 and c_2, and provided that ε is small enough. Of course, under our regularity assumption the exponential estimates have been replaced with polynomial estimates, and earlier examples show that exponential stability cannot possibly hold under such a weak regularity assumption (this is discussed in [MS04]). The proof will use once again the ideas of Lochak which, among other things, reduces the analytic part to its minimum and we will also follow the implementation of Marco and Sauzin in the Gevrey case.

4. As we recalled above, KAM theory for finitely differentiable Hamiltonian systems has been widely studied, and so we believe that Nekhoroshev's esti-

mates under weaker regularity assumptions have their own interest. Moreover, for obvious reasons, examples of unstable solutions (so-called Arnold diffusion) are more easily constructed in the non-analytic case, and it is a natural question to estimate the speed of instability (see [KL08a] and [KL08b] for examples of class C^k with a polynomial speed of diffusion). Finally, one of our motivations is to generalize these estimates using the method of [BN09], where Lochak's ideas are extended to deal with analytic but C^k-generic unperturbed Hamiltonians, with $k > 2n + 2$.

2 Main result

1. Let $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$, and consider a Hamiltonian function H defined on the domain

$$\mathcal{D}_R = \mathbb{T}^n \times B_R,$$

where B_R is the open ball of $\mathbb{R}^n$ around the origin of radius R, with respect to the supremum norm $|\,.\,|$. As usual, we shall occasionally identify H with a function defined on $\mathbb{R}^n \times B_R$ which is 1-periodic with respect to the first n variables.

We assume that H is of class C^k, for an integer $k \geq 2$, *i.e.* it is k-times differentiable and all its derivatives up to order k extend continuously to the closure $\overline{\mathcal{D}}_R$. We denote by $C^k(\mathcal{D}_R)$ the space of such functions, which is a Banach space with the norm

$$|H|_{C^k(\mathcal{D}_R)} = \sup_{0\leq l\leq k} \sup_{|\alpha|=l} \left(\sup_{x\in\overline{\mathcal{D}}_R} |\partial^\alpha H(x)| \right)$$

where $x = (\theta, I)$, $\alpha = (\alpha_1, \ldots, \alpha_{2n}) \in \mathbb{N}^{2n}$, $|\alpha| = \alpha_1 + \cdots + \alpha_{2n}$ and

$$\partial^\alpha = \partial_1^{\alpha_1} \ldots \partial_{2n}^{\alpha_{2n}}.$$

In the case where the Hamiltonian $H = h$ depends only on the action variables, we will simply write $|h|_{C^k(B_R)}$.

Our Hamiltonian $H \in C^k(\mathcal{D}_R)$ is assumed to be C^k-close to integrable, that is, of the form

$$\begin{cases} H(\theta, I) = h(I) + f(\theta, I) \\ |f|_{C^k(\mathcal{D}_R)} < \varepsilon \ll 1 \end{cases} \tag{$*$}$$

where h is the integrable part and f a small perturbation of size ε in the C^k topology.

We may assume that there exists a positive constant M such that

$$|h|_{C^k(B_R)} < M. \tag{B}$$

Moreover, we will require that h satisfy the following quasi-convexity assumption: there exists a positive constant m such that

$$\forall I \in B_R, \forall v \in \mathbb{R}^n, \quad \nabla h(I).v = 0 \Longrightarrow \nabla^2 h(I)v.v \geq m|v|^2. \tag{C}$$

2. Our main theorem is the following.

Theorem 2.1. *Let H be as in $(*)$, with $k \geq 3$, h satisfying (B) and (C), and set*

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

Then there exist ε_0, c_1 and c_2 such that if $\varepsilon \leq \varepsilon_0$, all solutions $(\theta(t), I(t))$ of H with $I(0) \in B_{R/2}$ satisfy

$$|I(t) - I(0)| \leq c_1 \varepsilon^b, \quad |t| \leq c_2 \varepsilon^{-a}.$$

First note that we have stated our theorem for H of class C^k, $k \geq 3$, but the statement also holds with no changes if H is of class $C^{k-1,1}$, $k \geq 3$, that is, H is of class C^{k-1} and its partial derivatives of order $k-1$ are Lipschitz continuous. For C^2 systems, these estimates also hold true but it is not useful since trivially all solutions satisfy

$$|I(t) - I(0)| \leq \varepsilon, \quad |t| \leq 1.$$

In fact for C^2 Hamiltonians for which the derivatives up to order 2 are not more than continuous, we believe that one cannot obtain better estimates.

Moreover, the time of stability obtained is "optimal" in the sense that one can construct examples of unstable orbits with a polynomial speed of diffusion, but we do not know what the optimal exponents should be. However, using the geometric arguments of [BM10] one can easily improve the stability exponent a in order to obtain

$$a = \frac{k-2}{2(n-1)} - \delta,$$

for $\delta > 0$ but arbitrarily small.

Let us finally point out that if H is C^∞, then it is an immediate consequence of the above result that the action variables are stable for an interval of time which is longer than any prescribed power of ε^{-1}, but even in this case exponential stability does not hold.

3. As in the analytic or Gevrey case, we can also state a refined result near resonances. Suppose Λ is a sub-module of $\mathbb{Z}^n$ of rank m, $d = n - m$ and let S_Λ be the corresponding resonant manifold, that is

$$S_\Lambda = \{I \in B_R \mid k.\nabla h(I) = 0, \ k \in \Lambda\}.$$

We can prove the following statement, which actually contains the previous one.

Theorem 2.2. *Under the previous hypotheses, assume* $d(I(0), S_\Lambda) \leq \sigma\sqrt{\varepsilon}$ *for some constant* $\sigma > 0$*, and set*

$$a_d = \frac{k-2}{2d}, \quad b_d = \frac{1}{2d}.$$

Then there exist ε_0'*,* c_1' *and* c_2' *such that if* $\varepsilon \leq \varepsilon_0'$*, one has*

$$|I(t) - I(0)| \leq c_1' \varepsilon^{b_d}, \quad |t| \leq c_2' \varepsilon^{-a_d}.$$

For $\Lambda = \{0\}$, $d = n$ and $S_\Lambda = B_{R/2}$, we recover Theorem 2.1 and therefore it will be enough to prove Theorem 2.2.

4. The constants ε_0, c_1 and c_2 depend only on h, more precisely they depend on k, n, R, M and m while the constants ε_0', c_1' and c_2' also depend on σ and Λ. However we will not give explicit values for them in order to avoid complicated and rather meaningless expressions. Hence we shall replace them by the symbol $\cdot$ when it is convenient: for instance, we shall write $u \lessdot v$ when there exists a positive constant c depending only on the previous parameters, but not on f, such that $u < cv$.

5. This paper is divided into two sections. The next section contains the analytical part of the proof, where we will construct a system of local coordinates for our Hamiltonian which is more convenient to study the evolution of the action variables. Then, in the last section we will conclude the proof using our convexity assumption and Dirichlet's theorem on simultaneous Diophantine approximation.

3 Analytical part

1. Given an action $I \in B_R$ and denoting by $\omega = \nabla h(I)$ its frequency, we know from classical averaging theory that the relevant part of the perturbation

$$f(\theta, I) = \sum_{k \in \mathbb{Z}^n} \hat{f}_k(I) e^{i2\pi k.\theta}$$

is given by those harmonics associated with integers $k \in \mathbb{Z}^n$ in resonance with ω, that is such that $k.\omega = 0$. Actually one can construct a symplectic, close-to-identity transformation Φ defined around I, such that

$$H \circ \Phi = h + g + \tilde{f}$$

where g contains only harmonics in resonance with ω and $\tilde{f}$ is a small remainder. These are usually called resonant normal forms, and to obtain them one has to deal with small divisors $k.\omega$ which involve technical estimates. If the system is analytic, the above remainder $\tilde{f}$ can be made exponentially small with respect to the inverse of the size of the perturbation, as was first

shown by Nekhoroshev. But for finitely differentiable systems one might guess that the remainder can only be polynomially small, even though this should be difficult (or at least technical) to prove using the usual approach.

2. It is a remarkable fact discovered by Lochak ([Loc92]) that to prove exponential estimates in the quasi-convex case with the analyticity assumption, it is enough to average along periodic frequencies, which are frequencies ω such that $T\omega \in \mathbb{Z}^n \setminus \{0\}$ for some $T > 0$ (see also [BN09] for an extension of this method for generic integrable Hamiltonians). These periodic frequencies correspond to periodic orbits of the unperturbed Hamiltonian, hence in this approach no small divisors arise. As a consequence this special resonant normal form is much easier to obtain. The aim of this section is to construct such a normal form, up to a polynomial remainder. This will be done in 3.3. But first we will recall some useful estimates concerning the C^k norm in 3.1, and then prove an intermediate statement in 3.2.

3.1 Elementary estimates

3. Let us begin by recalling some easy estimates. Given two functions $f, g \in C^k(\mathcal{D}_R)$, the product fg belongs to $C^k(\mathcal{D}_R)$ and by the Leibniz rule

$$|fg|_{C^k(\mathcal{D}_R)} \lesssim |f|_{C^k(\mathcal{D}_R)} |g|_{C^k(\mathcal{D}_R)}.$$

The Poisson Bracket $\{f, g\}$ belongs to $C^{k-1}(\mathcal{D}_R)$, and by its definition and the Leibniz rule one gets

$$|\{f, g\}|_{C^{k-1}(\mathcal{D}_R)} \lesssim |f|_{C^k(\mathcal{D}_R)} |g|_{C^k(\mathcal{D}_R)}.$$

The above implicit constants depend only on n and k (in fact in the first estimate one can trivially modify the definition of the C^k norm so as to have a constant equal to one, but this will not be important for us). These are very elementary facts, but we shall also need estimates concerning vector fields, canonical transformations and compositions.

4. First, given a vector-valued function $F \in C^k(\mathcal{D}_R, \mathbb{R}^l)$, $F = (F_1, \ldots, F_l)$ and $l \in \mathbb{N}$, we extend the norm component-wise, that is

$$|F|_{C^k(\mathcal{D}_R)} = \sup_{1 \le i \le l} |F_i|_{C^k(\mathcal{D}_R)}.$$

Now for a function $f \in C^k(\mathcal{D}_R)$, we define its Hamiltonian vector field X_f by

$$X_f = (\partial_I f, -\partial_\theta f)$$

where

$$\partial_I f = (\partial_{I_1} f, \ldots, \partial_{I_n} f), \quad \partial_\theta f = (\partial_{\theta_1} f, \ldots, \partial_{\theta_n} f).$$

Obviously $X_f \in C^{k-1}(\mathcal{D}_R, \mathbb{R}^{2n})$, and trivially

$$|X_f|_{C^{k-1}(\mathcal{D}_R)} \leq |f|_{C^k(\mathcal{D}_R)}.$$

Moreover, by classical theorems on ordinary differential equations, if X_f is of class C^{k-1} then so is the time-t map Φ_t^f of the vector field X_f, when it exists. Assuming $|\partial_\theta f|_{C^0(\mathcal{D}_R)} < r$ for some $r < R$ (for example $|X_f|_{C^0(\mathcal{D}_R)} < r$), then by the mean value theorem

$$\Phi^f = \Phi_1^f : \mathcal{D}_{R-r} \longrightarrow \mathcal{D}_R$$

is a well-defined C^{k-1}-embedding. In the case where f is integrable, one can choose $r = 0$.

In the sequel, we will need to estimate the C^k norm of Φ^f in terms of the C^k norm of the vector field X_f. More precisely we need the rather natural fact that Φ^f is C^k-close to the identity when X_f is C^k-close to zero. This is trivial for $k = 0$. In the general case, this follows by induction on k using on the one hand the relation

$$\Phi_t^f = \mathrm{Id} + \int_0^t X_f \circ \Phi_s^f ds,$$

and on the other the formula of Faà di Bruno (see [AR67] for example), which gives bounds of the form

$$|F \circ G|_{C^k} \lesssim |F|_{C^k} |G|_{C^k}^k$$

and also

$$|F \circ G|_{C^k} \lesssim |F|_{C^1} |G|_{C^k}^k + |F|_{C^k} |G|_{C^{k-1}}^k$$

for C^k vector-valued functions on appropriate domains (once again, the above implicit constants depend only on k). Let us state this as a lemma, for which we refer to [DH09], Lemma 3.15 and appendix C, for a detailed proof.

Lemma 3.1. *Let $X_f \in C^k(\mathcal{D}_R, \mathbb{R}^{2n})$, assume that $|X_f|_{C^0(\mathcal{D}_R)} < r$ and*

$$|X_f|_{C^k(\mathcal{D}_R)} < 1. \tag{1}$$

Then

$$|\Phi^f - \mathrm{Id}|_{C^k(\mathcal{D}_{R-r})} \lesssim |X_f|_{C^k(\mathcal{D}_R)}.$$

The above implicit constant depends only on k and R. Now if $H \in C^k(\mathcal{D}_R)$, under the above hypotheses we have $H \circ \Phi^f \in C^k(\mathcal{D}_{R-r})$ and the estimate

$$|H \circ \Phi^f|_{C^k(\mathcal{D}_{R-r})} \lesssim |H|_{C^k(\mathcal{D}_R)} |\Phi^f|_{C^k(\mathcal{D}_{R-r})}^k \tag{2}$$

follows trivially from the Faà di Bruno formula.

3.2 The linear case

Following [MS02], we change for a moment our setting and we consider a perturbation of a linear Hamiltonian, more precisely the Hamiltonian

$$\begin{cases} H(\theta, I) = l(I) + f(\theta, I) \\ |f|_{C^k(\mathcal{D}_\rho)} < \mu \ll 1 \end{cases} \tag{$**$}$$

where $\rho > 0$ is fixed and $l(I) = \omega.I$ is a linear Hamiltonian with a T-periodic frequency ω. Recall that this means that

$$T = \inf\{t > 0 \mid t\omega \in \mathbb{Z}^n \setminus \{0\}\}$$

is well-defined. In this context, our small parameter is μ.

In the proposition below, we will construct a "global" normal form for the Hamiltonian $(**)$, which we will use in the next section to produce a "local" normal form around periodic orbits for our original Hamiltonian $(*)$.

Proposition 3.2. *Consider H as in $(**)$ with $k \geq 2$, and assume*

$$T\mu \lessdot 1, \quad |\omega| \lessdot 1. \tag{3}$$

Then there exists a C^2 symplectic transformation

$$\Phi : \mathcal{D}_{\rho/2} \to \mathcal{D}_\rho$$

with $|\Phi - \mathrm{Id}|_{C^2(\mathcal{D}_{\rho/2})} \lessdot T\mu$ such that

$$H \circ \Phi = l + g + f,$$

with $\{g, l\} = 0$ and the estimates

$$|g|_{C^2(\mathcal{D}_{\rho/2})} \lessdot \mu, \quad |f|_{C^2(\mathcal{D}_{\rho/2})} \lessdot (T\mu)^{k-2}\mu$$

hold true.

First note that $\{g, l\} = 0$ means exactly that $\partial_\theta g.\omega = 0$, and expanding g in a Fourier series, one easily sees that it contains only harmonics associated with integers k satisfying $k.\omega = 0$. Therefore the above proposition gives indeed a resonant normal form, up to a polynomial remainder.

Note also that we need our transformed Hamiltonian $H \circ \Phi$, and hence our transformation Φ, to be at least of class C^2, simply because we need our transformed vector field to be of class C^1 to have existence and uniqueness of solutions (or course, a $C^{1,1}$ regularity would have been enough). This explains the factor $(k-2)$ in our stability exponent a.

Proof. First, as we are assuming $|\omega| \lessdot 1$, we have

$$|\Phi_t^l|_{C^k(\mathcal{D}_\rho)} \lessdot 1. \tag{4}$$

Our transformation Φ will be obtained by a finite composition of averaging transformations. Let us define

$$r = \frac{\rho}{2(k-2)}$$

and for $j \in \{0, \dots, k-2\}$, let

$$\rho_j = \rho - jr \geq \rho/2.$$

Then we claim that for any $j \in \{0, \dots, k-2\}$, there exists a C^{k-j} symplectic transformation $\Phi_j : \mathcal{D}_{\rho_j} \to \mathcal{D}_\rho$ with $|\Phi_j - \mathrm{Id}|_{C^{k-j}(\mathcal{D}_{\rho_j})} \lessdot T\mu$ such that

$$H \circ \Phi_j = l + g_j + f_j,$$

with g_j and f_j of class C^{k-j}, $\{g_j, l\} = 0$ and the estimates

$$|g_j|_{C^{k-j}(\mathcal{D}_{\rho_j})} \lessdot \mu, \quad |f_j|_{C^{k-j}(\mathcal{D}_{\rho_j})} \lessdot (T\mu)^j \mu.$$

The proposition follows easily by taking $\Phi = \Phi_{k-2}$, $g = g_{k-2}$ and $f = f_{k-2}$.

We will prove the claim by induction on $j \in \{0, \dots, k-2\}$. For $j = 0$, there is nothing to prove since we can simply write $H = l + g_0 + f_0$ with $g_0 = 0$, $f_0 = f$ and therefore Φ_0 is the identity. Now assume the claim is true for some $j \in \{0, \dots, k-3\}$, and consider

$$H_j = H \circ \Phi_j = l + g_j + f_j.$$

Let us define

$$[f_j] = \frac{1}{T} \int_0^T f_j \circ \Phi_t^l dt$$

and

$$\chi_j = \frac{1}{T} \int_0^T t(f_j - [f_j]) \circ \Phi_t^l dt.$$

We have

$$|[f_j]|_{C^{k-j}(\mathcal{D}_{\rho_j})} \leq |f_j \circ \Phi_t^l|_{C^{k-j}(\mathcal{D}_{\rho_j})}$$

hence by (2) and (4) we obtain

$$|[f_j]|_{C^{k-j}(\mathcal{D}_{\rho_j})} \lessdot |f_j|_{C^{k-j}(\mathcal{D}_{\rho_j})}$$

and with our hypotheses of induction, this gives

$$|[f_j]|_{C^{k-j}(\mathcal{D}_{\rho_j})} \lessdot (T\mu)^j \mu.$$

Similarly

$$|\chi_j|_{C^{k-j}(\mathcal{D}_{\rho_j})} \lessdot T(T\mu)^j\mu = (T\mu)^{j+1}$$

and since $T\mu \lessdot 1$,

$$|\chi_j|_{C^{k-j}(\mathcal{D}_{\rho_j})} \lessdot T\mu.$$

If we let Φ^{χ_j} be the time-one map of the Hamiltonian vector field generated by χ_j, then we will show that the map

$$\Phi_{j+1} = \Phi_j \circ \Phi^{\chi_j}$$

satisfies the assumptions.

Indeed, thanks to the condition $T\mu \lessdot 1$, we can ensure that Φ^{χ_j}, which is of class C^{k-j-1}, is a well-defined embedding

$$\Phi^{\chi_j} : \mathcal{D}_{\rho_{j+1}} \longrightarrow \mathcal{D}_{\rho_j}.$$

Moreover, as $|X_{\chi_j}|_{C^{k-j-1}(\mathcal{D}_{\rho_j})} \lessdot T\mu$ and using once again $T\mu \lessdot 1$, we can arrange condition (1) and apply Lemma 3.1 to obtain

$$|\Phi^{\chi_j} - \mathrm{Id}|_{C^{k-j-1}(\mathcal{D}_{\rho_{j+1}})} \lessdot T\mu.$$

Now

$$\begin{aligned}
|\Phi_{j+1} - \mathrm{Id}|_{C^{k-j-1}(\mathcal{D}_{\rho_{j+1}})} &= |\Phi_j \circ \Phi^{\chi_j} - \Phi^{\chi_j} + \Phi^{\chi_j} - \mathrm{Id}|_{C^{k-j-1}(\mathcal{D}_{\rho_{j+1}})} \\
&\leq |(\Phi_j - \mathrm{Id}) \circ \Phi^{\chi_j}|_{C^{k-j-1}(\mathcal{D}_{\rho_{j+1}})} \\
&+ |\Phi^{\chi_j} - \mathrm{Id}|_{C^{k-j-1}(\mathcal{D}_{\rho_{j+1}})} \\
&\lessdot |\Phi_j - \mathrm{Id}|_{C^{k-j-1}(\mathcal{D}_{\rho_j})} + |\Phi^{\chi_j} - \mathrm{Id}|_{C^{k-j-1}(\mathcal{D}_{\rho_{j+1}})}
\end{aligned}$$

where we have used (2) in the last line. By our hypotheses of induction, this eventually gives

$$|\Phi_{j+1} - \mathrm{Id}|_{C^{k-j-1}(\mathcal{D}_{\rho_{j+1}})} \lessdot T\mu.$$

Now by Taylor's formula with integral remainder, we can expand

$$H_{j+1} = H \circ \Phi_{j+1} = l + g_{j+1} + f_{j+1}$$

with

$$g_{j+1} = g_j + [f_j], \quad f_{j+1} = \int_0^1 \{g_j + f_j^t, \chi_j\} \circ \Phi_t^{\chi} dt$$

where $f_j^t = tf_j + (1-t)[f_j]$, as one can check by a standard calculation. Since $\{g_j, l\} = 0$ by our hypothesis of induction and obviously $\{[f_j], l\} = 0$, we have $\{g_{j+1}, l\} = 0$ together with the estimate

$$\begin{aligned}
|g_{j+1}|_{C^{k-j}(\mathcal{D}_{\rho_j})} &\leq |g_j|_{C^{k-j}(\mathcal{D}_{\rho_j})} + |[f_j]|_{C^{k-j}(\mathcal{D}_{\rho_j})} \\
&\lessdot \mu + (T\mu)^j\mu \\
&\lessdot \mu
\end{aligned}$$

using $T\mu \lessdot 1$. For the remainder, using (2) and the trivial estimate for the Poisson bracket we compute

$$\begin{aligned}
|f_{j+1}|_{C^{k-j-1}(\mathcal{D}_{\rho_{j+1}})} &\leq |\{g_j + f_j^t, \chi_j\} \circ \Phi_t^{\chi_j}|_{C^{k-j-1}(\mathcal{D}_{\rho_{j+1}})} \\
&\lessdot |\{g_j + f_j^t, \chi_j\}|_{C^{k-j-1}(\mathcal{D}_{\rho_j})} \\
&\lessdot |\{g_j, \chi_j\}|_{C^{k-j-1}(\mathcal{D}_{\rho_j})} + |\{f_j^t, \chi_j\}|_{C^{k-j-1}(\mathcal{D}_{\rho_j})} \\
&\lessdot |g_j|_{C^{k-j}(\mathcal{D}_{\rho_j})}|\chi_j|_{C^{k-j}(\mathcal{D}_{\rho_j})} + |f_j^t|_{C^{k-j}(\mathcal{D}_{\rho_j})}|\chi_j|_{C^{k-j}(\mathcal{D}_{\rho_j})} \\
&\lessdot \mu(T\mu)^{j+1} + \mu(T\mu)^j(T\mu)^{j+1} \\
&\lessdot \mu(T\mu)^{j+1}
\end{aligned}$$

using once again $T\mu \lessdot 1$. This concludes the proof. □

3.3 Normal form

Now let us come back to our original setting which is the Hamiltonian

$$\begin{cases} H(\theta, I) = h(I) + f(\theta, I) \\ |f|_{C^k(\mathcal{D}_R)} < \varepsilon \ll 1. \end{cases}$$

We will say that an action $I_* \in B_R$ is T-periodic if its frequency vector $\nabla h(I_*)$ is T-periodic. In the proposition below, we will fix a T-periodic action I_*, l will be the linear integrable Hamiltonian $l(I) = \omega.I$ associated with the periodic frequency $\omega = \nabla h(I_*)$ and we denote by

$$\Pi_I : \mathbb{T}^n \times B_R \to B_R$$

the projection onto the action space.

Proposition 3.3 (Normal form)**.** *Suppose H is as in $(*)$, with h satisfying (B). Under the previous hypotheses, let $\mu > 0$ be such that*

$$\varepsilon \lessdot \mu^2, \quad \mu \lessdot 1, \quad T\mu \lessdot 1. \tag{5}$$

Then there exists a C^2 symplectic transformation

$$\Phi : \mathbb{T}^n \times B(I_*, \mu) \to \mathbb{T}^n \times B(I_*, 2\mu)$$

with $|\Pi_I \Phi - \mathrm{Id}_I|_{C^0(B(I_*,\mu))} \lessdot T\mu^2$ such that

$$H \circ \Phi = h + g + \tilde{f},$$

with $\{g, l\} = 0$ and the estimates

$$|g + \tilde{f}|_{C^0(\mathbb{T}^n \times B(I_*,\mu))} \lessdot \mu^2, \quad |\partial_{\tilde{\theta}} \tilde{f}|_{C^0(\mathbb{T}^n \times B(I_*,\mu))} \lessdot (T\mu)^{k-2}\mu^2$$

hold true.

Let us immediately explain how such coordinates will be used in the sequel. If we denote them by $(\tilde{\theta}, \tilde{I})$, then writing down the equations of motion for $\tilde{H} = H \circ \Phi$, using the fact that $\partial_{\tilde{\theta}} g.\omega = 0$ and the mean value theorem, one shows that $\tilde{I}(t)$ remains close to the hyperplane orthogonal to ω, passing through $\tilde{I}(0)$, for an interval of time governed by the size of $\partial_{\tilde{\theta}} \tilde{f}$. Hence any potential drift has to occur along that hyperplane. Then this picture persists and gets only slightly distorted when we come back to the original coordinates (θ, I), since the projection of Φ onto action space is close to identity.

Proof. First note that since h satisfies (B), then

$$|\omega| = |\nabla h(I_*)| < M$$

and therefore the second part of condition (3) is satisfied.

Without loss of generality, we may assume $h(I_*) = 0$. To analyze our Hamiltonian H in a neighborhood of size μ around I^*, we translate and rescale the action variables using the map

$$\sigma_\mu : (\theta, I^\mu) \longmapsto (\theta, I) = (\theta, I_* + \mu I^\mu)$$

which sends the domain $\mathcal{D}_2 = \mathbb{T}^n \times B_2$ onto $\mathbb{T}^n \times B(I_*, 2\mu)$, and note that by the condition $\mu < 1$, we can assume that the latter domain is included in $\mathcal{D}_R$. Let

$$H_\mu = \mu^{-1}(H \circ \sigma_\mu)$$

be the rescaled Hamiltonian, so H_μ is defined on $\mathcal{D}_2$ and reads

$$H_\mu(\theta, I^\mu) = \mu^{-1} H(\theta, I_* + \mu I^\mu) = \mu^{-1} h(I_* + \mu I^\mu) + \mu^{-1} f(\theta, I_* + \mu I^\mu)$$

for $(\theta, I^\mu) \in \mathcal{D}_2$. Now using Taylor's formula we can expand h around I_* to obtain

$$\begin{aligned} h(I_* + \mu I^\mu) &= \mu\, \omega.I^\mu + \mu^2 \int_0^1 (1-t) \nabla^2 h(I_* + t\mu I^\mu) I^\mu.I^\mu dt \\ &= \mu\, \omega.I^\mu + \mu^2 h_\mu(I^\mu) \end{aligned}$$

where we set

$$h_\mu(I^\mu) = \int_0^1 (1-t) \nabla^2 h(I_* + t\mu I^\mu) I^\mu.I^\mu dt.$$

Therefore we can write

$$H_\mu = l + f_\mu$$

with

$$f_\mu = \mu h_\mu + \mu^{-1}(f \circ \sigma_\mu).$$

Now we know that $|f|_{C^k(\mathcal{D}_R)} < \varepsilon \lessdot \mu^2$ from the first part of condition (5) and $|h|_{C^k(\mathcal{D}_R)} \lessdot 1$ since we are assuming (B), so we obviously have

$$|f_\mu|_{C^k(\mathcal{D}_2)} \lessdot \mu.$$

Therefore conditions (3) are satisfied and we can apply Proposition 3.2, with $\rho = 2$, to the Hamiltonian $H_\mu = l + f_\mu$: there exists a C^2 symplectic transformation $\Phi_\mu : \mathcal{D}_1 \to \mathcal{D}_2$ with $|\Phi_\mu - \mathrm{Id}|_{C^2(\mathcal{D}_1)} \lessdot T\mu$ such that

$$H_\mu \circ \Phi_\mu = l + g_\mu + \tilde{f}_\mu$$

with $\{g_\mu, l\} = 0$ and the estimates

$$|g_\mu|_{C^2(\mathcal{D}_1)} \lessdot \mu, \quad |\tilde{f}_\mu|_{C^2(\mathcal{D}_1)} \lessdot (T\mu)^{k-2}\mu.$$

Moreover, if we set

$$\hat{f}_\mu = g_\mu - \mu h_\mu,$$

we have $\{\hat{f}_\mu, l\} = 0$ and $|\hat{f}_\mu|_{C^2(\mathcal{D}_1)} \lessdot \mu$, and so the transformed Hamiltonian can also be written as

$$H_\mu \circ \Phi_\mu = l + \mu h_\mu + \hat{f}_\mu + \tilde{f}_\mu.$$

Now scaling back to our original coordinates, we define $\Phi = \sigma_\mu \circ \Phi_\mu \circ \sigma_\mu^{-1}$, therefore

$$\Phi : \mathbb{T}^n \times B(I_*, \mu) \longrightarrow \mathbb{T}^n \times B(I_*, 2\mu)$$

and

$$\begin{aligned} H \circ \Phi &= \mu H_\mu \circ \Phi_\mu \circ \sigma_\mu^{-1} \\ &= \mu(l + \mu h_\mu + \hat{f}_\mu + \tilde{f}_\mu) \circ \sigma_\mu^{-1} \\ &= (\mu l + \mu^2 h_\mu) \circ \sigma_\mu^{-1} + \mu \hat{f}_\mu \circ \sigma_\mu^{-1} + \mu \tilde{f}_\mu \circ \sigma_\mu^{-1}. \end{aligned}$$

Observe that $(\mu l + \mu^2 h_\mu) \circ \sigma_\mu^{-1} = h$, so we may set

$$g = \mu \hat{f}_\mu \circ \sigma_\mu^{-1}, \quad \tilde{f} = \mu \tilde{f}_\mu \circ \sigma_\mu^{-1},$$

and write

$$H \circ \Phi = h + g + \tilde{f}.$$

It is obvious that $\{g, l\} = 0$ with

$$|g|_{C^0(\mathbb{T}^n \times B(I_*, \mu))} \leq \mu |\hat{f}_\mu|_{C^0(\mathcal{D}_1)} \lessdot \mu^2$$

and similarly

$$|\tilde{f}|_{C^0(\mathbb{T}^n \times B(I_*, \mu))} \leq \mu |\tilde{f}_\mu|_{C^0(\mathcal{D}_1)} \lessdot (T\mu)^{k-2}\mu^2 \lessdot \mu^2$$

so

$$|g+\tilde{f}|_{C^0(\mathbb{T}^n\times B(I_*,\mu))} \lessdot \mu^2.$$

Moreover, as $\partial_{\tilde{\theta}}\tilde{f} = \mu\partial_{\tilde{\theta}}\tilde{f}_\mu$ then

$$|\partial_{\tilde{\theta}}\tilde{f}|_{C^0(\mathbb{T}^n\times B(I_*,\mu))} \lessdot (T\mu)^{k-2}\mu^2$$

and finally

$$|\Pi_I\Phi - \mathrm{Id}_I|_{C^0(B(I_*,\mu))} \lessdot T\mu^2.$$

is trivial. This ends the proof. □

4 Proof of Theorem 2.2

1. Now we can complete the proof of our Theorem 2.2 in the spirit of Lochak, following three elementary steps that use successively some arithmetic (simultaneous Diophantine approximation), some analysis (normal forms around periodic orbits) and some geometry (quasi-convexity). The analysis has been done in the previous section, and the arithmetic and geometry are exactly the same as in the analytic case or the Gevrey case. Therefore instead of rewriting proofs which are well-known, we will merely explain the ideas and state the relevant results which can be found in [Loc92] and [MS02].

2. Let us begin with the arithmetic part, since in order to use our Proposition 3.3 (the normal form), we will need to show that any action $I_0 \in B_{R/2}$, which is close to some resonant surface S_Λ, can be approximated by a periodic action. Let $\omega_0 = \nabla h(I_0)$, then using the isoenergetic non-degeneracy of h, (which is easily implied by the quasi-convexity assumption), it is enough to approximate ω_0 by a periodic vector ω. If Λ has rank $n-1$, this is totally obvious, since necessarily we have $\Lambda = \omega^\perp \cap \mathbb{Z}^n$ for some periodic vector ω and so each action in S_Λ is periodic. Now in the case where Λ has rank $m = n-d$ with $d > 1$, a good approximation is given by a theorem of Dirichlet, which moreover gives an explicit bound on the period T.

Proposition 4.1. *Let $I_0 \in B_{R/2}$, Λ be a sub-module of $\mathbb{Z}^n$ of rank $m = n-d$, with $d > 1$ and Q be a real number such that*

$$Q \cdot\!> 1. \tag{6}$$

Then there exists a T-periodic action $I_* \in B_R$ such that

$$|I_0 - I_*| \lessdot \max\left(d(I_0, S_\Lambda), T^{-1}Q^{-\frac{1}{d-1}}\right)$$

and the period T satisfies

$$1 \lessdot T \lessdot Q.$$

The proof of the above proposition can be found in [MS02], Corollary 3.2.

3. Now it remains to explain how the quasi-convexity hypothesis (C) on h, together with the normal form obtained in Proposition 3.3, will enable us to control the variation $|I(t) - I_0|$, for an initial action I_0 close to some periodic action I_*. The idea goes as follows.

Let $\omega = \nabla h(I_*)$. We have already explained after Proposition 3.3 how in those new coordinates, the evolution of the actions in the direction generated by ω is controlled for some interval of time: if F is the hyperplane orthogonal to ω, then $\tilde{I}(t)$ remains close to the affine subspace $\tilde{I}_0 + F$ for an interval of time $|t| \leq \tau$ where τ is essentially given by the inverse of the size of $\tilde{f}$. Now by preservation of energy, for all time $\tilde{I}(t)$ remains close to the unperturbed energy hypersurface $E_0 = \{\tilde{I} \in B_R \mid h(\tilde{I}) = h(\tilde{I}_0)\}$, and as the latter is strictly convex by quasi-convexity of h, the connected component of $\tilde{I}_0$ in $E_0 \cap (\tilde{I}_0 + F)$ is in fact bounded and so is the variation $|\tilde{I}(t) - \tilde{I}_0|$ for $|t| \leq \tau$.

This idea is formalized in the proposition below. Once again, I_* is a T-periodic action and l is the linear Hamiltonian with periodic frequency $\omega = \nabla h(I_*)$.

Proposition 4.2. *Under the previous hypotheses, let $r > 0$, $\tau > 0$ and*

$$\tilde{H} = h + g + \tilde{f} \in C^2(\mathbb{T}^n \times B(I_*, r))$$

with h satisfying (C), $\{g, l\} = 0$ and the estimates

$$|g + \tilde{f}|_{C^0(\mathbb{T}^n \times B(I_*, r))} < r^2, \quad |\partial_{\tilde{\theta}} \tilde{f}|_{C^0(\mathbb{T}^n \times B(I_*, r))} < r^2 \tau^{-1}.$$

If

$$r \lessdot 1 \tag{7}$$

then for any initial condition $(\tilde{\theta}_0, \tilde{I}_0) \in \mathbb{T}^n \times B(I_*, r)$, the solution satisfies

$$|\tilde{I}(t) - \tilde{I}_0| \lessdot r, \quad |t| \leq \tau.$$

Once again, we refer to [MS02], Corollary 3.1, for a complete proof.

4. Let us now conclude the proof of Theorem 2.2. In a first step we will use Proposition 4.1 to find a periodic action close to our initial action, then in a second step we will apply Proposition 3.3 to find adapted coordinates and the third step will consist in applying Proposition 4.2 to control the evolution of the action variables in those coordinates, and hence in the original coordinates.

Proof of Theorem 2.2. Let $I_0 \in B_{R/2}$, let Λ be a sub-module of $\mathbb{Z}^n$ of rank $m = n - d$, and assume that $d(I_0, S_\Lambda) \lessdot \sqrt{\varepsilon}$ (note that this last assumption is void if Λ is trivial).

First step. In the case $d = 1$, any action $I \in S_\Lambda$ has a frequency $\nabla h(I)$ which is a multiple of some non zero vector $k_* \in \mathbb{Z}^n$, therefore we can choose a periodic action $I_* \in S_\Lambda$ so

$$|I_0 - I_*| = d(I_0, S_\Lambda) <\cdot \sqrt{\varepsilon}$$

and the period T trivially satisfies $T <\cdot 1$. In the case $d > 1$, we apply Proposition 4.1 with

$$Q =\cdot \varepsilon^{-\frac{d-1}{2d}}.$$

and the condition (6) gives a first smallness condition on ε. Observe that $Q^{-\frac{1}{d-1}} =\cdot \varepsilon^{\frac{1}{2d}}$, hence the periodic action I_* given by the proposition satisfies

$$|I_0 - I_*| <\cdot \max\left(d(I_0, S_\Lambda), T^{-1}Q^{-\frac{1}{d-1}}\right) <\cdot T^{-1}\varepsilon^{\frac{1}{2d}}$$

and the period T

$$1 <\cdot T <\cdot \varepsilon^{-\frac{d-1}{2d}}. \tag{8}$$

Second step. Having found a periodic action, we will now apply Proposition 3.3 with

$$\mu =\cdot T^{-1}\varepsilon^{\frac{1}{2d}}.$$

With this choice, for $d > 1$ the first part of condition (5) is satisfied thanks to the upper bound (8) on the period T (for $d = 1$, this is trivial). The other conditions in (5) give only further smallness conditions on ε. Applying the proposition, we have a C^2 symplectic transformation

$$\Phi : \mathbb{T}^n \times B(I_*, \mu) \to \mathbb{T}^n \times B(I_*, 2\mu)$$

with $|\Pi_I \Phi - \mathrm{Id}|_{C^0(B(I_*,\mu))} <\cdot T\mu^2$ such that

$$H \circ \Phi = h + g + \tilde{f}$$

with $\{g, l\} = 0$ and the estimates

$$|g + \tilde{f}|_{C^0(\mathbb{T}^n \times B(I_*,\mu))} <\cdot \mu^2, \quad |\partial_{\tilde{\theta}} \tilde{f}|_{C^0(\mathbb{T}^n \times B(I_*,\mu))} <\cdot (T\mu)^{k-2}\mu^2.$$

Let us write $\tilde{H} = H \circ \Phi$ and $(\tilde{\theta}, \tilde{I})$ are the new coordinates in $\mathbb{T}^n \times B(I_*, \mu)$.

Third step. Now we set

$$r =\cdot \mu, \quad \tau =\cdot (T\mu)^{k-2},$$

and we apply Proposition 4.2 to the Hamiltonian $\tilde{H}$. To do so, we need to impose condition (7) and this gives our last smallness condition on ε. Therefore we find

$$|\tilde{I}(t) - \tilde{I}_0| <\cdot \mu, \quad |t| <\cdot (T\mu)^{k-2},$$

and recalling that

$$\mu =\cdot T^{-1}\varepsilon^{\frac{1}{2d}} <\cdot \varepsilon^{\frac{1}{2d}}$$

this gives

$$|\tilde{I}(t) - \tilde{I}_0| <\cdot \varepsilon^{\frac{1}{2d}}, \quad |t| <\cdot \varepsilon^{\frac{k-2}{2d}}.$$

Now since

$$|\Pi_I \Phi - \mathrm{Id}|_{C^0(B(I_*,\mu))} <\cdot T\mu^2 <\cdot \varepsilon^{\frac{1}{2d}},$$

standard arguments give the conclusion

$$|I(t) - I_0| <\cdot \varepsilon^{\frac{1}{2d}}, \quad |t| <\cdot \varepsilon^{\frac{k-2}{2d}}.$$

This ends the proof. □

Acknowledgements. The author thanks Jean-Pierre Marco, Laurent Niederman and also H. Scott Dumas for a careful reading of a first version of the manuscript.

References

[Alb07] J. Albrecht, *On the existence of invariant tori in nearly-integrable Hamiltonian systems with finitely differentiable perturbations*, Regul. Chaotic Dyn. **12** (2007), no. 3, 281–320.

[AR67] R. Abraham and J. Robbin, *Transversal mappings and flows*, Benjamin, New-York, 1967.

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

[BM10] A Bounemoura and J.-P. Marco, *Improved exponential stability for near-integrable quasi-convex Hamiltonians*, Preprint (2010).

[BN09] A. Bounemoura and L. Niederman, *Generic Nekhoroshev theory without small divisors*, Preprint (2009).

[Bos86] J.-B. Bost, *Tores invariants des systemes dynamiques Hamiltoniens*, Séminaire Bourbaki **133-134** (1986), 113–157.

[DH09] A. Delshams and G. Huguet, *Geography of resonances and Arnold diffusion in a priori unstable Hamiltonian systems*, Nonlinearity **22** (2009), no. 8, 1997–2077.

[Féj04] J. Féjoz, *Démonstration du théorème d'Arnold sur la stabilité dy système planétaire (d'après Herman)*, Erg. Th. Dyn. Sys. **24** (2004), 1521–1582.

[Her86] M.-R. Herman, *Sur les courbes invariantes par les difféomorphismes de l'anneau*, Vol. 2. (French) With a correction to: On the curves invariant under diffeomorphisms of the annulus, Vol. 1 (French), Astérisque No. 144, 248 pp., 1986.

[KL08a] V. Kaloshin and M. Levi, *An example of Arnold diffusion for near-integrable Hamiltonians*, Bull. Amer. Math. Soc. (N.S.) **45** (2008), no. 3, 409–427.

[KL08b] _____, *Geometry of Arnold diffusion*, SIAM Rev. **50** (2008), no. 4, 702–720.

[Loc92] P. Lochak, *Canonical perturbation theory via simultaneous approximation.*, Russ. Math. Surv. **47** (1992), no. 6, 57–133.

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

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

[MS04] _____, *Wandering domains and random walks in Gevrey near-integrable systems*, Erg. Th. Dyn. Sys. **5** (2004), 1619–1666.

[Nek77] N.N. Nekhoroshev, *An exponential estimate of the time of stability of nearly integrable Hamiltonian systems*, Russian Math. Surveys **32** (1977), no. 6, 1–65.

[Nek79] _____, *An exponential estimate of the time of stability of nearly integrable Hamiltonian systems II*, Trudy Sem. Petrovs **5** (1979), 5–50.

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

[Pös82] J. Pöschel, *Integrability of Hamiltonian systems on Cantor sets*, Comm. Pure Appl. Math. **35** (1982), no. 5, 653–696.

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

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

[SZ89] D.A. Salamon and E. Zehnder, *KAM theory in configuration space*, Comment. Math. Helv. **64** (1989), no. 1, 84–132.