HAL
open science

Improved exponential stability for near-integrable quasi-convex Hamiltonians

Abed Bounemoura, Jean-Pierre Marco

To cite this version:

Abed Bounemoura, Jean-Pierre Marco. Improved exponential stability for near-integrable quasi-convex Hamiltonians. Nonlinearity, 2011, 24 (1), pp.97-112. hal-00470295v2

HAL Id: hal-00470295

https://hal.science/hal-00470295v2

Submitted on 7 Nov 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.

Improved exponential stability for near-integrable quasi-convex Hamiltonians

Abed Bounemoura * and Jean-Pierre Marco †

November 7, 2010

Abstract

In this article, we improve previous results on exponential stability for analytic and Gevrey perturbations of quasi-convex integrable Hamiltonian systems. In particular, this provides a sharper lower bound on the time of Arnold diffusion which we believe to be optimal.

1 Introduction

This paper deals with some 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, of size ε in some suitable topology defined by a norm $|\,.\,|$. If the system is analytic and satisfies a suitable quantitative transversality condition called *steepness*, Nekhoroshev ([Nek77], [Nek79]) proved that the action variables $I(t)$ are stable for an exponentially long time, in the sense that for ε sufficiently small, one has

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

for any initial action I_0. The constants c_1, c_2, c_3, a and b depend only h; one calls a and b the *stability exponents* and among them the value of a is of course the most important, as it specifies the time scale of stability. Nekhoroshev's estimates complement the well-known KAM theory ([Kol54],

*Laboratoire de Mathématiques d'Orsay et Institut de Mathématiques de Jussieu
†Institut de Mathématiques de Jussieu

see also [Pös01] for a nice survey) which gives, under some mild non degeneracy condition on h and for ε small enough, the existence of a constant c such that

$$|I(t) - I_0| \leq c\sqrt{\varepsilon}$$

for any time $t \in \mathbb{R}$, but only for a strict subset (of large relative measure) of the set of initial conditions. Of course KAM theory gives much more information; these stable solutions are in fact quasi-periodic and $\sqrt{\varepsilon}$-close to the corresponding unperturbed solutions. In particular, for $n = 2$ and in the case when h is isoenergetically non-degenerate, this stability property even holds for all solutions. On the contrary, for $n \geq 3$, following Arnold ([Arn64]) one can find examples of near-integrable Hamiltonian systems with a solution satisfying

$$|I(\tau) - I_0| \geq 1$$

for some time $\tau = \tau(\varepsilon) > 0$, no matter how small the perturbation is. Such instability is commonly referred to as Arnold diffusion. Obviously Nekhoroshev's estimates give a lower bound on the diffusion time $\tau(\varepsilon)$ (or equivalently an upper bound on the diffusion speed) which is exponentially large (or exponentially small when referring to the rate of diffusion).

This paper is concerned with the precise time scale at which stability breaks down and instability takes place, by which we mean the precise value of the exponent a, in the special case where the unperturbed Hamiltonian h is quasi-convex (that is, convex when restricted to its energy sub-levels).

The quasi-convex case, which is of both practical and theoretical interest, has been widely studied in Nekhoroshev theory, essentially for two reasons. First, the proof is much easier in this situation and a more refined result is available: the stability exponents may be chosen as

$$a = b = (2n)^{-1}.$$

These facts are best illustrated by the striking proof given by Lochak ([Loc92], see also [LN92] and [LNN94]), though they are also accessible via a more traditional approach as was shown by Pöschel ([Pös93]). Note also that these values have been generalized by Niederman in the steep case ([Nie04]) using both ideas of Lochak and Pöschel. Yet there is another reason for which the quasi-convex case is interesting, which is the so-called *stabilization by resonances*: if a solution starts close to a resonance of multiplicity m, $m < n$, then it possesses better stability properties, described by the "local" exponents

$$a_m = b_m = (2(n-m))^{-1}.$$

This is a quite surprising fact, as it shows that even though resonances are the main cause of diffusion, at the same time they improve finite time stability. However, this property certainly does not hold without some convexity assumption (in the steep case for instance).

The optimality of the exponent a, in connection with the minimal time of instability, was first studied by Bessi who introduced powerful variational methods to revisit Arnold's example and estimate the time of diffusion. In [Bes96] and [Bes97], he proved that the latter is of order $\exp\left(\varepsilon^{-\frac{1}{2}}\right)$ for $n = 3$ and $\exp\left(\varepsilon^{-\frac{1}{4}}\right)$ for $n = 4$. Moreover, in Bessi's example the solution passes close to a double resonance, and so the time is the best possible in this case, in view of the values of the local exponent a_2 for $n = 3$ and $n = 4$. Recently, using similar variational arguments, this result has been generalized to an arbitrary number of degrees of freedom n by Ke Zhang ([Zha09]), namely he constructed a special orbit passing close to a double resonance for which the time of diffusion is estimated by $\exp\left(\varepsilon^{-\frac{1}{2(n-2)}}\right)$.

Another approach has been proposed by Marco-Sauzin ([MS02]) and Lochak-Marco ([LM05]), following novel ideas of Herman. In [MS02] the authors show that Nekhoroshev's estimates extend to perturbations of quasi-convex Hamiltonians which are α-Gevrey regular, $\alpha \geq 1$, with the exponents

$$a = (2\alpha n)^{-1}, \quad b = (2n)^{-1}$$

and local exponents

$$a_m = (2\alpha(n-m))^{-1}, \quad b_m = (2(n-m))^{-1}.$$

Note that 1-Gevrey functions are exactly analytic functions, and basically when α ranges from one to infinity α-Gevrey functions interpolate between analytic and C^∞ functions. Therefore this result generalizes the estimates in the analytic case. Using a geometric mechanism different and more precise than Arnold's, in [MS02] the authors constructed a drifting orbit with time of order $\exp\left(\varepsilon^{-\frac{1}{2\alpha(n-2)}}\right)$ in the non-analytic case, that is when $\alpha > 1$. Adding some more technical ideas, it is shown in [LM05] that the example also works in the analytic case, but the time is estimated as $\exp\left(\varepsilon^{-\frac{1}{2(n-3)}}\right)$, which is only close to optimal (however refinements are certainly possible to reach the value $(2(n-2))^{-1}$ in this class of examples).

Therefore, if the unperturbed Hamiltonian is quasi-convex, the best exponent of stability a up to now satisfies

$$(2n)^{-1} \leq a < (2(n-2))^{-1}$$

in the analytic case and more generally

$$(2\alpha n)^{-1} \leq a < (2\alpha(n-2))^{-1}$$

in the Gevrey case. The goal of this paper is to improve the lower bound both in the analytic or Gevrey case, so as to have

$$(2(n-1))^{-1} - \delta \leq a < (2(n-2))^{-1}$$

and

$$(2\alpha(n-1))^{-1} - \delta \leq a < (2\alpha(n-2))^{-1}$$

for $\delta > 0$ but arbitrarily small (see Theorem 2.1 and Theorem 2.2 in the next section). We believe that this bound is optimal, in the sense that one could reach the value $(2(n-1))$ in Arnold diffusion, using a significantly different mechanism of instability.

2 Main results

1. In order to state our main results, let us now describe our setting more precisely, beginning with the analytic case. Let $B = B(0, R)$ be the open ball of $\mathbb{R}^n$ of radius $R > 0$, with respect to the supremum norm $|\,.\,|$, centered at the origin. Given $s > 0$, we let $\mathcal{A}_s(\mathcal{D})$ be the space of bounded real-analytic functions on $\mathcal{D} = \mathbb{T}^n \times B$ which extend as holomorphic functions on the complex domain

$$\mathcal{D}_s = \{(\theta, I) \in (\mathbb{C}^n/\mathbb{Z}^n) \times \mathbb{C}^n \mid |\mathcal{I}(\theta)| < s,\ d(I, B) < s\},$$

and which are continuous on the closure of $\mathcal{D}_s$. Here we denoted by $\mathcal{I}(\theta)$ the imaginary part of θ, by $|\,.\,|$ the supremum norm on $\mathbb{C}^n$, and by d the associated distance on $\mathbb{C}^n$. It is well-known that $\mathcal{A}_s(\mathcal{D})$ is a Banach space with its usual supremum norm $|\,.\,|_s$, where

$$|f|_s = \sup_{z \in \mathcal{D}_s} |f(z)|, \quad f \in \mathcal{A}_s(\mathcal{D}).$$

In the following, we shall denote by

$$B_s = \{I \in \mathbb{R}^n \mid d(I, B) < s\}$$

the real part of our domain $\mathcal{D}_s$ in action space. The geometric parameters n, R, s are assumed to be chosen once and for all in the following.

We now introduce the parameters related to the choice of the system. The integrable part $h : B_s \to \mathbb{R}$ will be assumed to be strictly quasi-convex: the gradient map ∇h does not vanish and there exists a positive number m such that

$$\nabla^2 h(I)v.v \geq m|v|^2 \qquad (QC(m))$$

holds for any $I \in B_s$ and any v orthogonal to $\nabla h(I)$ (with respect to the Euclidean scalar product). Moreover, the derivatives up to order 3 of h on B_s are assumed to be bounded: there exists $M > 0$ such that for all $I \in B_s$, one has

$$|\partial^k h(I)| \leq M, \quad 1 \leq |k_1| + \cdots + |k_n| \leq 3. \qquad (B(M))$$

Therefore we will consider systems of the form

$$\begin{cases} H(\theta, I) = h(I) + f(\theta, I), \\ h \in \mathcal{A}_s(\mathcal{D}),\ f \in \mathcal{A}_s(\mathcal{D}), \\ h \text{ satisfies } (QC(m)) \text{ and } (B(M)), \\ |f|_s < \varepsilon. \end{cases} \tag{$C(M, m, \varepsilon)$}$$

Note that we suppress of the geometric parameters in the notation. In the following we will call a *stable constant* (in the analytic case) any positive constant c which depends on the whole set of parameters, that is n, R, s, M, m, together with a parameter δ or ρ to be defined below, but not on a particular choice of H satisfying the condition $(C(M, m, \varepsilon))$.

2. The main result of the paper is the following.

Theorem 2.1. *Consider a real number δ satisfying*

$$0 < \delta \leq (2n(n-1))^{-1}.$$

Then there exist stable constants c_1, c_2, c_3 and ε_0 such that if $0 \leq \varepsilon \leq \varepsilon_0$, and if H satisfies $(C(M, m, \varepsilon))$, the following estimates

$$|I(t) - I_0| \leq c_1 \varepsilon^{\delta(n-1)}, \quad |t| \leq c_2 \exp\left(c_3 \varepsilon^{-\frac{1}{2(n-1)} + \delta}\right)$$

hold true for every initial action $I_0 \in B(0, R/2)$.

Moreover, consider a real number ρ satisfying $0 < \rho < R/2$. Then there exist stable constants c_4, c_5 and $\tilde{\varepsilon}_0$ such that if $0 \leq \varepsilon \leq \tilde{\varepsilon}_0$, then

$$|I(t) - I_0| \leq \rho, \quad |t| \leq c_4 \exp\left(c_5 \varepsilon^{-\frac{1}{2(n-1)}}\right)$$

for every $I_0 \in B(0, R/2)$.

Choosing our constant δ arbitrarily close to zero, our result ensures stability for a time scale which is arbitrarily close to $\exp\left(\varepsilon^{-\frac{1}{2(n-1)}}\right)$, therefore we improve the previous results of stability obtained independently by Lochak-Neishtadt ([Loc92] and [LN92]) and Pöschel ([Pös93]), which were believed to be optimal.

In fact in the extreme case where $\delta = (2n(n-1))^{-1}$, which in our situation gives the worst stability time (but of course the best radius of confinement), our result reads

$$|I(t) - I_0| \leq c_1 \varepsilon^{\frac{1}{2n}}, \quad |t| \leq c_2 \exp\left(c_3 \varepsilon^{-\frac{1}{2n}}\right)$$

and we recover the previous result of stability. Hence, when our parameter δ ranges from $(2n(n-1))^{-1}$ to zero, our theorem "interpolates" between previous stability results and what should be the optimal stability.

Indeed, in the other extreme case which corresponds to the second part of our theorem, our result does not give stability, since the radius of confinement can be arbitrarily small but no longer tends to 0 with ε. We believe that this is not an artefact of the method and that instability should occur at this precise time scale, at a time of order $\exp\left(\varepsilon^{-\frac{1}{2(n-1)}}\right)$. We plan to construct an example with an unstable orbit which has a drift of order one during such a time interval. This necessitates the use of more refined instability mechanism in the neighbourhood of double resonances, a topic which is also crucial in connection with the problem of genericity of Arnold diffusion.

3. Our result also holds if the Hamiltonian is only Gevrey regular. Let us recall that given $\alpha \geq 1$ and $L > 0$, a function $H \in C^\infty(\mathcal{D})$ is (α, L)-Gevrey if, using the standard multi-index notation, we have

$$|H|_{\alpha,L} = \sum_{l \in \mathbb{N}^{2n}} L^{|l|\alpha}(l!)^{-\alpha}|\partial^l H|_{\mathcal{D}} < \infty$$

where $|\,.\,|_{\mathcal{D}}$ is the usual supremum norm for functions on $\mathcal{D}$. The space of such functions, with the above norm, is a Banach space that we denote by $G^{\alpha,L}(\mathcal{D})$. Analytic functions are a particular case of Gevrey functions, as one can check that $G^{1,L}(\mathcal{D}) = \mathcal{A}_L(\mathcal{D})$.

Let us introduce the main condition on the Hamiltonian systems in the Gevrey case

$$\begin{cases} H(\theta, I) = h(I) + f(\theta, I), \\ h \in G^{\alpha,L}(\mathcal{D}),\ f \in G^{\alpha,L}(\mathcal{D}), \\ h \text{ satisfies } (QC(m)) \text{ and } (B(M)), \\ |f|_{\alpha,L} < \varepsilon. \end{cases} \qquad (C(\alpha, L, M, m, \varepsilon))$$

We now call a *stable constant* (in the Gevrey case) any positive constant c which depends on the whole set of parameters, that is α, L, n, R, M, m, together with a parameter δ or ρ which will be defined below, but not on a particular choice of H satisfying the condition $(C(\alpha, L, M, m, \varepsilon))$.

4. Our second result is the following Gevrey version of Theorem 2.1.

Theorem 2.2. *Consider a real number δ satisfying*

$$0 < \delta \leq (2\alpha n(n-1))^{-1}.$$

Then there exist stable constants c_1', c_2', c_3' and ε_0' such that if $0 \leq \varepsilon \leq \varepsilon_0'$, and if H satisfies $C(\alpha, L, M, m, \varepsilon)$, the following estimates

$$|I(t) - I_0| \leq c_1' \varepsilon^{\frac{2\delta}{5(n-1)}}, \quad |t| \leq c_2' \exp\left(c_3' \varepsilon^{-\frac{1}{2\alpha(n-1)}+\delta}\right),$$

hold true for every initial action $I_0 \in B(0, R/2)$.

Moreover, consider a real number ρ *satisfying* $0 < \rho < R/2$. *Then there exist stable constants* c_4', c_5' *and* $\tilde{\varepsilon}_0'$ *such that if* $0 \leq \varepsilon \leq \tilde{\varepsilon}_0'$, *then*

$$|I(t) - I_0| \leq \rho, \quad |t| \leq c_4' \exp\left(c_5' \varepsilon^{-\frac{1}{2\alpha(n-1)}}\right)$$

for every $I_0 \in B(0, R/2)$.

The same remarks as above apply in the Gevrey case. In particular δ can be chosen arbitrarily close to zero and our result ensures stability for an interval of time which is arbitrarily close to $\exp\left(\varepsilon^{-\frac{1}{2\alpha(n-1)}}\right)$. However, our radius of stability is worse than in the analytic case, so we do not fully recover the result obtained in [MS02], but of course the time of stability is the most important issue.

5. To avoid cumbersome expressions in the following, when there is no risk of confusion we will replace the stable constants with a dot. More precisely, an assertion of the form "there exists a stable constant c such that $f < c\,g$" will be simply replaced with "$f \lessdot g$", when the context is clear. *Such modifications will only concern assertions stating the existence of stable constants, and dealing with equalities or (strict or large) inequalities involving well-defined functions.*

6. In the rest of the paper, as usual and without loss of generality, we will only consider solutions starting at time $t = 0$ and evolving in positive time, and initial conditions will be denoted by $(I_0, \theta_0) = (I(0), \theta(0))$.

7. Finally, in the rest of this text all the norms will be denoted by $|\,.\,|$; this will always be the supremum norm for vectors and the induced norm for matrices, except for vectors in $\mathbb{Z}^n$ for which $|\,.\,|$ will stand for the ℓ^1–norm.

3 The analytic case

In this section, the geometric constants n, s, R together with the constants m and M are fixed once and for all.

The proof of Theorem 2.1 relies on two elementary facts. The first one, which we recalled in the introduction, is the stabilizing effect of resonances: on account of quasi-convexity, solutions of any Hamiltonian satisfying $(C(M, m, \varepsilon))$ and starting sufficiently close to a resonance are stable for a longer interval of time. Of course this concerns only some special solutions, but we need to consider all of them. Our second remark is that, to deal with the remaining solutions, one can take advantage of the geometry of resonances in the integrable system to obtain a confinement result. More precisely, using the isoenergetic non-degeneracy implied by our quasi-convexity assumption, we will show that solutions that avoid all resonances are necessarily stable for all time.

1. Let us begin by making the first point explicit, using Pöschel's approach to Nekhoroshev's theory. We shall denote by $\Omega = \nabla h(B)$ the space of frequencies. Let Λ be a submodule of $\mathbb{Z}^n$ of rank r, with $r \in \{1, \dots, n\}$. The resonant space associated with Λ is defined by

$$R_\Lambda = \{\omega \in \Omega \mid k.\omega = 0,\ \forall k \in \Lambda\}.$$

Given a real number $K \geq 1$, we will say that Λ is a K-submodule if it admits a $\mathbb{Z}$-basis $\{k^1, \dots, k^r\}$ satisfying $|k^i| \leq K$ for $i \in \{1, \dots, r\}$, where

$$|k^i| = |k^i_1| + \cdots + |k^i_n|.$$

As usual, it is enough to consider only maximal submodules, which are those that are not strictly contained in any other submodule of the same rank. Given such a submodule, we define its volume by

$$|\Lambda| = \sqrt{\det {}^tMM}$$

where M is any $n \times r$ matrix whose columns form a basis for Λ (this is easily seen to be independent of the choice of such a matrix). The following stability theorem is due to Pöschel.

Theorem 3.1 (Pöschel)**.** *Let Λ be a K-submodule of $\mathbb{Z}^n$ of rank r, with $r \in \{0, \dots, n-1\}$. Assume that $\varepsilon \geq 0$ and $K \geq 1$ satisfy*

$$\varepsilon |\Lambda|^2 K^{2(n-r)} \lessdot 1$$

and H satisfies $(C(M, m, \varepsilon))$. Then, if $\omega(0) = \nabla h(I_0)$, for any solution $(\theta(t), I(t))$ with $I_0 \in B(0, R/2)$ and $d(\omega(0), R_\Lambda) \lessdot \sqrt{\varepsilon}$ the following estimates

$$|I(t) - I_0| \lessdot \left(\varepsilon |\Lambda|^2\right)^{\frac{1}{2(n-r)}}, \quad t \lessdot \exp\left(\cdot \varepsilon |\Lambda|^2\right)^{-\frac{1}{2(n-r)}},$$

hold true.

As we recalled in the introduction, we use a dot in the various inequalities to abbreviate an assertion such as "there exists a stable constant c_i such that", located at the beginning of the statement.

The previous theorem gives exactly the content of Theorem 3 in [Pös93], to which we refer for a possible choice of stable constants. Note that Pöschel uses a more quantitative version of quasi-convexity, namely that there exist two positive numbers l and m such that at least one of the inequalities

$$|\nabla h(I).v| > l|v|, \quad \nabla^2 h(I)v.v \geq m|v|^2,$$

holds for any $I \in B_s$ and any $v \in \mathbb{R}^n$. It turns out that this notion is in fact equivalent to our condition $(QC(m))$.

We shall only need Pöschel's result in the special case where the K-submodule Λ has rank 1 and where the solution starts precisely *on* the associated resonant manifold. So we let

$$R_K = \bigcup_{\Lambda} R_\Lambda$$

where the union is taken over all (maximal) K-submodules of rank 1, so it is the set of simply resonant frequencies of order K. Moreover, the "volume" of a rank-one submodule $|\Lambda|$ is nothing but the Euclidean norm of one of its $\mathbb{Z}$-generators (since Λ has a rank equals to one, it has just two generators that differ only by their sign), so one gets the trivial estimate

$$1 \leq |\Lambda|^2 \leq K^2$$

from which we deduce the following lemma.

Lemma 3.2. *Let Λ be a K-submodule of $\mathbb{Z}^n$ of rank 1. Assume that $\varepsilon \geq 0$ and $K \geq 1$ satisfy*

$$\varepsilon K^{2n} \lessdot 1 \tag{1}$$

and H satisfies $(C(M, m, \varepsilon))$. Then for any solution $(\theta(t), I(t))$ such that $I_0 \in B(0, R/2)$ and $\omega(0) \in R_K$ the following estimates

$$|I(t) - I_0| \lessdot \left(\varepsilon K^2\right)^{\frac{1}{2(n-1)}}, \quad t \lessdot \exp\left(\cdot \varepsilon K^2\right)^{-\frac{1}{2(n-1)}},$$

hold true.

2. Let us now turn to our second remark, which is a simple geometric property of the integrable system based on the isoenergetic non-degeneracy. The latter condition is known to be implied by quasi-convexity, and it can be interpreted in various ways (see ([Sev06]) and references therein), but for subsequent arguments we will adopt the following form.

Lemma 3.3. *Suppose h satisfies $(QC(m))$. Then the map*

$$\begin{array}{rccc} \Psi_h & : & B \times \mathbb{R}_*^+ & \longrightarrow & \mathbb{R} \times (\mathbb{R}^n \setminus \{0\}) \\ & & (I, \lambda) & \longmapsto & (h(I), \lambda\omega(I)). \end{array}$$

is a local diffeomorphism in the neighbourhood of any point $(I_0, \lambda_0) \in B \times \mathbb{R}_*^+$.

In particular, there exist stable constants ρ_0 and C such that if $\rho < \rho_0$, then Ψ_h is a diffeomorphism restricted to the ball $B((I_0, \lambda_0), \rho) \subseteq B \times \mathbb{R}_*^+$, whose image contains the closed ball $\overline{B}(\Psi_h(I_0), C\rho) \subseteq \mathbb{R} \times \mathbb{R}^n$.

The second statement is a consequence of the first one (see [MS02] Lemma 3.17 for quantitative estimates on the stable constants ρ_0 and C), and the first statement is well-known, but we give a proof for convenience.

Proof. By the inverse function theorem, it is enough to prove that $d_{(I_0,\lambda_0)}\Psi_h$ is non singular at any point $(I_0,\lambda_0) \in B \times \mathbb{R}_*^+$. Given $u \in \mathbb{R}$, $v \in \mathbb{R}^n$, we easily compute

$$d_{(I_0,\lambda_0)}\Psi_h(v,u) = (\omega(I_0).v, u\omega(I_0) + \lambda_0 \nabla^2 h(I_0)v)$$

and we need to show that this vector is non zero if the vector $(v,u) \in \mathbb{R}^{n+1}$ is non zero. If either $\omega(I_0).v \neq 0$, in which case the first component is non zero, or $v = 0$ and hence the second component is non zero (since $u \neq 0$ and so $u\omega(I_0) \neq 0$), the statement is obvious. Otherwise, $\omega(I_0).v = 0$ and $v \neq 0$, since h satisfies $(QC(m))$ this gives

$$\begin{aligned}(u\omega(I_0) + \lambda_0\nabla^2 h(I_0)v).v &= \lambda_0 \nabla^2 h(I_0)v.v \\ &\geq \lambda_0 m|v|^2\end{aligned}$$

and therefore $u\omega(I_0) + \lambda_0\nabla^2 h(I_0)v \neq 0$. □

3. We can now take one step further in the dynamical consequences of the structure of simple resonances. We consider a Hamiltonian H satisfying $(C(M,m,\varepsilon))$. We will actually focus on those simple resonances for which the ratio of two frequencies becomes rational. The following elementary lemma will allow us to deal with this simple case.

Lemma 3.4. *Let I be a closed interval of length $l > 0$ contained in $[-1,1]$. Then there exists a rational number $p/q \in I \cap \mathbb{Q}$ satisfying*

$$|q| + |p| < 6l^{-1}.$$

Proof. Let us write $I = [x - l/2, x + l/2]$ for some suitable $x \in [-1,1]$, and let q be the smallest integer larger than l^{-1}, that is

$$l^{-1} \leq q < l^{-1} + 1.$$

We claim that there always exists an integer $p \in \mathbb{Z}$ such that

$$|x - p/q| \leq (2q)^{-1}.$$

Indeed, let $[qx] \in \mathbb{Z}$ be the integer part of qx. Then either $qx - [qx] \leq 2^{-1}$, in which case we can choose $p = [qx]$, or $2^{-1} < qx - [qx] < 1$, and then one can set $p = [qx] + 1$. Since $l^{-1} \leq q$, we have $q^{-1} \leq l$, and so

$$|x - p/q| \leq (2q)^{-1} \leq l/2$$

which means that $p/q \in I$. Moreover, as $I \subset [-1,1]$, then $|p| \leq q$ and

$$|q| + |p| \leq 2q.$$

Recall that $q < l^{-1}+1$, but since $I \subseteq [-1,1]$, we have $l \leq 2$, so that $1 \leq 2l^{-1}$ and therefore

$$q < 3l^{-1}.$$

This gives

$$|q| + |p| < 6l^{-1}$$

which concludes the proof. □

The following result is our main lemma. It essentially says that a (long) drifting orbit has to cross a simple resonance, since all other orbits are stable on the interval of time over which they are defined.

Lemma 3.5. *Consider $\varepsilon \geq 0$ and $K \geq 1$ such that*

$$K^{-1} \lessdot 1, \quad \varepsilon K \lessdot 1. \tag{2}$$

Let H be a Hamiltonian satisfying ($C(M,m,\varepsilon)$), $\tau \in \mathbb{R}^+ \cup \{+\infty\}$ and let $(\theta(t), I(t))$ be a solution defined on $[0,\tau[$ with $I_0 \in B(0,R/2)$. If

$$\omega(t) \notin R_K, \quad 0 \leq t < \tau,$$

then the inequality

$$|I(t) - I_0| \lessdot K^{-1}, \quad 0 \leq t < \tau$$

holds true.

Let us mention that if τ is the maximal time of existence of the solution within the initial domain $\mathbb{T}^n \times B$, then our stability estimate easily ensures that $\tau = +\infty$, which in turn implies stability for all time (see the proof of Theorem 2.1).

Proof. We will make conditions (2) explicit and prove that

$$|I(t) - I_0| < 6\, C^{-1}K^{-1}, \quad 0 \leq t < \tau,$$

when

$$K^{-1} < 6^{-1}C\rho_0, \quad \varepsilon K < 3,$$

where ρ_0 and C are the stable constants of lemma 3.3.

We will argue by contradiction, so we assume that there exists a time $\tilde{t}$ for which

$$|I(\tilde{t}) - I_0| \geq 6\, C^{-1}K^{-1}.$$

Consider the curve

$$\sigma(t) = (I(t), |\omega(t)|^{-1}) \in B \times \mathbb{R}^+$$

and let $\rho = 6\, C^{-1} K^{-1}$. Then

$$t^* = \inf\{t \in [0, \tau[\mid \sigma(t) \notin B(\sigma(0), \rho)\}$$

is well-defined, as the above set contains $\tilde{t}$. Now $\rho < \rho_0$, so we can apply Lemma 3.3: the restriction of Ψ_h to the open ball $B(\sigma(0), \rho)$ is a diffeomorphism whose image contains the closed ball $\overline{B}(\Psi_h(\sigma(0)), 6K^{-1})$. Considering a slightly larger ball over which Ψ_h remains a diffeomorphism, this easily implies that

$$\Psi_h(\sigma(t^*)) \notin B(\Psi_h(\sigma(0)), 6\, K^{-1}),$$

that is,

$$\big|\big(h(I(t^*)), |\omega(t^*)|^{-1}\omega(t^*)\big) - \big(h(I_0), |\omega(0)|^{-1}\omega(0)\big)\big| \geq 6\, K^{-1}.$$

Using the conservation of energy, one has

$$|h(I(t^*)) - h(I_0)| < 2\varepsilon < 6\, K^{-1}$$

so that necessarily

$$\big||\omega(t^*)|^{-1}\omega(t^*) - |\omega(0)|^{-1}\omega(0)\big| \geq 6\, K^{-1}.$$

Therefore there exists an index $i \in \{1, \ldots, n\}$ such that

$$\left|\frac{\omega_i(t^*)}{|\omega(t^*)|} - \frac{\omega_i(0)}{|\omega(0)|}\right| \geq 6\, K^{-1}$$

and this estimate means that the image of the interval $[0, t^*]$ under the continuous function

$$t \longmapsto \frac{\omega_i(t)}{|\omega(t)|} \in [-1, 1]$$

contains a non trivial interval I of length $l = 6\, K^{-1}$. Now we can apply Lemma 3.4 to find a rational number $p/q \in \mathbb{Q}$ in reduced form and a time $t' \in [0, t^*]$ such that

$$\frac{\omega_i(t')}{|\omega(t')|} = \frac{p}{q} \tag{3}$$

with

$$|p| + |q| < 6(6\, K^{-1})^{-1} = K. \tag{4}$$

But since $|\omega(t')| = |\omega_j(t')|$ for some $j \in \{1, \ldots, n\}$, and replacing p with $-p$ if $\omega_j(t')$ is negative, the equality (3) can be written as

$$q\omega_i(t') - p\omega_j(t') = 0. \tag{5}$$

Now let us write $k' = qe_i - pe_j \in \mathbb{Z}$, then from (4) and (5) we have

$$k'.\omega(t') = 0, \quad |k'| < K,$$

and since the submodule generated by k' is maximal, since p and q are co-prime, we find that $\omega(t') \in R_K$. This gives the desired contradiction. $\square$

4. We finally arrive to the proof of Theorem 2.1.

Proof of Theorem 2.1. We choose K of the form

$$K = K_0 \left(\frac{\varepsilon_0}{\varepsilon}\right)^{\gamma}$$

with suitable stable constants K_0 and ε_0 so that conditions (1) and (2) are satisfied if

$$0 \leq \varepsilon \leq \varepsilon_0, \quad 0 < \gamma \leq (2n)^{-1}.$$

With this threshold and these bounds on γ, both Lemma 3.2 and Lemma 3.5 can be applied. Let $(\theta_0, I_0) \in \mathbb{T}^n \times B(0, R/2)$ and T be the maximal time of existence within $\mathbb{T}^n \times B(0, R)$ of the solution $(\theta(t), I(t))$ starting at (θ_0, I_0). We have to distinguish two cases.

In the first case, we assume that $\omega(t) \in N_K$ for all $t < T$. Then we apply Lemma 3.5 with $\tau = T$ to get

$$|I(t) - I_0| \lessdot \varepsilon^{\gamma}, \quad 0 \leq t < T.$$

From this estimate we can deduce that the solution $I(t)$ belongs to some compact ball around I_0 which, taking ε_0 small enough (this is possible since $\gamma > 0$), is included in $B(0, R)$. Therefore this solution is defined for all time, that is $T = +\infty$, and so the previous estimate gives

$$|I(t) - I_0| \lessdot \varepsilon^{\gamma}, \quad 0 \leq t < T = +\infty.$$

In the second case, there exists a smallest time $0 \leq t^* < T$ such that $\omega(t^*)$ belongs to the set R_K. We apply once again Lemma 3.5 with $\tau = t^*$ so that

$$|I(t) - I_0| \lessdot \varepsilon^{\gamma}, \quad 0 \leq t \leq t^*.$$

Again, taking ε_0 small enough we can ensure that $I(t^*) \in B(0, R/2)$, then we can apply Lemma 3.2 to the solution $I_{t^*}(t) = I(t + t^*)$, whose initial frequency belongs to R_K, to obtain

$$|I_{t^*}(t) - I_{t^*}(0)| \lessdot \left(\varepsilon K^2\right)^{\frac{1}{2(n-1)}}, \quad 0 \leq t \lessdot \exp\left(\cdot \varepsilon K^2\right)^{-\frac{1}{2(n-1)}}.$$

Setting

$$a_\gamma = \frac{1 - 2\gamma}{2(n-1)}$$

this gives

$$|I_{t^*}(t) - I_{t^*}(0)| \lessdot \varepsilon^{a_\gamma}, \quad 0 \leq t \lessdot \exp(\cdot \varepsilon^{-a_\gamma}).$$

Since $t^* \geq 0$, we get in particular

$$|I(t) - I(t^*)| \lessdot \varepsilon^{a_\gamma}, \quad t^* \leq t \lessdot \exp(\cdot \varepsilon^{-a_\gamma}),$$

and we conclude that

$$|I(t) - I_0| \lessdot (\varepsilon^\gamma + \varepsilon^{a_\gamma}) \lessdot \max\{\varepsilon^\gamma, \varepsilon^{a_\gamma}\}, \quad 0 \le t \lessdot \exp(\cdot\varepsilon^{-a_\gamma}).$$

Now, using the condition $0 < \gamma \le (2n)^{-1}$ one has

$$(2n)^{-1} \le a_\gamma < (2(n-1))^{-1}$$

in particular

$$\gamma \le a_\gamma$$

and hence

$$\max\{\varepsilon^\gamma, \varepsilon^{a_\gamma}\} = \varepsilon^\gamma.$$

Therefore for all solutions, the estimates

$$|I(t) - I_0| \lessdot \varepsilon^\gamma, \quad 0 \le t \lessdot \exp(\cdot\varepsilon^{-a_\gamma}),$$

hold true.

Finally, to obtain our statement, just set

$$\delta = \gamma(n-1)^{-1}$$

so that

$$a_\gamma = (2(n-1))^{-1} - \delta$$

hence

$$|I(t) - I_0| \lessdot \varepsilon^{\delta(n-1)}, \quad 0 \le t \lessdot \exp\left(\cdot\varepsilon^{-(2(n-1))^{-1}+\delta}\right)$$

provided that

$$0 < \delta \le (2n(n-1))^{-1}.$$

This ends the proof of the first part of the statement. For the second part, choosing K in terms of ρ but independent of ε, the proof is similar and even simpler, so we do not repeat the details. □

4 The Gevrey case

In this section, we will prove Theorem 2.2. In fact, it will be enough to have a version of Lemma 3.2 in the Gevrey case, as the geometric considerations of the previous section still apply with no changes.

Here we shall use a result from [MS02], which follows the method introduced by Lochak ([Loc92]) from which the results of improved stability near resonances actually originate. In the latter approach, the notion of "order" of a resonance is more intrinsic, however it is also more difficult to compute.

Let Λ be a submodule of $\mathbb{Z}^n$ of rank r, with $r \in \{1, \dots, n\}$, and choose a basis $\{k^1, \dots, k^r\} \in \mathbb{Z}^n$ for Λ. We define the matrix L of size $r \times n$

with integer entries whose rows are given by the vectors $k^i = (k_1^i, \dots, k_n^i)$, $1 \le i \le r$, that is

$$L = \begin{pmatrix} k_1^1 & \cdots & k_n^1 \\ \vdots & & \vdots \\ k_1^r & \cdots & k_n^r \end{pmatrix} \in M_{r,n}(\mathbb{Z}).$$

Then it is an elementary result of linear algebra that there exists integers $d_1, \dots, d_r \in \mathbb{Z}$, satisfying the divisibility conditions $d_1 | \dots | d_r$, such that L is equivalent to the diagonal matrix

$$\Delta = \begin{pmatrix} d_1 & & 0 & 0 & \cdots & 0 \\ & \ddots & & & \vdots & \\ 0 & & d_r & 0 & \cdots & 0 \end{pmatrix} \in M_{r,n}(\mathbb{Z}).$$

Therefore one can write

$$L = B \Delta A \tag{6}$$

for some matrices $A \in GL(n, \mathbb{Z})$ and $B \in GL(r, \mathbb{Z})$.

The numbers d_i are called the invariant factors of the module, and for a maximal module one can show that these numbers are all equal to one. The above normal form result can be proved equivalently by elementary operations on rows and columns or using the structure of finitely generated modules over a principal domain.

One can easily check that ${}^tA$ sends the standard submodule (which is the one generated by the first r vectors of the canonical basis of $\mathbb{Z}^n$) to the submodule Λ. So quantitative information about the submodule is encoded in those matrices $A \in GL(n, \mathbb{Z})$.

Following Lochak, we define c_Λ (resp. c'_Λ) as the minimal value of the norm $|A^{-1}|$ (resp. of $|A|$) among all matrices $A \in GL(n, \mathbb{Z})$ satisfying the relation (6) (it is easy to see that those constants depend only on Λ and not on the choice of such a matrix). In the space $M_n(\mathbb{Z})$ we may choose the norm $|\,.\,|$ induced by the usual supremum norm for vectors, which is nothing but the maximum of the sums of the absolute values of the elements in each row.

With these definitions, one can state the following stability result in the Gevrey class.

Theorem 4.1 (Marco-Sauzin)**.** *Let Λ be a K-submodule of $\mathbb{Z}^n$ of rank r, with $r \in \{0, \dots, n-1\}$. Assume that $\varepsilon \ge 0$ satisfies*

$$\varepsilon c_\Lambda^{5(n-r)} \lessdot 1, \quad \varepsilon c_\Lambda^{3(n-r)} c_\Lambda'^{2(n-r)} \lessdot 1,$$

and H satisfies $(C(\alpha, L, M, m, \varepsilon))$. Then for any solution $(\theta(t), I(t))$ with $I_0 \in B(0, R/2)$ and $d(\omega(0), R_\Lambda) \lessdot \sqrt{\varepsilon}$ the following estimates

$$|I(t) - I_0| \lessdot c_\Lambda^{3/2} c'_\Lambda \varepsilon^{\frac{1}{2(n-r)}}, \quad t \lessdot \exp\left(\cdot c_\Lambda^{5(n-r)} \varepsilon\right)^{-\frac{1}{2\alpha(n-r)}}$$

hold true.

This is exactly the addendum to Theorem A in [MS02], to which we refer for a possible choice of stable constants. Note also that in [MS02] the names of these constants are a bit different: there c''_Λ stands for what we have called c_Λ, and another constant called c_Λ is introduced which is obviously equivalent to ours.

It will be sufficient to restrict ourselves to the case where the submodule is of rank one. Even in this case, the constants c_Λ and c'_Λ are in general very difficult to compute, so we will use only the obvious estimate

$$c_\Lambda \leq |A^{-1}|, \quad c'_\Lambda \leq |A|,$$

for some suitable matrix A satisfying (6) for a given K-submodule Λ.

Proposition 4.2. *Let Λ be a maximal K-submodule of rank 1. Then we have the estimate*

$$c_\Lambda \leq n!K^{n-1}, \quad c'_\Lambda \leq K.$$

Let us point out that these estimates are very rough, however it seems difficult to improve the exponent in K.

Proof. Let $k = (k_1, \ldots, k_n) \in \mathbb{Z}^n \setminus \{0\}$ be a generating vector of Λ, with $|k| \leq K$. Since Λ is maximal, the components of k are relatively prime and the invariant factor of Λ is equal to one, so

$$B = (1), \qquad \Lambda = (1\ 0 \ldots 0),$$

and a matrix $A \in GL(n, \mathbb{Z})$ satisfies $L = B\Delta A$ as in (6) if and only if its first row is equal to k.

1. To prove our estimates, it will be enough to show that one can choose A such that the ℓ^1–norm of each of its rows is bounded by K, that is $|A| \leq K$. Indeed, assuming the existence of such a matrix A, one immediately gets

$$c'_\Lambda \leq |A| \leq K.$$

Moreover, since the determinant of A is ± 1, A^{-1} is a matrix of cofactors, therefore the absolute value of each of its element is trivially bounded by $(n-1)!K^{n-1}$, which gives

$$c_\Lambda \leq |A^{-1}| \leq n!K^{n-1}.$$

So it remains to prove that one can construct such a matrix, and we will do this by induction on $n \geq 1$.

2. Let us first state an elementary remark, which is an easy consequence of the Euclidean division algorithm. Let $(x, y) \in \mathbb{Z}^2 \setminus \{(0, 0)\}$ with $d = \gcd(x, y)$. Then there exist $u, v \in \mathbb{Z}$ satisfying the equation

$$ux + vy = d$$

with the estimates

$$|u| \leq \frac{|y|}{d}, \quad |v| \leq \frac{|x|}{d}.$$

3. Let $K \geq 1$ be given. We can now state our induction hypothesis **H(n)** for $n \geq 1$.

H(n). *Let $k = (k_1, \ldots, k_n)$ be a vector in $\mathbb{Z}^n \setminus \{0\}$ with co-prime components such that $|k| \leq K$. Then there exists a matrix $A \in GL(n, \mathbb{Z})$ with first row equal to k, which satisfies $|A| \leq K$.*

The assertion **H(1)** is immediate since in this case $k = (\pm 1)$. Now for $n \geq 2$, assume that **H(n-1)** holds true and consider $k = (k_1, \ldots, k_n)$ in $\mathbb{Z}^n \setminus \{0\}$ with co-prime components and $|k| \leq K$.

We may suppose that $k_* = (k_1, \ldots, k_{n-1})$ is non zero (otherwise we consider $k_* = (k_2, \ldots, k_n)$) and we set $d = \gcd(k_1, \ldots, k_{n-1})$. So $d \geq 1$, the integers $d^{-1}k_1, \ldots, d^{-1}k_{n-1}$ are co-prime, and

$$|k_*| \leq \frac{K}{d}.$$

By **H(n-1)** we can find a matrix

$$\begin{pmatrix} d^{-1}k_1 & \cdots & d^{-1}k_{n-1} \\ l_{2,1} & \cdots & l_{2,n-1} \\ \vdots & & \vdots \\ l_{n-1,1} & \cdots & l_{n-1,n-1} \end{pmatrix} \in GL(n-1, \mathbb{Z}),$$

such that

$$|\sum_{j=1}^{n-1} l_{i,j}| \leq K$$

for each $i \in \{2, \ldots, n-1\}$.

Now since d and k_n are co-prime, one can find integers $u, v \in \mathbb{Z}$ such that

$$ud + vk_n = 1$$

and therefore define a matrix

$$A(u,v) = \begin{pmatrix} k_1 & \cdots & k_{n-1} & k_n \\ l_{2,1} & \cdots & l_{2,n-1} & 0 \\ \vdots & & \vdots & \vdots \\ l_{n-1,1} & \cdots & l_{n-1,n-1} & 0 \\ (-1)^{n-1}vd^{-1}k_1 & \cdots & (-1)^{n-1}vd^{-1}k_{n-1} & (-1)^{n-1}u \end{pmatrix}.$$

Expanding the determinant along the last column easily proves that $A(u,v) \in GL(n, \mathbb{Z})$.

As for the estimates, first assume that $k_n = 0$. Then $d = 1$ and we may choose $u = 1$ and $v = 0$, so obviously

$$|A(1,0)| \leq K.$$

If now k_n is non zero, then by our previous remark we can choose (u_*, v_*) so that

$$|u_*| \leq |k_n|, \quad |v_*| \leq d,$$

which proves that the ℓ^1–norm of the last row is bounded by K, and therefore $|A(u_*, v_*)| \leq K$. This ends the proof. □

With these estimates, one can deduce from Theorem 4.1 the following lemma.

Lemma 4.3. *Let Λ be a K-submodule of $\mathbb{Z}^n$ of rank 1. Assume that $\varepsilon \geq 0$ and $K \geq 1$ satisfy*

$$\varepsilon K^{5(n-1)^2} \lessdot 1 \tag{7}$$

and H satisfies $(C(\alpha, L, M, m, \varepsilon))$. Then for any solution $(\theta(t), I(t))$ such that $I_0 \in B(0, R/2)$ and $\omega(0) \in R_K$ the following estimates

$$|I(t) - I_0| \lessdot (\varepsilon K^{(n-1)(3n-1)})^{\frac{1}{2(n-1)}}, \quad t \lessdot \exp\left(\cdot \varepsilon K^{5(n-1)^2}\right)^{-\frac{1}{2\alpha(n-1)}},$$

hold true.

The proof of Theorem 2.2 is now completely similar to the proof of Theorem 2.1, but we shall repeat some details.

Proof of Theorem 2.2. We choose K of the form

$$K = K_0 \left(\frac{\varepsilon_0}{\varepsilon}\right)^{\gamma}$$

with suitable stable constants K_0 and ε_0 so that conditions (2) and (7) are satisfied if

$$0 \leq \varepsilon \leq \varepsilon_0, \quad 0 < \gamma \leq 5^{-1}(n-1)^{-2}.$$

Then we can apply both Lemma 3.2 and Lemma 3.5 in the same way as in the proof of Theorem 2.2, and setting

$$a_\gamma = \frac{1 - 5\gamma(n-1)^2}{2\alpha(n-1)}, \quad b_\gamma = \frac{1 - \gamma(n-1)(3n-1)}{2(n-1)},$$

we find that all solutions $(\theta(t), I(t))$ starting at (θ_0, I_0), with $I_0 \in B(0, R/2)$, satisfy

$$|I(t) - I_0| \lessdot \max\left\{\varepsilon^{\gamma}, \varepsilon^{b_\gamma}\right\}, \quad t \lessdot \exp(\cdot \varepsilon^{-a_\gamma}).$$

Next, using our condition

$$0 < \gamma \leq 5^{-1}(n-1)^{-2}$$

we have the bounds

$$5^{-1}(n-2)(n-1)^{-2} \leq b_\gamma \leq (2(n-1))^{-1}$$

hence $\gamma \leq b_\gamma$ so that

$$|I(t) - I_0| \lessdot \varepsilon^\gamma, \quad t \lessdot \exp(\cdot \varepsilon^{-a_\gamma}).$$

To conclude, just set

$$\delta = \frac{5}{2}\gamma(n-1)$$

so that

$$a_\gamma = (2\alpha(n-1))^{-1} - \delta$$

hence

$$|I(t) - I_0| \lessdot \varepsilon^{\frac{2}{5}\delta(n-1)^{-1}}, \quad |t| \lessdot \exp\left(\cdot \varepsilon^{-(2\alpha(n-1))^{-1}+\delta}\right)$$

provided that

$$0 < \delta \leq (2(n-1))^{-1}$$

from which we will only retain

$$0 < \delta \leq (2\alpha n(n-1))^{-1}.$$

This concludes the proof of the first part, and the arguments for the second part are analogous, choosing K that depends on ρ but is independent of ε. □

Acknowledgements. The authors thank Pierre Lochak for numerous discussions on the problem of optimality of the stability exponents, Henri Scott Dumas for careful reading, and A.B. also thanks Viet Loc Bui for his explanations of some arithmetic issues and Laurent Niederman for his support. Finally, the authors thank one of the referees for pointing out a mistake in a previous version of this paper.

References

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

[Bes96] U. Bessi, *An approach to Arnold's diffusion through the calculus of variations*, Nonlinear Anal., Theory Methods Appl. **26** (1996), no. 6, 1115–1135.

[Bes97] ———, *Arnold's example with three rotators*, Nonlinearity **10** (1997), no. 3, 763–781.

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

[LM05] P Lochak and J.P. Marco, *Diffusion times and stability exponents for nearly integrable analytic systems*, Central European Journal of Mathematics **3** (2005), no. 3, 342–397.

[LN92] P. Lochak and A.I. Neishtadt, *Estimates of stability time for nearly integrable systems with a quasiconvex Hamiltonian*, Chaos **2** (1992), no. 4, 495–499.

[LNN94] P. Lochak, A.I. Neistadt, and L. Niederman, *Stability of nearly integrable convex Hamiltonian systems over exponentially long times*, Kuksin, S. (ed.) et al., Seminar on dynamical systems. Basel: Birkhäuser. Prog. Nonlinear Differ. Equ. Appl. 12, 15-34 (1994)., 1994.

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

[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.

[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.

[Nie04] L. Niederman, *Exponential stability for small perturbations of steep integrable Hamiltonian systems*, Erg. Th. Dyn. Sys. **24** (2004), no. 2, 593–608.

[Pös93] J. Pöschel, *Nekhoroshev estimates for quasi-convex Hamiltonian systems*, Math. Z. **213** (1993), 187–216.

[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.

[Sev06] Mikhail B. Sevryuk, *Partial preservation of frequencies in KAM theory*, Nonlinearity **19** (2006), no. 5, 1099–1140.

[Zha09] Ke Zhang, *Speed of Arnold diffusion for analytic Hamiltonian systems*, Preprint (2009).