



**HAL**  
open science

# Generic super-exponential stability of invariant tori in Hamiltonian systems

Abed Bounemoura

► **To cite this version:**

Abed Bounemoura. Generic super-exponential stability of invariant tori in Hamiltonian systems. Ergodic Theory and Dynamical Systems, 2011, 31 (5), pp.1287-1303. hal-00441994v2

**HAL Id: hal-00441994**

**<https://hal.science/hal-00441994v2>**

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.

# Generic super-exponential stability of invariant tori in Hamiltonian systems.

Abed Bounemoura \*

June 7, 2010

## Abstract

In this article, we consider solutions starting close to some linearly stable invariant tori in an analytic Hamiltonian system and we prove results of stability for a super-exponentially long interval of time, under generic conditions. The proof combines classical Birkhoff normal forms and a new method to obtain generic Nekhoroshev's estimates developed by the author and L. Niederman in another paper. We will mainly focus on the neighbourhood of elliptic fixed points, since with our approach the other cases are completely similar.

## 1 Introduction and main results

In this paper, we are interested in the stability properties of some linearly stable invariant tori in analytic Hamiltonian systems. Let us begin by the case of elliptic fixed points.

1. As the problem is local, it is enough to consider a Hamiltonian  $H$  defined and analytic on an open neighbourhood of 0 in  $\mathbb{R}^{2n}$ , having the origin as a fixed point. Up to an irrelevant additive constant and expanding the Hamiltonian as a power series at the origin, we can write

$$H(z) = H_2(z) + V(z),$$

where  $z$  is sufficiently close to 0 in  $\mathbb{R}^{2n}$ ,  $H_2$  is the quadratic part of  $H$  at 0 and  $V(z) = O(\|z\|^3)$ . Recall that the fixed point is said to be elliptic if the spectrum of the linearized system is purely imaginary. Then it has the form  $\{\pm i\alpha_1, \dots, \pm i\alpha_n\}$ , for some vector  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$  which is called the normal (or characteristic) frequency. Due to the symplectic character of the equations, such equilibria are the only linearly stable fixed points. Now

---

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

we assume that the components of  $\alpha$  are all distinct so that we can make a symplectic linear change of variables that diagonalizes the quadratic part:

$$H(z) = \sum_{i=1}^n \frac{\alpha_i}{2} (z_i^2 + z_{n+i}^2) + V(z) = \alpha \cdot \tilde{I} + V(z),$$

where  $\tilde{I} = \tilde{I}(z)$  is the vector of “formal actions”, that is

$$\tilde{I}(z) = \frac{1}{2}(z_1^2 + z_{n+1}^2, \dots, z_n^2 + z_{2n}^2) \in \mathbb{R}^n.$$

Assuming the components of  $\alpha$  are all of the same sign, it is easy to see that  $H$  is a Lyapunov function so the fixed point is stable. But in the general case, one has to study the influence of the higher order terms  $V(z)$ , and we will explain how it can be done using classical perturbation theory.

**2.** Let us first note that, given a solution  $z(t)$  of  $H$ , if  $\tilde{I}(t) = \tilde{I}(z(t))$  then

$$|\tilde{I}(t)|_1 = \sum_{i=1}^n |\tilde{I}_i(t)|$$

is (up to a factor one-half) the square of the Euclidean distance of  $z(t)$  to the origin, so that Lyapunov stability can be proved if  $|\tilde{I}(t) - \tilde{I}(0)|_1$  does not vary much for all times.

Now in order to study the dynamics on a small neighbourhood of size  $\rho$  around the origin in  $\mathbb{R}^{2n}$ , it is more convenient to change coordinates by performing the standard scalings

$$z \longmapsto \rho z, \quad H \longmapsto \rho^{-2} H,$$

to have a Hamiltonian defined on a fixed neighbourhood of zero in  $\mathbb{R}^{2n}$ . Then, by analyticity, we extend the resulting Hamiltonian to a holomorphic function on some complex neighbourhood of zero in  $\mathbb{C}^{2n}$ . So eventually we will consider the following setting: we define the Euclidean ball in  $\mathbb{C}^{2n}$

$$\mathcal{D}_s = \{z \in \mathbb{C}^{2n} \mid \|z\| < s\}$$

of radius  $s$  around the origin, and if  $\mathcal{A}_s$  is the space of holomorphic functions on  $\mathcal{D}_s$  which are real valued for real arguments, endowed with its usual supremum norm  $|\cdot|_s$ , we consider

$$\begin{cases} H(z) = \alpha \cdot \tilde{I} + f(z) \\ H \in \mathcal{A}_s, |f|_s < \rho. \end{cases} \quad (A)$$

Let us emphasize that the small parameter  $\rho$ , which was originally describing the size of the neighbourhood of 0, now describes the size of the “perturbation”  $f$  on a neighbourhood of *fixed size*  $s$ . Without loss of generality, we may assume  $s > 3$ .

**3.** Probably the main tool to investigate stability properties is the construction of normal forms using averaging methods, and in this case these are the so-called Birkhoff normal forms. For an integer  $m \geq 1$ , assuming  $\alpha$  is non-resonant up to order  $2m$ , that is

$$k \cdot \alpha \neq 0, \quad k \in \mathbb{Z}^n, \quad 0 < |k|_1 \leq 2m,$$

there exists an analytic symplectic transformation  $\Phi_m$  close to identity such that  $H \circ \Phi_m$  is in Birkhoff normal form up to order  $2m$ , that is

$$H \circ \Phi_m(z) = h_m(\tilde{I}) + f_m(z),$$

where  $h_m$  is a polynomial of degree at most  $m$  in the  $\tilde{I}$  variables, and the remainder  $f_m$  is roughly of order  $\rho^{2m-1}$  (since before the scaling  $f_m(z)$  is of order  $\|z\|^{2m+1}$ , see [Bir66], or [Dou88] for a more recent exposition). The polynomials  $h_m$  are uniquely defined once  $\alpha$  is fixed, and are usually called the Birkhoff invariants. Therefore the transformed Hamiltonian is the sum of an integrable part  $h_m$ , for which the origin is trivially stable, since  $\tilde{I}(t)$  is constant for all times, and a smaller perturbation  $f_m$ . Moreover, if  $\alpha$  is non-resonant up to any order, we can even define a formal symplectic transformation  $\Phi_\infty$  and a formal power series  $h_\infty = \sum_{k \geq 1} h^k$ , with  $h_m = \sum_{k=1}^m h^k$ , such that

$$H \circ \Phi_\infty(z) = h_\infty(\tilde{I}).$$

In general the series  $h_\infty$  is divergent (this is a result of Siegel) and the convergence properties of the transformation  $\Phi_\infty$  are even more subtle (see [PM03]). However, Birkhoff normal forms at finite order are still very useful, not only because the ‘‘perturbation’’  $f_m$  is made smaller, but also because the ‘‘integrable’’ part  $h_m$ , for  $m \geq 2$ , is now non-linear and other classical techniques from perturbation theory can be used.

**4.** First, in the case  $n = 2$ , a complete result of stability follows from KAM theory. Indeed, if the frequency  $\alpha \in \mathbb{R}^2$  is non-resonant up to order 4, the Birkhoff normal form reads

$$H(z) = \alpha \cdot \tilde{I} + \beta \tilde{I} \cdot \tilde{I} + f_2(z),$$

with  $\beta$  a symmetric matrix of size  $n = 2$  and  $f_2$  a small perturbation. This time we consider the non-linear part  $h_2(\tilde{I}) = \alpha \cdot \tilde{I} + \beta \tilde{I} \cdot \tilde{I}$  as the integrable system, and if it is isoenergetically non degenerate, the persistence of two-dimensional tori in each energy level close to the fixed point implies Lyapunov stability (see [AKN06], [Arn63], or [Arn61] for other results).

However, for  $n \geq 3$ , it is believed that ‘‘generic’’ elliptic fixed points are unstable, although this is totally unclear for the moment (see [DLC83], [Dou88] and [KMV04]).

Therefore, for  $n \geq 3$ , stability results under general assumptions can only concern finite but hopefully long intervals of time, and this is the content

of the paper. More precisely, we will prove, under generic assumptions and provided  $\rho$  is sufficiently small, that for all initial conditions the variation  $|\tilde{I}(t) - \tilde{I}(0)|_1$  is of order  $\rho$  for  $t \in T(\rho)$ , where  $T(\rho)$  is an interval of time of order  $\exp(\exp(\rho^{-1}))$  (see Theorem 1.1 for a precise formulation). The interpretation in the original coordinates is the following: if a solution starts in a sufficiently small neighbourhood of the origin, it stays in some larger neighbourhood during an interval of time which is super-exponentially long with respect to the inverse of the initial distance to the origin. But first, let us describe previously known results on exponential stability, where there were basically two strategies.

**5.** In a first approach, one assumes a Diophantine condition on  $\alpha$ , that is there exist  $\gamma > 0$  and  $\tau > n - 1$  such that

$$|k \cdot \alpha| \geq \gamma |k|_1^{-\tau}, \quad k \in \mathbb{Z}^n \setminus \{0\},$$

but no conditions on the Birkhoff invariants. From the point of view of perturbation theory, the linear part is considered as the integrable system. In particular,  $\alpha$  is non-resonant up to any order, hence we can perform any finite number of Birkhoff normalizations, and since we have a control on the small divisors, we can precisely estimate the size of the remainder  $f_m$  (in terms of  $\gamma$  and  $\tau$ ). The usual trick is then to optimize the choice of  $m$  as a function of  $\rho$  in order to obtain an exponentially small remainder with respect to  $\rho^{-1}$ . Therefore the exponential stability is immediately read from the normal form, and this requires only an assumption on the linear part (see [GDF<sup>+</sup>89] or [DG96]). The above Diophantine condition has full Lebesgue measure. However, as we will see later, the threshold of the perturbation and the constants of stability are very sensitive to the Diophantine properties of  $\alpha$ , in particular the small parameter  $\gamma$ .

**6.** The second approach is fundamentally different, and it does not rely on the arithmetic properties of  $\alpha$ . Here, one just assumes that  $\alpha$  is non-resonant up to order 4, so that the Hamiltonian reduces to

$$H(z) = \alpha \cdot \tilde{I} + \beta \tilde{I} \cdot \tilde{I} + f_2(z).$$

In this case  $h_2(\tilde{I}) = \alpha \cdot \tilde{I} + \beta \tilde{I} \cdot \tilde{I}$  is considered as the integrable system ( $\beta$  being a symmetric matrix of size  $n$ ). Now we suppose that the non-linear part is convex, which is equivalent to  $\beta$  being sign definite. Under those assumptions, it was predicted and partially proved by Lochak ([Loc92] and [Loc95]), and completely proved independently by Niederman ([Nie98]) and Fassò, Guzzo and Benettin ([FGB98] and [GFB98]) that exponential stability holds. Their proofs are based on the implementation of Nekhoroshev's estimates in Cartesian coordinates, but they are radically different: the first one uses Lochak's method of periodic averagings and simultaneous Diophantine approximations, while the second one is based on Nekhoroshev's original

mechanism. The proof of Niederman was later clarified by Pöschel ([Pös99]). However, the method of Lochak was restricted to the convex case, and it was not clear how to remove this hypothesis to have a result valid in a more general context.

**7.** In this paper, using the method of [BN09] we are able to replace the convexity condition by a generic assumption. Then, combining both Birkhoff theory and Nekhoroshev theory as in [MG95], we will obtain the following result.

**Theorem 1.1.** *Suppose  $H$  is as in (A), with  $\alpha$  non-resonant up to any order. Then under a generic condition (G) on  $h_\infty$ , there exist positive constants  $a, a', c_1, c_2$  and  $\rho_0$  such that for  $\rho \leq \rho_0$ , every solution  $z(t)$  of  $H$  with  $|\tilde{I}(0)|_1 < 1$  satisfies*

$$|\tilde{I}(t) - \tilde{I}(0)|_1 < c_1 \rho, \quad |t| < \exp\left(\rho^{-a'} \exp(c_2 a' \rho^{-a})\right).$$

Denoting  $h_\infty = \sum_{k \geq 1} h^k$  and  $h_m = \sum_{k=1}^m h^k$ , let us explain our generic condition (G) on the formal power series  $h_\infty$ . In fact

$$(G) = \bigcup_{m \in \mathbb{N}^*} (G_m)$$

consists in countably many conditions, where  $(G_m)$  is a condition on  $h_m$ . The first condition  $(G_1)$  requires that  $h_1(I) = \alpha \cdot I$  with a  $(\gamma, \tau)$ -Diophantine vector  $\alpha$ . The other conditions  $(G_m)$ , for  $m \geq 2$ , are that each polynomial function  $h_m$  belongs to a special class of functions called  $SDM_{\gamma'}^{\tau'}$  which was introduced in [BN09] (SDM stands for ‘‘Simultaneous Diophantine Morse’’ functions, see Appendix A for a definition). In this appendix we will show that each condition  $(G_m)$  is of full Lebesgue measure in the finite dimensional space of polynomials of degree  $m$  with  $n$  variables, assuming  $\tau$  and  $\tau'$  are large enough. This is well-known for  $m = 1$ , it will be elementary for  $m = 2$  (see Theorem A.8) but for  $m > 2$  it requires the quantitative Morse-Sard theory of Yomdin ([Yom83], [YC04], see Theorem A.3 in the appendix). Let us point out that this would have not been possible if we had assumed  $h_m$ , for  $m \geq 2$ , to be steep in the sense of Nekhoroshev, as polynomials are generically steep only if their degrees are sufficiently large with respect to the number of degrees of freedom (see [LM88]).

Our condition (G) on the formal series  $h_\infty$  is therefore of ‘‘full Lebesgue measure at any order’’. From an abstract point of view, this condition defines a prevalent set in the space of formal power series, where prevalence is an analog of the notion of full Lebesgue measure in the context of infinite dimensional vector spaces. This will be proved in Appendix A, Theorem A.6. In Theorem 1.1, we can choose the exponents

$$a = (1 + \tau)^{-1}, \quad a' = 3^{-1}(2(n + 1)\tau')^{-n},$$

and our threshold  $\rho_0$  depends in particular on  $\gamma$  and  $\gamma'$ . Moreover our constants  $c_1$  and  $c_2$  also depend on  $\gamma$  but not on  $\gamma'$ , and we shall be a little more precise later on.

As we have already explained, the proof is based on a combination of Birkhoff normalizations up to an exponentially small remainder, which are well-known (a statement is recalled in Proposition 2.1 below), and Nekhoroshev's estimates for a generic integrable Hamiltonian near an elliptic fixed point (Theorem 2.2 below). The latter result is new, and it will follow rather easily from the new approach of Nekhoroshev theory in a generic case taken in [BN09].

**8.** As a direct consequence of our Nekhoroshev estimates near an elliptic fixed point, we can derive an exponential stability result more general than those obtained in [FGB98] and [Nie98]. Like in those papers, we only require  $\alpha$  to be non-resonant up to order 4, and after the scalings

$$z \mapsto \rho z, \quad H \mapsto \rho^{-4} H, \quad \alpha \mapsto \rho^2 \alpha,$$

we consider

$$\begin{cases} H(z) = \alpha \tilde{I} + \beta \tilde{I} \cdot \tilde{I} + f(z) \\ H \in \mathcal{A}_s, |f|_s < \rho. \end{cases} \quad (B)$$

However, instead of assuming that  $\beta$  is sign definite, our result applies to Lebesgue almost all symmetric matrices  $\beta$  without any condition on  $\alpha$ . Let  $S_n(\mathbb{R})$  be the space of symmetric matrices of size  $n$  with real entries.

**Theorem 1.2.** *Suppose  $H$  is as in (B). For Lebesgue almost all  $\beta \in S_n(\mathbb{R})$ , there exist positive constants  $a', b'$  and  $\rho_0$  such that, for  $\rho \leq \rho_0$ , every solution  $z(t)$  of  $H$  with  $|\tilde{I}(0)|_1 < 1$  satisfies*

$$|\tilde{I}(t) - \tilde{I}(0)|_1 < n(n^2 + 1)\rho^{-b'}, \quad |t| < \exp(\rho^{-a'}).$$

The above theorem is a direct consequence of Theorem 2.2 below, provided that  $h_2(\tilde{I}) = \alpha \tilde{I} + \beta \tilde{I} \cdot \tilde{I}$  belongs to  $SDM_{\gamma'}^{a'}$ . But we will prove in Appendix A that this happens for almost all symmetric matrices  $\beta$ , independently of  $\alpha$  (see Theorem A.8). Once again, let us also mention that this result is not possible in the steep case, as the quadratic part  $h^2(\tilde{I}) = \beta \tilde{I} \cdot \tilde{I}$  is steep if and only if  $\beta$  is sign definite. In the above statement one can choose the exponents

$$a' = b' = 3^{-1}(2(n+1)\tau')^{-n},$$

and the threshold  $\rho_0$  depends on  $\gamma'$ .

**9.** Let us add that in order to avoid useless expressions, we will only keep track of the small parameters  $\rho$ ,  $\gamma$  and  $\gamma'$  and replace any other positive constants by a dot ( $\cdot$ ) when it is convenient.

Moreover, in this text we shall use various norms for vectors  $v \in \mathbb{R}^n$  or  $v \in \mathbb{C}^n$ :  $|\cdot|$  will be the supremum norm,  $|\cdot|_1$  the  $\ell_1$ -norm and  $\|\cdot\|$  the Euclidean (or Hermitian) norm.

**10.** Let us now describe the plan of the paper. Section 2 is devoted to the proof of Theorem 1.1 and Theorem 1.2. In 2.1, we give a statement of the Birkhoff normal form up to an exponentially small remainder. In 2.2, we will explain how Nekhoroshev's estimates obtained in [BN09] generalize in the neighbourhood of elliptic fixed points, and how they imply Theorem 1.2. In 2.3, we will show how Theorem 1.1 follows from a simple combination of Birkhoff's estimates and Nekhoroshev's estimates, provided our assumption on  $h_\infty$  is satisfied. Then, in section 3, we will state similar results for invariant Lagrangian tori and more generally for invariant linearly stable isotropic reducible tori. Finally, an appendix is devoted to our genericity assumptions.

## 2 Proof of Theorem 1.1 and Theorem 1.2

In the sequel, we recall that we will use the “formal” actions

$$\tilde{I} = \tilde{I}(z) = \frac{1}{2}(z_1^2 + z_{n+1}^2, \dots, z_n^2 + z_{2n}^2) \in \mathbb{R}^n,$$

but one has to remember that these are nothing but notations for expressions in  $z \in \mathbb{R}^{2n}$ . Moreover, we will also need to use complex coordinates for the normal forms, and, abusing notations, we will also denote them by  $z \in \mathbb{C}^{2n}$ , but of course the solutions we consider are real.

### 2.1 Birkhoff's estimates

Here we consider a Hamiltonian as in (A), and we assume that the vector  $\alpha$  is  $(\gamma, \tau)$ -Diophantine. In this context, the following result is classical.

**Proposition 2.1.** *Under the previous assumptions, if  $\rho \ll \gamma$ , then there exist an integer  $m = m(\rho)$  and an analytic symplectic transformation*

$$\Phi_m : \mathcal{D}_{3s/4} \rightarrow D_s$$

such that

$$H \circ \Phi_m(z) = h_m(\tilde{I}) + f_m(z)$$

is in Birkhoff normal form, with a remainder  $f_m$  satisfying the estimate

$$|f_m|_{3s/4} \ll \rho \exp(-(\gamma\rho^{-1})^a), \quad a = (1 + \tau)^{-1}.$$

Moreover,  $|\Phi_m - \text{Id}|_{3s/4} \ll \gamma^{-1}\rho$  and the image of  $\Phi_m$  contains the domain  $\mathcal{D}_{s/2}$ .



For a proof, we refer to [GDF<sup>+</sup>89] and [DG96]. The analogous result for invariant Lagrangian tori can be found in [PW94] or [Pös93], and in [JV97] in the more general case of isotropic and reducible linearly stable invariant tori.

In the above proposition, one has to choose the integer  $m = m(\rho)$  of order  $(\gamma\rho^{-1})^{(\tau+1)^{-1}}$ . So letting  $\rho$  go to zero, the degree of the polynomial  $h_m$  goes to infinity, and this explains why in the proof of Theorem 1.1 we require a condition on the whole formal power series  $h_\infty$ .

## 2.2 Nekhoroshev's estimates and proof of Theorem 1.2

1. Here we consider the Hamiltonian

$$\begin{cases} H(z) = h(\tilde{I}) + f(z) \\ H \in \mathcal{A}_s, h \in \text{SDM}_{\gamma'}^{\tau'}, |f|_s < \varepsilon \end{cases} \quad (E)$$

and we have assumed that, on the real part of the domain, the derivatives up to order 3 of  $h$  are uniformly bounded by some constant  $M > 1$ . The definition of the set  $\text{SDM}_{\gamma'}^{\tau'}$  is recalled in Appendix A.

**Theorem 2.2.** *Let  $H$  be as in (E), with  $\tau' \geq 2$  and  $\gamma' \leq 1$ . Then there exists  $\varepsilon_0$  such that if  $\varepsilon \leq \varepsilon_0$ , for every solution  $z(t)$  with  $|\tilde{I}(0)| < 1$ , we have*

$$|\tilde{I}(t) - \tilde{I}(0)| < (n^2 + 1)\varepsilon^{b'}, \quad |t| < \exp(\varepsilon^{-a'}),$$

with the exponents  $a' = b' = 3^{-1}(2(n+1)\tau')^{-n}$ .

Theorem 1.2 is now an immediate consequence of this result and Theorem A.8 (see Appendix A).

*Proof of Theorem 1.2.* From Theorem A.8, we know that for almost all  $\beta \in S_n(\mathbb{R})$ , the Hamiltonian  $h(\tilde{I}) = \alpha\tilde{I} + \beta\tilde{I}\cdot\tilde{I}$  belongs to  $\text{SDM}^{\tau'}(B)$  with  $\tau' > n^2 + 1$ . So we can apply Theorem 2.2: for every solution  $z(t)$  with  $|\tilde{I}(0)| < 1$ , we have

$$|\tilde{I}(t) - \tilde{I}(0)| < (n^2 + 1)\varepsilon^{b'}, \quad |t| < \exp(\varepsilon^{-a'}),$$

with the exponents  $a' = b' = 3^{-1}(2(n+1)\tau')^{-n}$ . In particular, this gives

$$|\tilde{I}(t) - \tilde{I}(0)|_1 < n(n^2 + 1)\varepsilon^{b'}, \quad |t| < \exp(\varepsilon^{-a'}),$$

for every solution  $z(t)$  with  $|\tilde{I}(0)|_1 < 1$ . □

2. The statement of Theorem 2.2 is the analogue of the main statement of [BN09]. However, the difference is that here we are using Cartesian coordinates and not action-angle coordinates (*i.e.* symplectic polar coordinates), and we *cannot* use the latter since they become singular at the origin. So we *cannot* apply directly the main result of [BN09]. This is not a serious

issue when applying KAM theory in this context (see [Arn63] or [Pös82] for example), but this becomes problematic in Nekhoroshev theory (see [Loc92] or [Loc95] for detailed explanations). This result was only conjectured by Nekhoroshev in [Nek77], and it took a long time before it could be solved in the convex case ([Nie98],[FGB98]). Here we are able to solve this problem in the generic case. The reason is that even though we cannot apply the result of [BN09], we can use exactly the same approach, since the method of averagings along unperturbed periodic flows is intrinsic, *i.e.* independent of the choice of coordinates, a fact that was first used implicitly in [Nie98] and made completely clear in [Pös99].

The proof of such estimates usually requires an analytic part, which boils down to the construction of suitable normal forms, and a geometric part. The geometric part of [BN09] goes exactly the same way, so in the sequel we will restrict ourselves to indicating the very slight modifications in the construction of the normal forms.

**3.** Consider linearly independent periodic vectors  $\omega_1, \dots, \omega_n$  of  $\mathbb{R}^n$ , with periods  $(T_1, \dots, T_n)$ , that is

$$T_j = \inf\{t > 0 \mid t\omega_j \in \mathbb{Z}^n\}, \quad 1 \leq j \leq n.$$

Define the domains

$$\mathcal{D}_{r_j, s_j}(\omega_j) = \{z \in \mathcal{D}_{s_j} \mid |\nabla h(\tilde{I}) - \omega_j| < r_j\}, \quad 1 \leq j \leq n,$$

given two sequences  $(r_1, \dots, r_n)$  and  $(s_1, \dots, s_n)$  (recall that  $\mathcal{D}_{s_j}$  is the complex ball of radius  $s_j$ ). We will denote by  $l_j$  the linear Hamiltonian with frequency  $\omega_j$ , that is  $l_j(\tilde{I}) = \omega_j \cdot \tilde{I}$ .

The supremum norm of a function  $f$  defined on  $\mathcal{D}_{r_j, s_j}(\omega_j)$  will be simply denoted by

$$|f|_{r_j, s_j} = |f|_{\mathcal{D}_{r_j, s_j}(\omega_j)},$$

and for a Hamiltonian vector field  $X_f$ , we will write

$$|X_f|_{r_j, s_j} = \max_{1 \leq i \leq 2n} |\partial_{z_i} f|_{r_j, s_j}.$$

To obtain normal forms on these domains we will make the following assumptions  $(A_j)$ , for  $j \in \{1, \dots, n\}$ , where  $(A_1)$  is

$$\begin{cases} mT_1\varepsilon < r_1, \quad mT_1r_1 < s_1, \quad 0 < r_1 < s_1, \\ \mathcal{D}_{r_1, s_1}(\omega_1) \neq \emptyset, \quad s_1 < s, \end{cases} \quad (A_1)$$

and for  $j \in \{2, \dots, n\}$ ,  $(A_j)$  is

$$\begin{cases} mT_j\varepsilon < r_j, \quad mT_jr_j < s_j, \quad 0 < r_j < s_j, \\ \mathcal{D}_{r_j, s_j}(\omega_j) \neq \emptyset, \quad \mathcal{D}_{r_j, s_j}(\omega_j) \subseteq \mathcal{D}_{2r_{j-1}/3, 2s_{j-1}/3}(\omega_{j-1}). \end{cases} \quad (A_j)$$

With these assumptions, one can prove the following proposition.

**Proposition 2.3.** Consider  $H = h + f$  on the domain  $\mathcal{D}_{r_1, s_1}(\omega_1)$ , with  $|X_f|_{r_1, s_1} < \varepsilon$ , and let  $j \in \{1, \dots, n\}$ . For any  $i \in \{1, \dots, j\}$ , if  $(A_i)$  is satisfied, then there exists an analytic symplectic transformation

$$\Psi_j : \mathcal{D}_{2r_j/3, 2s_j/3}(\omega_j) \rightarrow \mathcal{D}_{r_1, s_1}(\omega_1)$$

such that

$$H \circ \Psi_j = h + g_j + f_j,$$

with  $\{g_j, l_i\} = 0$  for  $i \in \{1, \dots, j\}$ , and the estimates

$$|X_{g_j}|_{2r_j/3, 2s_j/3} < \varepsilon, \quad |X_{f_j}|_{2r_j/3, 2s_j/3} < e^{-m} \varepsilon.$$

Moreover, we have  $\Psi_j = \Phi_1 \circ \dots \circ \Phi_j$  with

$$\Phi_i : \mathcal{D}_{2r_i/3, 2s_i/3}(\omega_i) \rightarrow \mathcal{D}_{r_i, s_i}(\omega_i)$$

such that  $|\Phi_i - \text{Id}|_{2r_i/3, 2s_i/3} < r_i$ .

The proof is completely analogous to the corresponding one in [BN09], Appendix A, to which we refer for more details. In fact, here the proof is even simpler since one does not have to use “weighted” norms for vector fields. It relies on a finite composition of averagings along the periodic flows generated by  $l_j$ ,  $j \in \{1, \dots, n\}$ . The case  $j = 1$  is due to Pöschel ([Pös99]) and, for  $j > 1$ , the proof goes by induction using our assumption  $(A_j)$ ,  $j \in \{1, \dots, n\}$ .

Once we have this normal form, the rest of the proof in [BN09] goes exactly the same way: every solution  $z(t)$  of  $H$  with  $|\tilde{I}(0)| < 1$  satisfies

$$|\tilde{I}(t) - \tilde{I}(0)| < (n^2 + 1)\varepsilon^{b'}, \quad |t| < \exp(\varepsilon^{-a'}),$$

provided that  $\varepsilon \leq \varepsilon_0$ , with  $\varepsilon_0$  depending on  $n, s, M, \gamma'$  and  $\tau'$  and with the exponents

$$a' = b' = 3^{-1}(2(n+1)\tau')^{-n}.$$

### 2.3 Proof of Theorem 1.1

Now we can finally prove Theorem 1.1, by using successively Birkhoff’s estimates and Nekhoroshev’s estimates.

*Proof of Theorem 1.1.* Let  $H$  be as in (A), first assume that  $\rho < \rho_1$  with  $\rho_1 = \cdot \gamma$  so that using our assumption  $(G_1)$  we can apply Proposition 2.1: there exist an integer  $m = m(\rho)$  and an analytic symplectic transformation

$$\Phi_m : \mathcal{D}_{3s/4} \rightarrow \mathcal{D}_s$$

such that

$$H \circ \Phi_m(z) = h_m(\tilde{I}) + f_m(z)$$

is in Birkhoff normal form, with a remainder  $f_m$  satisfying the estimate

$$|f_m|_{3s/4} < \rho \exp\left(-(\gamma\rho^{-1})^{a-1}\right), \quad a = \tau + 1.$$

So let  $H_m = H \circ \Phi_m$ , and set

$$\varepsilon = \rho \exp\left(-(\gamma\rho^{-1})^{a-1}\right).$$

By our assumption  $(G_m)$ , for  $m \geq 2$ , the Hamiltonian  $H_m$ , which is defined on the domain  $\mathcal{D}_{3s/4}$ , satisfies  $(E)$ . Now assume that  $\varepsilon < \varepsilon_0$ , which gives another threshold  $\rho < \rho_2$ , with  $\rho_2$  also depending on  $\gamma'$ , and our final threshold is  $\rho_0 = \min\{\rho_1, \rho_2\}$ . We can apply Theorem 2.2: every solution  $z_m(t)$  of  $H_m$  with  $|\tilde{I}_m(0)|_1 < 1$  satisfies

$$|\tilde{I}_m(t) - \tilde{I}_m(0)|_1 < \varepsilon^{b'}, \quad |t| < \exp(\varepsilon^{-a'}),$$

with

$$a' = b' = 3^{-1}(2(n+1)\tau')^{-n}.$$

Recalling the definition of  $\varepsilon$ , this gives

$$|\tilde{I}_m(t) - \tilde{I}_m(0)|_1 < \rho^{b'} \exp(-b'(\gamma\rho^{-1})^a), \quad |t| < \exp\left(\rho^{-a'} \exp(a'(\gamma\rho^{-1})^a)\right).$$

However, one has

$$\rho^{b'} \exp(-b'(\gamma\rho^{-1})^a) < \gamma^{-1}\rho,$$

and as  $\Phi_m$  satisfies  $|\Phi_m - \text{Id}|_{3s/4} < \gamma^{-1}\rho$ , and its image contains the domain  $\mathcal{D}_{s/2}$ , a standard argument gives

$$|I(t) - I(0)|_1 < \gamma^{-1}\rho, \quad |t| < \exp\left(\rho^{-a'} \exp(a'(\gamma\rho^{-1})^a)\right),$$

for any solution  $z(t)$  of  $H$  with  $|\tilde{I}(0)|_1 < 1$ . □

### 3 Further results and comments

As we have already mentioned, the idea of combining both Birkhoff theory and Nekhoroshev theory to obtain super-exponential stability was discovered by Morbidelli and Giorgilli ([MG95]) in the context of Lagrangian Diophantine tori. Evidently, we can also state results in this context.

**1.** Consider a Hamiltonian system on a manifold which carries an invariant Lagrangian Diophantine torus, that is an invariant sub-manifold  $\mathcal{T}$  which is diffeomorphic to the standard torus  $\mathbb{T}^n$ , and whose induced flow is conjugated to a linear flow on  $\mathbb{T}^n$  with a Diophantine frequency. Since the torus is Lagrangian, one can locally reduce the situation to a Hamiltonian defined on  $T^*\mathbb{T}^n = \mathbb{T}^n \times \mathbb{R}^n$ , having  $\mathbb{T}^n \times \{0\}$  as the invariant torus. Moreover, by

invariance and transitivity of the torus, in the coordinates  $(\theta, I) \in \mathbb{T}^n \times \mathbb{R}^n$  we can write

$$H(\theta, I) = \omega \cdot I + F(\theta, I),$$

where  $\omega$  is a  $(\gamma, \tau)$ -Diophantine vector and  $F(\theta, I) = O(|I|^2)$ . After some scalings one is led to consider

$$\begin{cases} H(\theta, I) = \omega \cdot I + f(\theta, I) \\ H \in \mathcal{A}_s, |f|_s < \rho \end{cases} \quad (C)$$

where  $\mathcal{A}_s$  is the space of holomorphic functions on the domain

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

with  $\mathcal{I}(\theta)$  the imaginary part of  $\theta$ . Here one can also define polynomials  $h_m$  and a formal power series  $h_\infty$  and we can state the following result.

**Theorem 3.1.** *Suppose  $H$  is as in (C). Then, under a generic condition on  $h_\infty$ , there exist positive constants  $a, a', c_1, c_2$  and  $\rho_0$  such that for  $\rho \leq \rho_0$ , every solution  $(\theta(t), I(t))$  of  $H$  with  $|I(0)| < 1$  satisfies*

$$|I(t) - I(0)| < c_1 \rho, \quad |t| < \exp\left(\rho^{-a'} \exp(c_2 a' \rho^{-a})\right).$$

The assumption on  $h_\infty$  and the values of the constants  $a$  and  $a'$  are the same as in Theorem 1.1, as the proof is completely analogous. In fact, it is even simpler since we are using action-angle coordinates, and therefore we can immediately use Nekhoroshev's estimates obtained in [BN09] without any modifications.

However, it is important to note that one cannot obtain a statement similar to Theorem 1.2, simply because in this case a non-resonant condition up to a finite order does not allow to build the corresponding Birkhoff normal form.

If we compare this result with [MG95], our assumption is generic and we do not require any convexity. But of course the price to pay is that one has to consider the full set of Birkhoff invariants.

**2.** As a final result, one can also obtain similar estimates for the general case of a linearly stable lower-dimensional torus, under the common assumptions of isotropicity and reducibility (which were automatic for a fixed point or a Lagrangian torus). In that context, it is enough to consider a Hamiltonian defined in  $\mathbb{T}^k \times \mathbb{R}^k \times \mathbb{R}^{2l}$  (by isotropicity), of the form

$$H(\theta, I, z) = \omega \cdot I + \frac{1}{2} B z \cdot z + F(\theta, I, z).$$

Here  $B$  is a symmetric matrix (constant by reducibility) such that  $J_{2l} B$  has a purely imaginary spectrum ( $J_{2l}$  being the canonical symplectic structure

of  $\mathbb{R}^{2l}$ ), and  $F(\theta, I, z) = O(|I|^2, \|z\|^3)$ . In those coordinates, the invariant torus is simply given by  $I = 0$ ,  $z = 0$ , and this generalizes both the case of an elliptic fixed point (where the directions  $(\theta, I)$  are absent) and of a Lagrangian invariant torus (where the directions  $z$  are absent). If the spectrum  $\{\pm i\alpha_1, \dots, \pm i\alpha_l\}$  of  $J_{2l}B$  is simple, one can assume further that

$$H(\theta, I, z) = \omega \cdot I + \alpha \cdot \tilde{I} + F(\theta, I, z),$$

where  $\tilde{I}$  are the ‘‘formal actions’’ associated to the  $z$  variables. Therefore, after some appropriate scalings we can consider

$$\begin{cases} H(\theta, I, z) = \omega \cdot I + \alpha \cdot \tilde{I} + F(\theta, I, z) \\ H \in \mathcal{A}_s, |f|_s < \rho \end{cases} \quad (D)$$

where  $\mathcal{A}_s$  is the space of holomorphic functions on the domain

$$\mathcal{D}_s = \{(\theta, I, z) \in (\mathbb{C}^k / \mathbb{Z}^k) \times \mathbb{C}^k \times \mathbb{C}^{2l} \mid |\mathcal{I}(\theta)| < s, |I| < s, \|z\| < s\}.$$

Under a suitable Diophantine condition on the vector  $(\omega, \alpha) \in \mathbb{R}^{k+l}$ , one can define polynomials  $h_m$  and a formal series  $h_\infty$  depending on  $J = (I, \tilde{I})$ . Birkhoff’s exponential estimates in this more difficult situation have been obtained in [JV97]. Regarding Nekhoroshev’s estimates for a generic integrable Hamiltonian which depends both on actions and formal actions, they can be easily obtained by obvious modifications of our method. Therefore we can state the following result.

**Theorem 3.2.** *Suppose  $H$  is as in (D). Then under a generic condition on  $h_\infty$ , there exist positive constants  $a, a', c_1, c_2$  and  $\rho_0$  such that for  $\rho \leq \rho_0$ , every solution  $(\theta(t), I(t), z(t))$  of  $H$  with  $|J(0)| < 1$  satisfies*

$$|J(t) - J(0)| < c_1 \rho, \quad |t| < \exp\left(\rho^{-a'} \exp(c_2 a' \rho^{-a})\right).$$

Once again, the condition on  $h_\infty$  and the values of the exponents are the same.

**3.** Let us add that one could easily give similar estimates in the discrete case, that is for exact symplectic diffeomorphisms near an elliptic fixed point, an invariant Lagrangian torus or an invariant linearly stable isotropic reducible torus. Even if one has the possibility to re-write the proof in these settings, the easiest way is to use suspension arguments, as it is done qualitatively in [Dou88] or quantitatively in [KP94] (see also [PT97] for a different approach) and deduce stability results in the discrete case from the corresponding results in the continuous case.

To conclude, let us mention that important examples of invariant tori satisfying our assumptions (linearly stable, reducible, isotropic) are those given by KAM theory. However, the latter not only gives individual tori but

a whole Cantor family (see [Pös01] or [AKN06]). In this context, Popov has proved exponential stability estimates for the family of Lagrangian KAM tori, if the Hamiltonian is analytic or Gevrey ([Pop00] and [Pop04]). His proof relies on a KAM theorem with Gevrey smoothness on the parameters (in the sense of Whitney) and some kind of simultaneous Birkhoff normal form over the Cantor set of tori. We believe that our method should be useful in trying to extend those results to obtain super-exponential stability under generic conditions. But clearly this is a more difficult problem, and the first step is to obtain Nekhoroshev's estimates in Gevrey regularity for a generic integrable Hamiltonian, the quasi-convex case having been settled in [MS02].

## A Generic assumptions

In this appendix, we will show that our assumption  $(G)$  is generic, in the sense that it defines a prevalent set in the infinite dimensional space of formal power series.

1. Let us first recall the definition of Simultaneous Diophantine Morse functions (SDM in the following). Let  $G(n, k)$  be the set of all vector subspaces of  $\mathbb{R}^n$  of dimension  $k$ . We equip  $\mathbb{R}^n$  with the Euclidean scalar product, and given an integer  $L \in \mathbb{N}^*$ , we define  $G^L(n, k)$  as the subset of  $G(n, k)$  consisting of subspaces whose orthogonal complement can be generated by integer vectors with components bounded by  $L$ . In the sequel,  $B$  will be an arbitrary open ball of  $\mathbb{R}^n$ .

**Definition A.1.** *A smooth function  $h : B \rightarrow \mathbb{R}$  is said to be SDM if there exist  $\gamma' > 0$  and  $\tau' \geq 0$  such that for any  $L \in \mathbb{N}^*$ , any  $k \in \{1, \dots, n\}$  and any  $\Lambda \in G^L(n, k)$ , there exists  $(e_1, \dots, e_k)$  (resp.  $(f_1, \dots, f_{n-k})$ ), an orthonormal basis of  $\Lambda$  (resp. of  $\Lambda^\perp$ ), such that the function  $h_\Lambda$  defined on  $B$  by*

$$h_\Lambda(\alpha, \beta) = h(\alpha_1 e_1 + \dots + \alpha_k e_k + \beta_1 f_1 + \dots + \beta_{n-k} f_{n-k})$$

*satisfies the following: for any  $(\alpha, \beta) \in B$ ,*

$$\|\partial_\alpha h_\Lambda(\alpha, \beta)\| \leq \gamma' L^{-\tau'} \implies \|\partial_{\alpha\alpha} h_\Lambda(\alpha, \beta) \cdot \eta\| > \gamma' L^{-\tau'} \|\eta\|,$$

*for any  $\eta \in \mathbb{R}^k \setminus \{0\}$ .*

This definition is inspired by the steepness condition of Nekhoroshev and the quantitative Morse-Sard theory of Yomdin (see [BN09] for more explanations). It depends on a choice of coordinates adapted to the orthogonal decomposition  $\Lambda \oplus \Lambda^\perp$ , so for  $\Lambda \in G^L(n, k)$  and  $(\alpha, \beta) \in B$ ,  $\partial_\alpha h_\Lambda(\alpha, \beta)$  is a vector in  $\mathbb{R}^k$  and  $\partial_{\alpha\alpha} h_\Lambda(\alpha, \beta)$  is a symmetric matrix of size  $k$  with real entries.

**Remark A.2.** Note also that the definition can be stated as the following alternative: for any  $(\alpha, \beta) \in B$ , either we have  $\|\partial_\alpha h_\Lambda(\alpha, \beta)\| > \gamma L^{-\tau}$  or  $\|\partial_{\alpha\alpha} h_\Lambda(\alpha, \beta) \cdot \eta\| > \gamma L^{-\tau} \|\eta\|$  for any  $\eta \in \mathbb{R}^k \setminus \{0\}$ . Hence for a given function it is sufficient to verify that  $\|\partial_{\alpha\alpha} h_\Lambda(\alpha, \beta) \cdot \eta\| > \gamma L^{-\tau} \|\eta\|$  for any  $\eta \in \mathbb{R}^k \setminus \{0\}$ , and we will use this fact later (in Theorem A.8).

**2.** The set of SDM functions on  $B$  with respect to  $\gamma' > 0$  and  $\tau' \geq 0$  will be denoted by  $SDM_{\gamma'}^{\tau'}(B)$ , and we will also use the notation

$$SDM^{\tau'}(B) = \bigcup_{\gamma' > 0} SDM_{\gamma'}^{\tau'}(B).$$

The following theorem was proved in [BN09], and it relies on non trivial results from quantitative Morse-Sard theory ([Yom83],[YC04]).

**Proposition A.3** ([BN09]). Let  $\tau > 2(n^2 + 1)$ , and  $h \in C^{2n+2}(B)$ . Then for Lebesgue almost all  $\xi \in \mathbb{R}^n$ , the function  $h_\xi$ , defined by  $h_\xi(I) = h(I) - \xi \cdot I$  for  $I \in B$ , belongs to  $SDM^{\tau'}(B)$ .

Now let us recall the definition of a prevalent set ([HSY92], see also [OY05]).

**Definition A.4.** Let  $E$  be a completely metrizable topological vector space. A Borel subset  $S \subseteq E$  is said to be shy if there exists a Borel measure  $\mu$  on  $E$ , with  $0 < \mu(C) < \infty$  for some compact set  $C \subseteq E$ , such that  $\mu(x + S) = 0$  for all  $x \in E$ .

An arbitrary set is called shy if it is contained in a shy Borel subset, and finally the complement of a shy set is called prevalent.

For a finite dimensional vector space  $E$ , by an easy application of Fubini theorem, prevalence is equivalent to full Lebesgue measure. The following ‘‘genericity’’ properties can be checked ([OY05]): a prevalent set is dense, a set containing a prevalent set is also prevalent, and prevalent sets are stable under translation and countable intersection. Furthermore, we have an easy but useful criterion for a set to be prevalent.

**Proposition A.5** ([HSY92]). Let  $A$  be a Borel subset of  $E$ . Suppose there exists a finite dimensional subspace  $F$  of  $E$  such that, denoting  $\lambda_F$  the Lebesgue measure supported on  $F$ , the set  $x + A$  has full  $\lambda_F$ -measure for all  $x \in E$ . Then  $A$  is prevalent.

It is an obvious consequence of Proposition A.3 and Proposition A.5 that  $SDM^{\tau'}(B)$  is prevalent in  $C^{2n+2}(B)$  for  $\tau' > 2(n^2 + 1)$ .

**3.** Now let  $\mathcal{P}_\infty = \mathbb{R}[[X_1, \dots, X_n]]$  be the space of all formal power series in  $n$  variables with real coefficients. It is naturally a Fréchet space, as the



projective limit of the finite dimensional spaces  $\mathcal{P}_m$  consisting of polynomials in  $n$  variables of degree less than or equal to  $m$ . We define the subset

$$\mathcal{S}_\infty^{\tau'} = \{h_\infty \in \mathcal{P}_\infty \mid h_m \in \text{SDM}^{\tau'}(B), \forall m \geq 2\},$$

where  $h_m = \sum_{k=1}^m h^k$  if  $h_\infty = \sum_{k \geq 1} h^k$ , and we identify the polynomial  $h_m$  with the associated function defined on  $B$ . Let us also define

$$\mathcal{D}_\infty^\tau = \{h_\infty \in \mathcal{P}_\infty \mid h_1(X) = \alpha \cdot X, \alpha \in \mathcal{D}^\tau\},$$

where  $\mathcal{D}^\tau$  is the set of Diophantine vectors of  $\mathbb{R}^n$  with exponent  $\tau$ , and finally

$$\mathcal{G}_\infty^{\tau, \tau'} = \mathcal{D}_\infty^\tau \cap \mathcal{S}_\infty^{\tau'}.$$

The set  $\mathcal{G}_\infty^{\tau, \tau'}$  is the set of formal power series for which condition (G) holds.

**Theorem A.6.** *For  $\tau > n - 1$  and  $\tau' > 2(n^2 + 1)$ , the set  $\mathcal{G}_\infty^{\tau, \tau'}$  is prevalent in  $\mathcal{P}_\infty$ .*

*Proof.* As the intersection of two prevalent sets is prevalent, it is enough to prove that both sets  $\mathcal{D}_\infty^\tau$ , for  $\tau > n - 1$ , and  $\mathcal{S}_\infty^{\tau'}$ , for  $\tau' > 2(n^2 + 1)$ , are prevalent.

For the set  $\mathcal{D}_\infty^\tau$ , this is an easy consequence of the fact that  $\mathcal{D}^\tau$  is of full Lebesgue measure in  $\mathbb{R}^n$ , for  $\tau > n - 1$ , and Proposition A.5 with  $F = \mathcal{P}_1$ , the space of linear forms. For the set  $\mathcal{S}_\infty^{\tau'}$ , first note that we can write

$$\mathcal{S}_\infty^{\tau'} = \bigcap_{m \geq 2} \mathcal{S}_{\infty, m}^{\tau'},$$

where, for an integer  $m \geq 2$ ,

$$\mathcal{S}_{\infty, m}^{\tau'} = \{h_\infty \in \mathcal{P}_\infty \mid h_m \in \text{SDM}^{\tau'}(B)\}.$$

As a countable intersection of prevalent sets is prevalent, it is enough to prove that for each  $m \geq 2$ , the set  $\mathcal{S}_{\infty, m}^{\tau'}$  is prevalent in  $\mathcal{P}_\infty$ . But once again this is just a consequence of Proposition A.3 and Proposition A.5 with  $F = \mathcal{P}_1$  the space of linear forms.  $\square$

For  $m \geq 2$ , the set of polynomials  $h_m$  for which condition ( $G_m$ ) is satisfied is given by

$$\mathcal{S}_m^{\tau'} = \{h_m \in \mathcal{P}_m \mid h_m \in \text{SDM}^{\tau'}(B)\},$$

and the proof of the above theorem immediately gives the following result.

**Theorem A.7.** *For  $\tau' > 2(n^2 + 1)$ , the set  $\mathcal{S}_m^{\tau'}$  is of full Lebesgue measure in  $\mathcal{P}_m$ .*

4. Now in the special case  $m = 2$ , we can state a refined result which is due to Niederman ([Nie07]).

**Theorem A.8.** *For Lebesgue almost all  $\beta \in S_n(\mathbb{R})$ , the function*

$$h(I) = \alpha.I + \beta I.I$$

*belongs to  $SDM^{\tau'}(B)$  provided  $\tau' > n^2 + 1$ .*

In the above theorem, there is no condition on  $\alpha$ , and contrary to Proposition A.3, the proof does not rely on Morse-Sard theory. Let us denote by  $\lambda$  the one-dimensional Lebesgue measure and by  $I_k$  the identity matrix of size  $k$ . We shall use the following elementary lemma.

**Lemma A.9.** *Let  $k \in \{1, \dots, n\}$ ,  $\beta_k \in S_k(\mathbb{R})$  and  $\kappa > 0$ . Then there exists a subset  $\mathcal{C}_\kappa \subseteq \mathbb{R}$  such that*

$$\lambda(\mathcal{C}_\kappa) \leq 2k\kappa,$$

*and for any  $\xi \notin \mathcal{C}_\kappa$ , the matrix  $\beta_{k,\xi} = \beta_k - \xi I_k$  satisfies*

$$\|\beta_{k,\xi}.\eta\| > \kappa\|\eta\|,$$

*for any  $\eta \in \mathbb{R}^k \setminus \{0\}$ .*

Of course, the set  $\mathcal{C}_\kappa$  depends on the matrix  $\beta_k$ .

*Proof.* Let  $\{\lambda_1, \dots, \lambda_k\}$  be the eigenvalues of  $\beta_k$ , then in an orthonormal basis of eigenvectors for  $\beta_k$ , the matrix  $\beta_{k,\xi}$  is also diagonal, with eigenvalues  $\{\lambda_1 - \xi, \dots, \lambda_k - \xi\}$ . Then one has  $\|\beta_{k,\xi}.\eta\| > \kappa\|\eta\|$  for any  $\eta \in \mathbb{R}^k \setminus \{0\}$  provided that for all  $i \in \{1, \dots, k\}$ ,  $|\lambda_i - \xi| > \kappa$ , that is if  $\xi$  does not belong to

$$\mathcal{C}_\kappa = \bigcup_{i=1}^k [\lambda_i - \kappa, \lambda_i + \kappa].$$

The measure estimate  $\lambda(\mathcal{C}_\kappa) \leq 2k\kappa$  is trivial.  $\square$

With this lemma, the proof is now similar to that of Proposition A.3.

*Proof of Theorem A.8.* Let  $h(I) = \alpha.I + \beta I.I$ , and given  $\Lambda \in G^L(n, k)$ , we denote by  $\beta_\Lambda \in S_k(\mathbb{R})$  the matrix which represents the quadratic form  $\beta I.I$  restricted to the subspace  $\Lambda$ . Since the second derivative of  $h$  along any subspace is constant, then coming back to definition A.1 and using remark A.2,  $h \in SDM_{\gamma'}^{\tau'}$  if

$$\|\beta_\Lambda.\eta\| > \gamma' L^{-\tau'} \|\eta\|, \quad (1)$$

for any  $\Lambda \in G^L(n, k)$  and any  $\eta \in \mathbb{R}^k \setminus \{0\}$ . Let  $A_{\gamma'}^{\tau'}$  be the subset of  $S_n(\mathbb{R})$  whose elements contradict condition (1), that is

$$A_{\gamma'}^{\tau'} = \{\beta \in S_n(\mathbb{R}) \mid \|\beta_\Lambda.\eta\| \leq \gamma' L^{\tau'} \|\eta\|, \Lambda \in G^L(n, k), \eta \in \mathbb{R}^k \setminus \{0\}\},$$

and

$$A^{\tau'} = \bigcap_{\gamma' > 0} A_{\gamma'}^{\tau'}.$$

What we need to show is that  $A^{\tau'}$  has zero Lebesgue measure in  $S_n(\mathbb{R})$  provided  $\tau' > n^2 + 1$ .

Apply lemma A.9 to  $\beta_\Lambda \in S_k(\mathbb{R})$ , with  $\kappa = \gamma' L^{-\tau'}$ , to have a subset  $\mathcal{C}_{\gamma', \tau', L, \Lambda} \subseteq \mathbb{R}$  such that

$$\lambda(\mathcal{C}_{\gamma', \tau', L, \Lambda}) \leq 2k\gamma' L^{-\tau'}, \quad (2)$$

and for any  $\xi \notin \mathcal{C}_{\gamma', \tau', L, \Lambda}$ , the matrix  $\beta_{\Lambda, \xi} = \beta_\Lambda - \xi I_k$  satisfies

$$\|\beta_{\Lambda, \xi} \cdot \eta\| > \gamma' L^{-\tau'} \|\eta\|$$

for any  $\eta \in \mathbb{R}^k \setminus \{0\}$ . If we define

$$\mathcal{C}_{\gamma', \tau'} = \bigcup_{L \in \mathbb{N}^*} \bigcup_{k \in \{1, \dots, n\}} \bigcup_{\Lambda \in G^L(n, k)} \mathcal{C}_{\gamma', \tau', L, \Lambda},$$

then

$$\mathcal{C}_{\gamma', \tau'} = \{\xi \in \mathbb{R} \mid \beta_\xi = \beta - \xi I_n \in A_{\gamma'}^{\tau'}\}$$

and so

$$\mathcal{C}_{\tau'} = \bigcap_{\gamma' > 0} \mathcal{C}_{\gamma', \tau'} = \{\xi \in \mathbb{R} \mid \beta_\xi \in A^{\tau'}\}.$$

It remains to prove that the Lebesgue measure of  $\mathcal{C}_{\tau'}$  is zero, since by Fubini theorem, this will imply that the Lebesgue measure of  $A^{\tau'}$  is zero. By our estimate (2), we have

$$\begin{aligned} \lambda(\mathcal{C}_{\gamma', \tau'}) &\leq \sum_{L \in \mathbb{N}^*} \sum_{k=1}^n |G^L(n, k)| 2k\gamma' L^{-\tau'} \\ &\leq \sum_{L \in \mathbb{N}^*} \sum_{k=1}^n L^{n^2} 2k\gamma' L^{-\tau'} \\ &= 2 \left( \sum_{k=1}^n k \right) \left( \sum_{L \in \mathbb{N}^*} L^{n^2 - \tau'} \right) \gamma' \\ &= n(n+1) \left( \sum_{L \in \mathbb{N}^*} L^{n^2 - \tau'} \right) \gamma' \end{aligned}$$

and, since  $\tau' > n^2 + 1$ , the above series is convergent. Hence

$$\lambda(\mathcal{C}_{\tau'}) = \inf_{\gamma' > 0} \lambda(\mathcal{C}_{\gamma', \tau'}) = 0.$$

□

### *Acknowledgments.*

The author thanks Laurent Niederman for many useful and lengthy discussions on Nekhoroshev theory, Jacques Féjoz for a careful reading, Jean-Pierre Marco for his support and all the members of ASD team at “Observatoire de Paris”. The author is indebted to the referee for a very careful reading, and for pointing out some mistakes in a previous version.

## References

- [AKN06] V.I. Arnold, V.V. Kozlov, and A.I. Neishtadt, *Mathematical aspects of classical and celestial mechanics*, [Dynamical Systems III], Transl. from the Russian original by E. Khukhro, Third edition, Encyclopedia of Mathematical Sciences **3**, Springer-Verlag, Berlin, 2006.
- [Arn61] V.I. Arnol’d, *The stability of the equilibrium position of a Hamiltonian system of ordinary differential equations in the general elliptic case*, Sov. Math., Dokl. **2** (1961), 247–249.
- [Arn63] ———, *Small denominators and problems of stability of motion in classical and celestial mechanics*, Russ. Math. Surv. **18** (1963), no. 6, 85–191.
- [Bir66] G.D. Birkhoff, *Dynamical systems*, American Mathematical Society, Providence, R.I., 1966.
- [BN09] A. Bounemoura and L. Niederman, *Generic Nekhoroshev theory without small divisors*, Preprint (2009).
- [DG96] A. Delshams and P. Gutiérrez, *Estimates on invariant tori near an elliptic equilibrium point of a Hamiltonian system*, J. Differ. Equations **131** (1996), 277–303.
- [DLC83] R. Douady and P. Le Calvez, *Exemple de point fixe elliptique non topologiquement stable en dimension 4*, C. R. Acad. Sci. Paris **296** (1983), 895–898.
- [Dou88] R. Douady, *Stabilité ou instabilité des points fixes elliptiques*, Ann. Sci. Ec. Norm. Sup **21** (1988), no. 1, 1–46.
- [FGB98] F. Fassò, M. Guzzo, and G. Benettin, *Nekhoroshev-stability of elliptic equilibria of Hamiltonian systems*, Comm. Math. Phys. **197** (1998), no. 2, 347–360.
- [GDF<sup>+</sup>89] A. Giorgilli, A. Delshams, E. Fontich, L. Galgani, and C. Simó, *Effective stability for a Hamiltonian system near an elliptic equilibrium point, with an application to the restricted three body problem*, J. Differ. Equations **77** (1989), 167–198.

- [GFB98] M. Guzzo, F. Fassò, and G. Benettin, *On the stability of elliptic equilibria*, Math. Phys. Electron. J. **4** (1998), 16 pp., paper 1.
- [HSY92] B.R. Hunt, T. Sauer, and J.A. Yorke, *Prevalence: a translation-invariant “almost every” on infinite-dimensional spaces*, Bull. of the Amer. Math. Soc. **27** (1992), 217–238.
- [JV97] À. Jorba and J. Villanueva, *On the normal behaviour of partially elliptic lower dimensional tori of Hamiltonian systems*, Nonlinearity **10** (1997), 783–822.
- [KMOV04] V. Kaloshin, J. N. Mather, and E. Valdinoci, *Instability of resonant totally elliptic points of symplectic maps in dimension 4*, Loday-Richaud, Michèle (ed.), Analyse complexe, systèmes dynamiques, sommabilité des séries divergentes et théories galoisiennes. II. Volume en l’honneur de Jean-Pierre Ramis. Paris: Société Mathématique de France. Astérisque 297, 79-116, 2004.
- [KP94] S. Kuksin and J. Pöschel, *On the inclusion of analytic symplectic maps in analytic Hamiltonian flows and its applications*, Seminar on dynamical systems (1994), 96–116, Birkhäuser, Basel.
- [LM88] P. Lochak and C. Meunier, *Multiphase averaging for classical systems. With applications to adiabatic theorems. Transl. from the French by H. S. Dumas*, Applied Mathematical Sciences, 72, New York etc, Springer-Verlag. xi, 360 pp., 1988.
- [Loc92] P. Lochak, *Canonical perturbation theory via simultaneous approximation.*, Russ. Math. Surv. **47** (1992), no. 6, 57–133.
- [Loc95] Pierre Lochak, *Stability of Hamiltonian systems over exponentially long times: The near-linear case*, Dumas, H. S. (ed.) et al., Hamiltonian dynamical systems: history, theory, and applications. Proceedings of the international conference held at the University of Cincinnati, OH (USA), March 1992. New York, NY: Springer-Verlag. IMA Vol. Math. Appl. 63, 221-229, 1995.
- [MG95] A. Morbidelli and A. Giorgilli, *Superexponential stability of KAM tori*, J. Stat. Phys. **78** (1995), 1607–1617.
- [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.

- [Nie98] L. Niederman, *Nonlinear stability around an elliptic equilibrium point in a Hamiltonian system*, *Nonlinearity* **11** (1998), no. 6, 1465–1479.
- [Nie07] ———, *Generic exponential stability of quadratic integrable Hamiltonian systems*, Unpublished (2007).
- [OY05] W. Ott and J.A. Yorke, *Prevalence*, *Bull. of the Amer. Math. Soc.* **42** (2005), no. 3, 263–290.
- [PM03] R. Pérez-Marco, *Convergence or generic divergence of the Birkhoff normal form*, *The Annals of Mathematics* **157** (2003), no. 2, 557–574.
- [Pop00] G. Popov, *Invariant tori, effective stability, and quasimodes with exponentially small error terms. I. Birkhoff normal forms*, *Ann. Henri Poincaré* **1** (2000), no. 2, 223–248.
- [Pop04] ———, *KAM theorem for Gevrey Hamiltonians*, *Erg. Th. Dyn. Sys.* **24** (2004), no. 5, 1753–1786.
- [PT97] A.V. Pronin and D.V. Treschev, *On the inclusion of analytic maps into analytic flows*, *Regular and Chaotic Dynamics* **2** (1997), no. 2, 14–24.
- [PW94] A.D. Perry and S. Wiggins, *KAM tori are very sticky: rigorous lower bounds on the time to move away from an invariant Lagrangian torus with linear flow*, *Phys. D* **71** (1994), 102–121.
- [Pös82] J. Pöschel, *Integrability of Hamiltonian systems on Cantor sets*, *Commun. Pure Appl. Math.* **35** (1982), 653–696.
- [Pös93] ———, *Nekhoroshev estimates for quasi-convex Hamiltonian systems*, *Math. Z.* **213** (1993), 187–216.
- [Pös99] ———, *On Nekhoroshev’s estimate at an elliptic equilibrium*, *Internat. Math. Res. Notices* **4** (1999), 203–215.
- [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.
- [YC04] Y. Yomdin and G. Comte, *Tame geometry with application in smooth analysis*, *Lecture Notes in Mathematics*, Springer Verlag, Berlin, 2004.
- [Yom83] Y. Yomdin, *The geometry of critical and near-critical values of differentiable mappings*, *Math. Ann.* **264** (1983), 495–515.