Ergod. Th. & Dynam. Sys. (2007), **27**, 905–928
doi:10.1017/S0143385706000927

Prevalence of exponential stability among nearly integrable Hamiltonian systems

LAURENT NIEDERMAN
Université Paris XI, Topologie et Dynamique, UMR 8628 du CNRS, Bâtiment 425, 91405 Orsay Cedex, France
and
IMCCE, Astronomie et Systèmes Dynamiques, UMR 8028 du CNRS, 77 avenue Denfert-Rochereau, 75014 Paris, France
(e-mail: laurent.niederman@math.u-psud.fr)

(*Received* 21 *February* 2005 *and accepted in revised form* 18 *September* 2006)

Abstract. In the 1970s, Nekhorochev proved that, for an analytic nearly integrable Hamiltonian system, the action variables of the unperturbed Hamiltonian remain nearly constant over an exponentially long time with respect to the size of the perturbation, provided that the unperturbed Hamiltonian satisfies some generic transversality condition known as *steepness*. Recently, Guzzo has given examples of exponentially stable integrable Hamiltonians that are non-steep but satisfy a weak condition of transversality which involves only the affine subspaces spanned by integer vectors. We generalize this notion for an arbitrary integrable Hamiltonian and prove Nekhorochev's estimates in this setting. The point in this refinement lies in the fact that it allows one to exhibit a *generic* class of real analytic integrable Hamiltonians which are exponentially stable with *fixed* exponents. Genericity is proved in the sense of measure since we exhibit a prevalent set of integrable Hamiltonians which satisfy the latter property. This is obtained by an application of a quantitative Sard theorem given by Yomdin.

1. *Introduction*

One of the main problems in Hamiltonian dynamics is the stability of motions in nearly integrable systems (e.g. the n-body planetary problem). The main method of investigation is the construction of normal forms (see [**2**] or [**5**] for an introduction and a survey of these topics). This yields two kinds of theorem.

(i) The first kind of theorem proves results of stability over infinite times provided by the Kolmogorov–Arnold–Moser theory which are valid for solutions with initial conditions in a Cantor set of large measure, although no information is given on the other trajectories. Rüssmann ([**21**], see also [**3**] for a survey) has given a minimal non-degeneracy condition on the unperturbed Hamiltonian to ensure the persistence of invariant tori under perturbation. Namely, the image of the gradient map associated to the integrable Hamiltonian should not be included in a hyperplane and this condition is generic among real analytic real-valued functions.

(ii) The second kind (see, e.g., [**14, 15**]) has proved global results of stability over open sets of the following type.

Definition 1.1. (Exponential stability) Consider an open set $\Omega \subset \mathbb{R}^n$, an analytic integrable Hamiltonian $h : \Omega \longrightarrow \mathbb{R}$ and action-angle variables $(I, \varphi) \in \Omega \times \mathbb{T}^n$ where $\mathbb{T} = \mathbb{R}/\mathbb{Z}$.

For an arbitrary $\rho > 0$, let $\mathcal{O}_\rho$ be the space of analytic functions over a complex neighbourhood $\Omega_\rho \subset \mathbb{C}^{2n}$ of size ρ around $\Omega \times \mathbb{T}^n$ equipped with the supremum norm $\| \cdot \|_\rho$ over Ω_ρ.

We say that the Hamiltonian h is exponentially stable over an open set $\tilde{\Omega} \subset \Omega$ if there exist positive constants ρ, C_1, C_2, a, b and ε_0 which depend only on h and $\tilde{\Omega}$ such that:

(i) $h \in \mathcal{O}_\rho$;

(ii) for any function $\mathcal{H}(I, \varphi) \in \mathcal{O}_\rho$ such that $\|\mathcal{H} - h\|_\rho = \varepsilon < \varepsilon_0$, an arbitrary solution $(I(t), \varphi(t))$ of the Hamiltonian system associated to $\mathcal{H}$ with an initial action $I(t_0)$ in $\tilde{\Omega}$ is defined over a time $\exp(C_2/\varepsilon^a)$ and satisfies

$$\|I(t) - I(t_0)\| \leq C_1 \varepsilon^b \quad \text{for } |t - t_0| \leq \exp(C_2/\varepsilon^a) \qquad (\mathcal{E})$$

where a and b are called stability exponents.

Remark 1.2. Using a similar technique, the previous definition can be extended to an integrable Hamiltonian in the Gevrey class (see [**13**]).

In this paper, we prove that such a property of stability is generic according to the following theorem.

THEOREM 1.3. (Genericity of exponential stability) *Consider an arbitrary real analytic integrable Hamiltonian h defined on a neighbourhood of the closed ball $\bar{B}_R^{(n)}$ of radius R centred at the origin in $\mathbb{R}^n$.*

For almost any $\Omega \in \mathbb{R}^n$, the integrable Hamiltonian $h_\Omega(x) = h(I) - \Omega.I$ is exponentially stable with the exponents

$$a = \frac{b}{2 + n^2} \quad \textit{and} \quad b = \frac{1}{2(2 + (2n)^n)}.$$

In order to introduce the problem, we begin by a typical example of a *non-exponentially* stable integrable Hamiltonian: $h(I_1, I_2) = I_1^2 - I_2^2$. Indeed, a solution of the perturbed system governed by $h(I_1, I_2) + \varepsilon \sin(I_1 + I_2)$ with an initial action located on the first diagonal $(I_1(0) = I_2(0))$ admits a drift of the actions $(I_1(t), I_2(t))$ on a segment of length 1 over a timespan of order $1/\varepsilon$. Actually, with this example, we have the fastest possible drift of the action variables according to the magnitude ε of the perturbation.

The important feature in this example which has to be avoided in order to ensure exponential stability is the fact that the gradient $\nabla h(I_1, I_1)$ remains orthogonal to the first diagonal.

Equivalently, the gradient of the restriction of h on this first diagonal is identically zero.

https://doi.org/10.1017/S0143385706000927 Published online by Cambridge University Press

Nekhorochev [**14, 15**] introduced the class of *steep* functions where this problem is avoided. This property of steepness will be specified in §2 but this kind of function can be characterized by the following simple geometric criterion.

THEOREM 1.4. [**17**] *A real analytic real-valued function without critical points is steep if and only if its restriction to any proper affine subspace admits only isolated critical points.*

In this setting, Nekhorochev proved the following theorem.

THEOREM 1.5. [**14, 15**] *If h is real analytic, non-degenerate ($|\nabla^2 h(I)| \neq 0$ for any $I \in \Omega$) and steep then h is exponentially stable.*

The fundamental difference between our result of stability and the generic theorems of stability which can be ensured with Nekhorochev's original work is the *fixed* value of the exponents a and b in our Theorem 1.3.

Indeed, the set of steep functions is generic among sufficiently smooth functions. For instance, we have seen that the function $x^2 - y^2$ is not steep but it can be easily demonstrated that $x^2 - y^2 + x^3$ is steep and, usually, a given function can be transformed into a steep function by adding higher-order terms. Actually, let $J_r(n)$ be the space of r-jets of the $\mathcal{C}^\infty$ real-valued function of n variables; Nekhorochev [**14**] proved that the non-steep functions admit an r-jet in an algebraic set of $J_r(n)$ with a codimension that goes to infinity as r goes to infinity. The point is that Theorem 1.5 allows one to find a generic set of exponentially stable integrable Hamiltonians but with exponents of stability which are arbitrary *small* since they are related to the steepness indices (see Theorem 2.2). Therefore, one cannot obtain uniform exponents of stability for a generic set of integrable Hamiltonians.

Here, according to Theorem 1.3, *fixed* stability exponents are obtained on a *measure-theoretic* generic set. Actually, we exhibit a set of exponentially stable integrable Hamiltonians that are *prevalent* according to the terminology of Hunt *et al* [**8**] or Kaloshin [**9**]. The precise definition of prevalence will be given in the third section of this paper.

Different prevalent properties of dynamical systems have been proved in [**8–10, 18**] but, to the best of the author's knowledge, the only result of this kind for nearly integrable Hamiltonian systems is that of Perez-Marco [**19**] who proved that Birkhoff's normal forms are convergent or divergent for a generic set of nearly integrable Hamiltonians. Nevertheless, he uses a stronger notion of genericity than prevalence (see also §3 of the present paper).

The aim of the rest of this paper is to provide a proof of Theorem 1.3 and the paper is organized as follows. In §2, we state a result of exponential stability (Theorem 2.5) under a strictly weaker assumption than steepness, which involves only affine subspaces spanned by integer vectors. These affine subspaces will be called *rational subspaces*.

Actually, a necessary condition for exponential stability was given in [**17**] since for a real analytic integrable Hamiltonian which admits a restriction to a rational subspace with an accumulation of critical points, one can build arbitrary small perturbations which lead to a polynomial speed of drift of the action variables. On the other hand, the fact that these restrictions admit only isolated critical points is not a sufficient condition for exponential stability. However, there exists a large set of exponentially stable integrable Hamiltonian which are non-steep along an affine subspace spanned by *irrational* vectors. Indeed, Guzzo [**7**] has given such examples of integrable Hamiltonians: if $h(I_1, I_2) = I_1^2 - \delta I_2^2$ where δ is the square of a Diophantine number then its isotropic direction is the line directed by $(1, \sqrt{\delta})$ and this allows one to prove that h is exponentially stable.

We generalize this property for an arbitrary integrable Hamiltonian by introducing a condition of *Diophantine steepness* which is sufficient to ensure exponential stability (Theorem 2.5). This latter result is proved along the lines of a previous paper [**16**]; for the convenience of the reader its proof is given in Appendix A.

In §3, we show that a set of Diophantine steep functions with *fixed* indices is generic in a measure-theoretic sense ('prevalent') among sufficiently smooth functions defined over a relatively compact subset in $\mathbb{R}^n$.

Actually, by an application of the usual Sard's theorem, one can see easily that the Morse functions are prevalent in the Banach space $(\mathcal{C}^2(\bar{B}_R^{(n)}, \mathbb{R}), \|\cdot\|_{\mathcal{C}^2})$ where $\bar{B}_R^{(n)}$ is the closed ball of radius R centred at the origin in $\mathbb{R}^n$ (see §3).

Our prevalent set of Diophantine steep functions with fixed indices will be obtained by introducing the class of Diophantine Morse function (Definition 3.1.1). We prove its prevalence in $(\mathcal{C}^k(\bar{B}_R^{(n)}, \mathbb{R}), \|\cdot\|_{\mathcal{C}^k})$ for $k = 2n + 2$ in Corollary 3.2.7 thanks to reasonings similar to those used for the classical Morse functions but we have to substitute the Sard's theorem by a quantitative Morse–Sard theory developed by Yomdin [**22, 23**].

Moreover, we show that the Diophantine Morse functions are Diophantine steep with indices equal to *two*, and hence this class of functions yields the desired prevalent set of integrable Hamiltonians.

2. *Results of stability with a Diophantine steepness condition*

In order to specify the problem, we first give the original definition of a *steep* function and its consequences.

Definition 2.1. [**14, 15, 17**] Consider an open set Ω in $\mathbb{R}^n$; a real analytic function $f : \Omega \longrightarrow \mathbb{R}$ is said to be steep at a point $I \in \Omega$ along an affine subspace Λ which contains I if there exist constants $C > 0$, $\delta > 0$ and $p > 0$ such that along any continuous curve Γ in Λ connecting I to a point at a distance $r < \delta$ the norm of the projection of the gradient $\nabla f(x)$ onto the direction of Λ is greater than Cr^p at some point $\Gamma(t_*)$ with $\|\Gamma(t) - I\| \leq r$ for all $t \in [0, t_*]$.

The constants (C, δ) and p are respectively called the steepness coefficients and the steepness index.

Under the previous assumptions, the function f is said to be steep at the point $I \in \Omega$ if, for every $m \in \{1, \dots, n-1\}$, there exist positive constants C_m, δ_m and p_m such that f is steep at I along any affine subspace of dimension m containing I uniformly with respect to the coefficients (C_m, δ_m) and the index p_m.

Finally, a real analytic function f is steep over a domain $\mathcal{P} \subseteq \mathbb{R}^n$ with the steepness coefficients $(C_1, \dots, C_{n-1}, \delta_1, \dots, \delta_{n-1})$ and the steepness indices $(p_1, \dots, p_{n-1})$ if there are no critical points for f in $\mathcal{P}$ and f is steep at any point $I \in \mathcal{P}$ uniformly with respect to these coefficients and indices.

For instance, convex functions are steep with all the steepness indices equal to one. On the other hand, $f(x, y) = x^2 - y^2$ is a typical non-steep function but by adding a third-order term (e.g. y^3) we recover steepness. Moreover, this definition is minimal since a function can be steep along all subspaces of dimension lower than or equal to $m < n - 1$ and not steep for a subspace of dimension l greater than m (consider the function $f(x, y, z) = (x^2 - y)^2 + z$ at $(0, 0, 0)$ along all the lines and along the plane $z = 0$). Also, a quadratic form is steep if and only if it is sign definite. Then, one can prove the following theorem.

THEOREM 2.2. **[14, 15, 16]** *If h is real analytic, non-degenerate ($|\nabla^2 h(I)| \neq 0$ for any $I \in \mathcal{P}$) and steep then h is exponentially stable with the exponents*

$$a = b = \frac{1}{(2n-1)p_1 \cdots p_{n-1} + 1},$$

and hence a and b depend only on the steepness indices.

Now, we can state the weaker definition of a *Diophantine steep* function. For $m \in \{1, \ldots, n\}$, we denote by $\mathrm{Graff}_R(n, m)$ the m-dimensional affine Grassmannian over $\bar{B}_R^{(n)} \subset \mathbb{R}^n$ (i.e. the set of affine subspaces of dimension m in $\mathbb{R}^n$ which intersect the closed ball $\bar{B}_R^{(n)}$ of radius $R > 0$ around the origin) and $\mathrm{Graff}_R^K(n, m) \subset \mathrm{Graff}_R(n, m)$ is the set of rational subspaces of dimension m in $\mathbb{R}^n$ whose direction is spanned by integer vectors of length $\| \overrightarrow{k} \|_1 = |k_1| + \cdots + |k_n| \leq K$ for a given $K \in \mathbb{N}^*$.

Definition 2.3. A differentiable function f defined on a neighbourhood of $\bar{B}_R^{(n)} \subset \mathbb{R}^n$ is said to be (γ, τ)-Diophantine steep with two positive constants γ and τ if, for any $m \in \{1, \ldots, n\}$, there exist an index $p_m \geq 1$ and coefficients $C_m > 0$, $\delta_m > 0$ such that along any affine subspace $\Lambda_m \in \mathrm{Graff}_R^K(n, m)$ and any continuous curve Γ from $[0, 1]$ to $\Lambda_m \cap B_R$ with $\|\Gamma(0) - \Gamma(1)\| = r \leq \delta_m(\gamma/K^\tau)$, we have that

$$\text{there exists } t_* \in [0,1] \text{ such that } \begin{cases} \|\Gamma(0) - \Gamma(t)\| \leq r \quad \text{for all } t \in [0, t_*], \\ \|\mathrm{Proj}_{\overrightarrow{\Lambda}_m}(\nabla f(\gamma(t_*)))\| \geq C_m r^{p_m}, \end{cases} \tag{1}$$

where $\overrightarrow{\Lambda}_m$ is the direction of Λ_m.

Remark 2.4. (i) The space $\mathbb{R}^n$ is itself the only element of $\mathrm{Graff}_R^1(n, n)$. Therefore, along any arc in B_R of length $r \leq \delta_n \gamma$, there exists a point where the norm of the gradient ∇f is greater than or equal to $C_n r^{p_n}$ (the projection is reduced to the identity in this case).

(ii) With no loss of generality, we will assume that the coefficients $(C_1, \ldots, C_n)$ are equal to *one*. Indeed, the problem can always be reduced to this case by using steepness indices slightly greater than the optimal values.

We now describe the regularity of the perturbed Hamiltonian. Consider a nearly integrable Hamiltonian $\mathcal{H}(I, \varphi) = h(I) + \varepsilon f(I, \varphi)$, where $(I, \varphi) \in \mathbb{R}^n \times \mathbb{T}^n$ and $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ are action-angle variables of the integrable Hamiltonian h. We assume that $\mathcal{H}$ is *analytic* around a fixed complex neighbourhood $V_{r,s}\mathcal{P} \subset \mathbb{C}^{2n}$ of a real domain

https://doi.org/10.1017/S0143385706000927 Published online by Cambridge University Press

$\mathcal{P} = B_R \times \mathbb{T}^n \subset \mathbb{R}^n \times \mathbb{T}^n$ where B_R is the ball of radius R centred at the origin and

$$V_{r,s}\mathcal{P} = V_r(B_R) \times W_s(\mathbb{T}^n) = \{(I, \varphi) \in \mathbb{C}^{2n} \text{ such that } \mathrm{dist}(I, B_R) \leq r \text{ and } \mathrm{Re}(\varphi) \in \mathbb{T}^n; \mathrm{Max}_{j\in\{1,\dots,n\}}|\mathrm{Im}(\varphi_j)| \leq s\} \tag{2}$$

with $1 > r > 0$, $s > 0$ and the distance to B_R given by the Euclidean norm in $\mathbb{C}^n$.

Let $\|\cdot\|_{r,s}$ be the sup norm (L^∞) for real or vector-valued functions defined and bounded over $V_{r,s}\mathcal{P}$. We assume that $\|f\|_{r,s} \leq 1$ and that ε is a small parameter.

The Jacobian and the Hessian matrix are also assumed to be uniformly bounded with respect to the norm on the operators, also denoted by $\|\cdot\|$, induced by the Euclidean norm, i.e.

$$\text{there exists } M > 1, \text{ such that } \|\partial_I h(I)\|_{r,s} \leq M$$

and

$$\|\partial_I^2 h(I)\|_{r,s} \leq M \text{ for all } I \in V_r(B_R). \tag{3}$$

Under the previous assumptions, we can state the following result which will be proved in Appendix A.

THEOREM 2.5. *Let $\mathcal{H}(I,\varphi) = h(I) + \varepsilon f(I,\varphi)$ be a nearly integrable Hamiltonian analytic on the complex neighbourhood $V_{r,s}\mathcal{P} \subset \mathbb{C}^{2n}$ defined in (2) with an integrable part $h(I)$ which is (γ,τ)-Diophantine steep.*

Consider

$$\beta = \frac{1}{2(1 + n^n p_1 \cdots p_{n-1})}, \quad a = \frac{\beta}{1+\tau}, \quad b = \frac{\beta}{p_n}.$$

There exists a positive constant C which depends on n, M, R, s and τ but not on ε and γ such that for a small enough perturbation $\varepsilon \leq C\ \mathrm{Inf}(\gamma^{1/a}, \gamma^{1/b})$ and for any orbit of the perturbed system with initial conditions $(I(t_0), \varphi(t_0)) \in B_R \times \mathbb{T}^n$ far enough from the boundary of B_R, we have

$$\|I(t) - I(t_0)\| \leq (n+1)^2 \varepsilon^b \quad \text{for } |t| \leq \exp\left(\frac{s}{6}\varepsilon^{-a}\right).$$

Remark 2.6. (i) In this study, we do not look for accurate estimates of the exponents a and b as in [**16**], but we focus our attention on the most direct proof of the stability result in the Diophantine steep case. Actually, our ultimate goal is the existence of a uniform exponent of stability valid for a generic set of integrable Hamiltonians. The question of optimality is not very relevant in this problem since we use very general estimates and do not exploit the specificity of a given Hamiltonian.

(ii) The fact that the exponents of stability are *independent* of γ is crucial for our subsequent reasonings. On the other hand, the upper bound on the size of the perturbation in Theorem 2.5 depends on γ (this is reminiscent of KAM theory where the latter quantities depend on $\sqrt{\gamma}$).

The proof of Theorem 2.5 is based on reasonings already given in a previous paper [**16**] which rely on the construction of local resonant normal forms along each trajectory of the perturbed system together with the use of a simultaneous Diophantine approximation

https://doi.org/10.1017/S0143385706000927 Published online by Cambridge University Press

as in Lochak's proof [**11, 12**] of Nekhorochev's estimates. However, this study [**16**] is generalized in three directions. First, we substitute the original Nekhorochev's condition of steepness by our weak assumption of Diophantine steepness given above. Moreover, thanks to a construction of the non-resonant sets directly in the frequency space, we can remove the non-degeneracy condition on the frequency map ($|\nabla^2 h| \neq 0$) assumed in [**16**]. Finally, according to Remark 2.4, our integrable Hamiltonian h can admit critical points I (while $\nabla h(I) \neq 0$ was assumed in [**16**]) provided that h satisfies a global steepness condition on the full space $\mathbb{R}^n$. This last point is reminiscent of the notion of symmetrically steep (or S-steep) function considered by Nekhorochev [**14**].

3. *Genericity of Diophantine steepness among smooth functions*

First, any linear form $h(I) = \omega.I$ with a (γ, τ)-Diophantine vector $\omega \in \mathbb{R}^n$ is Diophantine steep with indices and coefficients equal to one. Hence, a linear form is almost always Diophantine steep while it cannot be steep according to Definition 2.1. At second order, one can prove that a quadratic form is almost always Diophantine steep with indices equal to two (it can be shown that for any quadratic form $q(I) = I^{t}AI$, for any $\tau > n^2$ and for almost all $\lambda \in \mathbb{R}$, there exists $\gamma > 0$ such that the modified quadratic form $q_\lambda(I) = I^{t}AI + \lambda\|I\|^2$ is (γ, τ)-Diophantine steep). We see that at first and second order, the set of Diophantine steep functions is much wider than the initial class of steep functions.

Starting from these examples, we look for a full measure set of Diophantine steep functions in the space of $\mathcal{C}^k$ real-valued functions defined on an open set in $\mathbb{R}^n$. Actually, a set in an infinite-dimensional space which is invariant by translation can be of zero measure only if it is a trivial set (see [**8**]). For this reason, Christensen [**4**], Hunt *et al* [**8**] and Kaloshin [**9**] have introduced a weak notion of a full measure set in an infinite-dimensional space called prevalence which corresponds to the usual property in a finite-dimensional space. In its simplest setting, a set $\mathcal{P}$ is said to be shy if there exists a *finite*-dimensional subspace F called a probe space such that any affine subspace of direction F intersects $\mathcal{P}$ along a zero measure set for the usual Lebesgue measure on this subspace. A set is prevalent if its complement is shy. Stronger notions of prevalence can be defined (see [**8, 9, 18, 19**]). For instance, Perez-Marco [**19**] considers sets which intersect *any* finite-dimensional affine subspace along a full measure set with respect to the finite-dimensional Lebesgue measure.

An example of a prevalent set is given by the Morse functions in the Banach space $(\mathcal{C}^2(\bar{B}_R^{(n)}, \mathbb{R}), \|\cdot\|_{\mathcal{C}^2})$ where $\bar{B}_R^{(n)}$ is the closed ball of radius R centred at the origin in $\mathbb{R}^n$. Indeed, for any function $f \in \mathcal{C}^2(\bar{B}_R^{(n)}, \mathbb{R})$, by an application of Sard's theorem on the gradient map ∇f one can prove that for almost any linear form $\omega \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ the function $f_\omega = f + \omega$ is Morse and the probe space is given by the linear forms. The modified function f_ω is called a morsification of f (see [**1**] and [**6**]).

Here we look for a set $\mathcal{P}$ of Diophantine steep functions with *fixed* indices (in order to obtain fixed exponents of stability according to Theorem 2.5) which is prevalent in $\mathcal{C}^k(\bar{B}_R^{(n)}, \mathbb{R})$ for a certain $k \in \mathbb{N}^*$. As mentioned previously, this will be obtained by introducing the class of Diophantine Morse functions (Definition 3.1.1) and proving its prevalence in $\mathcal{C}^{2n+2}(\bar{B}_R^{(n)}, \mathbb{R})$ thanks to the quantitative Morse–Sard theory developed by Yomdin [**22, 23**] together with reasonings similar to those used for the usual Morse

https://doi.org/10.1017/S0143385706000927 Published online by Cambridge University Press

functions. Moreover, we show that the Diophantine Morse functions are Diophantine steep with indices equal to two and these later ingredients yield our main theorem (Theorem 1.3). Finally, according to Yomdin, our estimates derived in Theorems 3.2.4 and 3.2.5 should be useful to locate the nearly critical points of a generic mapping (i.e. the problem of the 'organizing centre', see [**24**, p. 296]).

3.1. *Diophantine Morse functions.*

Definition 3.1.1. We denote by $\mathrm{Gr}(n,m)$ the set of all vectorial subspaces of dimension m in $\mathbb{R}^n$ and, for $K \in \mathbb{N}^*$, $\mathrm{Gr}_K(n,m) \subset \mathrm{Gr}(n,m)$ is the set of vectorial subspaces in $\mathbb{R}^n$ spanned by integer vectors of length $\|\overrightarrow{k}\|_1 = |k_1| + \cdots + |k_n| \leq K$; moreover $\mathrm{Gr}(n) = \bigcup_{m=1}^n \mathrm{Gr}(n,m)$ and $\mathrm{Gr}_K(n) = \bigcup_{m=1}^n \mathrm{Gr}_K(n,m)$.

A twice differentiable function $f \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$ defined on a neighbourhood of the closed ball $\bar{B}_R^{(n)} \subset \mathbb{R}^n$ of radius R centred at the origin is said to be (γ,τ)-Diophantine Morse with two positive constants γ and τ if, for any $K \in \mathbb{N}^*$, any $m \in \{1,\ldots,n\}$ and any $\Lambda \in \mathrm{Gr}_K(n,m)$, there exists $(e_1,\ldots,e_m)$ (respectively $(f_1,\ldots,f_{n-m})$), an orthonormal basis of Λ (respectively of $\Lambda^\perp$), such that the function

$$f_\Lambda(\alpha,\beta) := f(\alpha_1 e_1 + \cdots + \alpha_m e_m + \beta_1 f_1 + \cdots + \beta_{n-m} f_{n-m}), \tag{4}$$

which is twice differentiable on a neighbourhood of $\bar{B}_R^{(n)}$, satisfies

$$\text{for all } (\alpha,\beta) \in \bar{B}_R^{(n)} \text{ we have } \left\| \frac{\partial f_\Lambda}{\partial \alpha}(\alpha,\beta) \right\| > \frac{\gamma}{K^\tau}$$
$$\text{or } \left\| \frac{\partial^2 f_\Lambda}{\partial \alpha^2}\bigg|_{(\alpha,\beta)}(\eta) \right\| > \frac{\gamma}{K^\tau}\|\eta\| \quad (\text{for all } \eta \in \mathbb{R}^m).$$

The link between the Diophantine Morse functions and the Diophantine steep functions is given in the following theorem.

THEOREM 3.1.2. *With the previous notation, if a differentiable function $f \in \mathcal{C}^3(\mathbb{R}^n, \mathbb{R})$ defined on a neighbourhood of the closed ball $\bar{B}_{2R}^{(n)} \subset \mathbb{R}^n$ is (γ,τ)-Diophantine Morse for some positive constants γ and τ, then f is (γ,τ)-Diophantine steep over $\bar{B}_R$ with the coefficients $C_m = 1$, $\delta_m = 1/2M$ and the indices $p_m = 2$ for $m \in \{1,\ldots,n\}$.*

Remark 3.1.3. Our definition of Diophantine Morse function relies on the choice of an orthonormal basis in any subspaces $\Lambda \in \mathrm{Gr}_K(n)$ and the eigenvalues of the Hessian matrix which are extrinsic. However, the property of Diophantine steepness involves only the norm of the gradient ∇f since

$$\left\| \frac{\partial f_\Lambda}{\partial \alpha}(\alpha,\beta) \right\| = \|\mathrm{Proj}_\Lambda(\nabla f(\alpha_1 e_1 + \cdots + \alpha_m e_m + \beta_1 f_1 + \cdots + \beta_{n-m} f_{n-m}))\|$$

which does not depend of the considered orthonormal basis.

Proof. Consider $f \in \mathcal{C}^3(\mathbb{R}^n, \mathbb{R})$ with $\|f\|_{\mathcal{C}^3} \leq M$ for some $M \geq 1$ over $\bar{B}_{2R}^{(n)}$ such that

$$\left\| \frac{\partial f_\Lambda}{\partial \alpha}(\alpha, \beta) \right\| > \frac{\gamma}{K^\tau} \text{ or } \left\| \frac{\partial^2 f_\Lambda}{\partial \alpha^2}\bigg|_{(\alpha,\beta)}(\eta) \right\| > \frac{\gamma}{K^\tau}\|\eta\| \quad (\text{for all } \eta \in \mathbb{R}^m)$$

for all $(\alpha, \beta) \in \bar{B}_R^{(m)} \times \bar{B}_R^{(n-m)}$ with $\Lambda \in \mathrm{Gr}_K(n, m)$.

Then, for any continuous curve $\Gamma : [0, 1] \longrightarrow \bar{B}_R^{(m)}$ of length $r \leq \mathrm{Inf}((\gamma/2MK^\tau), 1)$, we have:

(i) either,

$$\left\| \frac{\partial f_\Lambda}{\partial \alpha}(\Gamma(0), \beta) \right\| > \frac{\gamma}{K^\tau} > r \geq r^2;$$

(ii) otherwise, for $\alpha \in \mathbb{R}^m$ such that $\|\alpha - \Gamma(0)\| < \gamma/2MK^\tau$ we have

$$\left\| \frac{\partial^2 f_\Lambda}{\partial \alpha^2}\bigg|_{(\alpha,\beta)} - \frac{\partial^2 f_\Lambda}{\partial \alpha^2}\bigg|_{(\Gamma(0),\beta)} \right\| < \frac{\gamma}{2K^\tau}$$

and

$$\left\| \frac{\partial^2 f_\Lambda}{\partial \alpha^2}\bigg|_{(\Gamma(0),\beta)}(\eta) \right\| > \frac{\gamma}{K^\tau}\|\eta\| \Longrightarrow \left\| \frac{\partial^2 f_\Lambda}{\partial \alpha^2}\bigg|_{(\alpha,\beta)}(\eta) \right\| > \frac{\gamma}{2K^\tau}\|\eta\| \quad \text{for all } \eta \in \mathbb{R}^m.$$

The mean value theorem gives

$$\left\| \frac{\partial f_\Lambda}{\partial \alpha}(\alpha, \beta) - \frac{\partial f_\Lambda}{\partial \alpha}(\Gamma(0), \beta) \right\| = \left\| \frac{\partial^2 f_\Lambda}{\partial \alpha^2}\bigg|_{(\alpha_*,\beta)}(\alpha - \Gamma(0)) \right\|$$

for some α_* on the segment that connects $\Gamma(0)$ and α, it implies

$$\left\| \frac{\partial f_\Lambda}{\partial \alpha}(\alpha, \beta) - \frac{\partial f_\Lambda}{\partial \alpha}(\Gamma(0), \beta) \right\| \geq \frac{\gamma}{2K^\tau}\|\alpha - \Gamma(0)\| \geq \|\alpha - \Gamma(0)\|^2.$$

Hence, we can ensure at least a variation of size r^2 on the norm

$$\left\| \frac{\partial f_\Lambda}{\partial \alpha}(\Gamma(t), \beta) \right\|$$

along any path Γ of length $r \leq \mathrm{Inf}(\gamma/2MK^\tau, 1)$.

Moreover, the choice of the orthonormal basis $(e_1, \ldots, e_m)$ gives

$$\frac{\partial f_\Lambda}{\partial \alpha}(\alpha, \beta) = \mathrm{Proj}_\Lambda(\nabla f(\alpha_1 e_1 + \cdots + \alpha_m e_m + \beta_1 f_1 + \cdots + \beta_{n-m} f_{n-m})).$$

Hence, for an arbitrary path $\tilde{\Gamma}$ of length $r \leq \mathrm{Inf}(\gamma/2MK^\tau, 1)$ in the affine subspace $x + \Lambda$ with $\Lambda \in \mathrm{Gr}_K(n, m)$ and $x \in \Lambda^\perp$, there exists $t_* \in [0, 1]$ such that

$$\|\mathrm{Proj}_\Lambda(\nabla f(\tilde{\Gamma}(t_*)))\| \geq r^2$$

and we can always choose this time t_* such that $\|\tilde{\Gamma}(t) - \tilde{\Gamma}(0)\| < r$ for all $t \in [0, t_*]$.

Finally, any rational subspace spanned by integer vectors of lengths bounded by $K \in \mathbb{N}^*$ can be seen as the sum $x + \Lambda$ for some $x \in \Lambda^\perp \cap \bar{B}_R^{(n)}$ with the direction $\Lambda \in \mathrm{Gr}_K(n)$.

Hence, the definition of Diophantine steepness for f over B_R is satisfied with the coefficients $C_m = 1$, $\delta_m = 1/2M$ and the index $p_m = 2$ for $m \in \{1, \ldots, n\}$. □

3.2. *Quantitative Morse–Sard theory and applications.* Now, an application of a quantitative version of Sard's theorem by Yomdin [**22**] allows us to show that, for a fixed $\tau > 0$ which is large enough, any sufficiently smooth function $f \in \mathcal{C}^p(\mathbb{R}^n, \mathbb{R})$ can be transformed into a (γ, τ)-Diophantine Morse function by adding almost any linear form.

We recall the main results of this Yomdin theory along the lines of a recent expository book of Yomdin and Comte [**23**]. For k, m and $n \in \mathbb{N}^*$ such that $m \leq n$, consider a mapping $g \in \mathcal{C}^{k+1}(\mathbb{R}^n, \mathbb{R}^m)$ defined on a neighbourhood of the closed ball $\bar{B}_R^{(n)} \subset \mathbb{R}^n$ for some radius $R > 0$ with the bound $\|g\|_{\mathcal{C}^{k+1}} = \mathcal{M} \geq 1$ for the usual $\mathcal{C}^{k+1}$-norm over $\mathcal{C}^{k+1}(\bar{B}_R^{(n)}, \mathbb{R}^m)$. With the previous assumptions, the quantity $R_k(g) = (\mathcal{M}/k!)R^{k+1}$ bounds the Taylor remainder term at order k over the closed ball $\bar{B}_R^{(n)}$.

For any matrix $A \in \mathcal{M}_{(m,n)}(\mathbb{R})$ with $1 \leq m \leq n$, the ordered singular values of A (i.e. the eigenvalues of $A^{\mathrm{t}}A$) are denoted by $0 \leq \lambda_1(A) \leq \cdots \leq \lambda_m(A)$ and, for any $x \in \bar{B}_R^{(n)}$, the singular values of $dg(x)$ are denoted by $\lambda_i(x)$ with $i \in \{1, \ldots, m\}$. In other words, $dg(x)$ maps the unit ball in $\mathbb{R}^n$ onto the ellipsoid of principal axes $0 \leq \lambda_1(x) \leq \cdots \leq \lambda_m(x)$ in $\mathbb{R}^m$. For $\lambda = (\lambda_1, \ldots, \lambda_m)$ with $0 \leq \lambda_1 \leq \cdots \leq \lambda_m$, the set $\Sigma(g, \lambda, \bar{B}_R^{(n)})$ of λ-critical points and the set $\Delta(g, \lambda, \bar{B}_R^{(n)})$ of λ-critical values are defined as

$$\Sigma(g, \lambda, \bar{B}_R^{(n)}) = \{x \in \bar{B}_R^{(n)} \text{ such that } \lambda_i(x) \leq \lambda_i, \text{ for } i = 1, \ldots, m\}$$

and

$$\Delta(g, \lambda, \bar{B}_R^{(n)}) = g(\Sigma(g, \lambda, \bar{B}_R^{(n)})).$$

Finally, for any relatively compact subset $\mathcal{A}$ in $\mathbb{R}^n$, we denote by $M(\varepsilon, \mathcal{A})$ the minimal number of closed balls of radius ε in $\mathbb{R}^n$ covering $\mathcal{A}$.

The cornerstone of the quantitative Sard theory is given in the following theorem.

THEOREM 3.2.1. [**22, 23**, Theorem 9.2] *With the previous notation and assumptions, with $\lambda_0 = 1$ and $\lambda = (\lambda_1, \ldots, \lambda_m)$, we have*

$$M(\varepsilon, \Delta(g, \lambda, \bar{B}_R^{(n)})) \leq c_e \sum_{j=0}^{m} \lambda_0\lambda_1 \cdots \lambda_j \left(\frac{R}{\varepsilon}\right)^j \quad \text{for } \varepsilon \geq R_k(g)$$

$$M(\varepsilon, \Delta(g, \lambda, \bar{B}_R^{(n)})) \leq c_i \sum_{j=0}^{m} \lambda_0\lambda_1 \cdots \lambda_j \left(\frac{R}{\varepsilon}\right)^j \left(\frac{R_k(g)}{\varepsilon}\right)^{(n-j)/(k+1)} \quad \text{for } \varepsilon \leq R_k(g)$$

where $c_i > 0$ and $c_e > 0$ depend only on n, m and k.

COROLLARY 3.2.2. *With the previous notation and assumptions, for any $\varepsilon \in]0, 1[$, we have*

$$M(\varepsilon, \Delta(g, \lambda, \bar{B}_R^{(n)})) \leq C \sum_{j=0}^{m} \lambda_0\lambda_1 \cdots \lambda_j \left(\frac{1}{\varepsilon}\right)^{(n+kj)/(k+1)}$$

where $C > 0$ depends only on $\mathcal{M}, R, n, m$ and k.

If $\mathrm{Neigh}_\varepsilon(\mathcal{A}) = \bigcup_{x \in \mathcal{A}} B(x, \varepsilon)$ for a set $\mathcal{A} \subset \mathbb{R}^m$ then $\mathrm{Neigh}_\varepsilon(\Delta(g, \lambda, \bar{B}_R^{(n)}))$ can be covered by $M(\varepsilon, \Delta(g, \lambda, \bar{B}_R^{(n)}))$ balls of radius 2ε and, for the m-dimensional Lebesgue measure, we have

$$\mathrm{Vol}(\mathrm{Neigh}_\varepsilon(\Delta(g, \lambda, \bar{B}_R^{(n)}))) \leq V(m)(2\varepsilon)^m M(\varepsilon, \Delta(g, \lambda, \bar{B}_R^{(n)}))$$

https://doi.org/10.1017/S0143385706000927 Published online by Cambridge University Press

where $V(m)$ is the volume of the m-dimensional unit ball. Finally,

$$\mathrm{Vol}(\mathrm{Neigh}_\varepsilon(\Delta(g,\lambda,\bar{B}_R^{(n)}))) \le \tilde{C}\sum_{j=0}^{m}\lambda_0\lambda_1\cdots\lambda_j\left(\frac{1}{\varepsilon}\right)^{[(n+kj)/(k+1)]-m}$$

for some constant $\tilde{C}$ which depends only on $\mathcal{M}$, R, n, m and k.

COROLLARY 3.2.3. *For $\delta \in]0,1[$ and $\varepsilon = \delta^{(k+1)/k}$, we denote*

$$\Delta_\delta = \Delta(g,(\delta,\mathcal{M},\ldots,\mathcal{M}),\bar{B}_R^{(n)}) \quad \text{and} \quad \tilde{\Delta}_\delta = \mathrm{Neigh}_\varepsilon(\Delta_\delta)$$

and we have the following bounds:

(i) $\mathrm{Vol}(\tilde{\Delta}_\delta) \le \bar{C}\delta^{(k+1-n)/k}$ *where $\bar{C} > 0$ depends only on $\mathcal{M}$, R, n, m and k;*

(ii) *for $k = 2n$, we have* $\mathrm{Vol}(\tilde{\Delta}_\delta) \le \bar{C}\delta^{(n+1)/2n}$.

Proof. Since $\varepsilon \in]0,1[$, we have

$$\begin{aligned}\mathrm{Vol}(\tilde{\Delta}_\delta) &\le \tilde{C}\varepsilon^{m-[n/(k+1)]} + \tilde{C}\sum_{j=1}^{m}\mathcal{M}^{j-1}\delta\varepsilon^{m-[(n+kj)/(k+1)]}\\ &\le \tilde{C}\left(\varepsilon^{m-[n/(k+1)]} + \frac{\mathcal{M}^m-1}{\mathcal{M}-1}\delta\varepsilon^{m-[(n+km)/(k+1)]}\right)\end{aligned}$$

and $m \ge 1$ implies

$$\mathrm{Vol}(\tilde{\Delta}_\delta) \le \bar{C}_1\varepsilon^{[(k+1)m-n]/(k+1)} + \bar{C}_2\delta\varepsilon^{(m-n)/(k+1)} \le \bar{C}_1\varepsilon^{(k+1-n)/(k+1)} + \bar{C}_2\delta\varepsilon^{(1-n)/(k+1)}.$$

Finally, the choice $\varepsilon = \delta^{(k+1)/k}$ yields $\mathrm{Vol}(\tilde{\Delta}_\delta) \le \bar{C}\delta^{(k+1-n)/k}$ and $k = 2n$ allows one to obtain the second estimate. □

THEOREM 3.2.4. *For $\kappa \in]0,1[$ and $g \in \mathcal{C}^{2n+1}(\bar{B}_R^{(n)},\mathbb{R}^m)$ with $\|g\|_{\mathcal{C}^{2n+1}} = \mathcal{M} \ge 1$, there exists a subset $\mathcal{C}_\kappa \subset \mathbb{R}^m$ such that*

$$\mathrm{Vol}(\mathcal{C}_\kappa) \le \bar{C}\sqrt{\kappa} \quad \textit{(with the constant } \bar{C} \textit{ considered in the previous corollary)}$$

and, for any $\omega \in \mathbb{R}^m\backslash\mathcal{C}_\kappa$, the function $g_\omega(x) = g(x) - \omega$ satisfies at any point $x \in \bar{B}_R^{(n)}$,

$$\|g_\omega(x)\| > \kappa \quad \textit{or} \quad \|dg_\omega(x)\zeta\| > \kappa\|\zeta\| \quad \textit{(for all } \zeta \in \mathbb{R}^n).$$

Proof. We choose $\mathcal{C}_\kappa = \tilde{\Delta}_\delta$ with $\delta = \kappa^{n/(n+1)}$, and hence

$$\mathrm{Vol}(\mathcal{C}_\kappa) \le \bar{C}\delta^{(n+1)/2n} = \bar{C}\sqrt{\kappa}.$$

Now, with our bound on $\|g\|_{\mathcal{C}^{2n+1}}$, we have $\lambda_i(x) \le \mathcal{M}$ for any $i \in \{2,\ldots,m\}$ and any $x \in \bar{B}_R^{(n)}$; hence

$$\begin{aligned}\Delta_\delta &= \{x \in \bar{B}_R^{(n)} \text{ such that } \lambda_1(x) \le \delta\}\\ &= \{x \in \bar{B}_R^{(n)} \text{ such that there exists } \zeta \in \mathbb{R}^n \text{ with } \|dg(x)\zeta\| \le \delta\|\zeta\|\}.\end{aligned}$$

Moreover $\varepsilon = \delta^{(2n+1)/2n} = \kappa^{(2n+1)/(2n+2)} > \kappa$ with $\kappa < 1$, so $\|g_\omega(x)\| \le \kappa$ implies $\|g_\omega(x)\| < \varepsilon$ and $g(x) \notin \Delta_\delta$ since $\mathrm{Dist}(\omega,\Delta_\delta) \ge \varepsilon$; hence $\|dg(x)\zeta\| > \delta\|\zeta\|$ for all $\zeta \in \mathbb{R}^n$.

Finally, $\delta = \kappa^{n/(n+1)} > \kappa$ yields

$$\|dg_\omega(x)\zeta\| = \|dg(x)\zeta\| > \delta\|\zeta\| > \kappa\|\zeta\| \quad (\text{for all } \zeta \in \mathbb{R}^m)$$

and we obtain the second estimate. □

https://doi.org/10.1017/S0143385706000927 Published online by Cambridge University Press

We now consider the constants $\gamma > 0$, $\tau > 0$ and an arbitrary function $f \in \mathcal{C}^{2n+2}(\mathbb{R}^n, \mathbb{R})$ defined on a neighbourhood of the closed ball $\bar{B}_R^{(n)} \subset \mathbb{R}^n$ with the bound $\|f\|_{\mathcal{C}^{2n+2}} = \mathcal{M} \geq 1$.

The previous theorem (Theorem 3.2.4) allows one to bound the measure of the set of values $\Omega \in \mathbb{R}^n$ such that the modified function $f(x) - \Omega.x$ is not (γ, τ)-Diophantine Morse.

More specifically, for any $(K, n, m) \in \mathbb{N}^3$ with $1 \leq m \leq n$ and any subspace $\Lambda \in \mathrm{Gr}_K(n, m)$, thanks to the choice of an orthonormal basis in Λ and $\Lambda^\perp$, the function f_Λ defined in (4) admits the upper bound $\|\partial_\alpha f_\Lambda\|_{\mathcal{C}^{2n+1}} \leq \|f\|_{\mathcal{C}^{2n+2}} = \mathcal{M}$ for the usual $\mathcal{C}^{2n+1}$-norm over $\mathcal{C}^{2n+1}(\bar{B}_R^{(n)}, \mathbb{R}^m)$.

THEOREM 3.2.5. *Consider $\nu \in \mathbb{N}^*$, $K \in \mathbb{N}^*$ and $\Lambda \in \mathrm{Gr}_K(n, m)$; there exists a subset $\mathcal{C}_\Lambda^{(\nu)} \subset \mathrm{B}_\nu^{(n)}$ where $\mathrm{B}_\nu^{(n)}$ is the open ball of radius ν centred at the origin in $\mathbb{R}^n$ with*

$$\mathrm{Vol}(\mathcal{C}_\Lambda^{(\nu)}) \leq \bar{C}_m^{(\nu)} \sqrt{\frac{\gamma}{K^\tau}}$$

where the constant $\bar{C}_m^{(\nu)}$ depends only on $n, m, \mathcal{M}, R$ and ν such that, for any $\Omega \in \mathcal{B}_\nu^{(n)} \backslash \mathcal{C}_\Lambda^{(\nu)}$, the modified function $f_\Omega(x) = f(x) - \Omega.x$ satisfies, at any point $x \in \bar{B}_R^{(n)}$,

$$\|\partial_\alpha f_{(\Lambda,\Omega)}(\alpha, \beta)\| \geq \frac{\gamma}{K^\tau} \quad \text{or} \quad \|\partial_\alpha^2 f_{(\Lambda,\Omega)}(\alpha, \beta)\eta\| \geq \frac{\gamma}{K^\tau}\|\eta\| \quad (\text{for all } \eta \in \mathbb{R}^m).$$

(The function $f_{(\Lambda,\Omega)}$ is defined with respect to f_Ω along the lines of f_Λ with respect to f in the definition of a Diophantine Morse function.)

Proof. We apply Theorem 3.2.4 with the constant $\kappa = \gamma/K^\tau$ on the function $g(\alpha, \beta) = \partial_\alpha f_\Lambda(\alpha, \beta) \in \mathcal{C}^{2n+1}(\mathbb{R}^n, \mathbb{R}^m)$ in order to obtain a nearly critical set $\mathcal{C}_\kappa \subset \mathbb{R}^m$.

Then, for $\Omega \in \mathbb{R}^n$ such that $\mathrm{Proj}_\Lambda(\Omega) = \omega_1 e_1 + \cdots + \omega_m e_m$ with $\omega = (\omega_1, \ldots, \omega_m) \notin \mathcal{C}_\kappa$, the function $f_\Omega(x) = h(x) - \Omega.x$ satisfies $\partial_\alpha f_{(\Lambda,\Omega)}(\alpha, \beta) = \partial_\alpha f_\Lambda(\alpha, \beta) - \omega = g_\omega(\alpha, \beta)$ and

$$\|g_\omega(\alpha, \beta)\| = \|\partial_\alpha f_{(\Lambda,\Omega)}(\alpha, \beta)\| \geq \frac{\gamma}{K^\tau} \quad \text{or} \quad \|dg_\omega(\alpha, \beta)\zeta\| \geq \frac{\gamma}{K^\tau}\|\zeta\| \quad (\text{for all } \zeta \in \mathbb{R}^n)$$

but the differential $\partial_\alpha^2 f_{(\Lambda,\Omega)}(\alpha, \beta) = \partial_\alpha^2 f_\Lambda(\alpha, \beta)$ is the restriction of dg to the subspace $\mathbb{R}^m \times \{0\} \subset \mathbb{R}^n$ and admits the same lower bound on its singular values as $dg = dg_\omega$.

Next, we consider the set

$$\mathcal{C}_\Lambda^{(\nu)} = \{\Omega \in \mathrm{B}_\nu^{(n)} \text{ such that } \mathrm{Proj}_\Lambda(\Omega) = \omega_1 e_1 + \cdots + \omega_m e_m \\ \text{with } \omega = (\omega_1, \ldots, \omega_m) \in \mathcal{C}_\kappa\}$$

and we have the estimate

$$\mathrm{Vol}(\mathcal{C}_\Lambda^{(\nu)}) = \mathrm{Vol}(\mathrm{Proj}_\Lambda^{-1}(\mathcal{C}_\kappa) \cap \mathrm{B}_\nu^{(n)}) \leq \mathrm{Vol}(\mathrm{B}_\nu^{(n)})\, \mathrm{Vol}(\mathcal{C}_\kappa) \leq V(n)\nu^n \bar{C} \sqrt{\frac{\gamma}{K^\tau}}$$

where $V(n)$ is the volume of the unit ball in $\mathbb{R}^n$ and $\bar{C}$ is the constant in Theorem 3.2.4 computed for a function $g \in \mathcal{C}^{2n+1}(\mathbb{R}^n, \mathbb{R}^m)$ which depends only on $\mathcal{M}, R, n$ and m.

Finally,

$$\mathrm{Vol}(\mathcal{C}_\Lambda^{(\nu)}) = \bar{C}_m^{(\nu)}\sqrt{\frac{\gamma}{K^\tau}},$$

where the constant $\bar{C}_m^{(\nu)}$ depends only on $n, m, \mathcal{M}, R$ and ν. □

THEOREM 3.2.6. *Consider an arbitrary constant $\tau > 2(n^2+1)$ and a function $f \in \mathcal{C}^{2n+2}(\bar{B}_R^{(n)}, \mathbb{R})$ defined on a neighbourhood of the closed ball $\bar{B}_R^{(n)} \subset \mathbb{R}^n$. Then, for almost any $\Omega \in \mathbb{R}^n$ there exists $\gamma > 0$ such that the function $f_\Omega(x) = f(I) - \Omega.I$ is (γ, τ)-Diophantine Morse over $\bar{B}_R^{(n)}$.*

Proof. For any $\nu \in \mathbb{N}^*$ and $K \in \mathbb{N}^*$, we consider the set $\mathcal{C}_K^{(\nu)} = \bigcup_{m=1}^n \bigcup_{\Lambda \in \mathrm{Gr}_K(n,m)} \mathcal{C}_\Lambda^{(\nu)}$; by an application of Theorem 3.2.5, we obtain

$$\mathrm{Vol}(\mathcal{C}_K^{(\nu)}) \le \sum_{m=1}^n \mathrm{Card}(\mathrm{Gr}_K(n,m))\bar{C}_m^{(\nu)}\sqrt{\frac{\gamma}{K^\tau}} \le \bigg(\sum_{m=1}^n \bar{C}_m^{(\nu)}\bigg)K^{n^2}\sqrt{\frac{\gamma}{K^\tau}}.$$

Now, for fixed $\gamma > 0$, the set $\mathcal{C}_\gamma^{(\nu)} = \bigcup_{K\in\mathbb{N}^*} \mathcal{C}_K^{(\nu)}$ satisfies

$$\mathrm{Vol}(\mathcal{C}_\gamma^{(\nu)}) \le \bigg(\sum_{m=1}^n \bar{C}_m^{(\nu)}\bigg)\bigg(\sum_{K\in\mathbb{N}^*} K^{n^2-\tau/2}\bigg)\sqrt{\gamma}$$

and this upper bound is convergent with our assumption on τ.

For $\mathcal{C}^{(\nu)} = \bigcap_{\gamma>0} \mathcal{C}_\gamma^{(\nu)}$ we have $\mathrm{Vol}(\mathcal{C}^{(\nu)}) = 0$ and $\mathcal{C} = \bigcup_{\nu\in\mathbb{N}^*} \mathcal{C}^{(\nu)}$ satisfies $\mathrm{Vol}(\mathcal{C}) = 0$.

Finally, for any $\Omega \in \mathbb{R}^n \backslash \mathcal{C}$, the function $f_\Omega(x) = f(x) - \Omega.x$ is (γ, τ)-Diophantine Morse over $\bar{B}_R^{(n)}$ for some $\gamma > 0$ and we can choose $\tau = 2n^2 + 3 > 2(n^2+1)$. □

COROLLARY 3.2.7. (Prevalence of the Diophantine Morse functions) *The set of $(\gamma, 2n^2+3)$-Diophantine Morse functions for some $\gamma > 0$ is prevalent in $\mathcal{C}^{2n+2}(\mathbb{R}^n, \mathbb{R})$.*

Proof. In the previous theorem (Theorem 3.2.6), we can choose $\tau = 2n^2 + 3 > 2(n^2+1)$ and, for almost any $\Omega \in \mathbb{R}^n \backslash \mathcal{C}$, the function $f_\Omega(x) = f(x) - \Omega.x$ is (γ, τ)-Diophantine Morse over $\bar{B}_R^{(n)}$ for some $\gamma > 0$. This is exactly from the definition of a prevalent set with the probe space given by the linear forms. □

3.3. *End of the proof of the main result (Theorem 1.3).* Returning to the dynamics, our result of exponential stability (Theorem 2.5) together with the prevalence of Diophantine Morse functions (Corollary 3.2.7) implies that for an arbitrary real analytic integrable Hamiltonian h and for almost all linear forms $\omega \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$, the modified Hamiltonian $h_\omega(x) = h(x) + \omega(x)$ is exponentially stable with fixed exponents of stability (since the latter quantities depend only on the steepness indices). Indeed, for almost any $\Omega \in \mathbb{R}^n$ there exists $\gamma > 0$ such that the integrable Hamiltonian $h_\Omega(I) = h(I) + \Omega.I$ is (γ, τ)-Diophantine Morse with $\tau = 3 + 2n^2$. Hence, according to Theorem 3.1.2 the integrable Hamiltonian h_Ω is $(\gamma, 3 + 2n^2)$-Diophantine steep with indices equal to two and finally Theorem 2.5 ensures that h_Ω is exponentially stable with the desired exponents. □

Acknowledgements. The author wishes to thank François Ledrappier for his careful reading and his remarks which significantly improved the paper.

https://doi.org/10.1017/S0143385706000927 Published online by Cambridge University Press

A. *Appendix. Exponential stability with a Diophantine steepness condition*

A.1. *Description of our proof.* Our proof is based on the following simple algebraic property. Let $\omega \in \mathbb{R}^n$ be a rational vector, i.e. ω is a multiple of a vector with integer components. In such a case, the scalar products $|k.\omega|$ for $k \in \mathbb{Z}^n$ such that $k.\omega \neq 0$ admit a lower bound $\ell > 0$. Then, let $\omega \in \mathbb{R}^n$ be a rational vector and $K \in \mathbb{N}^*$ a positive integer, there exists a small neighbourhood V of ω which depends on K such that $|k.\omega'| \geq \ell/2$ for any $\omega' \in V$ and all $k \in \mathbb{Z}^n \backslash \langle\omega\rangle^\perp$ with $\|k\|_1 = |k_1| + \cdots + |k_n| \leq K$. Moreover, if we find a second rational vector $\tilde{\omega} \in V$, then the scalar products $|k.\tilde{\omega}|$ admit a uniform lower bound for all $k \in \mathbb{Z}^n \backslash \{\langle\omega\rangle^\perp \cap \langle\tilde{\omega}\rangle^\perp\}$ and $\|k\|_1 \leq K$. If ω and $\widetilde{\omega}$ are linearly independent, we also have $\mathrm{Dim}(\langle\omega\rangle^\perp \cap \langle\tilde{\omega}\rangle^\perp) = n-2$.

Along these lines, we can ensure that if we find a sequence $(\omega_1, \ldots, \omega_n)$ of close enough rational vectors which are linearly independent (i.e. $(\omega_1, \ldots, \omega_n)$ form a basis of $\mathbb{R}^n$), then all the scalar products $|k.\omega_n|$ admit a uniform lower bound for $k \in \mathbb{Z}^n$ and $\|k\|_1 < K$ with $K \in \mathbb{N}^*$.

Now, consider a trajectory of the perturbed system starting at a time t_0 which admits an increasing sequence of times $t_0 \leq t_1 \leq \cdots \leq t_n$ for some constant $K \in \mathbb{N}^*$ such that each frequency vector $\nabla h(I(t_k))$ is close to a rational vector ω_k for each $k \in \{1, \ldots, n\}$. Assume that $(\omega_1, \ldots, \omega_n)$ is a basis of $\mathbb{R}^n$ composed of rational vectors which are close enough to one another to satisfy the previous algebraic property with the constant K. Then, $I(t_n)$ is located in a resonance-free area up to some finite order and a local integrable normal form can be built up to an exponentially small remainder allowing one to confine the actions.

Our result of stability (Theorem 2.5) is proved by contradiction in the following way. Assume that a solution of the perturbed system starting at an initial time t_0 admits a drift of the action variables over an exponentially long time. Then, the Diophantine steepness of the integrable Hamiltonian ensures that, for a small enough perturbation, the sequence of times $(t_1, \ldots, t_n)$ and the basis $(\omega_1, \ldots, \omega_n)$ can be build recursively. Hence the actions are confined, which gives the desired contradiction.

The closeness of $\nabla h(I(t_k))$ to ω_k for $k \in \{1, \ldots, n\}$ is given by an application of a classical theorem of Dirichlet for a simultaneous Diophantine approximation which yields a minimal rate of approximation of an arbitrary vector by a rational one. This last argument gives an upper bound on the order K of normalization which can be carried out and imposes our value of the stability exponents a and b.

A.2. *Normal forms.* In order to avoid cumbersome expressions, we use the notation $u \preccurlyeq\!* \, v$ (respectively $u \,{*}\!\preccurlyeq v$, $u \asymp\!* \, v$, $u \,{*}\!\asymp v$) if there exists $0 < \mathbf{C} \leq 1$ such that $u < \mathbf{C}v$ (respectively $u\mathbf{C} < v$, $u = \mathbf{C}v$ or $u\mathbf{C} = v$) and the constant $\mathbf{C}$ depends only on the dimension n, the bound M, the radius R, the analyticity width s and the exponent τ but not on the small parameters ε and γ. We consider the perturbed Hamiltonian $\mathcal{H}$ which is holomorphic over the domain $V_{r,s}\mathcal{P}$ defined in (2).

Let Λ be a sublattice of $\mathbb{Z}^n$ and $K \in \mathbb{N}^*$. A subset $\mathcal{D} \subset B_R \subset \mathbb{R}^n$ is said to be (α, K)-non-resonant modulo Λ if, at every point $I \in \mathcal{D}$, we have

$$|k.\nabla h(I)| = |k.\omega(I)| \geq \alpha \quad \text{for all } I \in \mathcal{D} \text{ and } k \in \mathbb{Z}^n_K \backslash \Lambda$$

where $\mathbb{Z}^n_K = \{k \in \mathbb{Z}^n \text{ such that } \|k\|_1 \leq K\}$ with a fixed $K \in \mathbb{N}^*$.

In the neighbourhood of such a set $\mathcal{D}$, the perturbed Hamiltonian $\mathcal{H}$ can be put in a Λ-resonant normal form $h + g + f_*$ where the Fourier expansion of g contains only harmonics in $\mathbb{Z}^n_K \cap \Lambda$ while the remainder f_* is a small general term.

More specifically, we will consider the set

$$V_{r,s}\mathcal{D} = V_r(\mathcal{D}) \times W_s(\mathbb{T}^n) = \{(I,\varphi) \in \mathbb{C}^{2n} \text{ such that } \operatorname{dist}(I,\mathcal{D}) \leq r \text{ and } \operatorname{Re}(\varphi) \in \mathbb{T}^n;\ \operatorname{Max}_{j\in\{1,\dots,n\}}|\operatorname{Im}(\varphi_j)| \leq s\}$$

equipped with the supremum norm $\|\cdot\|_{r,s}$ for real or vector-valued functions defined and bounded over $V_{r,s}\mathcal{D}$. Using this notation we obtain the following lemma.

LEMMA A.2.1. (Normal form [**20**]) *Suppose that $\mathcal{D} \subset B_R$ is (α, K)-non-resonant modulo Λ and that the following inequalities hold:*

$$\varepsilon \preccurlyeq * \frac{\alpha r}{K}, \quad r \preccurlyeq * \operatorname{Min}\left(\frac{\alpha}{K}, R\right), \quad \frac{6}{s} \leq K. \tag{A.1}$$

Then we can define an holomorphic, symplectic transformation $\Phi : V_{r_*,s_*}\mathcal{D} \longmapsto V_{r,s}\mathcal{D}$ where $r_* = r/2$, $s_* = s/6$ which is one-to-one and real-valued for real variables such that the pull-back of $\mathcal{H}$ by Φ is a Λ-resonant normal form $\mathcal{H} \circ \Phi = h + g + f_*$ up to a remainder f_* with

$$\|g\|_{r_*,s_*} * \preccurlyeq \varepsilon \quad \textit{and} \quad \|f_*\|_{r_*,s_*} * \preccurlyeq \varepsilon \exp\left(-\frac{sK}{6}\right).$$

Moreover, $\|\Pi_I \circ \Phi - \mathrm{Id}_I\|_{r_*,s_*} \preccurlyeq * r/6$ uniformly over $V_{r_*,s_*}\mathcal{D}$ where Π_I denotes the projection onto the action space and Id_I is the identity in the action space. Hence,

$$V_{r/3}\mathcal{D} \subset \Pi_I(\Phi(V_{r_*,s_*}\mathcal{D})) \subset V_{2r/3}\mathcal{D}.$$

COROLLARY A.2.2. *With the notation of the previous lemma, consider a solution of the normalized system governed by $\mathcal{H} \circ \Phi$ and a time $t_k \in \mathbb{R}$. Let λ_k be the affine subspace which contains $I(t_k)$ and whose direction $\Lambda \otimes \mathbb{R}$ is the vector space spanned by Λ; then*

$$\operatorname{dist}(I(t), \lambda_k) = \|I(t) - \lambda_k\| * \preccurlyeq \varepsilon \quad \textit{for } |t - t_k| \leq \exp\left(\frac{sK}{6}\right) \quad \textit{and} \quad |t| < \mathcal{T}_* \tag{A.2}$$

where $\mathcal{T}_*$ is the time of escape of $V_{r_*,s_*}\mathcal{D}$.

Proof. We denote by $\mathcal{Q}$ the orthogonal projection on $\langle\Lambda\rangle^\perp$. Since $H \circ \Phi$ is in Λ-resonant normal form, we have $(d/dt)\mathcal{Q}(I(t)) = -\mathcal{Q}(\partial_\varphi f_*)$ and by the Cauchy inequality

$$\|I(t) - \lambda_k\| \leq \|\mathcal{Q}(I(t) - I(t_k))\| \leq \|\mathcal{Q}(\partial_\varphi f_*(I,\varphi))\|_{r_*,s_*}|t - t_k| * \preccurlyeq 6\frac{\varepsilon}{s}$$

provided that $|t - t_k| \leq \exp(sK/6)$. □

A.3. *Nearly periodic tori, non-resonant areas and approximation.* A vector $\omega \in \mathbb{R}^n$ is said to be *rational* if there is $t > 0$ such that $t\omega \in \mathbb{Z}^n$, in which case $\mathcal{T} = \operatorname{Inf}\{t > 0/t\omega \in \mathbb{Z}^n\}$ is called the *period* of ω. Consider $\varrho > 0$ and $\omega \in \mathbb{R}^n \setminus \{0\}$ a rational vector of period $\mathcal{T}$, the set

$$\mathcal{B}_\varrho(\omega) = \{I \in B_R \text{ such that } \|\nabla h(I) - \omega\| < \varrho\}$$

is called a *nearly periodic torus*.

https://doi.org/10.1017/S0143385706000927 Published online by Cambridge University Press

THEOREM A.3.1. *For $K > 0$ and $\varrho > 0$ such that $2K\varrho\mathcal{T} < 1$, the set $\mathcal{B}_\varrho(\omega)$ is $(1/2\mathcal{T}, K)$-non-resonant modulo the $\mathbb{Z}$-module Λ spanned by $\mathbb{Z}^n_K \cap \langle\omega\rangle^\perp$.*

Proof. First we have the following lemma.

LEMMA A.3.2. *Let Ω be the hyperplane $\langle\omega\rangle^\perp$, then for all $k \in \mathbb{Z}^n \backslash \Omega$ we have $|k.\omega| \geq 1/\mathcal{T}$.*

Proof. We have $|k.\omega| = 1/\mathcal{T}|k.\mathcal{T}\omega| = 1/\mathcal{T}|k.\alpha|$ for some $\alpha \in \mathbb{Z}^n$ and $|k.\alpha| \neq 0$ since $k \notin \langle\alpha\rangle^\perp = \langle\omega\rangle^\perp$. Hence $|k.\alpha| \geq 1$ and $|k.\omega| \geq 1/\mathcal{T}$. □

Then $\dim(\Lambda) \leq n-1$ since $\omega \neq 0$ and for all $I \in \mathcal{B}_\varrho(\omega)$, for all $k \in \mathbb{Z}^n_K \backslash \Lambda$, we have

$$|k.\nabla h(I)| \geq |k.\omega| - |k\|\nabla h(I) - \omega\| \geq \frac{1}{\mathcal{T}} - K\varrho > \frac{1}{2\mathcal{T}},$$

according to our threshold in Theorem A.3.1. □

Now, for an integer $m \in \{1, \ldots, n\}$, consider a decreasing sequence of positive real numbers $\varrho_1 \geq \cdots \geq \varrho_m$ and m rational vectors $(\omega_1, \ldots, \omega_m)$ in $\mathbb{R}^n$ with respective periods $(\mathcal{T}_1, \ldots, \mathcal{T}_m)$ such that

$$\|\omega_{j+1} - \omega_j\| \leq \varrho_j \quad \text{for all } j \in \{1, \ldots, m-1\}.$$

We denote by Ω_j the hyperplanes $\langle\omega_j\rangle^\perp$ and by $\mathcal{I}_j$ the sets $\Omega_1 \cap \cdots \cap \Omega_j$, for $j \in \{1, \ldots, m\}$.

Consider a positive constant K so then the $\mathbb{Z}$-module (respectively the $\mathbb{R}$-vector space) spanned by $\mathbb{Z}^n_K \cap \mathcal{I}_j$ is denoted by Λ_j (respectively $\Lambda_j \otimes \mathbb{R}$).

LEMMA A.3.3. *With the previous notation, if*

$$2(m-j+1)K\varrho_j\mathcal{T}_j < 1 \quad (\textit{for all } j \in \{1, \ldots, m\}), \tag{A.3}$$

then the nearly periodic tori

$$\mathcal{B}_j = \{I \in B_R \text{ such that } \|\nabla h(I) - \omega_j\| < (m-j+1)\varrho_j\}$$

are $(1/2\mathcal{T}_j, K)$-non-resonant modulo Λ_j for $j \in \{1, \ldots, m\}$.

Proof. Consider $j \in \{2, \ldots, m\}$ and $I \in \mathcal{B}_j$; since the sequence $(\varrho_l)_{1 \leq l < j}$ is decreasing, we have $\|\omega_l - \omega_j\| \leq \varrho_l + \cdots + \varrho_{j-1} \leq (j-1-l)\varrho_l$ for all $l \in \{1, \ldots, j-1\}$ and the assumption $\|\nabla h(I) - \omega_j\| \leq (m-j+1)\varrho_j$ yields

$$\|\nabla h(I) - \omega_l\| \leq (m-l)\varrho_l \leq (m-l+1)\varrho_l \quad \text{for all } l \in \{1, \ldots, j-1\}.$$

Then, the argument of the previous lemma (Lemma A.3.2) ensures that, for all $I \in \mathcal{B}_j$, for all $l \in \{1, \ldots, j\}$ and for all $k \in \mathbb{Z}^n_K \backslash \Omega_l$, we have

$$|k.\nabla h(I)| \geq \frac{1}{\mathcal{T}_l} - (m-l+1)K\varrho_l \geq \frac{1}{2\mathcal{T}_l}$$

with the thresholds (A.3).

Hence, for all $k \in \mathbb{Z}^n_K \backslash \Lambda_j$, the scalar products $|k.\nabla h(I)|$ are lowered by $1/2\mathcal{T}_j$ for any $I \in \mathcal{B}_j$. □

https://doi.org/10.1017/S0143385706000927 Published online by Cambridge University Press

Below, we will need the following direct corollary of Dirichlet's theorem on the simultaneous Diophantine approximation (see Lochak [**11**]).

LEMMA A.3.4. *For any $x \in \mathbb{R}^n$ and any $Q \in \mathbb{N}^*$, there exists a rational vector x^* of period $\mathcal{T}$ which satisfies*

$$\|x^* - x\| \le \frac{\sqrt{n-1}}{\mathcal{T} Q^{1/(n-1)}} \quad \textit{with} \ \frac{1}{\|x\|_\infty} \le \mathcal{T} \le \frac{Q}{\|x\|_\infty} \tag{A.4}$$

for the Euclidean norm $\|\cdot\|$ and the maximum of the components $\|\cdot\|_\infty$.

Proof. We can renumber the indices in such a way that $x = \xi(\pm 1, x')$ for some $x' \in \mathbb{R}^{n-1}$ and $\xi = \|x\|_\infty$.

The question is now reduced to an approximation in $\mathbb{R}^{n-1}$. Indeed, Dirichlet's theorem yields $q \in \mathbb{N}^*$ and $l' \in \mathbb{Z}^{n-1}$ such that $1 \le q < Q$ and $\|qx' - l'\|_\infty \le Q^{-1/(n-1)}$.

If $x^* = \xi(\pm 1, l'/q)$, we have

$$\|x^* - x\|_\infty \le \frac{\xi}{q} Q^{-1/(n-1)} \Longrightarrow \|x^* - x\| \le \sqrt{n-1}\frac{\xi}{q} Q^{-1/(n-1)}$$

for the euclidean norm.

One can check easily that x^* is a rational vector of period $\mathcal{T} = q/\xi$ which satisfies the desired claim. □

A.4. *Fitted sequence.* We first prove that the existence of a sequence of rational vectors as described in the previous section along a trajectory implies that the actions are confined. Then, in the following sections, we show that a drift of the action variables implies the existence of such a sequence which gives a contradiction.

Hence, Theorem 2.5 of exponential stability would be proved.

We study the perturbed system governed by the Hamiltonian $\mathcal{H}$ which is holomorphic over the domain $V_{r,s}\mathcal{P}$ defined in (2).

For $1 \le m \le n$, let $(\omega_1, \ldots, \omega_m)$ be a sequence of rational vectors.

As previously, for $j \in \{1, \ldots, m\}$ we define $\Omega_j = \langle \omega_j \rangle^\perp$ and $\mathcal{I}_j = \Omega_1 \cap \cdots \cap \Omega_j$; also let Λ_j be the $\mathbb{Z}$-module spanned by $\mathbb{Z}^n_K \cap \mathcal{I}_j$, and finally $d_j = \dim(\Lambda_j \otimes \mathbb{R})$.

Definition A.4.1. (Fitted sequence) For $m \in \{1, \ldots, n\}$, consider an integer $K \ge 6^m/s$. A sequence of rational vectors $(\omega_1, \ldots, \omega_m)$ is called a fitted sequence of order K for a solution $(I(t), \varphi(t))$ with an initial time t_0 if there exists:

(1) an increasing sequence of times such that

$$t_0 \le t_1 \le \cdots \le t_m \le t_0 + \exp(sK/6);$$

(2) a decreasing sequence of radii $R \ge r_0 \ge \cdots \ge r_m$;

(3) a decreasing sequence of domains

$$\mathcal{D}_k = \{I \in \mathcal{D}_{k-1} \text{ such that } \|\nabla h(I) - \omega_{k-1}\| < 4M(m-k+1)r_k\} \quad \text{and} \quad \mathcal{P}_k = \mathcal{D}_k \times \mathbb{T}^n$$

for $k \in \{1, \ldots, m\}$ with $\mathcal{D}_0 = \mathrm{B}_R$;

https://doi.org/10.1017/S0143385706000927 Published online by Cambridge University Press

(4) holomorphic, symplectic transformations $\Phi_k : V_{r_*^{(k)},s_k}\mathcal{D}_k \longmapsto V_{r_k,s_{k-1}}\mathcal{D}_k$ where $(r_*^{(k)}, s_k) = (r_k/2, s/6^k)$ which are one-to-one and real-valued for real variables with

$$\|\Pi_I \circ \Phi_k - \mathrm{Id}_I\|_{r_*^{(k)},s_k} \preccurlyeq * \frac{r_k}{6} \quad \text{for } k \in \{1,\ldots,m\}.$$

Then, the application $\Psi_k^{-1} = \Phi_1 \circ \cdots \circ \Phi_k$ is defined over $\mathcal{D}_k \times \mathbb{T}^n$ and we assume that $\mathcal{H} \circ \Psi_k^{-1}$ is in Λ_k-resonant normal form up to a remainder of order $\varepsilon \exp(-sK/6)$ as in Lemma A.2.1. From now on, we will denote by $I^{(k)} = \pi_I \circ \Psi_k(I,\varphi)$ the averaged actions under the transformation Ψ_k and $\Delta_k = \|I^{(k)} - I\|_{r_*^{(k)},s_k}$; hence $\Delta_k \preccurlyeq * r_1 + \cdots + r_k$ for $k \in \{1,\ldots,m\}$. Finally, the following three properties should hold:

(i) $\|\nabla h(I^{(k-1)}(t_k)) - \omega_k\| \le M r_k$ for $k \in \{1,\ldots,m\}$ with $I^{(0)} = I$;

(ii) $\|I^{(k)}(t) - I^{(k)}(t_k)\| \le r_k$ for $t \in [t_k, t_{k+1}]$ with $k \in \{0,\ldots,m-1\}$;

(iii) the dimensions $(d_1,\ldots,d_m)$ satisfy: $d_1 > \cdots > d_j > \cdots > d_m = 0$.

THEOREM A.4.2. *Consider a trajectory which admits a fitted sequence of order K.*
If $I(t_0) \in B_{R/2}$ and the threshold $\varepsilon \preccurlyeq * r_m$ is satisfied, then we have

$$\|I(t) - I(t_0)\| \le (n+1)^2 r_0 \quad \textit{for } t_0 \le t \le t_0 + \exp(sK/6). \tag{A.5}$$

Remark A.4.3. Our construction goes forward in time but the same results are valid backward in time.

Proof. First, we have $\|I(t) - I(t_0)\| \le r_0 < (m+1)^2 r_0$ for $t \in [t_0, t_1]$ and

$$\Delta_k \preccurlyeq * r_1 + \cdots + r_k \Longrightarrow \Delta_k \preccurlyeq * k r_1 \quad \text{for } k \in \{1,\ldots,m\},$$

which implies that, for all $t \in [t_k, t_{k+1}]$,

$$\begin{aligned}\|I(t) - I(t_k)\| &\le \|I(t) - I^{(k)}(t)\| + \|I^{(k)}(t) - I^{(k)}(t_k)\| + \|I^{(k)}(t_k) - I(t_k)\| \\ &\le 2\Delta_k + r_k \preccurlyeq * 2kr_1 + r_k \quad (\text{for all } k \in \{1,\ldots,n-1\}).\end{aligned}$$

Finally, the domain of definition of $\mathcal{D}_m$ contains a neighbourhood of $I(t_m)$ and Ψ_m transform the considered Hamiltonian $\mathcal{H}$ to an integrable one up to a perturbation of magnitude $\varepsilon \exp(-sK/6)$ since d_m is equal to 0.

Hence, $\mathcal{H} \circ \Psi_m^{-1}(I^{(m)}, \varphi^{(m)}) = h_m(I^{(m)}) + f_m(I^{(m)}, \varphi^{(m)})$ and an application of Cauchy's inequality on the real domain $\mathcal{P}_m$ yields

$$\left\|\frac{\partial f_m}{\partial \varphi^{(m)}}\right\|_{\mathcal{P}_m} * \preccurlyeq \frac{\varepsilon}{s_m} \exp\left(-\frac{sK}{6}\right).$$

Then, for $t \in [t_m, t_0 + \exp(sK/6)]$, the threshold $\varepsilon \preccurlyeq * r_m$ implies

$$\|I^{(m)}(t) - I^{(m)}(t_m)\| \le |t - t_m| \left\|\frac{\partial f_m}{\partial \varphi^{(m)}}\right\|_{\mathcal{P}_m} \le r_m$$

and hence

$$\begin{aligned}\|I(t) - I(t_m)\| &\le \|I(t) - I^{(m)}(t)\| + \|I^{(m)}(t) - I^{(m)}(t_m)\| + \|I^{(m)}(t_m) - I(t_m)\| \\ &\le 2\Delta_m + r_m \preccurlyeq * 2mr_1 + r_m\end{aligned}$$

https://doi.org/10.1017/S0143385706000927 Published online by Cambridge University Press

$$\begin{aligned}\Longrightarrow \|I(t)-I(t_0)\| &\leq \|I(t)-I(t_m)\|+\|I(t_m)-I(t_{m-1})\|+\cdots+\|I(t_1)-I(t_0)\|\\ &\leq 2(\Delta_m+\cdots+\Delta_1)+r_m+r_{m-1}+\cdots+r_1+r_0\\ &\preccurlyeq\!* \,2[m+(m-1)+\cdots+1]r_1+mr_1+r_0 \preccurlyeq\!*\,(m+1)^2 r_0\end{aligned}$$

since $r_1 \geq r_0$, and this yields the required inequality since $m \leq n$. □

A.5. *Formal construction of a fitted sequence.* Here, we assume that there is a sufficiently high density of rational vectors and look for the relations which should be satisfied by the parameters $\varepsilon, \tau, \gamma, K, s$, the radii $(R, r_0, \ldots, r_m)$ and the periods $(\mathcal{T}_1, \ldots, \mathcal{T}_m)$ to ensure the existence of a sequence fitted to a trajectory which admits a drift of the action variables as a result of our Diophantine steepness condition.

LEMMA A.5.1. *Consider two constants $0 < \tau$, $0 < \gamma < 1$, an integer $K \geq 6^n/s$ and a solution of the perturbed system with some initial condition $(I(t_0), \varphi(t_0))$ in $B_{R/2} \times \mathbb{T}^n$ such that the action variables admit the following drift:*

$$\textit{there exists } t_* \in [t_0, t_0+\exp(cK)] \quad \text{with} \quad \|I(t_*)-I(t_0)\| = (n+1)^2 r_0$$

for $0 < r_0 < R/[2(n+1)^2]$.

A sequence $(\omega_1, \ldots, \omega_m)$ of rational vectors is a fitted sequence of order K for the considered solution if the radii $(r_1, \ldots, r_m)$ and the periods $(\mathcal{T}_1, \ldots, \mathcal{T}_m)$ satisfy the following relations:

(i) $r_1 < \gamma/K^\tau$;

(ii) $\varepsilon \preccurlyeq\!* \,1/2M(r_k/2)^{\rho_k}$ *for* $k \in \{1, \ldots, m-1\}$ *where we denote* $\rho_k = p_{d_k}$;

(iii) $r_{k+1} < (1/6M)(r_k/2)^{\rho_k}$ *for* $k \in \{1, \ldots, m-1\}$;

(iv) $8M(m-k+1)Kr_k\mathcal{T}_k < 1$ *for* $k \in \{1, \ldots, m\}$;

(v) *the thresholds (A.1) are satisfied with the parameters* ε, K, r_k *and* $\alpha_k = 1/2\mathcal{T}_k$.

Proof.

First step. We *assume* the existence of a $\mathcal{T}_1$-periodic rational vector ω_1 such that $\|\nabla h(I^{(k)}(t_1)) - \omega_1\| \leq Mr_1$ for a given time $t_1 \in [t_0, t_0+\exp(sK/6)]$ which will be determined explicitly in the next section.

Consider the domain $\mathcal{D}_1 = \{I \in \mathrm{B}_R$ such that $\|\nabla h(I) - \omega_1\| < 4Mmr_1\}$. With our threshold (iv) in Lemma A.5.1, $\mathcal{D}_1$ is $(1/2\mathcal{T}_1, K)$-non-resonant modulo Λ_1. Then the last condition of Lemma A.5.1 implies the existence of a normalization Φ_1 with respect to Λ_1 from $V_{r_*^{(1)},s_1}\mathcal{D}_1$ to $V_{r_1,s_0}\mathcal{D}_1$ and $\Psi_1 = \Phi_1^{-1}$ is the desired transformation.

Iterative step. Assume that an increasing sequence of times $t_0 \leq t_1 \leq \cdots \leq t_k < t_0+\exp(cK)$ and a sequence of periodic vectors $(\omega_1, \ldots, \omega_k)$ with respective periods $(\mathcal{T}_1, \ldots, \mathcal{T}_k)$ which satisfy the assumptions of a fitted sequence have been built up to order $k \in \{1, \ldots, n-1\}$. We denote the projection $\mathrm{Proj}_{\Omega_k} = \mathrm{Proj}_{\overrightarrow{\Lambda_k}}$ (respectively $\mathrm{Proj}_{\langle\omega_k\rangle}$) by $\mathcal{Q}_k$ (respectively $\tilde{\mathcal{Q}}_k$). According to Corollary A.2.2, one can see that the normalized actions $I^{(k)}$ satisfy

$$\|I^{(k)}(t)-I^{(k)}(t_k)-\mathcal{Q}_k(I^{(k)}(t)-I^{(k)}(t_k))\| = \|\tilde{\mathcal{Q}}_k(I^{(k)}(t)-I^{(k)}(t_k))\| *\!\preccurlyeq \varepsilon \tag{A.6}$$

for $t \in [t_k, \mathrm{Inf}(t_0+\exp(cK); t_k^*)]$ where t_k^* is the time of escape of $\mathcal{D}_k$.

If $\|\mathcal{Q}_k(I^{(k)}(t) - I^{(k)}(t_k))\| < r_k/2$ for all $t \in]t_k, t_0 + \exp(cK)[$, the inequality (A.6) and the threshold (ii) of Lemma A.5.1 imply

$$\|I^{(k)}(t) - I^{(k)}(t_k)\| \le \frac{1}{2M}\left(\frac{r_k}{2}\right)^{\rho_k} + \frac{r_k}{2} \le r_k \quad \text{since } r_k < 1,\ \rho_k \ge 1 \text{ and } M \ge 1.$$

Then, as in the proof of Theorem A.4.2, we have

$$\|I(t) - I(t_k)\| \le 2\Delta_k + r_k \preccurlyeq * \, 2(r_1 + \cdots + r_k) + r_k \quad \text{for all } t \in]t_k, t_0 + \exp(cK)[$$

which yields

$$\|I(t_*) - I(t_0)\| \le 2(\Delta_1 + \cdots + \Delta_k) + r_1 + \cdots + r_k + r_0 \preccurlyeq * \,(k+1)^2 r_0 < (n+1)^2 r_0$$

while we have assumed $\|I(t_*) - I(t_0)\| = (n+1)^2 r_0$. Hence, there is an escape time

$$t_{*k} \in]t_k, t_0 + \exp(cK)[\quad \text{with } \|\mathcal{Q}_k(I^{(k)}(t_{*k}) - I^{(k)}(t_k))\| = \frac{r_k}{2}.$$

Since $I^{(k)}(t_k) + \mathcal{Q}_k(I^{(k)}(t) - I^{(k)}(t_k))$ is a continuous path in the subspace $\lambda_k = I^{(k)}(t_k) + \Lambda_k \otimes \mathbb{R}$, the steepness of h yields $t_{k+1} \in [t_k, t_{*k}]$ such that

$$\begin{cases} \|\mathrm{Proj}_{\vec{\Lambda}_k}(\nabla h(I^{(k)}(t_k) + \mathcal{Q}_k(I^{(k)}(t_{k+1}) - I^{(k)}(t_k))))\| \ge \left(\dfrac{r_k}{2}\right)^{\rho_k}, \\ \|\mathcal{Q}_k(I^{(k)}(t) - I^{(k)}(t_k))\| \le \dfrac{r_k}{2} \quad \text{for all } t \in [t_k, t_{k+1}]. \end{cases} \tag{A.7}$$

Moreover, the inequality (A.6) and threshold (ii) in Lemma A.5.1 together with $\rho_k \ge 1$ imply that

$$\|\nabla h(I^{(k)}(t_{k+1})) - \nabla h(I^{(k)}(t_k) + \mathcal{Q}_k(I^{(k)}(t_{k+1}) - I^{(k)}(t_k)))\| \le \frac{M}{2M}\left(\frac{r_k}{2}\right)^{\rho_k} = \frac{1}{2}\left(\frac{r_k}{2}\right)^{\rho_k}$$

and

$$(A.7) \Longrightarrow \|\mathcal{Q}_k(\nabla h(I^{(k)}(t_{k+1})))\| = \|\mathrm{Proj}_{\vec{\Lambda}_k}(\nabla h(I^{(k)}(t_{k+1})))\| \ge \frac{1}{2}\left(\frac{r_k}{2}\right)^{\rho_k}.$$

In the same way, (A.6) and (A.7) allow one to prove that

$$\|\mathcal{Q}_k(I^{(k)}(t) - I^{(k)}(t_k))\| \le \frac{r_k}{2} \Longrightarrow \|I^{(k)}(t) - I^{(k)}(t_k)\| \le r_k \quad \text{for all } t \in [t_k, t_{k+1}].$$

Finally, we *assume* the existence of a $\mathcal{T}_{k+1}$-periodic rational vector ω_{k+1} such that $\|\nabla h(I^{(k)}(t_{k+1})) - \omega_{k+1}\| \le M r_{k+1}$ which implies that

$$\begin{aligned} \|\omega_k - \omega_{k+1}\| &\le \|\omega_k - \nabla h(I^{(k-1)}(t_k))\| + \|\nabla h(I^{(k-1)}(t_k)) - \nabla h(I^{(k)}(t_k))\| \\ &\quad + \|\nabla h(I^{(k)}(t_k)) - \nabla h(I^{(k)}(t_{k+1}))\| + \|\nabla h(I^{(k)}(t_{k+1})) - \omega_{k+1}\| \\ &\Longrightarrow \|\omega_k - \omega_{k+1}\| \le 3Mr_k + Mr_{k+1} \le 4Mr_k. \end{aligned} \tag{A.8}$$

Moreover, threshold (iii) of Lemma A.5.1 implies that

$$\|\mathcal{Q}_k(\nabla h(I^{(k)}(t_{k+1})) - \omega_{k+1})\| \le \|\nabla h(I^{(k)}(t_{k+1})) - \omega_{k+1}\| \le M r_{k+1} < \frac{M}{6M}\left(\frac{r_k}{2}\right)^{\rho_k}$$

https://doi.org/10.1017/S0143385706000927 Published online by Cambridge University Press

and

$$\|\mathcal{Q}_k(\omega_{k+1})\| \geq \|\mathcal{Q}_k(\nabla h(I^{(k)}(t_{k+1})))\| - \|\mathcal{Q}_k(\nabla h(I^{(k)}(t_{k+1})) - \omega_{k+1})\| \geq \frac{1}{3}\left(\frac{r_k}{2}\right)^{\rho_k},$$

and hence ω_{k+1} is not orthogonal to the previous rational vectors and $d_{k+1} < d_k$.

Consider the domain

$$\mathcal{D}_{k+1} = \{I \in \mathcal{D}_k \text{ such that } \|\nabla h(I) - \omega_{k+1}\| < 4M(m-k)r_{k+1}\},$$

the inequality (A.8), threshold (iv) in the definition of a fitted sequence and Lemma A.3.3 together with the distances $\varrho_k = 4Mr_k$ ensure that $\mathcal{D}_{k+1}$ is $(1/2\mathcal{T}_{k+1}, K)$-non-resonant modulo Λ_{k+1}. Finally, threshold (v) of Lemma A.5.1 and Lemma A.2.1 imply the existence of a normalization Φ_{k+1} with respect to Λ_{k+1} from $V_{r_*^{(k+1)},s_{k+1}}\mathcal{D}_{k+1}$ to $V_{r_{k+1},s_k}\mathcal{D}_{k+1}$ and the desired transformation is given by $\Psi_{k+1} = \Phi_{k+1}^{-1} \circ \Psi_k$. □

A.6. *Complete construction of a fitted sequence.* Here, we tackle the problem of Diophantine approximation of the frequency vectors $(\nabla h(I(t_1)), \nabla h(I^{(1)}(t_2)), \ldots, \nabla h(I^{(n-1)}(t_n)))$ which was the missing ingredient in the previous section.

LEMMA A.6.1. *Consider two constants $0 < \tau$, $0 < \gamma < 1$, an integer $K \geq 6^n/s$ and a solution of the perturbed system with some initial condition $(I(t_0), \varphi(t_0))$ in $B_{R/2} \times \mathbb{T}^n$ such that the action variables admit the following drift:*

$$\textit{there exists } t_* \in [t_0, t_0 + \exp(cK)] \quad \textit{with } \|I(t_*) - I(t_0)\| = (n+1)^2 r_0$$

for $0 < r_0 < R/[2(n+1)^2]$.

Assume that

$$0 < r_0 < \operatorname{Inf}\left(\gamma, \frac{R}{2(n+1)^2}\right); \quad r_1 \leq \frac{r_0^{\rho_0}}{4M}$$

where we denote $\rho_0 = p_n$. Then, in the previous construction of a fitted sequence, for all sequences of strictly positive constants $(Q_1, \ldots, Q_m)$ there exists a $\mathcal{T}_k$-periodic rational vector ω_k which satisfies

$$\|\nabla h(I^{(k-1)}(t_k)) - \omega_k\| \leq \frac{\sqrt{n-1}}{\mathcal{T}_k Q_k^{1/(n-1)}} \textit{ with } \frac{1}{M} \leq \mathcal{T}_k \leq \frac{4Q_k}{(n+1)r_0^{\rho_0}} \quad \textit{for } k \in \{1, \ldots, m\}. \tag{A.9}$$

Proof. With our drift of the action variable, the Diophantine steepness assumption together with $r_0 \leq \gamma \leq 1$ and $\rho_0 = p_n \geq 1$ allow one to find a time $t_1 \in [0, t_*]$ such that

$$\|\nabla h(I(t_1))\| \geq (\tfrac{1}{2}(n+1)^2 r_0)^{\rho_0} \geq \tfrac{1}{2}(n+1)^2 r_0^{\rho_0}$$

and $\|I(t) - I(t_0)\| \leq \frac{1}{2}(n+1)^2 r_0$ for all $t \in [t_0, t_1]$.

Moreover, with our bound $M \geq 1$ on the norm of the Hessian matrix, we obtain

$$\|\nabla h(I)\| \geq \frac{(n+1)^2}{4} r_0^{\rho_0} \quad \text{for all } \|I - I(t_1)\| \leq \frac{(n+1)^2}{4M} r_0^{\rho_0} \leq \frac{(n+1)^2}{4} r_0.$$

In the regular case, we remove this first step since we can use a uniform lower bound on the gradient $\nabla h(I)$.

https://doi.org/10.1017/S0143385706000927 Published online by Cambridge University Press

Now, we can ensure that for any $I \in B(I(t_1), [(n+1)^2/4M]r_0^{\rho_0})$ we have

$$\frac{n+1}{4} r_0^{\rho_0} \leq \frac{\|\nabla h(I)\|}{n} \leq \|\nabla h(I)\|_\infty \leq \|\nabla h(I)\| \leq M$$

and Lemma A.3.4 applied to $x = \nabla h(I(t))$ implies that, for any $Q > 0$, there exists a rational vector ω of period $\mathcal{T}$ which satisfies

$$\|\omega - \nabla h(I)\| \leq \frac{\sqrt{n-1}}{\mathcal{T} Q^{1/(n-1)}} \quad \text{with } \frac{1}{M} \leq \mathcal{T} \leq \frac{4Q}{(n+1)r_0^{\rho_0}}. \tag{A.10}$$

Then, a repeat of the arguments in the proof of Theorem A.4.2 shows that, for $k \in \{1, \ldots, n\}$,

$$\|I^{(k)}(t) - I(t_1)\| \leq (k+1)^2 r_1$$

if $t_k \leq t \leq t_{k+1}$ with $t_{m+1} = t_0 + \exp(sK/6)$ and $r_1 \leq r_0^{\rho_0}/4M$ implies that $I(t_k) \in B(I(t_1), [(n+1)^2/4M]r_0)$ for all $k \in \{1, \ldots, m\}$, and hence (A.10) implies (A.9). □

From now on, for a fitted sequence of length m, we will denote

$$\pi_0 = 1 \text{ and } \pi_k = p_{d_1} \cdots p_{d_k} \quad \text{where } d_j = \dim(\Lambda_j \otimes \mathbb{R}) \text{ and } k \in \{1, \ldots, m\}. \tag{A.11}$$

With the previous lemmas, one can prove the following theorem.

THEOREM A.6.2. *There exists a sufficiently small positive constant C such that for an arbitrary trajectory of the perturbed system which admits a drift of the action variables as in the previous lemmas, if*

$$\beta = \frac{1}{2(1 + n^n p_1 \ldots p_{n-1})}, \quad a = \frac{\beta}{1+\tau}, \quad b = \frac{\beta}{\rho_0} \quad \text{and} \quad \varepsilon < C\gamma^{1/a}, \varepsilon < C\gamma^{1/b}, \tag{A.12}$$

then one can find a fitted sequence of length $m \in \{1, \ldots, n-1\}$ for the considered orbit.

The parameters of this sequence are

$$K = \mathrm{E}[\varepsilon^{-a}] + 1 \text{ and } r_0 = \varepsilon^b;\, r_{k+1} = \frac{\sqrt{n-1}}{M} \frac{\varepsilon^{\beta n^k \pi_k}}{\mathcal{T}_{k+1}} \quad \text{for any } k \in \{0, \ldots, m-1\} \tag{A.13}$$

where $\mathrm{E}[x]$ is the integer part of $x \in \mathbb{R}$.

Proof. In order to build a fitted sequence, our parameter K, the radii $(r_1, \ldots, r_m)$ and the periods $(\mathcal{T}_1, \ldots, \mathcal{T}_m)$ should satisfy the following thresholds:

$$(i)\ 1 *\preccurlyeq K; \ldots$$

Summary of the thresholds

(i) $1 *\preccurlyeq K$;

(ii) $r_0 < \dfrac{R}{2(n+1)^2}$;

(iii) $0 < r_0 < \mathrm{Inf}\left(\gamma, \dfrac{R}{2(n+1)^2}\right)$;

(iv) $r_1 < \dfrac{\gamma}{K^\tau}$;

https://doi.org/10.1017/S0143385706000927 Published online by Cambridge University Press

(v) $\varepsilon \preccurlyeq\!* \, r_k^{\rho_k}$ for $k \in \{1, \ldots, m-1\}$;
and for $k \in \{1, \ldots, m\}$:
(vi) $r_k \preccurlyeq\!* \, r_{k-1}^{\rho_{k-1}}$;
(vii) $\varepsilon T_k K \preccurlyeq\!* \, r_k$;
(viii) $K T_k r_k \preccurlyeq\!* \, 1$.

Here, we apply Lemma A.6.1 with the bounds

$$Q_{k+1} = \varepsilon^{-(n-1)\beta n^k \pi_k} \quad \text{for } k \in \{0, \ldots, m-1\},$$

and hence we have the upper bounds $\mathcal{T}_{k+1} \leq [4/(n+1)]\varepsilon^{-\beta-(n-1)\beta n^k \pi_k}$ on the periods. With the choice of parameters (A.12) and (A.13), all the previous thresholds are satisfied and there exists a fitted sequence for the considered trajectory for a small enough perturbation. □

A.7. *End of the proof of Theorem 2.5.* Now, we check that the inequality $\varepsilon \preccurlyeq\!* \, r_m$ is satisfied with our choice of parameters (A.12) and (A.13); hence Theorem A.4.2 implies:

$$\|I(t) - I(t_0)\| \leq (n+1)^2 r_0 \quad \text{for } t_0 \leq t \preccurlyeq\!* \, t_0 + \exp(sK/6)$$

while we have assumed the existence of an escape time,

$$t_* \in [t_0, t_0 + \exp(cK)] \quad \text{with } \|I(t_*) - I(t_0)\| = (n+1)^2 r_0.$$

This contradiction ensures the confinement of the action variables over an exponentially long time, $\exp(sK/6)$, which is greater than $\exp(s\varepsilon^{-a}/6)$ with our choice of K. This fulfils the proof of Theorem 2.5. □

References

[1] V. I. Arnold, V. V. Goryunov, O. V. Lyashko and V. A. Vasil'ev. *Singularities Theory I.* Springer, Berlin, 1998.

[2] V. I. Arnold, V. V. Kozlov and A. I. Neishtadt. *Dynamical Systems III: Mathematical Aspects of Classical and Celestial Mechanics.* Springer, Berlin, 1993.

[3] H. W. Broer, G. B. Huitema and M. B. Sevryuk. *Quasi-periodicity in Families of Dynamical Systems: Order Amidst Chaos (Lecture Notes in Mathematics, 1645).* Springer, Berlin, 1996.

[4] J. P. R. Christensen. On sets of Haar measure zero in abelian Polish groups. *Israel. J. Math.* **13** (1973), 255–260.

[5] A. Giorgilli. On the problem of stability for near to integrable Hamiltonian systems. *Proc. Int. Congress of Mathematicians (Berlin, 1998) (Documenta Mathematica)* **3** (1998), 143–152.

[6] V. Guillemin and A. Pollack. *Differential Topology*. Prentice-Hall, Englewood Cliffs, 1974.

[7] M. Guzzo. Stability of the asteroid belt dynamical system. *Private Communication*, 2003.

[8] B. Hunt, T. Sauer and J. A. Yorke. Prevalence: a translation-invariant 'almost every' on infinite-dimensional spaces. *Bull. Am. Math. Soc.* **27**(2) (1992), 217–238.

[9] V. Y. Kaloshin. Some prevalent properties of smooth dynamical systems. *Proc. Steklov Inst. Math.* **213** (1996), 115–140.

[10] V. Y. Kaloshin and B. R. Hunt. A stretched exponential bound on growth of the number of periodic points for prevalent diffeomorphisms, part I. *Ann. of Math.* **165**(1) (2006), 89–170.

[11] P. Lochak. Canonical perturbation theory via simultaneous approximation. *Russian Math. Surveys* **47** (1992), 57–133.

[12] P. Lochak. Hamiltonian perturbation theory: periodic orbits, resonances and intermittency. *Nonlinearity* **6**(6) (1992), 885–904.

[13] J. P. Marco and D. Sauzin. Stability and instability for Gevrey quasi-convex near-integrable Hamiltonian systems. *Publ. Math. Inst. Hautes Études Sci.* **96** (2003), 199–275.

[14] N. N. Nekhorochev. An exponential estimate of the time of stability of nearly integrable Hamiltonian systems. *Russian Math. Surveys* **32** (1977), 1–65.

[15] N. N. Nekhorochev. An exponential estimate of the time of stability of nearly integrable Hamiltonian systems II. *Trudy Sem. Petrovs.* **5**, 5–50. Translation in *Topics in Modern Mathematics (Petrovskii Seminars, 5)*. Ed. O. A. Oleinik. Consultants Bureau, New York, 1985.

[16] L. Niederman. Exponential stability for small perturbations of steep integrable Hamiltonian systems. *Ergod. Th. & Dynam. Sys.* **24**(2) (2004), 593–608.

[17] L. Niederman. Exponential stability and subanalytic geometry. *Ann. Inst. Fourier* **56**(3) (2006), 795–813.

[18] W. Ott and J. A. Yorke. Prevalence. *Bull. Am. Math. Soc.* **42**(3) (2005), 263–290.

[19] R. Pérez-Marco. Convergence or generic divergence of the Birkhoff normal form. *Ann. of Math.* (2) **157** (2003), 557–574.

[20] J. Pöschel. Nekhorochev estimates for quasi-convex Hamiltonian systems. *Math. Z.* **213** (1993), 187–217.

[21] H. Rüssmann. Invariant tori in non-degenerate nearly integrable Hamiltonian systems. *Regul. Chaotic Dyn.* **6**(2) (2001), 119–204.

[22] Y. Yomdin. The geometry of critical and near-critical values of differentiable mappings. *Math. Ann.* **264** (1983), 495–515.

[23] Y. Yomdin and G. Comte. *Tame Geometry with Application in Smooth Analysis (Lecture Notes in Mathematics, 1645)*. Springer, Berlin, 2004.

[24] Y. Yomdin. Some quantitative results in singularity theory. *Ann. Polon. Math.* **87** (2005), 277–299.