PLISKA
STUDIA MATHEMATICA
BULGARICA

ПЛИСКА
БЪЛГАРСКИ
МАТЕМАТИЧЕСКИ
СТУДИИ

Pliska Stud. Math. Bulgar. **23** (2014), 189–209

PLISKA
STUDIA MATHEMATICA
BULGARICA

# ON THE INSTABILITY OF ACTION VARIABLES IN NON-CONVEX HAMILTONIAN SYSTEMS

Borislav Yordanov and Roumyana Yordanova

Abstract. We show the instability of action variables on resonant orbits of certain non-convex Hamiltonian systems with several degrees of freedom. Such orbits remain in the vicinity of resonant surfaces where the action variables can undergo changes $O(1)$ infinitely often although the size of perturbations $O(\varepsilon)$ can be arbitrarily small.

We also perform numerical simulations to compare the effects of two condition for instability in two four-dimensional examples with random parameters.

## 1. Introduction

We study the instability of actions in perturbations of non-convex integrable Hamiltonian systems with several degrees of freedom. This mechanism is called fast drift, or diffusion, and means that action variables of some orbits change by $O(1)$ during a period of time $O(1/\varepsilon)$ when the size of perturbations is measured by $\varepsilon > 0$. The main result states that such drifts regularly occur on resonant orbits if the frequencies are special and the integrable part violates the well known Diophantine (rational) steepness; the latter condition generalizes Nekhoroshev steepness [27] and guarantees stability over periods essentially longer than $O(1/\varepsilon)$ [28], [29], [6], [8]. Similar results have already been obtained for fast diffusion in non-convex Hamiltonian systems with two degrees of freedom [7]. This paper not only

2000 *Mathematics Subject Classification:* Primary 37J25, 37J40; Secondary 37C10, 37C40
*Key words:* Hamiltonian systems, instability, resonances, Arnold diffusion

considers higher dimensions but also uses a different approach; no resonant normal forms are obtained and no drifting orbits are constructed or approximated for large times. It is shown instead that the spectral projections onto low frequencies of the integrable and perturbed vector fields are $O(1)$ away when their distance is measured in the $L^\infty$ norm. The proof relies on ideas from basic ergodic theory and Fourier analysis.

To give the exact statement, let $H : \mathbb{R}^n \times \mathbb{T}^n \to \mathbb{R}$ be defined by

$$H(I, \theta) = h(I) + \varepsilon f(I, \theta), \quad (I, \theta) \in \mathbb{R}^n \times \mathbb{T}^n. \tag{1}$$

We consider real-valued functions $h, f \in C^2(\mathbb{R}^n \times \mathbb{T}^n)$ satisfying the following:

$(H1)$ $\quad h(I) = \omega \cdot I + \frac{1}{2}\Omega I \cdot I, \quad \omega \in \mathbb{R}^n,\ \Omega \in \mathbb{R}^{n\times n},\ \det(\Omega) \neq 0,$

$(H2)$ $\quad f(I, \theta) = \sum_{k\in\mathcal{N}} [a_k(I) \cos k \cdot \theta + b_k(I) \sin k \cdot \theta], \quad a_k,\ b_k \in C^2(\mathbb{R}^n),$

$(H3)$ $\quad \mathcal{N}$ is a finite subset of $\mathbb{Z}^n \setminus \{\vec{0}\},$ such that $\Omega k \cdot l = 0$ for all $k,\ l \in \mathcal{N}$.

Here $\mathbb{T}^n = \mathbb{R}^n/2\pi\mathbb{Z}^n$, $\omega \cdot I$ is the dot product in $\mathbb{R}^n$ and $\Omega$ is a symmetric matrix.

The Hamiltonian system corresponding to (1) has the form

$$\frac{d\theta}{dt} = \nabla_I h(I) + \varepsilon \nabla_I f(I, \theta), \quad \frac{dI}{dt} = -\varepsilon \nabla_\theta f(I, \theta). \tag{2}$$

A non-convex $h(I)$ makes it difficult to check which orbits $(I(t), \theta(t))$ exist for $t \in \mathbb{R}$, since the conservation law of energy $h(I(t)) + \varepsilon f(I(t), \theta(t)) = h(I(0)) + \varepsilon f(I(0), \theta(0))$ does not imply *a priori* bounds on $I(t)$; convex $h(I)$ yield such bounds for small $\varepsilon$.

### 1.1. Results

For $I_0 \in \mathbb{R}^n$ and $\delta > 0$, we define $B^n(I_0; \delta) = \{I \in \mathbb{R}^n : |I - I_0| < \delta\}$, where $|I| = (I \cdot I)^{1/2}$. The resonant surfaces, determined from $\nabla_I h(I) \cdot k = 0$ for $k \in \mathcal{N}$ and $h(I)$ in $(H1)$–$(H3)$, are $(n-1)$-dimensional planes:

$$\mathcal{R}_k = \{I \in \mathbb{R}^n :\ k \cdot \omega + k \cdot \Omega I = 0\}, \quad k \in \mathcal{N}. \tag{3}$$

We will show that the actions $I$ infinitely often deviate by $O(1)$ from their initial values if the orbit passes near $\mathcal{R}_k$. Below $\lambda$ stands for the Lebesgue measure on $\mathbb{R}$.

**Theorem 1.1.** *Let $H$ be the Hamiltonian (1) and assume that (H1)–(H3) hold. Given a resonance plane (3) with $k \in \mathcal{N}$, choose $I_0 \in \mathcal{R}_k$ with $a_k^2(I_0) + b_k^2(I_0) > 0$ and fix a neighborhood $B^n(I_0;\delta)$.*

*There exist $\varepsilon_0 > 0$ and $T_0 > 0$ with the following property: if all trajectories of system (2) with $(I(0),\theta(0)) \in B^n(I_0;\delta) \times \mathbb{T}^n$ are global for $\varepsilon < \varepsilon_0$, then the inequality $T > T_0/\varepsilon$ implies*

$$\sup_{I(0)\in B^n(I_0;\delta)} \frac{\lambda(\{t \in [-T,T] : |I(t) - I(0)| > c(I_0;\delta)\})}{2T} \geq d(I_0;\delta),$$

*where $c$ and $d$ are positive functions determined only by $h$ and $f$.*

The next result is an immediate consequence of Theorem 1.1.

**Corollary 1.** *Let the assumptions of Theorem 1.1 hold. If $\varepsilon$ is sufficiently small, there exists a trajectory with initial data $(I(0),\theta(0)) \in B^n(I_0;\delta)\times\mathbb{T}^n$, such that $I$ either blows up in finite time or exists at all times and satisfies*

$$|I(t_j) - I(0)| \geq c(I_0;\delta) \quad \textit{for} \ \lim_{j\to\infty} |t_j| = \infty.$$

**Remark 1.2.** *Examples of fast drift have long been known. Nekhoroshev [27] gives*

$$H(I,\theta) = \frac{1}{2}(I_1^2 - I_2^2) + \varepsilon \sin(\theta_1 - \theta_2).$$

*Notice that (H1)–(H3) hold with $\omega = (0,0)$, $\Omega = diag(1,-1)$ and $k = (1,-1)$. The resonant plane is $\mathcal{R}_{(1,-1)} = \{I \in \mathbb{R}^2 : I_1 + I_2 = 0\}$. It is easy to find the relations*

$$I_1(t) + I_2(t) = I_1(0) + I_2(0), \quad \theta_1(t) - \theta_2(t) = [I_1(0) + I_2(0)]t + \theta_1(0) - \theta_2(0).$$

*Then we solve the Hamiltonian system with resonant initial values $I_1(0)+I_2(0) = 0$ and $\theta_1(0) - \theta_2(0) = 0$. The drift of actions is evident from*

$$I_1(t) = -\varepsilon t + I_1(0), \quad I_2(t) = \varepsilon t + I_2(0), \quad t \in \mathbb{R}.$$

*The purpose of [7] and this paper is to show the same phenomenon for a larger class of perturbations $f(I,\theta)$ whose frequency sets $\mathcal{N}$ satisfies conditions like (H3).*

**Remark 1.3.** *We use the finiteness of $\mathcal{N}$ in (H3) only to simplify calculations. Without major changes, (H3) can include all sufficiently smooth periodic functions of $\theta$, i.e., infinitely many $a_k$, $b_k$ in (H2) which decay sufficiently fast as $|k| \to \infty$.*

### 1.2. Discussion

The description of instabilities in Theorem 1.1 is more precise than the claim "$|I(t) - I(0)| \geq O(1)$ at some $t = O(1/\varepsilon)$," called diffusion, but less precise than claims like "$|I(0)-I_1| = O(\varepsilon)$ and $|I(t)-I_2| = O(\varepsilon)$ for any $I_1 \neq I_2$ on a connected level surface of $h$," called strong diffusion. Arnold [2], [3] has conjectured that strong diffusion takes place for non-degenerate $h$ and generic, in a suitable sense, perturbations $f$.

Another important fact is that Theorem 1.1 needs no Diophantine condition on $\omega$; the resonant planes $\mathcal{R}_k$ exist for any $\omega$, as $\Omega$ is non-singular and $\Omega k \neq \vec{0}$ for $k \neq \vec{0}$. The key assumption is included in $(H3)$: $\Omega k \cdot l = 0$ for $k, l \in \mathcal{N}$. Then

$$h(I_0 + k_1 J_1 + k_2 J_2 + \cdots + k_n J_n), \quad I_0 \in \mathbb{R}^n, \quad \{k_1, k_2, \ldots, k_n\} \subset \mathcal{N},$$

will depend linearly on $(J_1, J_2, \ldots, J_n) \in \mathbb{R}^n$. In other words, the restriction of $h$ to hyperplanes generated by vectors from $\mathcal{N}$ will be a linear function. This violates rational steepness, similar to condition (A.1) in [7]; $h(I_0+k_1J_1+k_2J_2+\cdots+k_nJ_n)$ will be constant on the resonant planes of $h$ and will have only non-isolated critical points there. It is interesting to compare $(H3)$ with the weaker condition from [20]:

$(H4)$ $\quad$ $\mathcal{N}$ is a finite subset of $\mathbb{Z}^n \setminus \{\vec{0}\}$, such that $\Omega k \cdot k = 0$ for all $k \in \mathcal{N}$.

The above is necessary for the fast drift of resonant orbits in certain perturbations. We do not know, however, whether $(H4)$ is also sufficient in the case of arbitrary perturbation $f$ with a frequency set $\mathcal{N}$. There is some evidence to the contrary: numerical experiments show very slow drifts or quasiperiodic behavior of some resonant orbits when $(H3)$ is replaced by $(H4)$. In contrast, substantial drifts of pseudorandom resonant orbits always occur under condition $(H3)$.

### 1.3. Idea of proof

Let us assume, only for simplicity, that (2) admits global solutions for all initial data $(I(0), \theta(0)) = (I, \theta) \in \mathbb{R}^n \times \mathbb{T}^n$ and $\varepsilon \in [0, \varepsilon_0)$. The von Neumann ergodic theorem [30] states that the average

$$Q_\varepsilon u(I, \theta) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} u(I(t), \theta(t))\, dt \tag{4}$$

exists for all $u \in L^2(\mathbb{R}^n \times \mathbb{T}^n)$ and gives the projection onto the subspace of invariant functions along the trajectories of (2), i.e., $\forall t\ Q_\varepsilon u(I(t), \theta(t)) = Q_\varepsilon u(I, \theta)$.

Assume further that the actions are stable on all trajectories starting from $B^n(I_0;\delta) \times \mathbb{T}^n$:

$$\sup_{t\in\mathbb{R}} |I(t) - I| \leq c\varepsilon. \tag{5}$$

Then every $u \in C_0^2(B^n(I_0;\delta))$ will satisfy $\|Q_\varepsilon u - Q_0 u\|_\infty \leq c(u)\varepsilon$, since $Q_0 u = u$ whenever $u$ depends only on $I$. The question now becomes to estimate the distance between projections $Q_\varepsilon$ and $Q_0$. Theorem 1.1 will follow from a lower bound of the type $\|Q_\varepsilon u - Q_0 u\|_\infty \geq c(u)$ contradicting the stability of actions (5) as $\varepsilon \to 0$.

### 1.4. Other mechanisms of instability

The mechanism of Arnold [1] in *a priori* unstable systems is the most studied and best understood; see [23], [31], [10], [11], [13], [14], [15], [4], [18], [19], [12], [32], [17] and the references therein. This type of instability results from the transverse intersection of stable and unstable manifolds of different invariant tori.

Another major problem, which has been open until recently, is the strong diffusion for generic perturbation of convex and quasi-convex *a priori* stable systems. The first proofs are announced in [25] and [26] for convex $h$ and time-dependent $f$ in dimension $n = 2$ but the details are not published yet. Complete results for such Hamiltonian systems are given, independently, by [21], [9] and [24]. A weaker version of Arnold conjecture is verified in [5] for convex $h$ in arbitrary dimension. This setting is studied further in the resent preprint [22] which also sketches the proof of strong Arnold conjecture.

### 1.5. Outline of paper

The paper is divided into six sections. Sections 2 provides the necessary definitions and background material. The proof of Theorem 1.1 takes sections 3–5 where Proposition 1 in section 3 is the key result. Section 6 describes our numerical simulations intended to check two conjectures about the sufficiency of $(H4)$ and the effects of $(H3)$ and $(H4)$ on drift magnitudes.

## 2. Preliminary Facts

### 2.1. Flows and generators

We consider the Hamiltonian system (2) only for actions $I(0)$ in a neighborhood of resonance. A local choice of data is the best possible without further conditions on the growth of $f(I,\theta)$ as $|I| \to \infty$; for some $I(0)$, the orbit can go to infinity in finite time. Thus, we fix $I_0 \in \mathcal{R}_k$, $\delta > 0$ and assume that the solution $(I(t), \theta(t))$ of equation (2) exists for all initial values $(I(0), \theta(0)) \in B^n(I_0;\delta) \times \mathbb{T}^n$ and times

$t \in [-T, T]$ whenever $\varepsilon < \varepsilon_0$. This gives rise to a $C^1$-family of maps $\Phi_\varepsilon(t) : B^n(I_0; \delta) \times \mathbb{T}^n \to \mathbb{R}^n \times \mathbb{T}^n$, such that

$$\Phi_\varepsilon(t)(I, \theta) = (I(t), \theta(t)), \quad t \in [-T, T],$$

when $(I(0), \theta(0)) = (I, \theta)$. Of course, $\{\Phi_\varepsilon(t)\}_{t\in\mathbb{R}}$ defines a flow on $\mathbb{R}^n \times \mathbb{T}^n$ if all solutions to (2) exist at all times. We will not make any new assumptions about $f(I, \theta)$ to guarantee global existence as our primary goal will be to study instability.

In the integrable case $\varepsilon = 0$, the Hamiltonian system (2) becomes much simpler:

$$\frac{d\theta}{dt} = \nabla_I h(I), \quad \frac{dI}{dt} = 0.$$

The global solutions are given by $(I(t), \theta(t)) = (I, \theta + t\nabla_I h(I))$ and define a flow:

$$\Phi_0(t)(I, \theta) = (I, \theta + t\nabla_I h(I)), \quad t \in \mathbb{R}.$$

More explicitly, we have $\Phi_0(t)(I, \theta) = (I, \theta + t(\omega + \Omega I))$ when $h(I)$ satisfies $(H1)$.

Associated with equation (2) is the differential operator $L_\varepsilon = L_0 + \varepsilon M$ or

$$L_\varepsilon = \frac{1}{i}\nabla_I h(I) \cdot \nabla_\theta + \frac{\varepsilon}{i}[\nabla_I f(I, \theta) \cdot \nabla_\theta - \nabla_\theta f(I, \theta) \cdot \nabla_I]. \tag{6}$$

It is well known that $L_\varepsilon$ generates shifts along the orbits of $\Phi_\varepsilon(t)$. This fact can be expressed, when $H \in C^2(\mathbb{R}^n \times \mathbb{T}^n)$ and $u \in C^1(\mathbb{R}^n \times \mathbb{T}^n)$, as follows: for $|t| < T$,

$$\frac{1}{i}\frac{\partial}{\partial t} u(\Phi_\varepsilon(t)(I, \theta)) = L_\varepsilon u(\Phi_\varepsilon(t)(I, \theta)), \quad (I, \theta) \in B^n(I_0; \delta) \times \mathbb{T}^n. \tag{7}$$

If $\varepsilon = 0$, the above equation actually holds for all $(I, \theta) \in \mathbb{R}^n \times \mathbb{T}^n$ and $t \in \mathbb{R}$.

## 2.2. Approximate spectral projections

Let us denote the Lebesgue measure on $\mathbb{R}^n \times \mathbb{T}^n$ by $dId\theta$ and the corresponding Lebesgue spaces by $L^p(\mathbb{R}^n \times \mathbb{T}^n)$ with norms $\|\cdot\|_p$ for $p \in [1, \infty]$. In $L^2(\mathbb{R}^n \times \mathbb{T}^n)$, we use the sesquilinear inner product

$$\langle u, v\rangle = \int\int_{\mathbb{R}^n\times\mathbb{T}^n} u(I, \theta)\overline{v}(I, \theta)\, dId\theta.$$

The exact domain of $L_\varepsilon$ in (6) is irrelevant, since this operator is applied only to functions from $C_0^\alpha(\mathbb{R}^n \times \mathbb{T}^n)$ with $\alpha \geq 1$; these are $\alpha \geq 1$ times continuously differentiable in $(I, \theta)$, compactly supported in $I$ and periodic in $\theta$. It is easy to check that $L_\varepsilon$ is a symmetric operator: if $u, v \in C_0^1(\mathbb{R}^n \times \mathbb{T}^n)$, then

$$\langle L_\varepsilon u, v\rangle = \langle u, L_\varepsilon v\rangle. \tag{8}$$

Finally, $\operatorname{div}(L_\varepsilon) = 0$ implies that the measure $dId\theta$ is $\Phi_\varepsilon$-invariant whenever the flow is globally defined. In particular, $\operatorname{div}(L_0) = 0$ and $dId\theta$ is $\Phi_0$-invariant:

$$\int\int_{\mathbb{R}^n \times \mathbb{T}^n} u(\Phi_0(t)(I, \theta))\, dId\theta = \int\int_{\mathbb{R}^n \times \mathbb{T}^n} u(I, \theta)\, dId\theta. \tag{9}$$

We can now introduce two standard operators in ergodic theory. The average over the trajectories of $\{\Phi_\varepsilon(t)\}_{t \in [-T,T]}$, which approximates the projection (4) on the null space of $L_\varepsilon$ for large $T$, is given by

$$A_\varepsilon(T)u(I, \theta) = \frac{1}{2T} \int_{-T}^{T} u(\Phi_\varepsilon(t)(I, \theta))\, dt. \tag{10}$$

This is an operator $C_0^\alpha(\mathbb{R}^n \times \mathbb{T}^n) \to C^\alpha(B^n(I_0; \delta) \times \mathbb{T}^n)$ for $\alpha = 0, 1$.

Our second definition involves the exact spectral projection of $L_0$ on frequencies in $(-\varepsilon, \varepsilon)$. In fact, $L_0$ can be extended to a self-adjoint operator in $L^2(\mathbb{R}^n \times \mathbb{T}^n)$ using that $L_0$ generates a continuous group of isomorphisms $u(I, \theta) \mapsto u(I, \theta + t\nabla h(I))$ in $L^2(\mathbb{R}^n \times \mathbb{T}^n)$. Functional calculus for $L_0$ is available through $\psi \mapsto \operatorname{Op}(\psi)$ and

$$\operatorname{Op}(\psi)u(I, \theta) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{\psi}(t) u(\Phi_0(t)(I, \theta))\, dt. \tag{11}$$

Here $\hat{\psi}(t) = \int_{\mathbb{R}} \psi(s) e^{-its}\, ds$ means the Fourier transform. The class of admissible symbols includes $\psi \in C_0^\infty(\mathbb{R})$ whose transforms $\hat{\psi}$ decay fast at infinity. We can think of $\operatorname{Op}(\psi)$ as an operator representing $\psi(L_0)$. Let $\chi \in C_0^\infty(\mathbb{R})$ and assume

$$\chi(s) \geq 0, \quad \chi(-s) = \chi(s), \quad \chi(s) = 1 \ \text{ for } \ s \in [-1, 1]. \tag{12}$$

We can set $\chi_\varepsilon(s) = \chi(s/\varepsilon)$ and introduce the operator $P_\varepsilon = \chi_\varepsilon(L_0)$ or $P_\varepsilon = \operatorname{Op}(\chi_\varepsilon)$:

$$P_\varepsilon u(I, \theta) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{\chi}_\varepsilon(t) u(\Phi_0(t)(I, \theta))\, dt. \tag{13}$$

It is important that $P_\varepsilon : C_0^\alpha(\mathbb{R}^n \times \mathbb{T}^n) \to C_0^\alpha(\mathbb{R}^n \times \mathbb{T}^n)$ for $\alpha = 0, 1$.

**Remark 2.1.** *A simple calculation shows that, for* $u(I,\theta) = \sum_{k\in\mathbb{Z}^n} u_k(I)e^{ik\cdot\theta}$,

$$P_\varepsilon u(I,\theta) = \sum_{k\in\mathbb{Z}^n} u_k(I)\chi(k\cdot(\omega+\Omega I)/\varepsilon)e^{ik\cdot\theta}.$$

*Hence,* $P_\varepsilon$ *truly acts as a cutoff for non-resonant frequencies. It is not surprising that similar operators occur in averaging theory [16] and Melnikov theory [31].*

## 3. Distance Between Null Spaces of $L_0$ and $L_\varepsilon$

The projections onto null spaces of $L_0$ and $L_\varepsilon = L_0 + \varepsilon M$ are not $\varepsilon$-close when $T$ is much larger then $1/\varepsilon$. This is the meaning of the following lower bound.

**Proposition 1.** *Assume that* $u \in C_0^2(\mathbb{R}^n\times\mathbb{T}^n)$ *and* $L_0u = 0$*; that is,* $u$ *is a* $\Phi_0$*-invariant function. Let* $A_\varepsilon(T)$ *and* $P_\varepsilon$ *be defined in (10) and (13), respectively. If* $c_1 = \|\hat\chi\|_1/(2\pi)$ *and* $c_2 = \|s\hat\chi\|_1/(2\pi)$*, then*

$$\|u - A_\varepsilon(T)u\|_\infty \ge \frac{\|P_\varepsilon Mu\|_2^2 - c_1(\varepsilon T)^{-1}\|u\|_\infty\|P_\varepsilon Mu\|_1}{c_2\|P_\varepsilon Mu\|_1 + \|MP_\varepsilon^2 Mu\|_1}.$$

Proof. Given $u \in C_0^2(\mathbb{R}^n\times\mathbb{T}^n)$ and $v \in C_0^1(\mathbb{R}^n\times\mathbb{T}^n)$, we apply (8) to derive

$$\langle u - A_\varepsilon(T)u, L_\varepsilon v\rangle = \langle L_\varepsilon u - L_\varepsilon A_\varepsilon(T)u, v\rangle.$$

It follows from (7) that

$$L_\varepsilon A_\varepsilon(T)u = \frac{1}{2Ti}\left[u(\Phi_\varepsilon(T)(I,\theta)) - u(\Phi_\varepsilon(-T)(I,\theta))\right].$$

Last two identities combine into

$$\langle u - A_\varepsilon(T)u, L_\varepsilon v\rangle = \langle L_\varepsilon u, v\rangle - \frac{1}{2Ti}\langle u(\Phi_\varepsilon(T)(I,\theta)) - u(\Phi_\varepsilon(-T)(I,\theta)), v\rangle$$

and, evidently, imply the following inequality:

$$\|u - A_\varepsilon(T)u\|_\infty\|L_\varepsilon v\|_1 \ge |\langle L_\varepsilon u, v\rangle| - T^{-1}\|u\|_\infty\|v\|_1.$$

We choose $v = P_\varepsilon^2 L_\varepsilon u$ and divide by $\|L_\varepsilon P_\varepsilon^2 L_\varepsilon u\|_1$ to obtain

$$\|u - A_\varepsilon(T)u\|_\infty \ge \frac{\|P_\varepsilon L_\varepsilon u\|_2^2 - T^{-1}\|u\|_\infty\|P_\varepsilon^2 L_\varepsilon u\|_1}{\|L_\varepsilon P_\varepsilon^2 L_\varepsilon u\|_1}. \tag{14}$$

To further transform the right hand side, we use that $L_\varepsilon u = \varepsilon M u$, split

$$\|L_\varepsilon P_\varepsilon^2 L_\varepsilon u\|_1 \le \|L_0 P_\varepsilon^2 L_\varepsilon u\|_1 + \varepsilon \|M P_\varepsilon^2 L_\varepsilon u\|_1$$

and divide both the numerator and denominator of (14) by $\varepsilon^2$:

$$\|u - A_\varepsilon(T)u\|_\infty \ge \frac{\|P_\varepsilon M u\|_2^2 - (\varepsilon T)^{-1}\|u\|_\infty \|P_\varepsilon^2 M u\|_1}{\|(L_0/\varepsilon)P_\varepsilon^2 M u\|_1 + \|M P_\varepsilon^2 M u\|_1}.$$

Now we use functional calculus for $L_0$ and recall that $\hat{\chi}_\varepsilon(t) = \varepsilon\hat{\chi}(\varepsilon t)$. The estimates

$$\begin{aligned} \|P_\varepsilon^2 M u\|_1 &\le c_1 \|P_\varepsilon M u\|_1, \quad c_1 = \|\hat{\chi}\|_1/(2\pi), \\ \|(L_0/\varepsilon) P_\varepsilon^2 M u\|_1 &\le c_2 \|P_\varepsilon M u\|_1, \quad c_2 = \|s\hat{\chi}\|_1/(2\pi), \end{aligned}$$

are just substituted into the lower bound of $\|u - A_\varepsilon(T)u\|_\infty$ to finish the proof. □

The above result is not immediately applicable, since it does not bound uniformly the distance between $P_\varepsilon$ and $A_\varepsilon(T)$ as $\varepsilon \to 0$. In fact, we can have the ratio of leading terms $\|P_\varepsilon M u\|_2 / \|M P_\varepsilon^2 M u\|_1 \to 0$ as $\varepsilon \to 0$. To derive a uniform lower bound requires additional conditions, such as $(H1) - (H3)$.

## 4. Estimates

We need simple expressions for $P_\varepsilon M u$ and $M P_\varepsilon^2 M u$, with $u \in C_0^2(\mathbb{R}^n)$, to use the full strength of Proposition 1. The exponential form of $f(I,\theta)$ in more convenient:

$$\begin{aligned} f(I,\theta) &= \sum_{k\in\mathcal{M}} f_k(I) e^{ik\cdot\theta}, \quad \mathcal{M} = \mathcal{N} \cup (-\mathcal{N}), \\ f_k(I) &= \frac{1}{2}[a_k(I) - i b_k(I)], \quad k \in \mathcal{N}, \\ f_k(I) &= \frac{1}{2}[a_{-k}(I) + i b_{-k}(I)], \quad k \in -\mathcal{N}. \end{aligned} \tag{15}$$

The final preparation for applying Proposition 1 involves estimates of $\|P_\varepsilon M u\|_2$, $\|P_\varepsilon M u\|_1$ and $\|M P_\varepsilon^2 M u\|_1$ as $\varepsilon \to 0$. We readily obtain the next result.

**Proposition 2.** *Let $u \in C_0^2(\mathbb{R}^n)$ and $M = \frac{1}{i}(\nabla_I f(I,\theta)\cdot\nabla_\theta - \nabla_\theta f(I,\theta)\cdot\nabla_I)$. If $P_\varepsilon$ is defined in (13), then*

$$(i) \qquad \|P_\varepsilon M u\|_2^2 = \sum_{k\in\mathcal{M}} (2\pi)^n \int_{\mathbb{R}^n} |f_k(I) k\cdot\nabla_I u(I)|^2 \chi^2(k\cdot(\omega+\Omega I)/\varepsilon)\, dI,$$

$$(ii) \qquad \|P_\varepsilon M u\|_1 \le \sum_{k\in\mathcal{M}} (2\pi)^n \int_{\mathbb{R}^n} |f_k(I) k\cdot\nabla_I u(I)| \chi(k\cdot(\omega+\Omega I)/\varepsilon)\, dI$$

*and*

$$(iii) \qquad \|MP_\varepsilon^2 Mu\|_1 \le \sum_{k\in\mathcal{M}}\sum_{l\in\mathcal{M}}(2\pi)^n\int_{\mathbb{R}^n}|[\nabla_I f_k(I)\cdot l][f_l(I)l\cdot\nabla_I u(I)]|\chi^2(l\cdot(\omega+\Omega I)/\varepsilon)\,dI + \sum_{k\in\mathcal{M}}\sum_{l\in\mathcal{M}}(2\pi)^n\int_{\mathbb{R}^n}|f_k(I)k\cdot\nabla_I[f_l(I)l\cdot\nabla_I u(I)]|\chi^2(l\cdot(\omega+\Omega I)/\varepsilon)\,dI.$$

Various integrals over planes with resonances are expressed through the Radon transform: if $v\in C_0(\mathbb{R}^n)$, $s$ is a real number and $\xi$ is a unit vector, then

$$\mathcal{R}v(s,\xi)=\int_{\xi\cdot I=s}v(I)\,dS(I).$$

Here the surface measure is normalized so that $dI=dS(I)ds$.

**Proposition 3.** *Under the assumptions of Proposition 2, let supp*$(u)\subset B^n(I_0;\delta)$.

$$\|P_\varepsilon Mu\|_2^2 \ \ge\ \varepsilon(2\pi)^n\|\chi\|_2^2\sum_{k\in\mathcal{M}}\frac{1}{|\Omega k|}\mathcal{R}[|f_k(I)k\cdot\nabla_I u(I)|^2]\left(-\frac{\omega\cdot k}{|\Omega k|},\frac{\Omega k}{|\Omega k|}\right) -\varepsilon^2(2\pi)^n C_n(2\delta)^{n-1}\||s|^{\frac12}\chi\|_2^2\sum_{k\in\mathcal{M}}\frac{1}{|\Omega k|^2}\|\nabla_I|f_k(I)k\cdot\nabla_I u(I)|^2\|_\infty.$$

P r o o f. We use Proposition 2($i$) and simple estimates of the Radon transform. □

**Proposition 4.** *Under the assumptions of Proposition 2, let supp*$(u)\subset B^n(I_0;\delta)$.

$$(i)\qquad \|P_\varepsilon Mu\|_1 \ \le\ \varepsilon(2\pi)^n C_n\delta^{n-1}\|\chi\|_1\sum_{k\in\mathcal{M}}\frac{\|f_k(I)k\cdot\nabla_I u(I)\|_\infty}{|\Omega k|},$$

$$(ii)\ \|MP_\varepsilon^2 Mu\|_1 \ \le\ \varepsilon(2\pi)^n C_n\delta^{n-1}\|\chi\|_2^2\sum_{k,l\in\mathcal{M}}\frac{\|[\nabla_I f_k(I)\cdot l][f_l(I)l\cdot\nabla_I u(I)]\|_\infty}{|\Omega l|} +\varepsilon(2\pi)^n C_n\delta^{n-1}\|\chi\|_2^2\sum_{k,l\in\mathcal{M}}\frac{\|f_k(I)k\cdot\nabla_I[f_l(I)l\cdot\nabla_I u(I)]\|_\infty}{|\Omega l|}.$$

Proof. The two estimates follow from Proposition 2(*ii*), (*iii*). □

We can now state the lower bound on the distance between spectral projections.

**Proposition 5.** *Let* $u \in C_0^2(\mathbb{R}^n)$ *and* $M = \frac{1}{i}(\nabla_I f(I,\theta)\cdot\nabla_\theta - \nabla_\theta f(I,\theta)\cdot\nabla_I)$. *If* $A_\varepsilon(T)$, $P_\varepsilon$ *are defined in (10), (13), respectively, and supp*$(u) \subset B^n(I_0;\delta)$, *then*

$$\|u - A_\varepsilon(T)u\|_\infty \geq \frac{d_0 - d_1\varepsilon - c_1 d_2(\varepsilon T)^{-1}}{c_2 d_3 + d_4}.$$

*The new constants* $d_1,\ldots,d_4$, *depending on* $I_0$ *and* $\delta$, *are given by*

$$\begin{aligned}
d_0(I_0;\delta) &= (2\pi)^n\|\chi\|_2^2\sum_{k\in\mathcal{M}}\frac{1}{|\Omega k|}\mathcal{R}[|f_k(I)k\cdot\nabla_I u(I)|^2]\left(-\frac{\omega\cdot k}{|\Omega k|},\frac{\Omega k}{|\Omega k|}\right),\\
d_1(I_0;\delta) &= (2\pi)^n C_n\||s|^{\frac12}\chi\|_2^2(2\delta)^{n-1}\sum_{k\in\mathcal{M}}\frac{1}{|\Omega k|^2}\|\nabla_I|f_k(I)k\cdot\nabla_I u(I)|^2\|_\infty,\\
d_2(I_0;\delta) &= (2\pi)^n C_n\|\chi\|_1\delta^{n-1}\sum_{k\in\mathcal{M}}\frac{1}{|\Omega k|}\|f_k(I)k\cdot\nabla_I u(I)\|_\infty\|u(I)\|_\infty,\\
d_3(I_0;\delta) &= (2\pi)^n C_n\|\chi\|_1\delta^{n-1}\sum_{k\in\mathcal{M}}\frac{1}{|\Omega k|}\|f_k(I)k\cdot\nabla_I u(I)\|_\infty,\\
d_4(I_0;\delta) &= (2\pi)^n C_n\|\chi\|_1\delta^{n-1}\sum_{k,l\in\mathcal{M}}\frac{1}{|\Omega l|}\|[\nabla_I f_k(I)\cdot l][f_l(I)l\cdot\nabla_I u(I)]\|_\infty\\
&\quad +(2\pi)^n C_n\|\chi\|_1\delta^{n-1}\sum_{k,l\in\mathcal{M}}\frac{1}{|\Omega l|}\|f_k(I)k\cdot\nabla_I[f_l(I)l\cdot\nabla_I u(I)]\|_\infty.
\end{aligned}$$

Proof. We substitute the estimates of Propositions 3 and 4 into Proposition 1 and give short names to the long expressions involving $\chi$, $f_k$ and $u$. □

## 5. Proof of Theorem 1.1

Proof. The idea is to apply Proposition 5 with a suitable $u \in C_0^2(\mathbb{R}^n)$. We choose a cutoff function $v \in C_0^2(\mathbb{R}^n)$, such that

$$0 \leq v \leq 1, \quad v(I) = 1 \text{ if } I \in B^n(I_0;1/2), \quad v(I) = 0 \text{ if } I \in \mathbb{R}^n \setminus B^n(I_0;1).$$

Substitute $u(I) = (1+|I|^2)^{1/2}v((I-I_0)/\delta)$ in Proposition 5. Then

$$\|u - A_\varepsilon(T)u\|_\infty \geq \frac{d_0 - d_1\varepsilon - c_1 d_2(\varepsilon T)^{-1}}{c_2 d_3 + d_4}$$

with the appropriate constant $c_1$, $c_2$ and $d_0, \dots, d_4$. To streamline the proof, let us assume that $d_0(I;\delta) > 0$ is already known; we will verify this claim later. Set

$$\varepsilon_0 = \frac{d_0}{4d_1}, \quad T_0 = \frac{4c_1 d_2}{d_0}.$$

If $\varepsilon < \varepsilon_0$ and $T > T_0/\varepsilon$, the above lower bound becomes independent of $\varepsilon$ and $T$:

$$\|u - A_\varepsilon(T)u\|_\infty \geq c_0, \quad c_0 = \frac{d_0}{2(c_2 d_3 + d_4)}.$$

A more transparent form, showing the drift of $I(t)$ from $I(0)$, is

$$\sup_{I(0)\in B^n(I_0;\delta)} \left| u(I(0)) - \frac{1}{2T}\int_{-T}^{T} u(I(t))\,dt \right| \geq c_0.$$

It is actually sufficient to use the following weaker inequality:

$$\sup_{I(0)\in B^n(I_0;\delta)} \frac{1}{2T}\int_{-T}^{T} |u(I(t)) - u(I(0))|\,dt \geq c_0. \tag{16}$$

To estimates the relative measure of large deviations, we consider the set

$$\mathcal{E}(T;c) = \{t \in [-T,T] : |I(t) - I(0)| > c\}.$$

The definition of $u$ and $u(I(t)) \leq (1 + (|I_0| + \delta)^2)^{1/2}$ yield

$$|u(I(t)) - u(I(0))| \leq (1 + |I_0| + \delta), \quad |u(I(t)) - u(I(0))| \leq \|\nabla_I u\|_\infty |I(t) - I(0)|.$$

We will choose one estimate for each $t \in [-T,T]$ depending on $|I(t) - I(0)|$:

$$\begin{aligned} |u(I(t)) - u(I(0))| &\leq (1 + |I_0| + \delta), \quad t \in \mathcal{E}(T;c), \\ |u(I(t)) - u(I(0))| &\leq \|\nabla_I u\|_\infty c, \quad t \in [-T,T] \setminus \mathcal{E}(T;c). \end{aligned}$$

Substituting these inequalities into (16) leads to

$$\sup_{I(0)\in B^n(I_0;\delta)} \left( \frac{1}{2T}\int_{\mathcal{E}(T;c)} (1 + |I_0| + \delta)dt + \frac{1}{2T}\int_{[-T,T]\setminus\mathcal{E}(T;c)} \|\nabla_I u\|_\infty c\,dt \right) \geq c_0$$

and, after some trivial simplifications, to

$$\left( \sup_{I(0)\in B^n(I_0;\delta)} \frac{\lambda(\mathcal{E}(T;c))}{2T} \right) (1 + |I_0| + \delta) + \|\nabla_I u\|_\infty c \geq c_0.$$

If we set $c = c_0/(2\|\nabla_I u\|_\infty)$, the above inequality becomes

$$\sup_{I(0)\in B^n(I_0;\delta)} \frac{\lambda(\mathcal{E}(T;c))}{2T} \quad \geq \quad \frac{c_0}{2(1+|I_0|+\delta)\|\nabla_I u\|_\infty}.$$

This is the necessary lower bound with the following constants $c$ and $d$:

$$c = \frac{c_0}{2\|\nabla_I u\|_\infty}, \quad d = \frac{c_0}{2(1+|I_0|+\delta)\|\nabla_I u\|_\infty}.$$

It remains to verify the starting claim that $d_0(I_0;\delta) > 0$. The definition of $d_0$ in Proposition 5, which involves Radon transforms or integrals over $\mathcal{R}_k$, shows that

$$d_0(I_0;\delta) \geq \frac{(2\pi)^n\|\chi\|_2^2}{|\Omega k|}\mathcal{R}[|f_k(I)k\cdot\nabla_I u(I)|^2]\left(-\frac{\omega\cdot k}{|\Omega k|}, \frac{\Omega k}{|\Omega k|}\right).$$

We recall that $f_k$ are given by (15) and $u(I) = (1+|I|^2)^{1/2}$ in $B^n(I_0;\delta/2)$. Thus, the function in the Radon transform satisfies

$$|f_k(I)k\cdot\nabla_I u(I)|^2 = \frac{[a_k^2(I)+b_k^2(I)]|k\cdot I|^2}{1+|I|^2}, \quad I\in B^n(I_0;\delta/2),$$

and the lower bound simplifies to

$$d_0(I_0;\delta) \geq \frac{(2\pi)^n\|\chi\|_2^2}{|\Omega k|}\int_{\mathcal{R}_k\cap B^n(I_0;\delta/2)} \frac{[a_k^2(I)+b_k^2(I)]|k\cdot I|^2}{1+|I|^2}\,dS(I).$$

Here the integration takes place over a relatively open set $\mathcal{R}_k\cap B^n(I_0;\delta/2)\neq\emptyset$. In addition, the integrand is continuous and not identically 0: for small $\delta$,

$$a_k^2(I)+b_k^2(I) > 0 \text{ and } |k\cdot I| > 0 \text{ a.e. near } I = I_0.$$

(The latter inequality holds since $k\perp\Omega k$ and $\{I : k\cdot I = 0\}$ is not parallel to $\mathcal{R}_k$.) Hence, the Radon transform is positive and so is the lower bound on $d_0(I_0;\delta)$. □

## 6. Numerical Experiments

In the final section we answers two interesting questions about $(H3)$ and $(H4)$:

1. Does $(H4)$ imply fast drifts for all resonant orbits?
2. Does $(H3)$ yield longer drifts than $(H4)$?

Here we mean only orbits starting from a simple resonance and evolving up to times $O(1/\varepsilon)$. These questions are not addressed by Theorem 1.1, since no answers can be given without further conditions on $H(I,\theta)$ in (1).

Thus we resort to statistical analysis to find out which answers to (1) and (2) are more likely and what additional properties of the system guarantee rigorous results. The first step is to generate pseudorandom orbits for two Hamiltonians $H_i(I,\theta)$, ($i = 3$ and 4) with identical $h(I)$ but different pseudorandom non-integrable parts $f_3(I,\theta)$ and $f_4(I,\theta)$ which satisfy, respectively, $(H3)$ and $(H4)$. Let us recall these conditions: if $\mathcal{N} \subset \mathbb{Z}^n \setminus \{\vec{0}\}$ is a finite set, then

$$(H3) \qquad \Omega k \cdot l = 0 \text{ for all } k,\ l \in \mathcal{N},$$
$$(H4) \qquad \Omega k \cdot k = 0 \text{ for all } k \in \mathcal{N}.$$

In fact, $(H4)$ means that every line of fast drift $k$ is parallel to its resonant plane $\mathcal{R}_k$ but not necessarily parallel to all resonant planes defined, in the case $\omega = \vec{0}$, by

$$\mathcal{R}_l = \{I \in \mathbb{R}^n : \ l \cdot \Omega I = 0\}, \quad l \in \mathcal{N}.$$

Having approximated $N$ pseudorandom orbits $\{(I^i(t), \theta^i(t)) \ : \ t \in [0,T]\}$, where $i = 1, \ldots, N$, we compute the maximum drift for each:

$$I^i_{\max} = \max_{t \in [0,T]} |I^i(t) - I^i(0)|.$$

The results are summarized in two histograms and a scatter plot of $(H3)$ versus $(H4)$ below. It turns out that the affirmative answers to both (1) and (2) strongly prevail for the type of perturbations considered here. However, (1) and (2) are false as conjectures. The outcomes seem to depend on the relative distances of $I(0)$ to other resonances and the relative sizes of resonant coefficients compared to $\varepsilon$.

### 6.1. Examples of $H_i$

To introduce the two Hamiltonians, we specify

$$n = 4, \quad \omega = \vec{0}, \quad \Omega = \mathrm{diag}(1, 1, -1, -1),$$

and the following two sets of frequencies:

$$\mathcal{N}_3 \ = \ \{(1,0,1,0),\ (0,1,0,1)\},$$
$$\mathcal{N}_4 \ = \ \{(1,0,1,0),\ (5,0,3,4)\}.$$

It is easy to see that $\mathcal{N}_i$ satisfies $(Hi)$ for $i = 3, 4$. There is nothing special about our choices except that $h$ has small integer coefficients and $\mathcal{N}_3$ has the smallest possible cardinality. The common resonant plane of $\mathcal{N}_3$ and $\mathcal{N}_4$ is

$$\mathcal{R}_{(1,0,1,0)} = \{I \in \mathbb{R}^4 : I_1 - I_3 = 0\}, \tag{17}$$

but there are two additional ones:

$$\mathcal{R}_{(0,1,0,1)} = \{I \in \mathbb{R}^4 : I_2 - I_4 = 0\},$$
$$\mathcal{R}_{(5,0,3,4)} = \{I \in \mathbb{R}^4 : 5I_1 - 3I_3 - 4I_4 = 0\}.$$

We will work with linear perturbations of $h(I) = \left(I_1^2 + I_2^2 - I_3^3 - I_4^2\right)/2$, i.e.,

$$H_3(I,\theta) = h(I) + \varepsilon f_3(I,\theta), \quad H_4(I,\theta) = h(I) + \varepsilon f_4(I,\theta),$$

where

$$\begin{aligned} f_3(I,\theta) &= (\alpha_1 I_1 + \alpha_2 I_2 + \alpha_3 I_3 + \alpha_4 I_4)\cos(\theta_1 + \theta_3) \\ &\quad + (\beta_1 I_1 + \beta_2 I_2 + \beta_3 I_3 + \beta_4 I_4)\cos(\theta_2 + \theta_4), \\ f_4(I,\theta) &= (\alpha_1 I_1 + \alpha_2 I_2 + \alpha_3 I_3 + \alpha_4 I_4)\cos(\theta_1 + \theta_3) \\ &\quad + (\beta_1 I_1 + \beta_2 I_2 + \beta_3 I_3 + \beta_4 I_4)\cos(5\theta_1 + 3\theta_3 + 4\theta_4). \end{aligned}$$

The random coefficients $\alpha, \beta \in [-1/2, 1/2]^4$ are uniformly distributed. Such $f_i$ lead to linear Hamiltonian systems with respect to $I \in \mathbb{R}^4$ and exclude any possibility for finite time blow-up of actions.

### 6.2. Conditions for simple resonance and non-degenerate coefficient

We fix the initial angles $\theta(0) = (\pi/4, \pi/4, \pi/4, \pi/4)$ but allow random initial actions $I(0) \in [-2, 2]^4$ uniformly distributed near a point on the resonant surface $\mathcal{R}_{(1,0,1,0)}$. It is important to start from a simple resonance and have a non-degenerate coefficient at the resonant frequency. The accuracy of computations is estimated by comparing resonant orbits with typical non-resonant ones which are expected to be either quasiperiodic or slowly drifting. These preparations are summarized below.

| | Random | Fixed | Conditions |
|---|---|---|---|
| Resonant | $\alpha, \beta, I_2, I_4$ | $\theta, I_1, I_3$ | $R,\ ND$ |
| Non-Resonant | $\alpha, \beta, I_2, I_4$ | $\theta, I_1, I_3$ | $NR,\ ND$ |

The abbreviated conditions are as follows:

$$
\begin{array}{rl}
(R) & I_1 = 1, \quad I_3 = 1; \\
(NR) & |I_1 - 1| \le 1/2, \quad |I_3 + 1| \le 1/2; \\
 & |I_2 - I_4| \ge \delta, \\
(ND) & |5I_1 - 3I_3 - 4I_4| \ge \delta, \\
 & |\alpha_1 I_1 + \alpha_2 I_2 + \alpha_3 I_3 + \alpha_4 I_4| \ge \delta,
\end{array}
$$

where $\delta$ is a suitable parameter to be specified later. Clearly, $\delta$ measures the initial distances to other resonances and the initial size of resonant coefficients in $f_i$.

## 6.3. Parameters

The choices of small parameters $\varepsilon$, $\delta$ and interval of time $[0, T]$ are far from unique. We begin with $\varepsilon = 10^{-2}$. Fast drifts occur at times $O(1/\varepsilon)$, so we generate orbits up to $T = 2 \times 10^2$, for instance. The remaining parameter is set to $\delta = 10^{-1}$, as the non-degeneracy conditions mean that $\delta$ is much larger than $\varepsilon$.

## 6.4. Discussion

Simulations are performed in MATLAB using an explicit Runge-Kutta formula implemented by ode45 function. We run $N = 10^3$ simulations with parameters described in section 8.3 and calculate the maximum drift in each run. Figure 1 represents the scatter plot of the maximum drifts $\{I^i_{\max} : i = 1, 2, \dots, 10^3\}$ for the two Hamiltonian systems $H_3$ and $H_4$ in section 8.1.

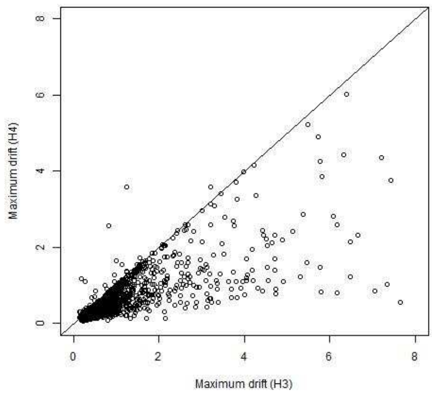


*Figure 1* – Scatter plot of maximum drifts from 1000 simulations for the Hamiltonian systems satisfying $H3$ ($x$ axis) and $H4$ ($y$ axis).

The graph illustrates that the maximum drift for ($H3$) is longer than the maximum drift for ($H4$) in more than ninety percent of all runs. This is consistent with our expectations, since no orbit can leave a resonant plane when ($H3$) holds, while some orbits can leave some resonant planes when only ($H4$) holds.

The histograms in Figure 2 and Figure 3 confirm that questions (1) and (2) usually have affirmative answers. However, Figure 2 also shows that the simulations do not find drifts ($I_{\max} < 10^{-1} = \delta$) in about one percent of all runs.

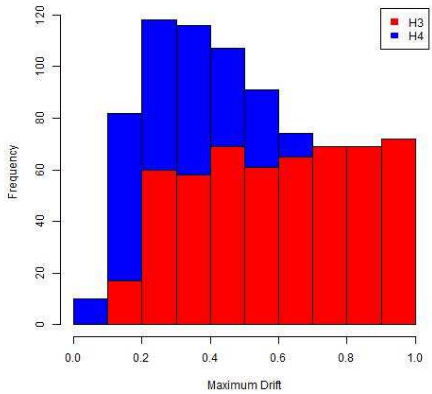


*Figure 2* – “Zoom in" histograms of the maximum drift for Hamiltonian systems satisfying $H3$ (red) and $H4$ (blue).

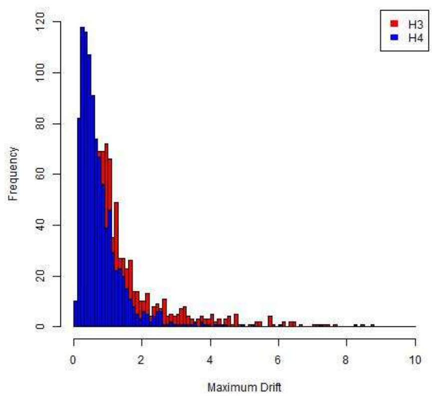


*Figure 3* – Histograms of the maximum drift for Hamiltonian systems satisfying $H3$ (red) and $H4$ (blue).

## References


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

[2] V. I. Arnold, *A stability problem and ergodic properties of classical dynamical systems*, Proc. Internat. Congr. Math. (Moscow, 1966). Izdat. "Mir", Moscow, 1968, 387–392.

[3] V. I. Arnold, *Mathematical problems in classical physics*, Trends and perspectives in applied mathematics. Vol. 100. Appl. Math. Sci. Springer, New York, 1994, 1–20.

[4] P. Bernard, *The dynamics of pseudographs in convex Hamiltonian systems*, Journal of the American Math. Society 21 (2008), no. 3, 615–669.

[5] P. Bernard, V. Kaloshin, and K. Zhang, *Arnold diffusion in arbitrary degrees of freedom and crumpled 3-dimensional normally hyperbolic invariant cylinders*, Preprint (2011), arXiv:1112.2773.

[6] A. Bounemoura, *Optimal stability and instability for near-linear hamiltonians*, Annales Henri Poincarŕe 13 (2012), no. 4, 857-–868.

[7] Abed Bounemoura and Vadim Kaloshin,*Generic fast diffusion for a class of non-convex Hamiltonians with two degrees of freedom*, preprint, November 11, 2013.

[8] A. Bounemoura and L. Niederman, *Generic Nekhoroshev theory without small divisors*, Ann. Inst. Fourier 62 (2012), no. 1, 277—324.

[9] C.-Q. Cheng, *Arnold diffusion in nearly integrable hamiltonian systems*, Preprint (2013), arXiv:1207.4016v2.

[10] C.-Q. Cheng and J. Yan, *Existence of diffusion orbits in a priori unstable Hamil- tonian systems*, J. Differ. Geom. 67 (2004), no. 3, 457—517.

[11] C.-Q. Cheng and J. Yan, *Arnold diffusion in Hamiltonian systems: a priori unstable case*, J. Differ. Geom. 82 (2009), no. 2, 229–277.

[12] A Delshams, M Gonchenko, P Gutiйrrez, *Exponentially small asymptotic estimates for the splitting of separatrices to whiskered tori with quadratic and cubic frequencies*, E. R. A. in Math. Sciences 21, (2014), 41–61.

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

[14] A. Delshams, R. de la Llave, and T.M. Seara, *A geometric mechanism for diffusion in Hamiltonian systems overcoming the large gap problem: heuristics and rigorous verification on a model*, Mem. Amer. Math. Soc. 179 (2006), no. 844.

[15] Amadeu Delshams, Rafael de la Llave, Tere M. Seara, *Instability of high dimensional Hamiltonian Systems: Multiple resonances do not impede diffusion*, preprint (2013), arXiv:1306.4614.

[16] Dodson, M. M., Rynne, B. P., Vickers, J. A. G., *Averaging in multifrequency systems*, Nonlinearity 2 (1989), 137—148.

[17] C. Efthymiopoulos, M. Harsoula, *The speed of Arnold diffusion*, Physica D: Nonlinear Phenomena 251, (2013) 19–38.

[18] M. Gidea and C. Robinson, *Shadowing orbits for transition chains of invariant tori alternating with Birkhoff zones of instability*, Nonlinearity 20 (2007), no. 5, 1115—1143.

[19] M. Gidea and C. Robinson, *Obstruction argument for transition chains of tori interspersed with gaps*, Discrete Contin. Dyn. Syst., Ser. S 2 (2009), no. 2, 393–416.

[20] Massimiliano Guzzo, Elena Lega, Claude Froeschlй, *Diffusion and stability in perturbed non convex integrable systems*, Nonlinearity, 19 (2006), no 5, 1049–1067.

[21] V. Kaloshin and K. Zhang, *A strong form of Arnold diffusion for two and a half degrees of freedom*, preprint (2012), arXiv:1212.1150.

[22] V. Kaloshin and K. Zhang, *Partial averaging and dynamics of the dominant Hamiltonian, with applications to Arnold diffusion*, preprint (2014), arxiv.org/pdf/1410.1844.

[23] Pierre Lochak, Jean-Pierre Marco, *Diffusion times and stability exponents for nearly integrable analytic systems*, CEJM 3 (2005), no. 3, 342–397.

[24] J.-P. Marco, *Generic hyperbolic properties of classical systems on the torus* $T^2$, Preprint (2012).

[25] J.N. Mather, *Arnold diffusion I: Announcement of results*, J. of Math. Sciences 124 (2004), no. 5, 5275–5289.

[26] J.N. Mather, *Arnold diffusion by variational methods*, Essays in mathematics and its applications, Springer, Heidelberg, 2012, 271—285.

[27] Nekhoroshev N.N., *Exponential estimates of the stability time of near-integrable Hamiltonian systems*, Russ. Math. Surv., 32 (1977), 1–65.

[28] L. Niederman, *Hamiltonian stability and subanalytic geometry*, Ann. Inst. Fourier 56 (2006), no. 3, 795—813.

[29] L. Niederman, *Prevalence of exponential stability among nearly integrable Hamiltonian systems*, Erg. Th. Dyn. Sys. 27 (2007), no. 3, 905—928.

[30] J. von Neumann, *Proof of the quasi-ergodic hypothesis*, Proc. Nat. Acad. Sci. USA , 18 (1932), 70—82.

[31] D. Treschev, *Evolution of slow variables in a priori unstable Hamiltonian systems*, Nonlinearity, 17 (2004), no. 5, 1803–1841.

[32] Ke Zhang, *Speed of Arnold diffusion for analytic Hamiltonian systems*, Inventiones mathematicae 186 (2011) no. 2, 255–290.

*Borislav Yordanov*
*Institute of Mathematics and Informatics*
*Acad. G. Bonchev Str. Bl. 8*
*1113 Sofia, Bulgaria*
*e-mail:*byordanov@math.bas.bg

*Roumyana Yordanova*
*Institute of Mathematics and Informatics*
*Acad. G. Bonchev Str. Bl. 8*
*1113 Sofia, Bulgaria*
*e-mail:*ryordanova@math.bas.bg