The Steep Nekhoroshev's Theorem and optimal stability exponents

M. Guzzo
Dipartimento di Matematica
Università degli Studi di Padova
Via Trieste 63 - 35121 Padova, Italy

L. Chierchia
Dipartimento di Matematica e Fisica
Università degli Studi Roma Tre
Largo San L. Murialdo 1 - 00146 Roma, Italy

G. Benettin
Dipartimento di Matematica
Università degli Studi di Padova
Via Trieste 63 - 35121 Padova, Italy

June 11, 2014

Abstract

Revising Nekhoroshev's geometry of resonances, we provide a fully constructive and quantitative proof of Nekhoroshev's theorem for steep Hamiltonian systems proving, in particular, that the exponential stability exponent can be taken to be $1/(2n\alpha_1 \cdots \alpha_{n-2})$ (α_i's being Nekhoroshev's steepness indices and $n \geq 3$ the number of degrees of freedom). On the base of a heuristic argument, we conjecture that the new stability exponent is optimal.

1 Introduction and results

A. *Motivations.* In 1977-1979 N.N. Nekhoroshev published a fundamental theorem ([22, 23]) about the "exponential stability" (i.e., "stability of action variables over times exponentially long with the inverse of the perturbation size") of nearly–integrable, real–analytic Hamiltonian systems with Hamiltonian given, in standard action–angle coordinates, by

$$H(I,\varphi) = h(I) + \varepsilon f(I,\varphi), \qquad (I,\varphi) \in U \times \mathbb{T}^n \ , \tag{1}$$

where: $U \subseteq \mathbb{R}^n$ is an open region, $\mathbb{T}^n = \mathbb{R}^n/(2\pi\mathbb{Z})^n$ is the standard flat n–dimensional torus and ε is a small parameter. The integrable limit $h(I)$ is assumed to satisfy a geometric condition, called by Nekhoroshev "steepness" (the definition is recalled in (3) below). Under such assumptions, Nekhoroshev's states his theorem as follows[1]:

Let H in (1) be real–analytic with h steep. Then, there exist positive constants a, b and ε_0 such that for any $0 \leq \varepsilon < \varepsilon_0$ the solution (I_t, φ_t) of the (standard) Hamilton equations

[1] Compare [22, p. 4 and p. 8]; see also[22, p. 30] for a more detailed and precise statement.

for $H(I,\varphi)$ satisfies

$$|I_t - I_0| \leq \varepsilon^b$$

for any time t satisfying

$$|t| \ \leq \ \frac{1}{\varepsilon} \ \exp\left(\frac{1}{\varepsilon^a}\right) \ .$$

Furthermore, a and b can be taken as follows:

$$a = \frac{2}{12\zeta + 3n + 14} \quad , \qquad b = \frac{3a}{2\alpha_{n-1}} \tag{2}$$

where

$$\zeta = \Big[\alpha_1 \cdot \Big(\alpha_2 \Big(\dots (\alpha_{n-3}(n\alpha_{n-2} + n - 2) + n - 3) + \dots \Big) + 2\Big) + 1\Big] - 1,$$

and α_i are the steepness indices of h.

Usually, a and b are called the "stability exponents". Clearly, the most relevant quantity in this theorem is the stability exponent a appearing in the exponential, which gives the dominant time–scale for the stability of the action variables. The exponential stability exponent a depends only on the number n of degrees of freedom and on the values of the first $n-2$ steepness indices α_i, $i \leq n-2$. Notice that, for any fixed n, the "best" exponents a, b in (2) are obtained in the special case $\alpha_1 = \dots = \alpha_{n-1} = 1$, corresponding to convex (or quasi–convex) $h(I)$ (which is the simplest instance of steep function). Actually, for any values of the steepness indices α_i, the parameter ζ defined in (3) grows faster than[2] $n(n-1)/2$. The hypotheses of Nekhoroshev's theorem, as pointed out by Nekhoroshev himself, are qualitatively optimal, and, in particular, non–steep Hamiltonians are in general non exponentially–stable [22, §11]. Furthermore, Nekhoroshev proved that steepness is a generic (in C^∞ category) property [21]. Finally, several interesting problems (e.g., in Celestial Mechanics, compare below) are steep but do not satisfy simpler assumptions (such as quasi–convexity). For all these reasons it seems natural and important to try to optimize the exponential stability exponents, especially with respect to the number n of the degrees of freedom which, in applications, typically range from $n = 3$ (restricted three-body problems) up to several tens (planetary problems); this has been done, up to now, under simplifying assumptions but not in the general steep case. This paper is devoted to the general case.

Before stating our result, let us briefly review the main extensions, applications and improvements concerning Nekhoroshev's theorem.

Various extensions have been discussed, so as to cover the degeneracies of the Hamilton function which are usually met in some important mechanical systems (fast rotations of the Euler–Poinsot rigid body [1, 2, 3]; the planetary N–body problem [22, 24, 9]; restricted three body problems [8], elliptic equilibria [10, 25, 15, 32]). Furthermore, steepness could be used, in non-convex systems, to study the long–term stability in problems such as the

[2]For any fixed sequence $\alpha_j \geq 1$, $j = 1, 2, \dots$, by considering a sequence of steep Hamiltonians h_n with n degrees of freedom and steepness indices $\alpha_1, \dots, \alpha_{n-1}$, the sequence of corresponding parameters $\zeta_n := \zeta$ satisfies $\zeta_n - \zeta_{n-1} \geq (n-1)\alpha_1 \cdots \alpha_{n-2}$, and the sequence of stability exponents $a_n := a$ satisfies $a_n^{-1} - a_{n-1}^{-1} \geq 6(n-1)\alpha_1 \cdots \alpha_{n-2}$.

Lagrangian equilibrium points L4-L5 of the restricted three body problem [4], asteroids of the Main Belt [20, 29, 16] and the Solar System [33, 30].
As far as improvements of the theoretical stability bounds (i.e., improvements on the stability exponent a), quite complete results have been achieved in the special case of convex and quasi–convex functions h: the proof of the theorem has been significantly simplified (see [11, 5, 6]) and the stability exponent improved up to $a = (2n)^{-1}$, ([19], [17], [31]; see [7] for exponents which are intermediate between $a = (2n)^{-1}$ and $a = (2(n-1))^{-1}$): such exponents (in the convex case) are nearly optimal, compare [35]. These improvements have been obtained by exploiting specific geometric properties of the convex and quasi–convex cases, which allow to use conservation of energy in order to obtain topological confinement of the actions ([6]). In fact, in the convex case, the analysis of the geometry of resonances, that is, the geometry of the manifolds

$$\{I \in U : k \cdot \omega(I) = 0\} \ , \quad \text{with} \quad \omega(I) = \nabla h(I) \text{ and } k \in \mathbb{Z}^n \backslash \{0\} \ ,$$

is greatly simplified, since the frequency map $I \mapsto \omega(I)$ is a diffeomorphism; on the other hand, in the general steep case, the Hamilton function cannot be used anymore in order to obtain topological confinement, and the geometry of resonances is significantly more complicate, due to possible folds and other degeneracies of the frequency map. Furthermore, while new different proofs of Nekhoroshev's theorem have appeared (compare [27], which is based on the method of simultaneous Diophantine approximations introduced in [17]), *no improvements on the original Nekhoroshev's stability exponents, in the general steep case, are yet available*[3], and *no conjecture about their optimal values has been discussed up to now* [4].

In this paper, we revisit and extend Nekhoroshev's geometric analysis obtaining, in particular, for[5] $n \geq 3$, the *new stability exponents* $a = 1/(2np_1)$ and $b = a/\alpha_{n-1}$ with p_1 being the product of the first $(n-2)$ steepness indices[6].
The new stability exponents represent an essential improvement with respect to Eq. (2); in particular, the dependence of a^{-1} on the number of degrees of freedom improves from quadratic to linear. It is also remarkable that, for $\alpha_1 = \ldots = \alpha_n = 1$ (quasi–convex case), we obtain the "optimal" stability exponents proved in [19, 17, 31], without using the local inversion of the frequency map, nor the Hamiltonian as a Lyapunov function. On the base of a simple heuristic argument, we conjecture that the new stability exponent is optimal: the details will be given in point **E** below, here we only mention that the root $1/p_1$ comes in by a natural and necessary iterative dimensional argument (related to a power–law scaling of the amplitudes of the resonance domains) and the tangential non–degeneracies given by the steepness property, while the constant in front of p_1 in the simplest and less degenerate case is the $2n$ coming from the quasi–convex case. Putting these two things together one sees that the exponent $1/(2np_1)$ is, roughly speaking, "necessary".
A precise and fully quantitative formulation of our result is given in the following paragraph.

---

[3]In the paper [26] there is a statement concerning improved values for the stability exponents, however, the proof appears to have a serious gap and such values are not justified; see [28].

[4]The need for, at least, a conjecture about the optimal value of the stability exponents in the general steep case was stressed also in [18], page 81.

[5]The cases $n \leq 2$ are, in general, totally stable and therefore are not included in our analysis.

[6]This result has been announced in [12].

B. *Statement of the result.* A C^1 function $h(I)$ is said to be steep in $U \subseteq \mathbb{R}^n$ with steepness indices $\alpha_1, \dots, \alpha_{n-1} \geq 1$ and (strictly positive) steepness coefficients $C_1, \dots, C_{n-1}$ and r, if $\inf_{I \in U} \|\omega(I)\| > 0$ and, for any $I \in U$, for any j–dimensional linear subspace $\Lambda \subseteq \mathbb{R}^n$ orthogonal to $\omega(I)$ with $1 \leq j \leq n-1$, one has[7]

$$\max_{0 \leq \eta \leq \xi} \ \min_{u \in \Lambda : \|u\| = \eta} \|\pi_\Lambda \omega(I+u)\| \geq C_j \xi^{\alpha_j} \qquad \forall\, \xi \in (0, r], \tag{3}$$

where π_Λ denotes the orthogonal projection over Λ.
To deal properly with initial data near the boundary, we will use the following notation: for any $\eta > 0$ and any $D \subseteq \mathbb{R}^n$, we let $D - \eta := \{I \in D : \overline{B(I,\eta)} \subseteq D\}$, where

$$B(I,\eta) = \{I' \in \mathbb{R}^n : \|I' - I\| < \eta\}$$

is the real Euclidean ball centered in I of radius η and $\overline{B(I,\eta)}$ its closure.

Theorem 1. *Let H in* (1) *be real–analytic with h steep in U with steepness indices α_1,...,α_{n-1} and let*

$$p_1 := \prod_{k=1}^{n-2} \alpha_k \ , \qquad a := \frac{1}{2np_1} \quad , \qquad b := \frac{a}{\alpha_{n-1}} \ . \tag{4}$$

Then, there exist positive constants $\varepsilon_0, R_0, T, c > 0$ such that for any $0 \leq \varepsilon < \varepsilon_0$ the solution (I_t, φ_t) of the Hamilton equations for $H(I,\varphi)$ with initial data (I_0, φ_0) with $I_0 \in U - 2R_0\varepsilon^b$ satisfies

$$\|I_t - I_0\| \leq R_0 \varepsilon^b \tag{5}$$

for any time t satisfying:

$$|t| \leq \frac{T}{\sqrt{\varepsilon}} \exp\left(\frac{c}{\varepsilon^a}\right) . \tag{6}$$

C. *Quantitative formulation.* Next, we provide explicit estimates for the parameters ε_0, R_0, T, c appearing in Theorem 1.

To do this, we need to introduce some notations. Given "extension parameters" $\eta, \sigma > 0$ and any set D, we let the "extended complex domains" be defined by:

$$D_\eta = \bigcup_{I' \in D} \{I \in \mathbb{C}^n : \ \ \|I - I'\| \leq \eta\} \quad \text{and} \quad \mathbb{T}^n_\sigma = \{\varphi \in \mathbb{C}^n/(2\pi\mathbb{Z})^n : \ \ |\operatorname{Im} \varphi_i| \leq \sigma\}.$$

For any real action–angle function $u(I,\varphi)$ analytic in $D_\eta \times \mathbb{T}^n_\sigma$, with Fourier harmonics $u_k(I)$, we denote its Fourier–norm

$$|u|_{\eta,\sigma} = \sum_{k \in \mathbb{Z}^n} |u_k|_{D_\eta}\, e^{|k|\sigma} \ ,$$

where $|\,.\,|_{D_\eta}$ denotes the sup–norm in D_η; if it needs to be specified, we shall also use the heavier notation $|u|_{D;\eta,\sigma}$.

[7] For any vector $u \in \mathbb{C}^n$ we denote by $\|u\| := \sqrt{\sum_i |u_i|^2}$ its hermitean norm and by $|u| = \sum_i |u_1|$.

Let H be real–analytic in $U \times \mathbb{T}^n$ with h steep in U with steepness indices $\alpha_1, \dots, \alpha_{n-1}$ and steepness coefficients $C_1, \dots, C_{n-1}$ and r. Without loss of generality, we can take the extension parameter in action space to be equal to the steepness coefficient r and we can find positive constants $s, \underline{\omega}, \overline{\omega}$ and M such that:

- $h(I)$ is real analytic on an open set which contains U_r;
- $f(I, \varphi)$ is real analytic on an open set which contains $U_r \times \mathbb{T}^n_s$;
- For any $I \in U$, we have:
$$\underline{\omega} \leq \|\omega(I)\| \leq \overline{\omega}$$
and, for any $I_1, I_2 \in U_r$, we have:
$$\|\omega(I_1) - \omega(I_2)\| \leq M\|I_1 - I_2\|.$$

Now, for $1 \leq j \leq n-2$, let

$$p_j := \prod_{k=j}^{n-2} \alpha_k \ , \quad q_j := np_j - j \ , \quad \beta_j := \alpha_j + j(\alpha_j - 1) \ , \tag{7}$$

and define the parameters

$$\kappa_j \ := \ \frac{\underline{\omega}}{M}\Big(\frac{C_j}{\underline{\omega}}\Big)^{\frac{1}{\alpha_j}} + 4\Big(2\frac{2\overline{\omega} + Mr}{\underline{\omega}}\Big)^{\frac{1}{\alpha_j}} \ , \tag{8}$$

$$E \ := \ \max\left(\max_{j \leq n-2}\left(\frac{(4M\kappa_j)^{\alpha_j}\ 6^{q_j(\alpha_j-1)}}{C_j\Big(\frac{\underline{\omega}}{2\sqrt{2}}\Big)^{\alpha_j-1}}\right)^{\frac{1}{\beta_j}} , \ 4\right) . \tag{9}$$

Then, in Theorem 1, one can take

$$\begin{aligned}
\varepsilon_* &:= \frac{1}{2^8}\frac{1}{6^{4np_1-5}E^{2np_1-1}}\frac{\underline{\omega}^2}{M|f|_{r,s}} \\
\varepsilon_0 &:= \varepsilon_* \ \min\left(\Big(\frac{6\sqrt{2}}{n}\frac{Mr}{\underline{\omega}}\Big)^{\frac{1}{b}}, \Big(\frac{18\sqrt{2}}{n}\Big)^{\frac{1}{b}}, \Big(\frac{r}{4n\kappa_{n-1}}\Big)^{\frac{1}{b}}\Big(\frac{12\sqrt{2}EC_{n-1}}{\underline{\omega}}\Big)^{\frac{1}{a}}, \Big(\frac{s}{6}\Big)^{\frac{1}{a}}, 1\right) \\
c &:= \varepsilon_*^a\ \frac{s}{6} \\
R_0 &:= \frac{r\ n\ \mu_0}{\varepsilon_*^b} \ , \text{ with } \ \mu_0 := \max\left(\frac{1}{24\sqrt{2}}\frac{\underline{\omega}}{Mr}, \frac{1}{6^2 2\sqrt{2}}, \frac{\kappa_{n-1}}{r}\Big(\frac{\underline{\omega}}{12\sqrt{2}EC_{n-1}}\Big)^{\frac{1}{\alpha_{n-1}}}\right) \\
T &:= \frac{s}{24\sqrt{2}}\frac{\underline{\omega}}{M(6E)^{\frac{1}{a}}\sqrt{\varepsilon_*}|f|_{r,s}}.
\end{aligned} \tag{10}$$

D. *On the proof.* The proof of Nekhoroshev's theorem, in its various settings, can be split into:
a *geometric part*, devoted to the analysis of distribution of small divisors in action–space;
an *analytic part*, devoted to the construction of normal forms;
a *stability argument* yielding the confinement of the actions.

While the analytic part is obtained by adapting averaging methods to an analytic setting, the heart of Nekhoroshev's theorem resides in its geometric part. The geometric part of the steep case presented in [22, 23] still needed a deep revisitation, which is performed here and leads, in particular, to substantially improved stability exponents.

The proof of Theorem 1 will be obtained by deeply revisiting the geometric part of ([22, 23]). The essential improvement are the following.
First, we extend Pöschel's Geometric Lemma (see [31]) to allow for a more general power–law scaling of the amplitudes of the resonance domains. In this way, we allow for a definition of the resonance domains which depends on the Euclidean volume (of a minimal cell) of the lattice generating the resonance, and is compatible with steepness indices $\alpha_i > 1$. In contrast with the convex case, the analog of Pöschel's Geometric Lemma is here far to accomplish the geometric part of the theorem. In fact, motions with initial conditions characterized by a given resonance, may move along preferential planes of the action space, called fast drift planes. In particular, one needs to extend in the action space with fast drift planes the resonant domains obtained by a pull back from the frequency space: eventual degeneracies of the frequency map, which are typical of the steep non–convex case, may produce topologically complicate sets. Nevertheless, a regularity of the distribution of these extended resonant sets must be proved: this is needed in order to grant the non overlapping of resonant domains of the same multiplicity. In [22, 23], the non–overlapping is granted simply by construction of the resonance domains, but the price paid was an overestimate of resonant domains with the consequence of a strong n^{-2} scaling of the stability exponent (2). Here, we do not grant the overlapping by construction but, with a careful analysis of the topology of these sets, we obtain a better balance between optimal definition of resonant domains and their non–overlapping. Finally, our geometric construction is fully compatible with the usual analytic part and stability argument, such as the so called *resonance trap* of [22, 23], and its improved version introduced in [5].

E. *Heuristic discussion.* The aim of this section is to provide some evidence that the result we obtained, for general steep Hamiltonians, is natural and cannot be substantially improved; in particular, the stability exponents a and b cannot be taken better. The reader is here assumed to have some familiarity with the traditional geometric construction entering Nekhoroshev theorem in the convex case.

The essence of Nekhoroshev theorem is to take care of the resonant motions, namely motions taking place in a neighborhood of the resonant set

$$k \cdot \omega(I) = 0 \quad , \quad k \in \mathbb{Z}^n \quad , \quad |k| \leq K := K(\epsilon),$$

with suitable $K(\epsilon) > 0$. The reduction to a finite set of resonances, up to a cut–off K, is possible thanks to the assumed analyticity of the perturbation: indeed the disregarded harmonics produce a drift of the actions with exponentially small speed of order $\exp(-\sigma K)$; choosing eventually the cut–off to be an inverse power of ϵ, precisely $K \sim \epsilon^{-a}$, the drift is negligible up to exponentially large times. Nonresonant motions are easily shown to be confined by constructing a nonresonant normal form in which all harmonics up to the cut–off are averaged out. The challenge of Nekhoroshev theorem is to show that resonant motions are also confined, within the same exponentially large times, if steepness is assumed.

The nonresonant construction is possible in a *nonresonant domain,* namely a domain which projects, in the frequency space, to a set where all small divisors satisfy an inequality of the form

$$|k \cdot \omega| \geq \lambda_1 \sim K\sqrt{\epsilon} \qquad \forall k \in \mathbb{Z}^n \backslash 0 : \quad |k| \leq K \ . \tag{11}$$

The product $K\sqrt{\epsilon}$ at the r.h.s. needs perhaps an explanation. As is well known, by an elementary perturbation step (no matter how it is performed) a factor ϵ is gained in front of the perturbation, but a factor proportional to the square of the small divisor gets lost: so, $\sqrt{\epsilon}$ is a kind of hard–core of the resonant region, and the extra factor K in (11) provides the effective gain of the step. Thanks to this extra factor, standard analytic techniques lead, with a number of steps proportional to K, to the wanted exponential estimates.

Consider now a resonance of multiplicity $j \geq 1$, i.e. a resonance related to a lattice Λ generated by $j \leq n-1$ independent vectors $k^{(1)}, \ldots, k^{(j)}$, such that $|k^{(i)}| \leq K$ for any $i \leq j$. (The case of complete resonance, $j = n$, corresponds to $\omega \simeq 0$ and is not interesting.) The exact resonance, in the frequency space, is the linear manifold of codimension j satisfying

$$k \cdot \omega = 0 \qquad \forall k \in \Lambda \ ;$$

the corresponding set in the action space, in lack of invertibility of the frequency map, could be complicated: however in this discussion, oriented to show that our result cannot be substantially improved, we shall assume $\omega(I)$ is locally invertible and in the action space too the exact resonance is a manifold of codimension j.

The influence of the resonance is assumed to extend, still in the frequency space, up to a distance λ_j depending in an essential way on the multiplicity j, larger for larger j. (Fine tuning requires that the size attributed to the resonant region also depends on the invidual lattice Λ, through the minimal volume of its cells, and not only on its dimension. This is important in the proof, but is not essential in this heuristic discussion, and will be disregarded.) Understanding the correct dependence of λ_j on ϵ is a delicate matter.

Suppose we are inside a resonance Λ, but far from other resonances. Then, standard resonant averaging shows that the motion, up to a negligible diffusion, is flattened on the so–called "plane of fast drift" $I + \langle\Lambda\rangle$, namely the plane parallel to Λ containing the initial datum I. The speed of fast drift is relatively large, namely of order ϵ. Unless during such motion new resonances are met, steepness directly provides confinement. Indeed, *the essence of steepness is that, moving away from I along the plane of fast drift,* (or any other linear manifold), *the frequency $\omega(I)$ must change, following essentially a power law in the distance from I.* Consequently, if no new resonances enter the game, resonant motions, no matter how complicated, cannot extend over a long distance, remaining inside the resonant region. The maximal distance ΔI the actions can travel inside a resonance, staying close to the fast drift plane, depends very much on the steepness indices. For convex Hamiltonians, the fast drift plane intersects transversely the resonance manifold, and $\Delta I \sim \Delta\omega$; in the general steep case instead the fast drift plane may be tangent to the local manifold, and it turns out that ΔI and $\Delta\omega$ are related by

$$\Delta I \sim \Delta\omega^{1/\alpha_j} \ . \tag{12}$$

A fast motion of course might exit the resonant region by loosing a resonance relation: in such a case however it enters a less resonant region, and the argument repeats till,

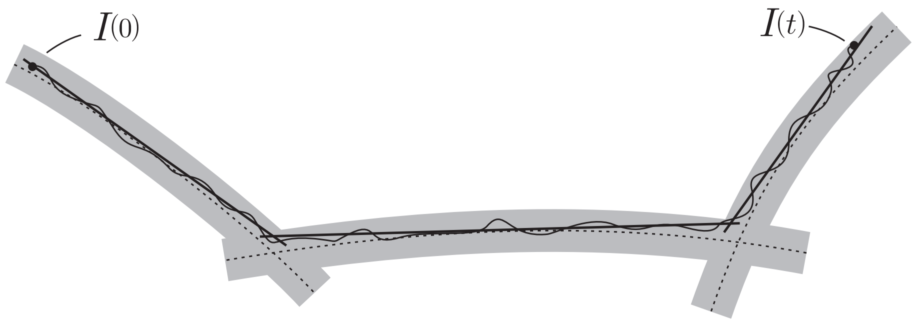

Figure 1: Illustrating the possible mechanism of fast diffusion through crossing of fast drift planes, if resonances are not well separated. Dashed curves: the resonant manifolds. Gray neighbours: the corresponding resonant regions. Lines: the planes of fast drift. Irregular curve: a possible motion of actions from $I(0)$ to $I(t)$, in a time of order $1/\epsilon$.

possibly, the nonresonant region is reached, where fast motions cannot further extend (this is the so–called "trapping" argument introduced by Nekhoroshev; only in the convex case trapping can be replaced by energy conservation [6]).

The actions can instead travel unboundedly large distances, if resonances of the same multiplicity are not well separated, and during the motion on the plane of fast drift $I + \langle\Lambda\rangle$ a new resonance is met. The possible scenario (see figure 1) is that the fast drift plane approaches a resonant manifold of larger multiplicity, there crosses some fast drift plane associated to a different resonance, and starts following it. The phenomenon could repeat: the new fast drift plane might cross, after approaching a convenient resonance of larger multiplicity, a third fast drift plane, and so on and so forth from resonance to resonance, till ΔI gets large. In lack of steepness, a similar diffusion with speed ϵ inside the web of resonances is possible and, remarkably, is compatible with the presence, in the phase space, of a large set of KAM tori. A similar fast diffusion inside the web of resonances should not be confused with the Arnol'd diffusion, which is a *much* slower quite different phenomenon.

Such a possible scenario has been described already by Nekhoroshev, who provided examples of resonances acting as "channels of superconductivity" where large variations of the actions in a time $1/\epsilon$ do occur [22, 23]. Numerical results, in turn, show the phenomenon of fast diffusion rather clearly [13, 34]: if the transversality between resonances and planes of fast drift is too weak, transitions from a resonance to a different resonance of the same multiplicity, through a resonance of higher multiplicity, are indeed observed, the higher multiplicity resonance behaving, so to speak, as a highways crossing. In such conditions, a sequence of similar transitions is numerically observed to produce a variation ΔI of the actions of order one in a polynomial time scale. In conclusion: the above scenario is realistic, the problem of crossing of fast drift planes is real, resonances of the same multiplicity need to be well separated, as a necessary condition to prevent fast diffusion and assure stability of the actions over exponentially long times.

So, let us now discuss which is the minimal separation between resonances, which is necessary in order to exclude the above crossing and prove long–term stability of the actions. We shall reason as follows: we assume that, in the convex case, the Nekhoroshev–Pöschel's geometric construction, leading to $a = b = 1/(2n)$, is optimal, and discuss which is the "unavoidable worsening" one should expect in the general steep case. The expression (4) of the stability exponents a and b will naturally follow. The claim is that our values of a and b cannot be improved, unless the corresponding estimates in the convex case are also improved.

The separation of resonances of the same multiplicity is obtained, both in the convex and in the general steep case, by appropriately tuning the sizes $\lambda_1, \dots, \lambda_{n-1}$ of the resonant regions of different multiplicity. In the convex case the appropriate scaling is

$$\lambda_{j+1} \sim K\lambda_j \ , \qquad \lambda_{n-1} \sim K^{-1} \ ,$$

with $K \sim \epsilon^{1/(2n)}$. The latter relation assures indeed that the largest resonant zones, of multiplicity $n-1$, are still as small as ϵ^b, with $b = 1/(2n)$; the factor K in the former one takes into account the fact that, by raising K, the angle between different resonant manifolds possibly gets small.

In the general steep case the above scaling is replaced by

$$\lambda_{j+1} \sim K^{A_j} \lambda_j^{1/\alpha_j} \ , \qquad A_j \geq 1 \ , \tag{13}$$

still with $\lambda_{n-1} \sim K^{-1}$. The exponent $1/\alpha_j$ of λ_j in (13) is unavoidable, for the possible tangency between resonant manifolds and fast–drift planes. The exponent A_j is a quantity that for the moment we leave undetermined. Having established that, on the one side, λ_1 must satisfy (11), while on the other side we need $\lambda_{n-1} \sim K^{-1}$, the value of K cannot be too large: more precisely, from the above scaling law (13) it follows that the dependence of K on ϵ cannot be better than

$$K \sim \epsilon^{-a} \ , \qquad a = \frac{1}{2A\,\alpha_1 \cdots \alpha_{n-2}} \ ,$$

with

$$A = 1 + A_{n-2} + \frac{A_{n-3}}{\alpha_{n-2}} + \frac{A_{n-4}}{\alpha_{n-3}\alpha_{n-2}} + \ldots + \frac{A_1}{\alpha_2 \cdots \alpha_{n-2}} + \frac{1}{\alpha_1 \cdots \alpha_{n-2}} \ . \tag{14}$$

In the convex case $\alpha_1 = \cdots = \alpha_{n-1} = 1$, $A_1 = \cdots A_{n-2} = 1$, the "optimal" value $A = n$ is recovered. In the general steep case one should obviously conjecture that such a value cannot improve, that is $A \geq n$. Computing A, with rigorous estimates everywhere, is not a trivial task: *remarkably, however, relevant compensations occur in the sum* (14), *and for all possible values of the steepness indices we are able to recover the value $A = n$, as in the convex case.* Correspondingly, we get $a = 1/(2n\alpha_1 \cdots \alpha_{n-2})$. Concerning b, its expression $b = a/\alpha_{n-1}$ directly follows from $\Delta\omega \sim \lambda_{n-1} \sim K^{-1}$ (the largest oscillation clearly occurs for $j = n-1$), using (12) to relate ΔI and $\Delta\omega$. So, the values of a and b appear to be the optimal generalizations of the convex case.[8]

F. The paper is organized as follows. The main part (i.e., the geometric analysis) is presented in § 2: in § 2.1 we introduce several auxiliary parameters (needed to measure various covering sets, small divisors, cut-offs in Fourier space, time scales, etc.) and

[8] As pointed out, e.g., in [31] (see [7] for a more refined discussion), in the convex case it is always possible to improve one exponent by worsening the other. The same argument repeats identical in the general steep case .

point out the relevant relations among them (relations, which, although based on simple calculus, are proven, for completeness in Appendix B). In § 2.2, merging and extending the geometric analysis of [22] and [31], we introduce a covering in action–space formed by (a suitable scale of) resonant and non–resonant regions. Section 2.3 is the heart of the paper, where the relevant analytic properties of the resonant and non–resonant regions are proven; the section is divided into three lemmata: the first is about geometric estimates concerning resonant domains; the second deals with small divisor estimates and the third one is a non–overlapping result for resonant regions corresponding to resonances of the same dimension. In § 3 we recall briefly Pöschel's normal form theory [31] and show how it can be used in our setting. In the final section 4 we put all pieces together and prove Theorem 1 with the constants listed in **C** above. In Appendix A we briefly review the notion of angles between linear spaces and, as mentioned above, Appendix B is an elementary check of the main relations among the auxiliary parameters.

Acknowledgement. We are indebted with L. Niederman for kindly providing Ref [28] prior to publication.

2 Geometry of resonances

2.1 Auxiliary parameters

In the proof of the Theorem 1 several auxiliary parameters will occur; in this section we define such parameters and point out some compatibility relations expressed as in inequalities, which will be needed in the following (we recall that ε_*, R_0, q_j are defined in (7) and (10)).

$$K := \left(\frac{\varepsilon_*}{\varepsilon}\right)^a \tag{15}$$

$$R := 2R_0\varepsilon^b \tag{16}$$

$$\rho := \frac{R}{2n} \ , \tag{17}$$

$$\widehat{\omega} := \frac{\omega}{2\sqrt{2}} \tag{18}$$

$$q_n := 0 \ , \quad q_{n-1} := 1 \ , \quad a_{n-1} := 1 \ , \quad a_j := q_j - q_{j+1} \quad (1 \le j \le n-2) \ . \tag{19}$$

Notice that the q_j's are strictly decreasing since $a_j \ge 1$, indeed:

$$a_{n-2} = n(\alpha_{n-2} - 1) + 1 \ ; \quad a_j = np_{j+1}(\alpha_j - 1) + 1 \ , \ (1 \le j \le n-3) \ . \tag{20}$$

Let Λ be any maximal K–lattice over $\mathbb{Z}^n$ of dimension[9] $1 \le j \le n-1$, $|\Lambda|$ its volume, and set:

$$\lambda_j := \frac{\widehat{\omega}}{(AK)^{q_j}} \ , \qquad \text{where } A := 6E \tag{21}$$

$$r_j := \kappa_j \Big(\frac{\lambda_j}{C_j}\Big)^{\frac{1}{\alpha_j}} \tag{22}$$

$$\delta_\Lambda := \frac{\lambda_j}{|\Lambda|} \tag{23}$$

$$r_\Lambda := \frac{\delta_\Lambda}{M} \tag{24}$$

$$\gamma_\Lambda := (EK)^{a_j}\delta_\Lambda \tag{25}$$

$$R_\Lambda := \frac{\gamma_\Lambda}{4MK}, \tag{26}$$

Finally, we set

$$r_0 := \frac{\lambda_1}{2MK} \tag{27}$$

$$T_0 := \frac{sr_0}{5\varepsilon|f|_{r,s}} e^{K\frac{s}{6}} \ , \qquad T_\Lambda := \frac{e\, s}{24}\frac{r_\Lambda}{\varepsilon|f|_{r,s}} e^{K\frac{s}{6}} \ , \quad T_j := \min_{\Lambda:\dim\Lambda = j} T_\Lambda \tag{28}$$

$$T_{\rm exp} := \min_{i=0,\dots,n-1} T_i. \tag{29}$$

It is then easy to check (see Appendix B) that under the assumption of Theorem 1, namely, $0 \le \varepsilon < \varepsilon_0$, for any maximal K–lattice of dimension $1 \le j \le n-1$ (unless otherwise specified) one has:

$$A \ge \max\Bigg(\max_{j\in\{1,\dots,n-1\}} \Big((E^{a_j}+1)^2 + 1 \Big)^{\frac{1}{2a_j}} \ , \ \Big(\frac{4}{E^{a_j}} + 2\Big)^{\frac{1}{a_j}} \Bigg) \tag{30}$$

$$Ks \ge 6 \tag{31}$$

$$r_\Lambda \le \min\Big(\frac{\rho}{2}, R_\Lambda\Big) \ , \tag{32}$$

$$\delta_\Lambda \le \min\Big(\frac{\omega}{r}\frac{\rho}{4} \ , \ \frac{\omega}{2r}(\rho - r_\Lambda) \ , \ \widehat{\omega} \Big), \qquad (j \le n-2) \tag{33}$$

$$KM\kappa_j \Big(\frac{\delta_\Lambda}{C_j}\Big)^{\frac{1}{\alpha_j}} \le \frac{1}{4}\gamma_\Lambda \ , \qquad (j \le n-2) \tag{34}$$

$$R_\Lambda \le r \ , \tag{35}$$

$$\varepsilon|f|_{r,s} \le \min\Bigg(\frac{1}{2^8}\frac{\lambda_1 r_0}{K} \ , \ \frac{\gamma_\Lambda r_\Lambda}{2^9 K} \ , \ \frac{\gamma_\Lambda R_\Lambda}{2^9 K} \Bigg) \tag{36}$$

$$\max_{0\le i\le n-1} r_i \le \rho \tag{37}$$

[9] We recall that a "maximal K–lattice" Λ is a lattice which admits a basis of vectors $\tilde{k} \in \mathbb{Z}^n$ with $|\tilde{k}| := \sum_{i=1}^n |\tilde{k}_i| \le K$, and it is not properly contained in any other lattice of the same dimension; the volume $|\Lambda|$ of the lattice Λ is defined as the euclidean volume of the parallelepiped spanned by a basis for Λ; (see [31]). Notice that for any K–lattice of dimension j, one has $1 \le |\Lambda| \le K^j$.

$$r_0 \leq r \tag{38}$$

$$\sum_{j=0}^{n-1} r_j \leq \frac{R}{2} \tag{39}$$

$$R \leq \frac{r}{2} \ , \tag{40}$$

$$T_{\exp} \geq \frac{T}{\sqrt{\varepsilon}} \exp\left(\frac{Ks}{6}\right) \ . \tag{41}$$

2.2 Resonant and non–resonant domains

Fix $I_0 \in U - R$ and consider the set:

$$B := B(I_0, R) \subseteq U.$$

In order to prove the stability of all motions with initial actions I_0, we need to cover the domain B with open domains where suitable normal forms adapted to the local resonance properties may be constructed. We here introduce resonant zones and resonant blocks as in [31], but, since we do not require any local inversion for the frequency map $\omega(I)$ (as it is typical of steepness [22, 23], see also [14]), these domains are directly defined in the action–space, without using any pull–back from a frequencies space. Then, we define suitable extensions, in the spirit of the original construction of [22] (see also [5]; see figure 2).

We first define the resonant zones and blocks depending on the parameter $K \geq 1$, representing a cut–off for the resonance order, and also on the parameters $0 < \lambda_1 < \ldots < \lambda_{n-1} < \widehat{\omega}$ defined above. As in [31], we consider only the resonances defined by

$$k \cdot \omega(I) = 0$$

with k in some maximal K–lattice $\Lambda \subseteq \mathbb{Z}^n$. We define the **resonant zone**

$$\mathcal{Z}_\Lambda := \{I \in B : \quad \|\pi_{\langle\Lambda\rangle}\omega(I)\| < \delta_\Lambda\}, \tag{42}$$

where $\langle\Lambda\rangle$ denotes the real vector space spanned by the lattice Λ, and the **resonant block**

$$B_\Lambda := \mathcal{Z}_\Lambda \backslash \mathcal{Z}_{j+1} \quad , \quad j = \dim \Lambda, \tag{43}$$

where:

$$\mathcal{Z}_i := \cup_{\{\Lambda': \ \dim \Lambda' = i\}} \mathcal{Z}_{\Lambda'}.$$

We also define $\mathcal{Z}_0 := B$ and the **non–resonant block** B_0 by

$$B_0 := \mathcal{Z}_0 \backslash \mathcal{Z}_1.$$

We remark that, since $\|\omega(I)\| \geq \underline{\omega} > \widehat{\omega} \geq \delta_\Lambda$ for any $I \in B$, the completely resonant zone $\mathcal{Z}_{\mathbb{Z}^n}$ is empty and so is $\mathcal{Z}_n$. This implies

$$B_\Lambda = \mathcal{Z}_\Lambda \ , \qquad \forall \ \Lambda \text{ s.t.} \quad \dim \Lambda = n-1 \ . \tag{44}$$

Furthermore, if one defines

$$B_j := \cup_{\{\Lambda' : \dim \Lambda' = j\}} B_{\Lambda'} \ ,$$

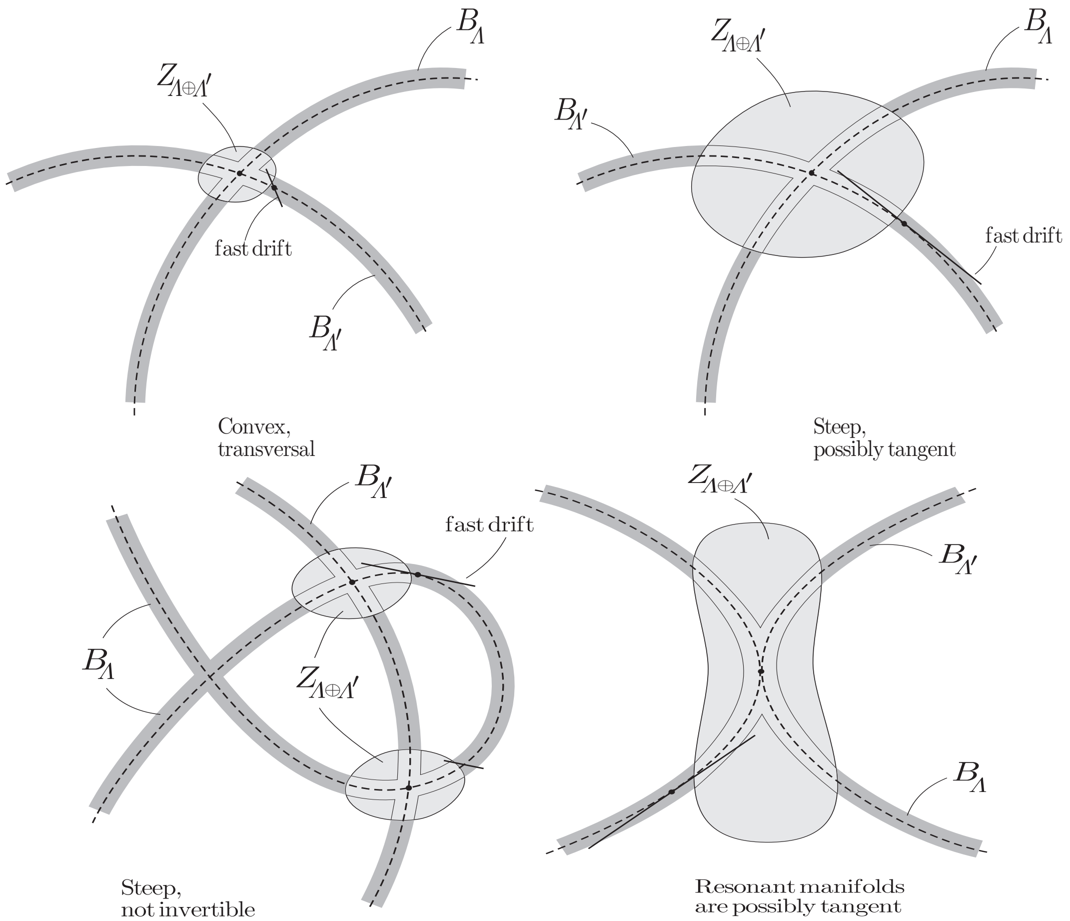

Figure 2: Sketch of the resonant sets: two resonances related to the lattices Λ and Λ' of the same multiplicity j are sketched, together with the resonance related to the lattice $\Lambda \oplus \Lambda'$ of multiplicity $j+1$, identified at the crossing of the two previous resonances, and some fast drift lines. Precisely, dashed curves represent the resonant manifolds of multiplicity j; dark gray areas denote resonant blocks of multiplicity j; light gray areas represent resonant zones of multiplicity $j+1$; by λ we denote some j dimensional fast-drift lines. Different situations are illustrated: in the top-left panel we plot the situation which is typical of the quasi–convex cases, with the fast drift planes which are transverse to the resonance; in the other panels we illustrate some situations which one may encounter in the more generic steep cases, where the fast drift planes may be tangent to the resonant manifolds, and also self-intersections (bottom-left panel) and tangencies of resonances (bottom–right panel) are possible. In all the represented cases, the fast drift lines $\lambda = (I+ < \Lambda >) \cap \mathcal{Z}_\Lambda$ or $\lambda = (I+ < \Lambda' >) \cap \mathcal{Z}'_\Lambda$ through points $I \in B_{\Lambda \bigoplus \Lambda'}$ do not intersect other resonant blocks of multiplicity $j+1$.

one sees immediately that

$$B_j = \mathcal{Z}_j \setminus \mathcal{Z}_{j+1},$$

so that, for any $1 \le j \le n-1$, we have:

$$B = B_0 \cup B_1 \cup ... \cup B_{j-1} \cup \mathcal{Z}_j, \tag{45}$$

and, in particular,

$$B = B_0 \cup B_1 \cup ... \cup B_{n-1} \ . \tag{46}$$

Next, following Nekhoroshev, we introduce **discs**

$$\mathcal{D}^{\rho}_{\Lambda,\eta}(I) := \left(\left(\bigcup_{I' \in I + \langle\Lambda\rangle} B(I',\eta) \right) \cap \mathcal{Z}_\Lambda \cap (B-\rho) \right)^{I} \subseteq \mathcal{Z}_\Lambda \cap (B-\rho), \qquad (47)$$

where $I + \langle\Lambda\rangle$ (called by Nekhoroshev, "fast drift plane") denotes the plane through I parallel to $\langle\Lambda\rangle$, $(C)^{I}$ denotes the connected component of a set C which contains I and η is any positive number less or equal than ρ. The **extended resonant blocks** are then defined by[10]:

$$B^{\rho}_{\Lambda,r_\Lambda} := \bigcup_{I \in B_\Lambda \cap (B-\rho)} \mathcal{D}^{\rho}_{\Lambda,r_\Lambda}(I) \subseteq \mathcal{Z}_\Lambda \cap (B-\rho), \qquad (48)$$

and the **extended non–resonant block** by:

$$B^{\rho}_{0} := B_0 \cap (B-\rho).$$

We remark that the set $B-\rho$ is not empty since $\rho < R$, and for any lattice Λ with $\dim \Lambda = n-1$, we have, by (44), (48) and footnote 10,

$$B^{\rho}_{\Lambda,r_\Lambda} = B_\Lambda \cap (B-\rho) \ , \qquad (\dim \Lambda = n-1) \ . \qquad (49)$$

2.3 Geometric properties of the resonant domains

• *Geometric estimates for resonant domains*

For any maximal K–lattice Λ, we need to estimate the diameter of the intersection of the fast drift planes $I + \langle\Lambda\rangle$ with the resonant zones:

Lemma 2.1 *For any $I' \in B_\Lambda \cap (B-\rho)$ and $I'' \in \mathcal{C}^{\rho}_{\Lambda,r_\Lambda}(I')$ we have:*

$$\|I' - I''\| \le \kappa_j \Big(\frac{\delta_\Lambda}{C_j}\Big)^{\frac{1}{\alpha_j}} \le r_j \ . \qquad (50)$$

Proof We divide the proof of this lemma in three steps.
Step 1. Let $\tilde\delta, \tilde\rho > 0$ be such that

$$\tilde\delta \le \min\Big(\frac{\tilde\rho}{r}, \frac{1}{\sqrt{2}}\Big)\underline{\omega} \ , \qquad (51)$$

and define

$$\mathcal{Z}_\Lambda(\tilde\delta) = \{ I \in B : \ \ \|\pi_{\langle\Lambda\rangle}\omega(I)\| < \tilde\delta \} \ . \qquad (52)$$

Let us also denote by $\langle\omega\rangle$ the linear space generated by $\omega(I)$; by $\langle\omega\rangle^{\perp}$ the linear space orthogonal to $\omega(I)$ and by $\Lambda_\omega = \pi_{\langle\omega\rangle^{\perp}}\langle\Lambda\rangle$ the linear space obtained by projecting every vector u of $\langle\Lambda\rangle$ on $\langle\omega\rangle^{\perp}$.
The first step will consist in proving that:

[10]Notice that, if $I \in B_\Lambda$, then $I \in \mathcal{D}^{\rho}_{\Lambda,\eta}(I)$ so that $B_\Lambda \cap (B-\rho) \subseteq B^{\rho}_{\Lambda,r_\Lambda}$.

For any $I \in \mathcal{Z}_\Lambda(\tilde\delta) \cap (B-\tilde\rho)$ *and any* $I' \in \Big((I+\langle\Lambda\rangle) \cap \mathcal{Z}_\Lambda(\tilde\delta) \cap (B-\tilde\rho) \Big)^I$ *one has:*

$$\| I - I' \| < 4 \Big(\frac{2\overline{\omega} + Mr}{\underline{\omega}} \frac{\tilde\delta}{C_j} \Big)^{\frac{1}{\alpha_j}} \quad . \tag{53}$$

Fix $I' \in \Big((I+\langle\Lambda\rangle) \cap \mathcal{Z}_\Lambda(\tilde\delta) \cap (B-\tilde\rho) \Big)^I$, with $I' \neq I$ (if $I' = I$ there is nothing to prove). Then, there exists a curve[11] $u(t) \in \langle\Lambda\rangle$ such that $u(0) = 0$, $u(1) = I' - I$, and for any t, $I + u(t) \in \Big(\big(I + \langle\Lambda\rangle \big) \cap \mathcal{Z}_\Lambda(\tilde\delta) \cap (B-\tilde\rho) \Big)^I$. In particular, $\|\pi_{\langle\Lambda\rangle}\omega(I+u(t))\| < \tilde\delta$.

The proof of (53) will be based on the following claims (i)÷(vii).

(i) Λ_ω *is a vector space of dimension* j.
Proof of (i): Clearly, if $u_1, \dots, u_j$ is a basis for $\langle\Lambda\rangle$, then any vector in Λ_ω can be represented as a linear combination of $\pi_{\langle\omega\rangle^\perp} u_1$, ..., $\pi_{\langle\omega\rangle^\perp} u_j \in \Lambda_\omega$. We prove that the vectors $\pi_{\langle\omega\rangle^\perp} u_1, \dots, \pi_{\langle\omega\rangle^\perp} u_j \in \Lambda_\omega$ are linearly independent, so that $\dim \Lambda_\omega = j$. First, we remark that the only vector u of $\langle\Lambda\rangle$ satisfying: $\pi_{\langle\omega\rangle^\perp} u = 0$ is $u = 0$. In fact, if there exists $u \neq 0$ such that $u \in \langle\Lambda\rangle$ and $\pi_{\langle\omega\rangle^\perp} u = 0$, then $\omega(I) \in \langle\Lambda\rangle$, and therefore we have:

$$\|\omega(I)\| = \|\pi_{\langle\Lambda\rangle}\omega(I)\| < \tilde\delta \le \frac{\underline{\omega}}{\sqrt{2}},$$

which is not possible since for any $I \in B$ we assumed $\|\omega(I)\| > \underline{\omega}$. Now, let us consider $c_1, \dots, c_j$ such that: $\sum_{i=1}^j c_i \pi_{\langle\omega\rangle^\perp} u_i = 0$. Then, $\pi_{\langle\omega\rangle^\perp} \sum_i c_i u_i = 0$, and therefore $\sum_i c_i u_i = 0$. But, since the u_i are linearly independent, it follows $c_1, \dots, c_j = 0$.

(ii) *For any* $u \in \langle\Lambda\rangle$, *we have* $\pi_{\Lambda_\omega} u = \pi_{\langle\omega\rangle^\perp} u$.
Proof of (ii): We first compute:

$$\pi_{\langle\omega\rangle^\perp} u = \pi_{\Lambda_\omega} \pi_{\langle\omega\rangle^\perp} u + \pi_{\Lambda_\omega^\perp} \pi_{\langle\omega\rangle^\perp} u \quad . \tag{54}$$

Since $\pi_{\langle\omega\rangle^\perp} u \in \Lambda_\omega$, we have $\pi_{\Lambda_\omega^\perp} \pi_{\langle\omega\rangle^\perp} u = 0$, so that (54) becomes:

$$\pi_{\langle\omega\rangle^\perp} u = \pi_{\Lambda_\omega} \pi_{\langle\omega\rangle^\perp} u \quad . \tag{55}$$

But, $\pi_{\Lambda_\omega} u = \pi_{\Lambda_\omega}(\pi_{\langle\omega\rangle^\perp} u + \pi_{\langle\omega\rangle} u) = \pi_{\Lambda_\omega}\pi_{\langle\omega\rangle^\perp} u + \pi_{\Lambda_\omega}\pi_{\langle\omega\rangle} u$ and since $\Lambda_\omega \subseteq \langle\omega\rangle^\perp$, we have $\pi_{\Lambda_\omega}\pi_{\langle\omega\rangle} u = 0$, and therefore:

$$\pi_{\Lambda_\omega} u = \pi_{\Lambda_\omega}\pi_{\langle\omega\rangle^\perp} u. \tag{56}$$

¿From equations (55) and (56) we get (ii).

(iii) *The angle*[12] *between* $\langle\Lambda\rangle$ *and* Λ_ω *is equal to the angle between* $\omega(I)$ *and* $\langle\Lambda\rangle^\perp$, *in formulae:*

$$\langle\Lambda\rangle \angle \Lambda_\omega = \omega(I) \angle \langle\Lambda\rangle^\perp. \tag{57}$$

Proof of (iii): By (ii) we have: $\langle\Lambda\rangle \angle \Lambda_\omega = \max_{u \in \langle\Lambda\rangle, u \neq 0} u \angle \pi_{\Lambda_\omega} u = \max_{u \in \langle\Lambda\rangle, u \neq 0} u \angle \pi_{\langle\omega\rangle^\perp} u = \langle\Lambda\rangle \angle \langle\omega\rangle^\perp$, and using (x) of Appendix A, we obtain $\langle\Lambda\rangle \angle \Lambda_\omega = \langle\Lambda\rangle \angle \langle\omega\rangle^\perp = \langle\omega\rangle \angle \langle\Lambda\rangle^\perp$.

[11]Notice that the set $\Big((I+\langle\Lambda\rangle) \cap \mathcal{Z}_\Lambda(\tilde\delta) \cap (B-\tilde\rho) \Big)^I$ is open in the relative topology of $I + \langle\Lambda\rangle$ and therefore is arc–connected in $I + \langle\Lambda\rangle$.

[12]The notion of angle between linear spaces is briefly reviewed in Appendix A.

(iv) *For any t, one has* $\|\pi_{\Lambda_\omega}\omega(I+u(t))\| < \dfrac{2\,\overline{\omega}\,\tilde{\delta}}{\underline{\omega}}$.

Proof of (iv): We start with

$$\begin{aligned}
\|\pi_{\Lambda_\omega}\omega(I+u(t))\| &= \sqrt{\|\omega(I+u(t))\|^2 - \|\pi_{{\Lambda_\omega}^\perp}\omega(I+u(t))\|^2} \\
&= \|\omega(I+u(t))\|\sqrt{1-|\cos(\omega(I+u(t))\,\angle\,{\Lambda_\omega}^\perp)|^2} \\
&= \|\omega(I+u(t))\||\sin(\omega(I+u(t))\,\angle\,{\Lambda_\omega}^\perp)| \\
&\leq \overline{\omega}|\sin(\omega(I+u(t))\,\angle\,{\Lambda_\omega}^\perp)|
\end{aligned}$$

and then we produce an upper bound estimate of the angle $\omega(I+u(t))\,\angle\,{\Lambda_\omega}^\perp$. By using property (ix) of Appendix A, we first obtain:

$$\omega(I+u(t))\,\angle\,{\Lambda_\omega}^\perp \leq \omega(I+u(t))\,\angle\,\langle\Lambda\rangle^\perp + \langle\Lambda\rangle^\perp\,\angle\,{\Lambda_\omega}^\perp. \tag{58}$$

Now, recalling that $\langle\Lambda\rangle$ and Λ_ω have the same dimension (claim (i) above), we see that by properties (x) and (xi) of Appendix A, $\langle\Lambda\rangle^\perp\,\angle\,{\Lambda_\omega}^\perp = \Lambda_\omega\,\angle\,\langle\Lambda\rangle = \langle\Lambda\rangle\,\angle\,\Lambda_\omega = \omega(I)\,\angle\,\langle\Lambda\rangle^\perp$. ¿From (58), we therefore obtain:

$$\omega(I+u(t))\,\angle\,{\Lambda_\omega}^\perp \leq \omega(I+u(t))\,\angle\,\langle\Lambda\rangle^\perp + \omega(I)\,\angle\,\langle\Lambda\rangle^\perp \quad . \tag{59}$$

Then, since:

$$\begin{aligned}
|\sin\Big(\omega(I+u(t))\,\angle\,\langle\Lambda\rangle^\perp\Big)| &= \frac{\|\pi_{\langle\Lambda\rangle}\omega(I+u(t))\|}{\|\omega(I+u(t))\|} < \frac{\tilde{\delta}}{\underline{\omega}} \\
|\sin\Big(\omega(I)\,\angle\,\langle\Lambda\rangle^\perp\Big)| &= \frac{\|\pi_{\langle\Lambda\rangle}\omega(I)\|}{\|\omega(I)\|} < \frac{\tilde{\delta}}{\underline{\omega}},
\end{aligned} \tag{60}$$

and $\tilde{\delta}/\underline{\omega} \leq 1/\sqrt{2}$, both angles are strictly smaller than $\pi/4$, their sum is strictly smaller than $\pi/2$, and since $\sin(x)$ is monotone in $[0,\pi/2]$, from (59) and standard trigonometry, we obtain:

$$\begin{aligned}
|\sin\omega(I+u(t))\,\angle\,{\Lambda_\omega}^\perp| &\leq |\sin(\omega(I+u(t))\,\angle\,\langle\Lambda\rangle^\perp + \omega(I)\,\angle\,\langle\Lambda\rangle^\perp)| \\
&\leq |\sin(\omega(I+u(t))\,\angle\,\langle\Lambda\rangle^\perp)| + |\sin(\omega(I)\,\angle\,\langle\Lambda\rangle^\perp)| < 2\,\frac{\tilde{\delta}}{\underline{\omega}}.
\end{aligned}$$

We therefore obtain: $\|\pi_{\Lambda_\omega}\omega(I+u(t))\| \leq \overline{\omega}|\sin\omega(I+u(t))\,\angle\,{\Lambda_\omega}^\perp| < \dfrac{2\,\overline{\omega}\,\tilde{\delta}}{\underline{\omega}}$.

(v) $\|\pi_{\langle\omega\rangle}u(t)\| < \dfrac{\tilde{\delta}}{\underline{\omega}}\|u(t)\|$.

Proof of (v): Since $u(t) \in \langle\Lambda\rangle$ and $I \in \mathcal{Z}_\Lambda(\tilde{\delta})$, we have:

$$\|\pi_{\langle\omega\rangle}u(t)\| = \frac{|\omega(I)\cdot u(t)|}{\|\omega(I)\|} = \frac{|\pi_{\langle\Lambda\rangle}\omega(I)\cdot u(t)|}{\|\omega(I)\|} < \frac{\tilde{\delta}}{\underline{\omega}}\|u(t)\| \quad .$$

(vi) $I + \pi_{\langle\omega\rangle^\perp}u(t) \in B$.

Proof of (vi): Since $I, I+u(t) \in B-\tilde{\rho}$, we have $\|u(t)\| \leq 2r$ and, using (40), we obtain $\|u(t)\| \leq r$. Then, from (v) and (51), we have:$\|\pi_{\langle\omega\rangle} u(t)\| < \frac{\tilde{\delta}}{\underline{\omega}}\|u(t)\| \leq \frac{\tilde{\delta}}{\underline{\omega}} r \leq \tilde{\rho}$. Therefore, $I + \pi_{\langle\omega\rangle^\perp} u(t) \in B$.

(vii) $\xi := \|\pi_{\langle\omega\rangle^\perp}(I'-I)\| \in (0, r]$.
Proof of (vii): Let us first assume $\xi = 0$, that is $I' - I \in \langle\omega\rangle$ so that

$$I' - I = \omega(I) \frac{\|I'-I\|}{\|\omega(I)\|}.$$

Since $I' - I \in \langle\Lambda\rangle$ and $I' \neq I$, this would imply also $\omega(I) \in \langle\Lambda\rangle$, and therefore:

$$\|\omega(I)\| = \|\pi_{\langle\Lambda\rangle}\omega(I)\| < \tilde{\delta} \leq \frac{\underline{\omega}}{\sqrt{2}},$$

which is not possible since for any $I \in B$ we have $\|\omega(I)\| > \underline{\omega}$. Therefore we have $\xi > 0$. Then, we have

$$\xi = \|\pi_{\langle\omega\rangle^\perp}(I'-I)\| \leq \|I'-I\| = \|u(1)\| \leq r.$$

Now, we are ready to complete the proof of (53). Since $0 < \xi \leq r$, let $0 \leq \eta_* \leq \xi$ the η which realizes the maximum in the definition of the steepness index of dimension j, that is:

$$\min_{u \in \Lambda_\omega:\ \|u\|=\eta_*} \|\pi_{\Lambda_\omega}\omega(I+u)\| > C_j \xi^{\alpha_j}. \tag{61}$$

The curve $\pi_{\langle\omega\rangle^\perp} u(t) \in \Lambda_\omega$ joins I and $I + \pi_{\langle\omega\rangle^\perp}(I'-I)$, and therefore

$$[0, \xi] \subseteq \cup_{t\in[0,1]} \|\pi_{\langle\omega\rangle^\perp} u(t)\|\ ,$$

so that there exists $t_* \in [0,1]$ such that $\|\pi_{\langle\omega\rangle^\perp} u(t_*)\| = \eta_*$. From (61) it follows:

$$\|\pi_{\Lambda_\omega}\omega(I + \pi_{\langle\omega\rangle^\perp} u(t_*)\| > C_j \xi^{\alpha_j}.$$

But using claims (iv) and (v) we also obtain:

$$\begin{aligned}
\|\pi_{\Lambda_\omega}\omega(I + \pi_{\langle\omega\rangle^\perp} u(t_*)\| &\leq \|\pi_{\Lambda_\omega}\omega(I + u(t_*))\| + M\|\pi_{\langle\omega\rangle} u(t_*)\| < 2\frac{\overline{\omega}}{\underline{\omega}}\tilde{\delta} + M\frac{\tilde{\delta}}{\underline{\omega}}\|u(t_*)\| \\
&< \frac{2\overline{\omega} + Mr}{\underline{\omega}}\,\tilde{\delta}
\end{aligned}$$

so that

$$C_j \xi^{\alpha_j} < \frac{2\overline{\omega} + Mr}{\underline{\omega}}\,\tilde{\delta},$$

and therefore

$$\|\pi_{\langle\omega\rangle^\perp}(I'-I)\| = \xi < \Big(\frac{2\overline{\omega} + Mr}{\underline{\omega}}\,\frac{\tilde{\delta}}{C_j}\Big)^{\frac{1}{\alpha_j}}.$$

Using again (v), we obtain:

$$\|I'-I\| \leq \|\pi_{\langle\omega\rangle^\perp}(I'-I)\| + \|\pi_{\langle\omega\rangle}(I'-I)\|$$

$$
\begin{aligned}
&< \Big(\frac{2\overline{\omega}+Mr}{\underline{\omega}}\frac{\tilde{\delta}}{C_j}\Big)^{\frac{1}{\alpha_j}} + \frac{\tilde{\delta}}{\underline{\omega}}\|I'-I\| \\
&\le \Big(\frac{2\overline{\omega}+Mr}{\underline{\omega}}\frac{\tilde{\delta}}{C_j}\Big)^{\frac{1}{\alpha_j}} + \frac{1}{\sqrt{2}}\|I'-I\|, \qquad (62)
\end{aligned}
$$

that is:

$$
\|I'-I\| < \frac{1}{1-\frac{1}{\sqrt{2}}}\Big(\frac{2\overline{\omega}+Mr}{\underline{\omega}}\frac{\tilde{\delta}}{C_j}\Big)^{\frac{1}{\alpha_j}} < 4\Big(\frac{2\overline{\omega}+Mr}{\underline{\omega}}\frac{\tilde{\delta}}{C_j}\Big)^{\frac{1}{\alpha_j}}.
$$

This finishes the proof of (53).

Step 2. Next, we prove that:
For any $I \in \mathcal{Z}_\Lambda \cap (B-\rho)$ and any $I' \in \mathcal{D}^{\rho}_{\Lambda,r_\Lambda}(I)$, we have:

$$
\|I-I'\| \le r_\Lambda + 4\Big(\frac{2\overline{\omega}+Mr}{\underline{\omega}}\frac{\delta_\Lambda+Mr_\Lambda}{C_j}\Big)^{\frac{1}{\alpha_j}}. \qquad (63)
$$

Fix $I' \in \mathcal{D}^{\rho}_{\Lambda,r_\Lambda}(I)$. Since $\mathcal{D}^{\rho}_{\Lambda,r_\Lambda}(I)$ is open and connected, there exists a curve $I+u(t) \in \mathcal{D}^{\rho}_{\Lambda,r_\Lambda}(I)$, $t \in [0,1]$, such that $I+r(0) = I$, $I+u(1) = I'$. Since $\mathcal{D}^{\rho}_{\Lambda,r_\Lambda}(I) \subseteq \mathcal{Z}_\Lambda$, we have: $\|\pi_{\langle\Lambda\rangle}\omega(I+u(t))\| < \delta_\Lambda$ for any $t \in [0,1]$, and also $\|\pi_{\langle\Lambda\rangle^\perp}u(t)\| \le r_\Lambda$. In fact, since $I+u(t) \in \mathcal{D}^{\rho}_{\Lambda,r_\Lambda}(I) \subseteq \cup_{\tilde{I}\in I+\langle\Lambda\rangle}B(\tilde{I},r_\Lambda)$, there exists a curve $u'(t) \in \langle\Lambda\rangle$ such that $\|u(t)-u'(t)\| \le r_\Lambda$, and therefore

$$
\|\pi_{\langle\Lambda\rangle^\perp}u(t)\| = \|\pi_{\langle\Lambda\rangle^\perp}(u(t)-u'(t))\| \le \|u(t)-u'(t)\| \le r_\Lambda\ .
$$

Then, we define $u''(t) := \pi_{\langle\Lambda\rangle}u(t)$, so that $u''(0) = \pi_{\langle\Lambda\rangle}u(0) = 0$, $I+u''(t) \in I+\langle\Lambda\rangle$, and

$$
\|u''(t)-u(t)\| = \|\pi_{\langle\Lambda\rangle}u(t)-u(t)\| = \|\pi_{\langle\Lambda\rangle^\perp}u(t)\| \le r_\Lambda.
$$

Therefore, on the one hand we have $I+u''(t) \in B-\rho+r_\Lambda$, on the other hand:

$$
\begin{aligned}
\|\pi_{\langle\Lambda\rangle}\omega(I+u''(t))\| &\le \|\pi_{\langle\Lambda\rangle}\omega(I+u(t))\| + \|\pi_{\langle\Lambda\rangle}(\omega(I+u''(t))-\omega(I+u(t)))\| \\
&\le \|\pi_{\langle\Lambda\rangle}\omega(I+u(t))\| + \|\omega(I+u''(t))-\omega(I+u(t))\| \\
&\le \|\pi_{\langle\Lambda\rangle}\omega(I+u(t))\| + M\|u''(t)-u(t)\| \le \delta_\Lambda + Mr_\Lambda.
\end{aligned}
$$

Therefore, for any $t \in [0,1]$, we have:

$$
I+u''(t) \in \Big((I+\langle\Lambda\rangle)\cap\mathcal{Z}_\Lambda(\delta_\Lambda+Mr_\Lambda)\cap(B-(\rho-r_\Lambda))\Big)^I.
$$

We use, now, (53) (step 1) with

$$
\tilde{\delta} := \delta_\Lambda + Mr_\Lambda \qquad \text{and} \qquad \tilde{\rho} := \rho - r_\Lambda\ .
$$

In fact, $I \in \mathcal{Z}_\Lambda \subseteq \mathcal{Z}_\Lambda(\tilde{\delta})$; $I \in B-\rho \subseteq B-\tilde{\rho}$; from (33) it follows:

$$
\tilde{\delta} \le \min\Big(\frac{\underline{\omega}}{r}\tilde{\rho}\ ,\ \frac{\underline{\omega}}{\sqrt{2}}\Big).
$$

Therefore, we have:

$$\|u''(t)\| \le 4\Big(\frac{2\overline{\omega}+Mr}{\underline{\omega}}\frac{\tilde{\delta}}{C_j}\Big)^{\frac{1}{\alpha_j}} = 4\Big(\frac{2\overline{\omega}+Mr}{\underline{\omega}}\frac{\delta_\Lambda+Mr_\Lambda}{C_j}\Big)^{\frac{1}{\alpha_j}},$$

for any $t\in[0,1]$. In particular we have:

$$\|I'-I\| = \|u(1)\| \le \|u''(1)-u(1)\| + \|u''(1)\| \le r_\Lambda + 4\Big(\frac{2\overline{\omega}+Mr}{\underline{\omega}}\frac{\delta_\Lambda+Mr_\Lambda}{C_j}\Big)^{\frac{1}{\alpha_j}}.$$

Step 3. We may conclude the proof of the lemma. ¿From (32) and (33) we obtain

$$\delta_\Lambda + Mr_\Lambda = 2\delta_\Lambda \le \frac{\underline{\omega}\rho}{2r} \le \frac{\underline{\omega}}{r}(\rho - r_\Lambda),$$

so that applying (63) and using again (32), we have:

$$\|I'-I''\| \le r_\Lambda + 4\Big(\frac{2\overline{\omega}+Mr}{\underline{\omega}}\frac{\delta_\Lambda+Mr_\Lambda}{C_j}\Big)^{\frac{1}{\alpha_j}} \le \frac{\delta_\Lambda}{M} + 4\Big(2\frac{2\overline{\omega}+Mr}{\underline{\omega}}\frac{\delta_\Lambda}{C_j}\Big)^{\frac{1}{\alpha_j}}.$$

Then, since $\alpha_j \ge 1$ and (recall (33)) $\delta_\Lambda/\underline{\omega} < 1$, we have $(\delta_\Lambda/\underline{\omega}) \le (\delta_\Lambda/\underline{\omega})^{\frac{1}{\alpha_j}}$, from which the first inequality in (50) follows at once; the second inequality follows from the fact that $\delta_\Lambda \le \lambda_j$ and from the definition of r_j. $\square$

• <u>*Small divisor estimates*</u>

We recall ([31]) that a set $\tilde{B} \subseteq B$ is γ–K non resonant modulo Λ if we have $|k\cdot\omega(I)| \ge \gamma$ for any $k \in \mathbb{Z}^n\backslash\Lambda$ such that $|k| \le K$; we will say that $\tilde{B}\subseteq B$ is γ–non resonant if $|k\cdot\omega(I)| \ge \gamma$ for any $k\in\mathbb{Z}^n\backslash\{0\}$ such that $|k|\le K$. The following result is a generalization of the Geometric Lemma in [31].

Lemma 2.2 (i) *For any maximal K–lattice Λ, the resonant block B_Λ is γ_Λ–K non resonant modulo Λ, while the non resonant block B_0 is λ_1–K non resonant.*
(ii) *If $j = n-1$, the extended block $B^\rho_{\Lambda,r_\Lambda}$ is γ_Λ–K non–resonant modulo Λ; if $j \le n-2$, the extended block $B^\rho_{\Lambda,r_\Lambda}$ is $\gamma_\Lambda/2$–K non–resonant modulo Λ.*

Proof of (i): Let us first consider $I\in B_0$, so that $I \notin \mathcal{Z}_1$. For any $k\in\mathbb{Z}^n$, with $|k|\le K$, let us denote by $\tilde{k}$ the vector which generates the maximal one dimensional K–lattice containing k. Since $I\notin\mathcal{Z}_1$ we have:

$$\|\pi_{\langle\tilde{k}\rangle}\omega(I)\| \ge \frac{\lambda_1}{\|\tilde{k}\|} \ge \frac{\lambda_1}{\|k\|},$$

and consequently $|k\cdot\omega(I)| = \|k\|\|\pi_{\langle\tilde{k}\rangle}\omega(I)\| \ge \lambda_1$. Therefore, B_0 is λ_1–K non resonant.

Now, consider a maximal K–lattice Λ, with $j := \dim\Lambda \in \{1,\dots,n-1\}$ and let $I\in B_\Lambda$. As in [31], let $k\notin\Lambda$ with $|k|\le K$ and denote by Λ_+ the maximal K–lattice generated by Λ and k (since Λ is maximal, $\dim\Lambda_+ = j+1$). For the purpose of this proof, let us denote

$$\pi := \pi_{\langle\Lambda\rangle}\ , \quad \pi_\perp := \pi_{\langle\Lambda\rangle^\perp} = \mathrm{Id} - \pi\ , \qquad \pi_+ := \pi_{\langle\Lambda_+\rangle}\ ,$$

where Id denotes the identity map. Since $\pi\pi_+ = \pi$, it is easy to check that

$$\pi_\perp k \cdot \Big(\pi_+\omega(I) - \pi\omega(I)\Big) + \pi k \cdot \pi\omega(I) = k \cdot \omega(I) \ .$$

Thus, since the vectors $\pi_\perp k$ and $\pi_+\omega(I) - \pi\omega(I) = \pi_\perp\pi_+\omega(I)$ are proportional, and $|\Lambda_+| \le |\Lambda| \, \|\pi_\perp k\|$, we obtain

$$\begin{aligned} |k\cdot\omega(I)| &\ge \left|\pi_\perp k \cdot \big(\pi_+\omega(I) - \pi\omega(I)\big)\right| - \left|\pi k \cdot \pi\omega(I)\right| \\ &= \|\pi_\perp k\| \, \|\pi_+\omega(I) - \pi\omega(I)\| - \left|\pi k \cdot \pi\omega(I)\right| \\ &\ge \frac{|\Lambda_+|}{|\Lambda|}\sqrt{\|\pi_+\omega(I)\|^2 - \|\pi\omega(I)\|^2} - \|\pi k\|\|\pi\omega(I)\|. \end{aligned}$$

Using $\|\pi k\| \le \|k\| \le |k| \le K$, $\|\pi\omega(I)\| < \delta_\Lambda$, $\|\pi\omega(I)\| \ge \delta_{\Lambda_+}$ we obtain:

$$|k\cdot\omega(I)| \ge \frac{|\Lambda_+|}{|\Lambda|}\sqrt{\frac{\lambda_{j+1}^2}{|\Lambda_+|^2} - \frac{\lambda_j^2}{|\Lambda|^2}} - K\delta_\Lambda \ .$$

Using again $|\Lambda_+| \le |\Lambda| K$, and $K \le K^{a_j}$, we obtain:

$$\begin{aligned} |k\cdot\omega(I)| &\ge \frac{1}{|\Lambda|}\Big(\sqrt{\lambda_{j+1}^2 - K^2\lambda_j^2} - K\lambda_j\Big) \ge \delta_\Lambda\Bigg(\sqrt{\Big(\frac{\lambda_{j+1}}{\lambda_j}\Big)^2 - K^2} - K\Bigg) \\ &\ge \delta_\Lambda\Big(\sqrt{(AK)^{2a_j} - K^2} - K\Big) \ge \delta_\Lambda\left(\sqrt{(AK)^{2a_j} - K^{2a_j}} - K^{a_j}\right) \\ &= \delta_\Lambda K^{a_j}\Big(\sqrt{A^{2a_j} - 1} - 1\Big), \end{aligned} \tag{64}$$

so that, by (30), we finally get:

$$|k\cdot\omega(I)| \ge \delta_\Lambda K^{a_j}\Big(\sqrt{A^{2a_j} - 1} - 1\Big) \ge E^{a_j} K^{a_j}\delta_\Lambda = \gamma_\Lambda.$$

□

Proof of (ii): If $j = n-1$, the conclusion follows directly from lemma 2.2-(i) and (49). Let us therefore consider, for any $j = 1, \dots, n-2$, $I \in B^\rho_{\Lambda, r_\Lambda}$ and $I' \in B_\Lambda \cap (B - \rho)$ such that $I \in \mathcal{C}^\rho_{\Lambda, r_\Lambda}(I')$. By (33) and (63) we get

$$\|I' - I\| \le r_\Lambda + 4\Big(\frac{2\overline{\omega} + Mr}{\underline{\omega}} \frac{\delta_\Lambda + Mr_\Lambda}{C_j}\Big)^{\frac{1}{\alpha_j}}.$$

Using also lemma 2.2-(i), for any $k \in \mathbb{Z}^n \backslash \Lambda$ with $|k| \le K$, we have

$$|k\cdot\omega(I)| \ge |k\cdot\omega(I')| - KM\|I - I'\| \ge \gamma_\Lambda - KM\Big(r_\Lambda + 4\Big(\frac{2\overline{\omega} + Mr}{\underline{\omega}}\frac{\delta_\Lambda + Mr_\Lambda}{C_j}\Big)^{\frac{1}{\alpha_j}}\Big). \tag{65}$$

But, since, by (32) and (25), $KMr_\Lambda \le \gamma_\Lambda/4$ from (34) there follows

$$\begin{aligned} 4KM\Big(\frac{2\overline{\omega} + Mr}{\underline{\omega}}\frac{\delta_\Lambda + Mr_\Lambda}{C_j}\Big)^{\frac{1}{\alpha_j}} &\le 4KM\Big(2\frac{2\overline{\omega} + Mr}{\underline{\omega}}\frac{\delta_\Lambda}{C_j}\Big)^{\frac{1}{\alpha_j}} \\ &\le KM\kappa_j\Big(\frac{\delta_\Lambda}{C_j}\Big)^{\frac{1}{\alpha_j}} \le \frac{\gamma_\Lambda}{4}, \end{aligned} \tag{66}$$

which, together with (65) yields $|k\cdot\omega(I)| \ge \gamma_\Lambda/2$. □

• *Non overlapping of extended blocks and zones*

Lemma 2.3 *For any maximal K–lattices $\Lambda \neq \Lambda'$ of the same dimension $j = 1, \dots, n-1$, we have*

$$\overline{B^{\rho}_{\Lambda,r_\Lambda}} \cap \mathcal{Z}_{\Lambda'} = \emptyset.$$

Proof. Let $\Lambda \neq \Lambda'$ be maximal K–lattices of the same dimension $j \leq n-1$ and consider $I \in \overline{B^{\rho}_{\Lambda,r_\Lambda}}$: we have to prove that $I \notin \mathcal{Z}_{\Lambda'}$, i.e.,

$$\|\pi_{\langle\Lambda'\rangle}\omega(I)\| \geq \delta_{\Lambda'}. \tag{67}$$

We divide the proof in two steps: the case $j \leq n-2$ and the case $j = n-1$.

Step 1. ($1 \leq j \leq n-2$). The argument follows from the following claims (i)÷(vi).

(i) *For any $\eta > 0$, there exists $I' \in B_\Lambda \cap (B - \rho)$ such that $\|I - I'\| \leq \kappa_j \left(\frac{\delta_\Lambda}{C_j}\right)^{\frac{1}{\alpha_j}} + \eta$.*

Proof of (i): Since $I \in \overline{B^{\rho}_{\Lambda,r_\Lambda}}$, there exists $I'' \in B^{\rho}_{\Lambda,r_\Lambda}$ such that $\|I''-I\| < \eta$; (by definition of $B^{\rho}_{\Lambda,r_\Lambda}$) there exists $I' \in B_\Lambda \cap (B-\rho)$ such that $I'' \in \mathcal{D}^{\rho}_{\Lambda,r_\Lambda}(I')$. Then, (i) immediately follows from (50).

(ii) $\|\pi_{\langle\Lambda'\rangle}\omega(I')\| \geq E^{a_j} K^{a_j-1}\delta_\Lambda$.

Proof of (ii): Since $\Lambda \neq \Lambda'$, there exists $k \in \Lambda'$ such that $k \notin \Lambda$ and $|k| \leq K$. Therefore we have $\|\pi_{\langle\Lambda'\rangle}\omega(I')\| \geq |k \cdot \omega(I')|/\|k\|$ and since $I' \in B_\Lambda$, (ii) follows from Lemma 2.2.

(iii) $\|\pi_{\langle\Lambda'\rangle}\omega(I)\| \geq \frac{1}{2}E^{a_j}K^{a_j-1}\delta_\Lambda$.

Proof of (iii): Choose $\eta \leq \frac{\gamma_\Lambda}{4KM}$. Then, by using (34), (i) and (ii), we obtain

$$\begin{aligned}\|\pi_{\langle\Lambda'\rangle}\omega(I)\| &\geq \|\pi_{\langle\Lambda'\rangle}\omega(I')\| - M\|I - I'\| \geq E^{a_j}K^{a_j-1}\delta_\Lambda - M\eta - M\kappa_j\left(\frac{\delta_\Lambda}{C_j}\right)^{\frac{1}{\alpha_j}} \\ &\geq E^{a_j}K^{a_j-1}\delta_\Lambda - \frac{\gamma_\Lambda}{2K} = \frac{1}{2}E^{a_j}K^{a_j-1}\delta_\Lambda.\end{aligned}$$

Now, observe that, if we have $\frac{1}{2}E^{a_j}K^{a_j-1}\delta_\Lambda \geq \delta_{\Lambda'}$, then (67) follows at once. Therefore, *let us henceforth assume that*

$$\frac{1}{2}E^{a_j}K^{a_j-1}\delta_\Lambda < \delta_{\Lambda'} \qquad \text{i.e.} \qquad \frac{|\Lambda'|}{|\Lambda|} < \frac{2}{E^{a_j}K^{a_j-1}}\ . \tag{68}$$

(iv) $\|\pi_{\langle\Lambda'\rangle}\omega(I')\| \geq A^{a_j}K^{a_j-1}\delta_{\Lambda'} - 2\delta_\Lambda$.

Proof of (iv): Since $\Lambda \neq \Lambda'$, we consider $k \in \Lambda$ such that $k \notin \Lambda'$ and $|k| \leq K$, and we denote by Λ'' the maximal K–lattice of dimension $j+1$ which contains Λ' and k. For the purpose of the proof of (iv) let us denote:

$$\pi := \pi_{\langle\Lambda\rangle}\ , \qquad \pi' := \pi_{\langle\Lambda'\rangle}\ , \qquad \pi'' := \pi_{\langle\Lambda''\rangle}\ .$$

First, since $I' \in B_\Lambda$, we have

$$\|\pi'\omega(I')\| \geq \|\pi'(\mathrm{Id} - \pi)\omega(I')\| - \|\pi'\pi\omega(I')\| \geq \|\pi'(\mathrm{Id} - \pi)\omega(I')\| - \delta_\Lambda. \tag{69}$$

Then, since $I' \in B_\Lambda$, $I' \notin \mathcal{Z}_{\Lambda''}$ and we have

$$\|\pi''\omega(I')\| \geq \delta_{\Lambda''}. \tag{70}$$

Let us consider the vector $\nu = \pi_{\langle\Lambda'\rangle^\perp} k$. We remark that $\nu \in \langle\Lambda''\rangle\backslash 0$. In fact, on the one hand $k \notin \langle\Lambda'\rangle$, so that $\nu \neq 0$; on the other hand, since $\nu = k - \pi_{\langle\Lambda'\rangle} k$ is the sum of $k \in \langle\Lambda\rangle$ and of $-\pi' k \in \langle\Lambda'\rangle$, we have also $\nu \in \langle\Lambda''\rangle$. Therefore, since ν is orthogonal to $\langle\Lambda'\rangle$, we have:

$$\pi_{\langle\Lambda''\rangle} = \pi_{\langle\Lambda'\rangle} + \pi_{\langle\nu\rangle}. \tag{71}$$

Moreover, we have:

$$\frac{|\nu \cdot k|}{\|\nu\|} \geq \frac{|\Lambda''|}{|\Lambda'|}. \tag{72}$$

In fact, on the one hand we have

$$\frac{|\nu \cdot k|}{\|\nu\|} = \frac{|\pi_{\langle\Lambda'\rangle^\perp} k \cdot k|}{\|\pi_{\langle\Lambda'\rangle^\perp} k\|} = \|\pi_{\langle\Lambda'\rangle^\perp} k\|,$$

on the other hand we have

$$\|\pi_{\langle\Lambda'\rangle^\perp} k\| \geq \frac{|\Lambda''|}{|\Lambda'|}.$$

¿From (71), we obtain:

$$\begin{aligned}
\|\pi'(\mathrm{Id} - \pi)\omega(I')\|^2 &= \|\pi'\pi''(\mathrm{Id} - \pi)\omega(I')\|^2 \\
&= \|\pi''(\mathrm{Id} - \pi)\omega(I')\|^2 - \|\pi_{\langle\nu\rangle}\pi''(\mathrm{Id} - \pi)\omega(I')\|^2 \\
&= \|\pi''(\mathrm{Id} - \pi)\omega(I')\|^2 - \frac{|\nu \cdot \pi''(\mathrm{Id} - \pi)\omega(I')|^2}{\|\nu\|^2}.
\end{aligned} \tag{73}$$

We notice that:

$$\pi''(\mathrm{Id} - \pi)\omega(I') \neq 0. \tag{74}$$

In fact, first we have

$$\begin{aligned}
\|\pi''(\mathrm{Id} - \pi)\omega(I')\| &\geq \|\pi''\omega(I')\| - \|\pi''\pi\omega(I')\| \geq \delta_{\Lambda''} - \delta_\Lambda \\
&\geq \frac{\lambda_j}{|\Lambda''|}\left(A^{a_j}K^{a_j} - \frac{|\Lambda''|}{|\Lambda|}\right),
\end{aligned}$$

then, using (72), (68), we obtain

$$\frac{|\Lambda''|}{|\Lambda|} = \frac{|\Lambda''|}{|\Lambda'|}\frac{|\Lambda'|}{|\Lambda|} \leq \|k\|\frac{|\Lambda'|}{|\Lambda|} < \frac{2\|k\|}{E^{a_j}K^{a_j-1}} \leq \frac{2K}{E^{a_j}K^{a_j-1}},$$

and therefore we have:

$$\|\pi''(\mathrm{Id} - \pi)\omega(I')\| > \frac{\lambda_j}{|\Lambda''|}\left(A^{a_j}K^{a_j} - \frac{2K}{E^{a_j}K^{a_j-1}}\right).$$

Finally, since $K \geq 1$, $a_j \geq 1$, and using also (30), we have

$$\begin{aligned}
\|\pi''(\mathrm{Id} - \pi)\omega(I')\| &> \frac{\lambda_j}{|\Lambda''|}\left(A^{a_j}K - \frac{2K}{E^{a_j}}\right) = \frac{\lambda_j K}{|\Lambda''|}\left(A^{a_j} - \frac{2}{E^{a_j}}\right) \\
&\geq \frac{\lambda_j K}{|\Lambda''|}\left(2 + \frac{2}{E^{a_j}}\right) > 0.
\end{aligned}$$

Therefore, from (73), (74), we have:

$$\|\pi'(\mathrm{Id}-\pi)\omega(I')\| = \|\pi''(\mathrm{Id}-\pi)\omega(I')\|\sqrt{1-\frac{(\nu\cdot\pi''(\mathrm{Id}-\pi)\omega(I'))^2}{\|\nu\|^2\|\pi''(\mathrm{Id}-\pi)\omega(I')\|^2}},$$

and, since

$$\pi''(\mathrm{Id}-\pi)\omega(I')\cdot k = (\mathrm{Id}-\pi)\omega(I')\cdot k = 0,$$

we obtain:

$$\|\pi'(\mathrm{Id}-\pi)\omega(I')\| \geq \|\pi''(\mathrm{Id}-\pi)\omega(I')\| \min_{u\in k^{\perp},\|u\|=1}\sqrt{1-\frac{(\nu\cdot u)^2}{\|\nu\|^2}}. \tag{75}$$

We remark that the maximum of $|\nu\cdot u| = |\pi_{\langle k^{\perp}\rangle}\nu\cdot u|$, for $u\in k^{\perp}$ and $\|u\|=1$, is obtained for u parallel to $\pi_{\langle k^{\perp}\rangle}\nu$, that is for $u=\pi_{\langle k^{\perp}\rangle}\nu/\|\pi_{\langle k^{\perp}\rangle}\nu\|$. Therefore, we have:

$$\max_{u\in k^{\perp},\|u\|=1}|\nu\cdot u| = \|\pi_{\langle k^{\perp}\rangle}\nu\|$$

and correspondingly:

$$\begin{aligned}\min_{u\in k^{\perp},\|u\|=1}\sqrt{1-\frac{(\nu\cdot u)^2}{\|\nu\|^2}} &= \sqrt{1-\frac{\|\pi_{\langle k^{\perp}\rangle}\nu\|^2}{\|\nu\|^2}} = \sqrt{\frac{\|\nu\|^2-\|\pi_{\langle k^{\perp}\rangle}\nu\|^2}{\|\nu\|^2}} = \frac{\|\pi_{\langle k\rangle}\nu\|}{\|\nu\|}\\ &= \frac{|\nu\cdot k|}{\|\nu\|\|k\|}.\end{aligned}$$

Therefore, from (75) and (72) we obtain

$$\|\pi'(\mathrm{Id}-\pi)\omega(I')\| \geq \|\pi''(\mathrm{Id}-\pi)\omega(I')\|\frac{|\nu\cdot k|}{\|\nu\|\|k\|} \geq \|\pi''(\mathrm{Id}-\pi)\omega(I')\|\frac{|\Lambda''|}{|\Lambda'|\|k\|}.$$

Then, since: $\|\pi''(\mathrm{Id}-\pi)\omega(I')\| \geq \|\pi''\omega(I')\| - \|\pi''\pi\omega(I')\| \geq \delta_{\Lambda''}-\delta_{\Lambda}$, we obtain:

$$\|\pi'(\mathrm{Id}-\pi)\omega(I')\| \geq (\delta_{\Lambda''}-\delta_{\Lambda})\frac{|\Lambda''|}{|\Lambda'|\|k\|} \geq \frac{\lambda_{j+1}}{|\Lambda'|\|k\|}-\delta_{\Lambda}\frac{|\Lambda''|}{|\Lambda'|\|k\|} \geq A^{a_j}K^{a_j-1}\delta_{\Lambda'}-\delta_{\Lambda}, \tag{76}$$

and using (69) we obtain (iv).

We now are ready to finish the proof of (67) in the case $j\leq n-2$. ¿From inequalities (iv) and (i), we obtain:

$$\|\pi'\omega(I)\| \geq A^{a_j}K^{a_j-1}\delta_{\Lambda'}-2\delta_{\Lambda}-M\Big(\kappa_j\Big(\frac{\delta_{\Lambda}}{C_j}\Big)^{\frac{1}{\alpha_j}}+\eta\Big). \tag{77}$$

Using (34) and choosing $\eta\leq\frac{\gamma_{\Lambda}}{4KM}$, we obtain:

$$\|\pi'\omega(I)\| \geq A^{a_j}K^{a_j-1}\delta_{\Lambda'}-2\delta_{\Lambda}-\frac{1}{2}E^{a_j}K^{a_j-1}\delta_{\Lambda}. \tag{78}$$

Since we are assuming (68) and since $K^{a_j-1}\geq 1$, we obtain

$$\begin{aligned}\|\pi'\omega(I)\| &\geq A^{a_j}K^{a_j-1}\delta_{\Lambda'}-2\delta_{\Lambda}-\frac{1}{2}E^{a_j}K^{a_j-1}\delta_{\Lambda}\\ &> A^{a_j}K^{a_j-1}\delta_{\Lambda'}-\frac{4}{E^{a_j}K^{a_j-1}}\delta_{\Lambda'}-\delta_{\Lambda'} > \Big(A^{a_j}-\frac{4}{E^{a_j}}-1\Big)\delta_{\Lambda'},\end{aligned}$$

which, by (30), yields (67).

Step 2. We now consider maximal K–lattices $\Lambda \neq \Lambda'$ of the same dimension $j = n-1$. Since $B^{\rho}_{\Lambda,r_\Lambda} = B_\Lambda \cap (B-\rho)$, we have $I \in \overline{B}_\Lambda$ and

$$\|\pi'\omega(I)\| \geq E^{a_j} K^{a_j-1}\delta_\Lambda. \tag{79}$$

In fact, since $\Lambda \neq \Lambda'$, there exists $k \in \Lambda'$ such that $k \notin \Lambda$ and $|k| \leq K$. Therefore we have $\|\pi'\omega(I)\| \geq |k\cdot\omega(I)|/\|k\|$ and since $I \in \overline{B}_\Lambda$, by lemma 2.2 we have

$$\|\pi'\omega(I)\| \geq \frac{|k\cdot\omega(I)|}{\|k\|} \geq E^{a_j} K^{a_j-1}\delta_\Lambda.$$

We also have:

$$\|\pi'\omega(I)\| \geq A^{a_j} K^{a_j-1}\delta_{\Lambda'} - 2\delta_\Lambda \quad . \tag{80}$$

First, since $I \in \overline{B}_\Lambda$, we have

$$\|\pi'\omega(I)\| \geq \|\pi'(\mathrm{Id}-\pi)\omega(I)\| - \|\pi'\pi\omega(I)\| \geq \|\pi'(\mathrm{Id}-\pi)\omega(I)\| - \delta_\Lambda. \tag{81}$$

Then, since $\Lambda \neq \Lambda'$, we consider $k \in \Lambda$ such that $k \notin \Lambda'$ and $|k| \leq K$. In particular, since $I \in \overline{B}_\Lambda$, we have

$$\|\omega(I)\| \geq \widehat{\omega}. \tag{82}$$

Let us consider the vector $\nu = \pi_{\langle\Lambda'\rangle^\perp} k$. Since ν is orthogonal to $\langle\Lambda'\rangle$, we have:

$$\mathrm{Id} = \pi' + \pi_{\langle\nu\rangle}. \tag{83}$$

Moreover, we have:

$$\|\pi_{\langle\Lambda'\rangle^\perp} k\| = \frac{|\nu\cdot k|}{\|\nu\|} \geq \frac{1}{|\Lambda'|}. \tag{84}$$

In fact, since the K–lattice $\langle\Lambda', k\rangle$ is generated by Λ' and k is properly contained in $\mathbb{Z}^n$, we have

$$|\Lambda'|\|\pi_{\langle\Lambda'\rangle^\perp} k\| \geq |\langle\Lambda', k\rangle| \geq 1.$$

¿From (83), we obtain:

$$\|\pi(\mathrm{Id}-\pi)\omega(I)\|^2 = \|(\mathrm{Id}-\pi)\omega(I)\|^2 - \|\pi_{\langle\nu\rangle}(\mathrm{Id}-\pi)\omega(I)\|^2,$$

and since:

$$\|\pi_{\langle\nu\rangle}(\mathrm{Id}-\pi)\omega(I)\| = \frac{|\nu\cdot(\mathrm{Id}-\pi)\omega(I)|}{\|\nu\|},$$

and:

$$\|(\mathrm{Id}-\pi)\omega(I)\| \geq \widehat{\omega} - \delta_\Lambda \geq \widehat{\omega} - \lambda_{n-1} > 0,$$

we have:

$$\|\pi'(\mathrm{Id}-\pi)\omega(I)\| = \|(\mathrm{Id}-\pi)\omega(I)\|\sqrt{1 - \frac{(\nu\cdot(\mathrm{Id}-\pi)\omega(I))^2}{\|\nu\|^2\|(\mathrm{Id}-\pi)\omega(I)\|^2}}.$$

Then, since $(\mathrm{Id}-\pi)\omega(I)\cdot k=0$, we have

$$\|\pi'(\mathrm{Id}-\pi)\omega(I')\| \geq \|(\mathrm{Id}-\pi)\omega(I')\| \min_{u\in k^\perp, \|u\|=1} \sqrt{1-\frac{(\nu\cdot u)^2}{\|\nu\|^2}}. \tag{85}$$

We remark that the maximum of $|\nu\cdot u| = |\pi_{\langle k^\perp\rangle}\nu\cdot u|$, for $u\in k^\perp$ and $\|u\|=1$, is obtained for u parallel to $\pi_{\langle k^\perp\rangle}\nu$, that is for $u=\pi_{\langle k^\perp\rangle}\nu/\|\pi_{\langle k^\perp\rangle}\nu\|$. Therefore, we have:

$$\max_{u\in k^\perp, \|u\|=1} |\nu\cdot u| = \|\pi_{\langle k^\perp\rangle}\nu\|$$

and correspondingly:

$$\begin{aligned} \min_{u\in k^\perp, \|u\|=1} \sqrt{1-\frac{(\nu\cdot u)^2}{\|\nu\|^2}} &= \sqrt{1-\frac{\|\pi_{\langle k^\perp\rangle}\nu\|^2}{\|\nu\|^2}} = \sqrt{\frac{\|\nu\|^2-\|\pi_{\langle k^\perp\rangle}\nu\|^2}{\|\nu\|^2}} = \frac{\|\pi_{\langle k\rangle}\nu\|}{\|\nu\|} \\ &= \frac{|\nu\cdot k|}{\|\nu\|\|k\|}. \end{aligned}$$

Therefore, from (85), we obtain:

$$\|\pi'(\mathrm{Id}-\pi)\omega(I)\| \geq \|(\mathrm{Id}-\pi)\omega(I)\| \frac{|\nu\cdot k|}{\|\nu\|\|k\|},$$

and from (84) we obtain also:

$$\|\pi'(\mathrm{Id}-\pi)\omega(I)\| \geq \|(\mathrm{Id}-\pi)\omega(I)\| \frac{1}{|\Lambda'|\|k\|}.$$

Then, since: $\|(\mathrm{Id}-\pi)\omega(I)\| \geq \|\omega(I)\| - \|\pi\omega(I)\| \geq \widehat{\omega}-\delta_\Lambda$, we obtain:

$$\|\pi'(\mathrm{Id}-\pi)\omega(I)\| \geq (\widehat{\omega}-\delta_\Lambda)\frac{1}{|\Lambda'|\|k\|} \geq A^{a_j}K^{a_j-1}\delta_{\Lambda'}-\delta_\Lambda, \tag{86}$$

and using (81) we obtain (80).

If $E^{a_j}K^{a_j-1}\delta_\Lambda \geq 2\delta_{\Lambda'}$, using (79), there is nothing more to prove. Therefore, we assume:

$$E^{a_j}K^{a_j-1}\delta_\Lambda < 2\delta_{\Lambda'}\ .$$

Then, using (80), we obtain:

$$\|\pi'\omega(I)\| \geq A^{a_j}K^{a_j-1}\delta_{\Lambda'}-\delta_\Lambda > A^{a_j}K^{a_j-1}\delta_{\Lambda'} - \frac{2}{E^{a_j}K^{a_j-1}}\delta_{\Lambda'}. \tag{87}$$

Since $K^{a_j-1}\geq 1$, we have:

$$\|\pi'\omega(I)\| > \Big(A^{a_j}-\frac{2}{E^{a_j}}\Big)\delta_{\Lambda'},$$

and using (30) we obtain (67).

□

3 Normal forms and dynamics in resonant blocks

The geometric construction of § 2 together with normal form theory allows to have some control of the dynamics in the extended blocks. We shall use normal form theory in the version given by Pöschel in [31]; see, in particular, the "Normal Form Lemma" at p. 192 of [31] (which we shall use with parameters $p = q = 2$); notice that the constant M used in [31] is an upper bound on the derivative of $\omega(I)$, which is used only as Lipschitz constant, so that our notation is consistent with that used in [31].
In fact, the following lemma holds

Lemma 3.1 (i) *Let* (I_t, φ_t) *be the solution of the Hamilton equations with initial condition*[13] $(\bar{I}_0, \varphi_0) \in B_0^\rho \times \mathbb{T}^n$*. Then,*

$$\|I_t - \bar{I}_0\| \le r_0 \tag{88}$$

for all times[14] $|t| \le T_0$*.*
(ii) *Let* Λ *be a maximal* K*–lattice of dimension* $j \in \{1, \dots, n-1\}$*, and let* (I_t, φ_t) *be the solution of the Hamilton equations with initial data* $(\bar{I}_0, \varphi_0) \in (B_\Lambda \cap (B - (j+1)\rho)) \times \mathbb{T}^n$*. Let* $\tau_{\rm e}$ *be the (possibly infinite) exit time from*[15] $B^\rho_{\Lambda, r_\Lambda}$*. Then, if* $|\tau_{\rm e}| \ge T_\Lambda$*, we have* $\|I_t - \bar{I}_0\| \le r_j$ *for any time* $|t| < T_\Lambda$*; otherwise, there exists* $0 \le i \le j-1$ *such that* $I_{\tau_{\rm e}} \in B_i \cap (B - j\rho)$*.*

Proof of (i): The non–resonant block B_0 is λ_1 non–resonant (see Lemma 2.2). Let us consider as extension vector (r_0, s). Because of the definition of r_0, (38), (31) and the first inequality in (36), we can apply the normal form lemma in [31] in B_0. It then follows at once (88) for all times $|t| \le T_0$ with T_0 as in (28). □

Proof of (ii): Let us first assume that $|\tau_{\rm e}| \ge T_\Lambda$ and consider the extension vector (r_Λ, s). By Lemma 2.2–(ii), the domain $B^\rho_{\Lambda, r_\Lambda}$ is $\gamma_\Lambda/2$–K non–resonant modulo Λ. Thus, since $r_\Lambda \le r$ (by (35)), and because of (31), the definition of R_Λ and the third inequality in (36), we can apply the Normal Form Lemma in [31] (with $p = q = 2$), in $B^\rho_{\Lambda, r_\Lambda}$. Thus, there exists a canonical transformation:

$$\begin{aligned} \phi: \quad (B^\rho_{\Lambda, r_\Lambda})_{\frac{r_\Lambda}{2}} \times \mathbb{T}^n_{\frac{s}{6}} &\to (B^\rho_{\Lambda, r_\Lambda})_{r_\Lambda} \times \mathbb{T}^n_s \\ (I', \varphi') &\longmapsto (I, \varphi) = \phi(I', \varphi') \end{aligned} \tag{89}$$

conjugating H to its resonant normal form:

$$H_\Lambda = H \circ \phi = h + \varepsilon g + \varepsilon f_* \tag{90}$$

with g a real–analytic functions having the Fourier expansion

$$g = \sum_{k \in \Lambda} g_k \exp(i k \cdot \varphi) \ , \tag{91}$$

[13] I.e. $I_t|_{t=0} = \bar{I}_0$: we are using here a slight abuse of notation in order not to confuse the point I_0 in the statement of Theorem 1 with the arbitrary point $\bar{I}_0$ used here.

[14] Recall the definition of T_0 in (28).

[15] I.e., $\tau_{\rm e}$ is such that $I_t \in B^\rho_{\Lambda, r_\Lambda}$ for $|t| < |\tau_{\rm e}|$ and $I_{\tau_{\rm e}} \notin B^\rho_{\Lambda, r_\Lambda}$.

and the "remainder" f_* satisfying the exponential bound:

$$|f_*|_{B^\rho_{\Lambda,r_\Lambda};\frac{r_\Lambda}{2},\frac{s}{6}} \leq e^{-K\frac{s}{6}}|f|_{r,s}. \tag{92}$$

Also, for any $(I',\varphi') \in (B^\rho_{\Lambda,r_\Lambda})_{\frac{r_\Lambda}{2}} \times \mathbb{T}^n$, by the third inequality in (36), one has:

$$\|I'-I\| \leq \frac{8K}{\gamma_\Lambda}\varepsilon|f|_{r,s} \leq \frac{1}{2^6}r_\Lambda,$$

so that $\phi^{-1}(B^\rho_{\Lambda,r_\Lambda} \times \mathbb{T}^n) \subseteq (B^\rho_{\Lambda,r_\Lambda})_{\frac{r_\Lambda}{2^6}} \times \mathbb{T}^n$. Finally, using the second inequality in (36) we have also:

$$\|I'-I\| \leq \frac{8K}{\gamma_\Lambda}\varepsilon|f|_{r,s} \leq \frac{1}{2^6}r_\Lambda\ .$$

Therefore, since $I_t \in B^\rho_{\Lambda,r_\Lambda}$ for any $|t| < |\tau_{\rm e}|$, we may define $(I'_t,\varphi'_t) = \phi^{-1}(I_t,\varphi_t)$, and using the specific form of Hamiltonian (90), we have

$$\|\pi_{\langle\Lambda^\perp\rangle}(I'_t - I'_0)\| \leq \varepsilon\|\int_0^t \frac{\partial f_*}{\partial\varphi}(I'_t,\varphi'_t)dt\| \leq \varepsilon|t| \sup_{(B^\rho_{\Lambda,r_\Lambda})_{\frac{r_\Lambda}{2^6}}\times\mathbb{T}^n} \|\frac{\partial f_*}{\partial\varphi}\|.$$

By Cauchy estimate (see Lemma B.3 of [31]) and by (92), we have:

$$\sup_{(B^\rho_{\Lambda,r_\Lambda})_{\frac{r_\Lambda}{2^6}}\times\mathbb{T}^n} \|\frac{\partial f_*}{\partial\varphi}\| \leq \frac{6}{es}|f_*|_{B^\rho_{\Lambda,r_\Lambda};\frac{r_\Lambda}{2},\frac{s}{6}} \leq \frac{6}{es}e^{-K\frac{s}{6}}|f|_{r,s},$$

so that, for any $|t| < T_\Lambda$, we have:

$$\|\pi_{\langle\Lambda^\perp\rangle}(I'_t - I'_0)\| \leq \frac{6\varepsilon}{es}|t|e^{-K\frac{s}{6}}|f|_{r,s} \leq \frac{1}{4}r_\Lambda.$$

As a consequence, the motion I_t has the representation:

$$I_t = \bar{I}_0 + v(t) + d(t) \tag{93}$$

with $v(t) \in \langle\Lambda\rangle$ with $v(0)=0$, and $\|d(t)\| < \frac{3}{4}r_\Lambda$: indeed, we can write

$$I_t = \bar{I}_0 + (I_t - I'_t) + (I'_t - I'_0) + (I'_0 - \bar{I}_0),$$

and take $v(t) = \pi_{\langle\Lambda\rangle}(I'_t - I'_0)$ and $d(t) = (I_t - I'_t) + \pi_{\langle\Lambda\rangle^\perp}(I'_t - I'_0) + (I'_0 - \bar{I}_0)$.
Therefore, $I_t \in B^\rho_{\Lambda,r_\Lambda} \subseteq \mathcal{Z}_\Lambda \cap (B-\rho)$ and because of the representation (93) the distance between I_t and the space $\bar{I}_0 + \langle\Lambda\rangle$ is smaller than $\frac{3}{4}r_\Lambda$. Furthermore, I_t is connected to $\bar{I}_0$ in the set $\left(\cup_{I'\in\bar{I}_0+\langle\Lambda\rangle} B(I',\frac{3}{4}r_\Lambda)\right) \cap \mathcal{Z}_\Lambda \cap (B-\rho)$ so that $I_t \in \mathcal{C}^\rho_{\Lambda,\frac{3}{4}r_\Lambda}(\bar{I}_0) \subseteq \mathcal{C}^\rho_{\Lambda,r_\Lambda}(\bar{I}_0)$. Thus, by Lemma 2.1 we have $\|I_t - \bar{I}_0\| \leq r_j$ for any $|t| < T_\Lambda$, as claimed.

Let us now assume that the exit time $\tau_{\rm e}$ satisfies: $0 < |\tau_{\rm e}| < T_\Lambda$. Since for any time $|t| < |\tau_{\rm e}|$, we have $I_t \in \mathcal{C}^\rho_{\Lambda,\frac{3}{4}r_\Lambda}(\bar{I}_0)$, we have also: $I_{\tau_{\rm e}} \in \overline{\mathcal{C}^\rho_{\Lambda,\frac{3}{4}r_\Lambda}(\bar{I}_0)}$. As a consequence (again Lemma 2.1), we have $\|I_t - \bar{I}_0\| \leq r_j < \rho$ for any $|t| \leq |\tau_{\rm e}|$and since $\bar{I}_0 \in B - (j+1)\rho$, we also have

$$I_{\tau_{\rm e}} \in B - j\rho\ . \tag{94}$$

Since $I_t \in \mathcal{C}^{\rho}_{\Lambda,\frac{3}{4}r_\Lambda}(\bar{I}_0)$, the distance between I_t and $\bar{I}_0 + \langle\Lambda\rangle$ is strictly smaller than $\frac{3}{4}r_\Lambda$, and the distance between $I_{\mathcal{T}_e}$ and $\bar{I}_0 + \langle\Lambda\rangle$ is smaller or equal than $\frac{3}{4}r_\Lambda$. Finally, since $I_t \in \mathcal{C}^{\rho}_{\Lambda,\frac{3}{4}r_\Lambda}(\bar{I}_0)$, we have $I_t \in \mathcal{Z}_\Lambda$, that is: $\|\pi_{\langle\Lambda\rangle}\omega(I_t)\| < \delta_\Lambda$. As a consequence, since $I_{\mathcal{T}_e} \notin B^{\rho}_{\Lambda,r_\Lambda}$, the only possibility is: $\|\pi_{\langle\Lambda\rangle}\omega(I_{\mathcal{T}_e})\| = \delta_\Lambda$. But this means that $I_{\mathcal{T}_e} \notin \mathcal{Z}_\Lambda$. On the other hand, by Lemma 2.3, $I_{\mathcal{T}_e}$ cannot belong to any $\mathcal{Z}_{\Lambda'}$ for any maximal K–lattice $\Lambda' \neq \Lambda$ of the same dimension j; therefore $I_{\mathcal{T}_e} \notin \mathcal{Z}_j$, whence, by (45), there must exist an $i \in 0, ..., j-1$ such that $I_{\mathcal{T}_e} \in B_i$, which, together with (94), concludes the proof of the lemma. $\square$

4 The resonance trap argument and conclusion of the proof

We are now in position to conclude the proof of the theorem, proving (5) and (6).

In view of (46), the are are two alternatives[16]:

(a) either $I_0 \in B_0 \cap (B - n\rho)$;

(b) or $I_0 \in B_\Lambda \cap (B-(j+1)\rho)$ for some maximal K–lattice of dimension $j \in \{1, ..., n-1\}$.

In case (a), by Lemma 3.1–(i), by (37), the definition of T_0 and $T_{\exp}$ ((28), (29)) and (41), the theorem is proved.
In case (b), by Lemma 3.1–(ii), there are two alternatives:

(b1) either $\|I_t - I_0\| \le r_j \overset{(37)}{\le} \rho$, for $|t| \le T_{\exp} \le T_\Lambda$

(b2) or there exist a t_1 such that $\|I_t - I_0\| \le \rho$ for all $|t| \le |t_1|$ and $I_{t_1} \in B_i \cap (B - j\rho)$ for some $i \in \{0, ..., j-1\}$.

In case (b1), by (28), (29) and (41), and recalling that $\rho = R_0\varepsilon^b/n$, the theorem is proved. In case (b2) we iterate the above scheme. Hence, after $0 \le k \le n-1$ steps we see that the action–trajectory I_t ends up either in a "trapping resonant region" $B_\Lambda \subseteq B_i$ where it gets stuck for exponentially long times or it will end up in B_0 where also gets stuck for exponentially long time. Since in such k steps I_t moves at most by $k\rho$ we see that in the (possible) fast drift we have $\|I_t - I_0\| \le k\rho \le (n-1)\rho$ to which we have to add the displacement in the trapping region which is again at most ρ. Thus, for times $|t| \le T_{\exp}$ we have $\|I_t - I_0\| \le n\rho = R_0\varepsilon^b$ as claimed. $\square$

We remark that, in the case (b) above, I_t may visit several blocks in the time $T_{\exp}$; let us denote by j^* their minimal multiplicity, and $t^* < T_{\exp}$ be such that $I_{t^*} \in B_{\Lambda^*} \cap (B - (j^*+1)\rho)$ with $\dim\Lambda^* = j^*$. Then, we have $I_0 \in B_{\Lambda^*}$ (and, therefore, $I_t \in B_{\Lambda^*}$ for all $|t| \le T_{\exp}$). In fact, since the geometry of resonances of the Hamilton function $-H$ is identical to the geometry of resonances of H, if we consider the solution (I'_t, φ'_t) of the Hamilton equations of $-H$ with $I'_0 = I_{t^*}$, and apply Lemma (3.1), we obtain that $I_0 = I'_{-t*} \in B_\Lambda \cap B_{\Lambda^*}$.

[16] Recall that I_0 is the center of $B \subseteq U$, which is a sphere of radius $r = 2R_0\varepsilon^b = 2n\rho$.

A Angles between linear spaces

In this appendix $n \geq 2$, $u, v, w, z...$ denote vectors in $\mathbb{R}^n$ and $L_1, L_2, L', ...$ linear vector subspaces of $\mathbb{R}^n$ of dimension $m \in \{1, ..., n-1\}$; π_L denotes the orthogonal projection onto the linear space L and Arccos : $[-1, 1] \to [0, \pi]$ denotes the principal branch of the inverse real cosine.

Definition A.1 Let . The angle between u and v is defined as

$$u \angle v = \begin{cases} \text{Arccos}\, \dfrac{u \cdot v}{\|u\|\,\|v\|} \ , & \text{if } u, v \neq 0 \\ \frac{\pi}{2} \ . & \text{otherwise} \ . \end{cases}$$

Definition A.2 The angle between L_1 and L_2 is defined as

$$L_1 \angle L_2 := \max_{u \in L_1 \setminus \{0\}} u \angle \pi_{L_2} u \ .$$

We, next, list a few elementary properties of angles between linear spaces, whose simple proof is left to the reader (for the proof of items x and xi , see, also, [Nekhoroshev79, p. 45]).

i. $u \angle v \in [0, \pi]$ and $u \angle v \in [0, \pi/2]$ if and only if $u \cdot v \geq 0$; $L_1 \angle L_2 \in [0, \pi/2]$.

ii. $L_1 \angle L_2 = \frac{\pi}{2}$ if and only if[17] $L_1 \cap L_2^{\perp} \neq \{0\}$.
$L_1 \angle L_2 < \frac{\pi}{2}$ if and only if $\{u \in L_1 : \pi_{L_2} u = 0\} = \{0\}$.

iii. $u \angle \pi_L u = \text{Arccos}\, \dfrac{\|\pi_L u\|}{\|u\|}$, $\quad \forall\, u \neq 0$.

iv. $u \angle \pi_L u + u \angle \pi_{L^{\perp}} u = \frac{\pi}{2}$, $\quad \forall\, u \neq 0$.

v. $u \angle \pi_L u = \min_{v \in L \setminus \{0\}} u \angle v$.

vi. $\cos L_1 \angle L_2 = \displaystyle\min_{u \in L_1 \setminus \{0\}} \max_{v \in L_2 \setminus \{0\}} \frac{u \cdot v}{\|u\|\,\|v\|}$.

vii. For any u and v one has[18] $u \angle v = v \angle u$.

viii. If $u \neq 0 \neq v$, $u \angle v$ coincides with the (Euclidean) length of the shortest geodesic (equivalently, shortest curve) on the unit sphere $S^{n-1} := \{\xi \in \mathbb{R}^n : \|\xi\| = 1\}$ having as end–points the projections of u and v on S^{n-1}.

ix. $u \angle v \leq u \angle w + w \angle v$. Also: $L_1 \angle L_2 \leq L_1 \angle L_3 + L_3 \angle L_2$.

x. $L_1 \angle L_2 = L_2^{\perp} \angle L_1^{\perp}$.

xi. If $\dim L_1 = \dim L_2$, then $L_1 \angle L_2 = L_2 \angle L_1$.

B Parameter Relations

For completeness, in this appendix, we prove the elementary inequalities (30)÷(41). Recall the definitions of the parameters given in (8)÷(29).
First, we observe that from these definitions and the hypothesis $0 \leq \varepsilon \leq \varepsilon_0$, it follows easily:

$$E \geq 4 \ ; \quad A := 6E \geq 24 \ , \quad K := \Big(\frac{\varepsilon_*}{\varepsilon}\Big)^a \geq \Big(\frac{\varepsilon_*}{\varepsilon_0}\Big)^a \geq 1 \ ; \quad 1 \leq |\Lambda| \leq K^j \ ; \tag{B.1}$$

$$\delta_\Lambda \leq \lambda_j \ ; \quad \rho = \frac{r\mu_0}{K^{1/\alpha_{n-1}}} \ ; \quad r = 2n\rho \ ; \tag{B.2}$$

$$q_1 + 1 = \frac{1}{2a} \geq n \ ; \quad q_j \geq 2 \quad (j \leq n-2) \ ; \quad q_j\Big(1 - \frac{1}{\alpha_j}\Big) = a_j - 1 - j\Big(1 - \frac{1}{\alpha_j}\Big). \tag{B.3}$$

---

[17] $L^{\perp} := \{u \in \mathbb{R}^n : u \cdot v = 0 \ , \ \forall\, v \in L\}$.

[18] But, in general, $L_1 \angle L_2 \neq L_2 \angle L_1$: for example, if $n = 3$, $L_1 = \{(0, t, t) : t \in \mathbb{R}\}$ and $L_2 = \{x_3 = 0\}$, then $L_1 \angle L_2 = \pi/4$, while $L_2 \angle L_1 = \pi/2$.

(30): It follows immediately from (B.1).

(31): It follows from $\dfrac{\varepsilon_0}{\varepsilon_*} \le \left(\dfrac{s}{6}\right)^{\frac{1}{a}}$.

(32): To get the 1st inequality observe that $q_j \ge 1 \ge 1/\alpha_{n-1}$, $r\mu_0 \ge \underline{\omega}/(24\sqrt{2}\,M)$ so that

$$\begin{aligned} r_\Lambda &= \frac{\underline{\omega}}{2\sqrt{2}}\,\frac{1}{(AK)^{q_j}}\,\frac{1}{|\Lambda|}\,\frac{1}{M} \le \frac{\underline{\omega}}{2\sqrt{2}}\,\frac{1}{A}\,\frac{1}{K^{\frac{1}{\alpha_{n-1}}}}\,\frac{1}{M} \overset{(B.1)}{\le} \frac{\underline{\omega}}{48\sqrt{2}}\,\frac{1}{K^{\frac{1}{\alpha_{n-1}}}}\,\frac{1}{M} = \frac{1}{2}\,\frac{\underline{\omega}}{24\sqrt{2}M}\left(\frac{\varepsilon}{\varepsilon_*}\right)^b \\ &\le \frac{r\mu_0}{2}\left(\frac{\varepsilon}{\varepsilon_*}\right)^b = \frac{\rho}{2}\ . \end{aligned}$$

As for the 2nd inequality we have: $r_\Lambda \le r_\Lambda\,\dfrac{(EK)^{a_j}}{4K} = R_\Lambda$.

(33), *first inequality*: Using: $q_j \ge 2 \ge 1/\alpha_{n-1}$, (B.1) and $\mu_0 \ge 1/(6^2 2\sqrt{2}) > 2/(\sqrt{2}(24)^2)$, one finds

$$\delta_\Lambda = \frac{\underline{\omega}}{2\sqrt{2}}\,\frac{1}{(AK)^{q_j}}\,\frac{1}{|\Lambda|} \le \frac{\underline{\omega}}{2\sqrt{2}}\,\frac{1}{24^2}\,\frac{1}{K^{\frac{1}{\alpha_{n-1}}}} \le \frac{\underline{\omega}}{4}\,\frac{\mu_0}{K^{\frac{1}{\alpha_{n-1}}}} = \frac{\underline{\omega}\rho}{r\,4}\ .$$

(33), *second inequality*: $\delta_\Lambda \overset{(33)}{\le} \dfrac{\underline{\omega}}{r}\,\dfrac{\rho}{4} \overset{(32)}{\le} \dfrac{\underline{\omega}}{2r}\,(\rho - r_\Lambda)$.

(33), *third inequality*: $\delta_\Lambda = \dfrac{\underline{\omega}}{2\sqrt{2}}\,\dfrac{1}{(AK)^{q_j}} \le \widehat{\omega}$.

(34): Using: the definitions given, the inequality $|\Lambda| \le K^j$ and last equality in (B.3), one has:

$$\begin{aligned} \frac{KM\kappa_j\left(\frac{\delta_\Lambda}{C_j}\right)^{\frac{1}{\alpha_j}}}{\frac{1}{4}\gamma_\Lambda} &= \left(\frac{1}{C_j}\right)^{\frac{1}{\alpha_j}}\,(4M\kappa_j)\,K^{1-a_j+q_j\left(1-\frac{1}{\alpha_j}\right)}\,|\Lambda|^{1-\frac{1}{\alpha_j}}\,\left(\frac{2\sqrt{2}}{\underline{\omega}}A^{q_j}\right)^{1-\frac{1}{\alpha_j}}\,\frac{1}{E^{a_j}} \\ &\le \left(\frac{1}{C_j}\right)^{\frac{1}{\alpha_j}}\,(4M\kappa_j)\,K^{q_j\left(1-\frac{1}{\alpha_j}\right)+1-a_j+\left(1-\frac{1}{\alpha_j}\right)}\,\left(\frac{2\sqrt{2}}{\underline{\omega}}A^{q_j}\right)^{1-\frac{1}{\alpha_j}}\,\frac{1}{E^{a_j}} \\ &= \left(\frac{(4M\kappa_j)^{\alpha_j}\;6^{np_j-j)(\alpha_j-1)}}{C_j\,\left(\frac{\underline{\omega}}{2\sqrt{2}}\right)^{\alpha_j-1}}\;\frac{1}{E^{\alpha_j+j(\alpha_j-1)}}\right)^{\frac{1}{\alpha_j}} \le 1\ , \end{aligned}$$

where last inequality comes from the definition of E.

(35): Using: $q_j \ge 1$, $q_j - a_j \ge 0$, $E \ge 1$, $a/b = \alpha_{n-1} \ge 1$, one has:

$$\begin{aligned} \frac{R_\Lambda}{r} &= \frac{1}{8\sqrt{2}}\,\frac{1}{6^{q_j}}\,\frac{\underline{\omega}}{Mr}\,\frac{1}{K^{q_j-a_j+1}}\,\frac{1}{E^{q_j-a_j}}\,\frac{1}{|\Lambda|} \le \frac{1}{48\sqrt{2}}\,\frac{\underline{\omega}}{Mr}\,\frac{1}{K} \overset{(B.1)}{\le} \frac{1}{48\sqrt{2}}\,\frac{\underline{\omega}}{Mr}\left(\frac{\varepsilon_0}{\varepsilon_*}\right)^a \\ &\overset{(10)}{\le} \frac{1}{48\sqrt{2}}\,\frac{\underline{\omega}}{Mr}\Big(\min\big(1, \frac{6\sqrt{2}}{n}\,\frac{Mr}{\underline{\omega}}\big)\Big)^{\frac{a}{b}} \le \frac{1}{8n} < 1. \end{aligned}$$

(36), *first inequality:* Using: the definitions given, $q_1 + 1 = \frac{1}{2a} = np_1$, $\varepsilon K^{1/a} = \varepsilon_*$, the definition of ε_* and E, one has:

$$\frac{2^8 K}{\lambda_1 r_0}\,\varepsilon\,|f|_{r,s} = \frac{2^4}{6^{2np_1-3}}\,\frac{1}{E} \le \frac{2^4}{6^3}\,\frac{1}{4} = \frac{1}{54} < 1\ ,$$

where in the inequality we used $n \ge 3$, $p_1 \ge 1$ and $E \ge 4$.

(36), *second inequality:* By the first inequality in (36), we see that the second inequality is implied by

$$\frac{1}{2^8}\frac{\lambda_1 r_0}{K} \le \frac{\gamma_\Lambda r_\Lambda}{2^9 K} \ .$$

Now, using: the definitions given, $|\Lambda| \le K^j$, $q_1 - q_j = n(p_1 - p_j) + (j-1) \ge 0$ and the relation $a_j + 1 + 2(q_1 - q_j - j) = a_j - 1 + 2n(p_1 - p_j) \ge 0$, one has:

$$\frac{1}{2^8}\frac{\lambda_1 r_0}{K}\frac{2^9 K}{\gamma_\Lambda r_\Lambda} = \frac{1}{E^{a_j}}\frac{1}{A^{2(q_1-q_j)}}\frac{|\Lambda|^2}{K^{a_j+1+2(q_1-q_j)}} \le \frac{1}{E^{a_j}}\frac{1}{A^{2(q_1-q_j)}}\frac{1}{K^{a_j+1+2(q_1-q_j-j)}} \le \frac{1}{E} < 1.$$

(36), *third inequality:* It follows immediately from the 2nd inequality in (32) and from the second inequality in (36).

(37): $\dfrac{r_0}{\rho} = \Big(4\sqrt{2}\,\dfrac{Mr}{\underline{\omega}}\,\mu_0\,A^{q_1}\,K^{q_1-\frac{1}{\alpha_{n-1}}}\Big)^{-1} < 1$ since $q_1 \ge 2 > 1/\alpha_{n-1}$, $K \ge 1$, and $\mu_0 \ge \frac{Mr}{\underline{\omega}}\frac{6\sqrt{2}}{n}$.

$\dfrac{r_{n-1}}{\rho} = K^{-\frac{q_{n-1}-1}{\alpha_{n-1}}}\,\dfrac{\kappa_{n-1}}{r\mu_0}\Big(\dfrac{\underline{\omega}}{C_{n-1}}\,\dfrac{1}{2\sqrt{2}A^{q_{n-1}}}\Big)^{\frac{1}{\alpha_{n-1}}} \le \Big(\dfrac{6E}{A}\Big)^{\frac{1}{\alpha_{n-1}}} = 1$, where in the inequality we used $q_{n-1} \ge 2 > 1/\alpha_{n-1}$, $K \ge 1$, $\dfrac{1}{\mu_0} \le \dfrac{r}{\kappa_{n-1}}\Big(\dfrac{12\sqrt{2}EC_{n-1}}{\underline{\omega}}\Big)^{\frac{1}{\alpha_{n-1}}}$ and $q_{n-1} \ge 1$.

Now, let $1 \le j \le n-2$ and define $\alpha_j' := \alpha_j - 1$, $b_j := \frac{q_j}{\beta_j} - 1$ (recall (7)) and observe that

$$b_j \ge \frac{\alpha_j}{\beta_j} > 0\ , \qquad \frac{1}{\alpha_j} - 1 + \frac{q_j\rho_j'}{\alpha_j\beta_j} = \Big(1 - \frac{1}{\alpha_j}\Big)\, b_j\ . \tag{B.4}$$

Observe also that from the definitions of κ_j, E and μ_0 it follows that

$$\kappa_j \ge \frac{\underline{\omega}}{M}\Big(\frac{C_j}{\underline{\omega}}\Big)^{\frac{1}{\alpha_j}} = \frac{\underline{\omega}^{1-\frac{1}{\alpha_j}}\,C_j^{\frac{1}{\alpha_j}}}{M}\ , \qquad E \ge \left(\frac{(M\kappa_j)^{\alpha_j}}{C_j\underline{\omega}^{\alpha_j'}}\right)^{\frac{1}{\beta_j}}\ , \qquad \mu_0 \ge \frac{\underline{\omega}}{r\,24\sqrt{2}\,M}\ . \tag{B.5}$$

Then, noticing also that $\dfrac{24\sqrt{2}}{6^{\frac{q_j}{\alpha_j}}} \le \dfrac{24\sqrt{2}}{6^2} < 1$, we obtain

$$\begin{aligned}\frac{r_j}{\rho} &= K^{-\frac{q_j-1}{\alpha_{n-1}}}\,\frac{\kappa_j}{r\mu_0}\Big(\frac{\underline{\omega}}{C_j}\,\frac{1}{2\sqrt{2}(6E)^{q_j}}\Big)^{\frac{1}{\alpha_j}} \le \frac{M}{\underline{\omega}}\,\kappa_j\Big(\frac{\underline{\omega}}{C_j}\Big)^{\frac{1}{\alpha_j}}\left(\frac{C_j\underline{\omega}^{\alpha_j'}}{(M\kappa_j)^{\alpha_j}}\right)^{\frac{q_j}{\alpha_j\beta_j}}\\ &\overset{\text{(B.4)}}{=} \left(\frac{\underline{\omega}^{1-\frac{1}{\alpha_j}}}{M}\,\frac{C_j^{\frac{1}{\alpha_j}}}{\kappa_j}\right)^{b_j} \overset{\text{(B.5)}}{\le} 1\ .\end{aligned}$$

(38): Using $q_1 + 1 = \frac{1}{2a}$, $6E \ge 1$ and $\varepsilon \le \varepsilon_0 \le \varepsilon_* \cdot \Big(\min\big(1, \frac{6\sqrt{2}}{n}\frac{Mr}{\underline{\omega}}\big)\Big)^{\frac{1}{b}}$ one finds:

$$\frac{r_0}{r} = \frac{\underline{\omega}}{4\sqrt{2}(6E)^{np_1-1}Mr}\Big(\frac{\varepsilon}{\varepsilon_*}\Big)^{\frac{1}{2}} \le \frac{\underline{\omega}}{4\sqrt{2}Mr}\min\left(1, \frac{6\sqrt{2}}{n}\frac{Mr}{\underline{\omega}}\right) \le 1.$$

(39): Since $\rho = R/(2n)$, (39) follows at one from (37).

(40): Set $x = \dfrac{n}{6\sqrt{2}}\dfrac{\underline{\omega}}{Mr}$, $y = \dfrac{n}{18\sqrt{2}}$, $z = \dfrac{4n\kappa_{n-1}}{r}\Big(\dfrac{\underline{\omega}}{12\sqrt{2}EC_{n-1}}\Big)^{\frac{1}{\alpha_{n-1}}}$. Then, from the definitions given and the hypothesis $\varepsilon \le \varepsilon_0$ it follows:

$$\frac{2R}{r} = 4n\mu_0\Big(\frac{\varepsilon}{\varepsilon_*}\Big)^b \le 4n\mu_0\Big(\frac{\varepsilon_0}{\varepsilon_*}\Big)^b = \max(x,y,z)\min(x^{-1},y^{-1},z^{-1},1) \le 1.$$

(41)**:** From the definitions given (and since $\frac{e}{24} > \frac{1}{10}$) it follows:

$$T_{\rm exp}\frac{\sqrt{\varepsilon}}{Te^{\frac{Ks}{6}}} = 6A^{\frac{1}{a}}\sqrt{\frac{\varepsilon_*}{\varepsilon}}\min\left(\frac{1}{10}\frac{1}{(AK)^{q_1}}\frac{1}{K}\ ,\ \frac{e}{24}\min_{\substack{1\leq j\leq n-1\\ \Lambda:\dim\Lambda=j}}\left(\frac{1}{|\Lambda|}\frac{1}{(AK)^{q_j}}\right)\right)$$
$$\geq \frac{3}{5}A^{\frac{1}{a}}\sqrt{\frac{\varepsilon_*}{\varepsilon}}\min_{1\leq j\leq n-1}\frac{1}{A^{q_j}}\frac{1}{K^{q_j+1}} = \frac{3}{5}KA^{np_1+1} \geq \frac{3}{5}A^4 > 1\ .$$

References

[1] Benettin G., Fassò F., *Fast rotations of the rigid body: a study by Hamiltonian perturbation theory. Part I*, Nonlinearity **9**, 137-186 (1996).

[2] Benettin G., Fassò F., Guzzo M., *Fast rotations of the rigid body: a study by Hamiltonian perturbation theory. Part II*, Nonlinearity **10**, 1695–1717 (1997).

[3] Benettin G., Fassò F., Guzzo M., *Long term stability of proper rotations of the perturbed Euler rigid body*, Communications in Mathematical Physics **250**, 133–160 (2004).

[4] Benettin G., Fassò F., Guzzo M., *Nekhoroshev stability of L4 and L5 in the spatial restricted three body problem*, Regul. Chaotic Dyn. **3**, 56–72 (1998).

[5] Benettin G., Galgani L., Giorgilli A., *A proof of Nekhoroshev's theorem for the stability times in nearly integrable Hamiltonian systems*, Cel. Mech. **37**: 1–25 (1985).

[6] Benettin G., Gallavotti G., *Stability of motions near resonances in quasi–integrable Hamiltonian systems*, J. Stat. Phys. **44**: 293–338 (1985).

[7] Bounemoura A., Marco, J.P., *Improved exponential stability for near-integrable quasi-convex Hamiltonians.* Nonlinearity **24**, no. 1, 97–112 (2011).

[8] Celletti A., Ferrara L., *An application of Nekhoroshev theorem to the restricted three-body problem*, Celestial Mechanics and Dynamical Astronomy **64**, p. 261–272 (1996)

[9] Chierchia L., Pinzari G., *Planetary Birkhoff normal forms* , Journal of Modern Dynamics **5** (4), 623–664 (2011)

[10] Fassó, F., Guzzo, M. Benettin G., *Nekhoroshev-Stability of Elliptic Equilibria in Hamiltonian Systems*, Communications in Mathematical Physics **197**: 347–360 (1998).

[11] Gallavotti, G., *Quasi–Integrable Mechanical Systems*, in "Critical phenomena, Random Systems, Gauge Theories", K. Osterwalder and R. Stora editors, Les Houches, Session XLIII, 1984 (North–Holland, Amsterdam 1986).

[12] Guzzo M., Chierchia L. and Benettin G., *The steep Nekhoroshev's Theorem and optimal stability exponents (an announcement).* To appear in Rend. Mat. Acc. Lincei.

[13] Guzzo M., Lega E., Froeschlé C., *Diffusion and stability in perturbed non-convex integrable systems*, Nonlinearity, **19**, 1049–1067 (2006).

[14] Guzzo M., Lega E., Froeschlé C., *First numerical investigation of a conjecture by N.N. Nekhoroshev about stability in quasi-integrable systems*, Chaos, **21**, Issue 3, paper 033101, 1–13 (2011).

[15] Guzzo M., Fassò F. Benettin G., *On The Stability of Elliptic Equilibria*, Mathematical Physics Electronic Journal, **4**, paper 1, 1–16 (1998).

[16] Lhotka Ch., Efthymiopoulos C., Dvorak R., *Nekhoroshev stability at L_4 or L_5 in the elliptic–restricted three–body problem – application to Trojan asteroid.* Monthly Notices of the Royal Astronomical Society **384**, Issue 3, pp. 1165–1177 (2008).

[17] Lochak P., *Canonical perturbation theory via simultaneous approximations.* Russ. Math. Surv. **47**, 57–133 (1992).

[18] Lochak P., Marco J.-P. and Sauzin D., *On the Splitting of Invariant Manifolds in Multidimensional Near-Integrable Hamiltonian Systems,* Memoirs of the AMS, 775, 2003.

[19] Lochak P., Neishtadt, A., *Estimates in the theorem of N.N. Nekhoroshev for systems with quasi-convex Hamiltonian*, Chaos **2**, 495–499 (1992).

[20] Morbidelli, A. Guzzo, M., *The Nekhoroshev theorem and the asteroid belt dynamical system.* Celestial Mechanics and Dynamical Astronomy **65**: 107–136, (1997).

[21] Nekhoroshev, N. N., *Stable lower estimates for smooth mappings and for the gradients of smooth functions.* (Russian) Mat. Sb. (N.S.) **90** (132) (1973), 432–478, 480.

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

[23] Nekhoroshev, N.N., *Exponential estimates of the stability time of near–integrable Hamiltonian systems* **2**, Trudy Sem. Petrovs., **5**, 5–50, (1979).

[24] Niederman, L., *Stability over exponentially long times in the planetary problem.* Nonlinearity **9** no. 6, 1703–1751, (1996).

[25] Niederman, L., *Nonlinear stability around an elliptic equilibrium point in a Hamiltonian system*, Nonlinearity **11** 1465–1479 (1998).

[26] Niederman, L., *Exponential stability for small perturbations of steep integrable Hamiltonian systems.* Ergodic Theory Dynam. Systems **24** no. 2, 593–608, (2004).

[27] Niederman, L., *Prevalence of exponential stability among nearly integrable Hamiltonian systems.* Ergodic Theory Dynam. Systems **27** no. 3, 905–928, (2007).

[28] Niederman, L., *Corrigendum on the article: Exponential stability for small perturbations of steep integrable Hamiltonian systems (Erg. Th. Dyn. Sys. (2004), 24, p. 593-608)*, to appear in Ergodic Theory Dynam. Systems

[29] Pavlovic, R., Guzzo, M., *Fulfillment of the conditions for the application of the Nekhoroshev theorem to the Koronis and Veritas asteroid families.* Monthly Notices of the Royal Astronomical Society, **384**, 1575–1582 (2008).

[30] Pinzari, G., *Aspects of the Planetary Birkhoff Normal Form*, Regular and Chaotic Dynamics, **18**, 6, 860–906 (2013).

[31] Pöschel, J., *Nekhoroshev estimates for quasi–convex hamiltonian systems.* Math. Z. **213**, pag. 187 (1993).

[32] Pöschel, J., *On Nekhoroshev's Estimate at an Elliptic Equilibrium*, Int. Math. Res. Not. **4**, 203–215 (1999).

[33] Sansottera M., Locatelli U. Giorgilli A., *On the stability of the secular evolution of the planar Sun-Jupiter-Saturn-Uranus system*, Mathematics and Computers in Simulation, in press, (2011).

[34] Todorović N., Guzzo M., Lega E. and Froeschlé Cl., *A numerical study of the stabilization effect of steepness*, Celest Mech Dyn Astr, **110**, 389–398 (2011).

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