Version 3.1, July 25, 1992
Math. Z. **213** (1993) 187–216

Nekhoroshev Estimates for Quasi-convex Hamiltonian Systems

JÜRGEN PÖSCHEL

0 Introduction

This paper is concerned with the stability of motions in nearly integrable, time-independent hamiltonian systems of n degrees of freedom, with hamiltonian

$$H = h(I) + f_\epsilon(I,\theta)$$

in n-dimensional action-angle variables I,θ depending on a small parameter ϵ such that $f_\epsilon \sim \epsilon$. The equations of motions are

$$\dot I = -\partial_\theta H, \qquad \dot\theta = \partial_I H$$

in usual vector notation, where the dot denotes differentiation with respect to the time variable and ∂ indicates partial differentiation.

For $\epsilon = 0$ these equations reduce to the integrable equations

$$\dot I = 0, \qquad \dot\theta = \omega(I) \stackrel{\text{def}}{=} \partial_I h.$$

All motions are linear, and the actions in particular stay put for all times. For arbitrarily small $\epsilon \neq 0$, however, it is a well known fact that the dynamics may be extremely complicated, and that the actions may vary unboundedly over unbounded time intervals even in nondegenerate systems. For example, the two degree of freedom hamiltonian

$$H = \frac{1}{2}(I_1^2 - I_2^2) + \epsilon \sin(\theta_1 - \theta_2)$$

admits the special solution $I(t) = (-\epsilon t, \epsilon t)$, $\theta(t) = -\frac{1}{2}(\epsilon t^2, \epsilon t^2)$, for which we have

$$\|I(t) - I(0)\| = \|I(t)\| = \sqrt{2}\,\epsilon t$$

for all t [20].

It was a remarkable achievement by Nekhoroshev to observe that nonetheless, generically, the variation of the actions of *all orbits* remains *small* over a finite, but *exponentially long* time interval. More precisely, he proved that for sufficiently small ϵ one has

$$\|I(t) - I(0)\| \le R_* \epsilon^b \quad \text{for} \quad |t| \le T_* \exp\left(\epsilon^{-a}\right)$$

for *all* orbits, provided the hamiltonian H is *real analytic,* and the unperturbed hamiltonian h meets certain generic transversality conditions known as *steepness* [19,20,21,11].

These bounds are referred to as the *stability radius* and the *stability time* of the system, respectively. Their most important constituents are their *stability exponents* a and b. They depend on the number of degrees of freedom, n, and the so called steepness indices and tend to zero as n tends to infinity.

We will not enter into a discussion of steepness here, however. Rather, it is the purpose of this paper to give simple and explicit expressions for these stability bounds for the important special cases of convex and quasi-convex systems, which are the 'steepest' among all steep systems. That is, the unperturbed hamiltonian h is assumed to be convex on all of its domain, or on each of its level sets, respectively. The former arise typically in classical mechanics as the hamiltonians of nearly integrable conservative systems. The importance of the latter stems from the fact that time-dependent perturbations of convex systems turn into time-independent perturbations of quasi-convex systems after passing to an extended phase space in the usual fashion.

Such systems admit the simplest and best stability estimates. In particular, we will obtain

$$a = \frac{1}{2n}.$$

Besides that we have $b = a$ on all of phase space, and even $b = 1/2$ in an open subdomain of large measure, which conforms with KAM-theory. In addition, the exponent a *improves* in small neighbourhoods of resonances, up to the value

$$a = \frac{1}{2}$$

near periodic orbits.

In his original paper [20] Nekhoroshev was mainly concerned with the generic steep case, so his estimates are not particularly sharp. In the eighties, Benettin, Gallavotti, Galgani and Giorgilli investigated the convex case more carefully in order to obtain bounds relevant for stability problems in Celestial and Statistical Mechanics, for example. In a number of papers, for instance [2,4,8,9], they refined the original proof to finally obtain a stability exponent around

$$a = \frac{1}{n^2 + n}.$$

Their work also helped to make these results more widely known and accessible.

Recently, Lochak [12,13] presented a novel approach to the convex case enabling him to obtain any exponent a with

$$a < \frac{1}{2n + 1}.$$

His analysis centers around periodic orbits, which represent the worst type of resonances in the system. He observes that on account of convexity suitable neighbourhoods of suitable periodic orbits are stable for exponentially long times. Covering the whole phase space by such neighbourhoods with the help of a standard result from the theory of diophantine approximations he then obtains the general estimate. In his words, "the important property that emerges is that, the more resonant the initial condition, the more stable the corresponding trajectory will be" in convex systems [13].

In this paper, however, we follow the traditional lines in order to obtain the results indicated above. First, normal forms are constructed on various subdomains of phase space known as 'resonance blocks' which are described as neighbourhoods of 'admissible resonances'. In the second step, these normal forms lead to stability estimates, which turn out to be rather straightforward in the convex case. Thirdly, a geometric argument is given how to cover the whole phase space by such resonance blocks.

Of these three steps only the last, geometric one is essentially new. The idea is to describe resonance zones *not* through small divisor conditions, because their geometry is somewhat complicated. Instead, they are given as tubular neighbourhoods of exact resonances, for which the required small divisor estimates are subsequently verified.

This leads to a simpler and essentially optimal covering of action space by resonance blocks. In turn, this results in a substantial improvement of the stability exponent.

After a previous version of this paper was completed the author learned from Anatolij Neishtadt a simple argument which allowed to sharpen the Normal Form Lemma and the Resonant Stability Lemma below. The same remark enabled Lochak, in a forthcoming joint paper with Neishtadt [15], to obtain the stability exponent $a = 1/2n$ along his lines as well. Yet another application of the same idea may be found in [5].

Before proceeding to the details I like to take the opportunity to thank Sergej Kuksin for numerous discussions during his stay in Bonn, which kept me working on this subject. I am also indebted to Giancarlo Benettin, Amadeo Delshams, Antonio Giorgilli, Pierre Lochak and Anatolij Neishtadt (in alphabetical order) for their many remarks and comments, which helped to improve this paper to its present state. Finally, I thank the Forschungsinstitut für Mathematik at the ETH Zürich for its hospitality while this work was completed, and the Deutsche Forschungsgemeinschaft for their financial support through a Heisenberg grant.

1 Set Up and Main Result

Set Up. We consider a nearly integrable hamiltonian

$$H = h(I) + f_\epsilon(I, \theta),$$

that is real analytic in the action-angle variables

$$I \in P \subseteq \mathbb{R}^n, \qquad \theta \in \mathbb{T}^n = \mathbb{R}^n / 2\pi\mathbb{Z}^n$$

and depends on a small parameter ϵ in an arbitrary fashion. More specifically, the hamiltonian is assumed to be real analytic on a fixed complex neighbourhood of $P \times \mathbb{T}^n$ of the form

$$V_{r_0,s_0} P = V_{r_0} P \times W_{s_0} \mathbb{T}^n \subseteq \mathbb{C}^n \times \mathbb{C}^n,$$

where for an arbitrary subset D of $\mathbb{R}^n$, $V_r D = \{ I : \|I - D\| < r \} \subseteq \mathbb{C}^n$ denotes the complex neighbourhood of radius r around D with respect to the euclidean norm $\| \cdot \|$, and $W_{s_0} \mathbb{T}^n = \{ \theta : |\theta| = \max_i |\mathrm{Im}\, \theta_i| < s_0 \}$ denotes the complex strip of width s_0 around the torus $\mathbb{T}^n$. The set P may be any point set. Later on, we will also use the notation $U_r D = V_r D \cap \mathbb{R}^n$ to denote real neighbourhoods of a subset D of $\mathbb{R}^n$.

We assume $n \geq 2$ in order to exclude the trivial case of one degree of freedom systems, which are always integrable.

For the sake of convenience the perturbation parameter ϵ is assumed to be chosen in such a way that

$$|f_\epsilon|_{P,r_0,s_0} \leq \epsilon$$

for all sufficiently small $\epsilon \geq 0$ in the following exponentially weighted norm. If u is analytic on $V_{r,s}D$ with Fourier expansion $\sum_k u_k(I)\, e^{ik\cdot\theta}$, then

$$|u|_{D,r,s} = \sup_{I \in V_r D} \sum_{k \in \mathbb{Z}^n} |u_k(I)|\, e^{|k|s},$$

where $|k| = |k_1| + \cdots + |k_n|$. The basic properties of these norms are collected in Appendix B. In particular, they are related to the usual sup-norms by

$$|u|_{D,r,s} \leq |u|_{D,r,s} \leq \coth^n \sigma \;\; |u|_{D,r,s+2\sigma}$$

for all $\sigma > 0$.

The hessian of the integrable hamiltonian h,

$$Q(I) = \partial_I^2 h(I),$$

is assumed to be uniformly bounded with respect to the operator norm induced by the euclidean norm, also denoted by $\|\cdot\|$:

$$\sup_{I \in V_{r_0} P} \|Q(I)\| \leq M < \infty.$$

For the symmetric operator Q this norm amounts to the spectral radius norm. These assumptions will be in effect throughout the following discussion.

Main Result. To state the main result we need to quantify the notion of convexity. Let l and m be positive numbers. The integrable hamiltonian h is said to be m*-convex*, if the inequality

$$\langle Q(I)\xi, \xi \rangle \geq m\|\xi\|^2, \qquad \xi \in \mathbb{R}^n,$$

holds at every point I in $U_{r_0}P$. More generally, h is said to be l, m*-quasi-convex*, if at every point I in $U_{r_0}P$ at least one of the inequalities

$$|\langle \omega(I), \xi \rangle| > l\|\xi\|, \qquad \langle Q(I)\xi, \xi \rangle \geq m\|\xi\|^2$$

holds for each $\xi \in \mathbb{R}^n$. Thus, uniform m-convexity is required not only on the hyperplane orthogonal to $\omega(I)$ but also outside a cone around $\omega(I)$ whose angle decreases with $\|\omega(I)\|/l$. This amounts to strict convexity when $\|\omega(I)\| \le l$.

Theorem 1. *Suppose h is l, m-quasi-convex, and*

$$|f_\epsilon|_{P,r_0,s_0} \le \epsilon \le \epsilon_0 = \frac{mr_0^2}{2^{10}A^{2n}},$$

where $r_0 \le 4l/m$ and $A = 11M/m$. Then for every orbit with initial position (I_0, θ_0) in $P \times \mathbb{T}^n$ one has

$$\|I(t) - I_0\| \le R_0 \left(\frac{\epsilon}{\epsilon_0}\right)^a \quad \textit{for} \quad |t| \le T_0 \exp\left(\frac{s_0}{6}\left(\frac{\epsilon_0}{\epsilon}\right)^a\right)$$

except when $\|\omega(I_0)\| \le mr_0/8$ in which case $\|I(t) - I_0\| \le r_0$ for all t. The parameters are

$$a = \frac{1}{2n}, \qquad R_0 = \frac{r_0}{A}, \qquad T_0 = A^2 \frac{s_0}{\Omega_0},$$

where $\Omega_0 = \sup_{\|I-I_0\| \le R_0} \|\omega(I)\|$.

In this statement the analyticity radius r_0 plays the rôle of an external parameter, which may be adjusted in some suitable way. For example, replacing r_0 by the smaller quantity $r_0\,(\epsilon/\epsilon_0)^{(1-\mu)/2}$ with $0 \le \mu \le 1$ one immediately obtains the following more general version of Theorem 1.

Theorem 1*. *Under the same hypotheses as in Theorem 1 one has for every orbit*

$$\|I(t) - I_0\| \le R_0 \left(\frac{\epsilon}{\epsilon_0}\right)^{\mu a + (1-\mu)/2} \quad \textit{for} \quad |t| \le T_0 \exp\left(\frac{s_0}{6}\left(\frac{\epsilon_0}{\epsilon}\right)^{\mu a}\right)$$

for every μ in $0 \le \mu \le 1$, except when $\|\omega(I_0)\| \le mr_0/8 \cdot (\epsilon/\epsilon_0)^{(1-\mu)/2}$, in which case one has perpetual stability in a slightly larger ball.

Quasi-convex systems are iso-energetically nondegenerate, as one easily verifies. Hence, KAM-theory applies as well, at least for sufficiently small ϵ, providing a Cantor family of invariant tori, on which orbits are indeed perpetually stable. In two degree of freedom systems this results in perpetual stability of all other orbits as well, due to their well known confinement between invariant tori. So in this case the Nekhoroshev estimate is of minor interest. Instabilities such as Arnold diffusion may only occur when the number of degrees of freedom is at least three, and here the Nekhoroshev estimate provides an upper bound for the speed of such processes.

The rest of this paper is almost entirely devoted to the proof of Theorem 1. The following three sections describe its three basic steps: the construction of normal forms, the stability estimates based on convexity, and the geometry of resonances. The subsequent section finishes the proof. In addition, it contains further results on the exceptional stability of orbits very close to resonances and perturbations of linear integrable systems.

2 Normal Forms

At a given point in phase space the relevant part of a perturbation of an integrable system consists of those Fourier terms that correspond to resonances, or near resonances, among the frequencies of the unperturbed motion at that point. All other terms may be transformed away to obtain a *resonant normal form* that allows a more detailed analysis of the evolution of the system.

Let Λ be a sublattice of $\mathbb{Z}^n$. We are considering subsets D of P where the frequencies of the integrable system,

$$\omega(I) = \partial_I h(I),$$

satisfy nonresonance conditions up to a certain order except for those integer vectors lying in Λ. To make this quantitative, a subset $D \subseteq P$ is said to be *α, K-nonresonant modulo Λ*, if at every point I in D,

$$|k \cdot \omega(I)| \geq \alpha \qquad \text{for all } k \in \mathbb{Z}^n_K \backslash \Lambda\,,$$

where $\mathbb{Z}^n_K = \{ k : |k| \leq K \}$. For the time being, α and K will be regarded as independent parameters, but eventually they will be chosen as suitable functions of ϵ.

In the vicinity of such a set D the perturbed hamiltonian $h + f_\epsilon$ will be put into a *Λ-resonant normal form $h + g + f_*$ up to f_**. That means,

$$g = \sum_{k \in \Lambda} g_k(I)\, e^{ik \cdot \theta}$$

contains only Λ-resonant terms, while f_* is a general term. The aim is to make f_* exponentially small in K.

A special situation arises when Λ is the trivial sublattice Θ of $\mathbb{Z}^n$ containing only 0. In this case the set D is said to be *completely α, K-nonresonant*. In the corresponding normal form, g is independent of the angles θ and thus *integrable*. Naturally, the analysis of this case is simpler than in the presence of resonances. This will become apparent in the next section.

The Lemma. To state the Normal Form Lemma a few notations are required. We write T_K and P_Λ for the projections of Fourier series onto their partial sums containing only terms with $k \in \mathbb{Z}^n_K$ and $k \in \Lambda$, respectively. Moreover, we write p and q for two positive numbers related by $p^{-1} + q^{-1} = 1$. Later on, we let $q = 9$ and $p = 9/8$, but this choice is quite arbitrary.

From now on we write s for s_0 for brevity.

Normal Form Lemma. *Suppose that $D \subseteq P$ is α, K-nonresonant modulo Λ. If*

$$\epsilon \le \frac{1}{2^7 q} \frac{\alpha r}{K}, \qquad r \le \min\left(\frac{\alpha}{pMK}, r_0\right),$$

and $Ks \ge 6$, then there exists a real analytic, symplectic coordinate transformation $\Psi\colon V_{r_*,s_*}D \to V_{r,s}D$, where $r_* = r/2$ and $s_* = s/6$, such that $H \circ \Psi = h + g + f_*$ is in Λ-resonant normal form up to f_* with

$$|g - g_0|_{r_*,s_*} \le 16c\epsilon^2, \qquad |f_*|_{r_*,s_*} \le e^{-Ks/6}\epsilon,$$

where $g_0 = P_\Lambda T_K f_\epsilon$ and $c = 2qK/\alpha r$. Moreover,

$$\|\Pi_I \Psi - I\| \le c\epsilon r$$

uniformly on $V_{r_*,s_*}D$, where Π_I denotes the projection onto I-coordinates.

A more detailed estimate of Ψ is given at the end of this section. Incidentally, Λ need not be a sublattice of $\mathbb{Z}^n$ for the Normal Form Lemma to be true. Any subset of $\mathbb{Z}^n$ is allowed that is symmetric about the origin. This greater generality, however, will not be needed here.

Proof of the Lemma. We prove the Normal Form Lemma by a finite iterative procedure reminiscent of the KAM-scheme. At each step – described in the Iterative Lemma below – a coordinate change is performed that takes the hamiltonian closer to normal form. Choosing the number of steps and their sizes carefully their product produces a normal form with an exponentially small remainder.

A significant feature of the lemma is the dependence of its smallness condition on K alone, rather than its square. This is accomplished by making the first step in the iteration *large*, while all subsequent steps are uniformly *small*. Such a scheme first appeared in [18] in a related averaging problem, and the author is indebted to Anatolij Neishtadt for communicating this idea.

Alternatively, one could work with a suitable vectorfield norm rather than a hamiltonian one [5,6]. This approach avoids the inconvenience of different step sizes and is of particular use in other related problems, but it also sacrifices the compactness of the hamiltonian formulation. Anyhow, the one big step would now be required beforehand to pass from the hamiltonian to the vectorfield norm.

Yet another approach avoiding successive coordinate transformations is described for example in [3,8,10]. There, the Lie transform method is employed which works in a fixed coordinate system in a manner similar to the comparison of coefficients in power series expansions.

From now on we write $|\cdot|_{r,s}$ for $|\cdot|_{D,r,s}$ and $V_{r,s}$ for $V_{r,s}D$ for brevity, since D is fixed. We also write $|\cdot|_{\mathcal{P}}$, $|\cdot|_{\mathcal{P}}$ for the phase space norm defined by $|(I,\theta)|_{\mathcal{P}} = \max(\|I\|, |\theta|)$ and its induced operator norm, respectively.

Iterative Lemma. *Suppose $D \subseteq P$ is α, K-nonresonant modulo Λ, and the hamiltonian $H = h + g + f$ is real analytic on $V_{r,s}$. If*

$$|f|_{r,s} < \frac{\alpha\rho\sigma}{2q},$$

with $2\rho < r \le \alpha/pMK$ and $2\sigma < s$, then there exists a real analytic, symplectic transformation $\Phi\colon V_{r-2\rho,s-2\sigma} \to V_{r,s}$ so that $H \circ \Phi = h + g_+ + f_+$ with $g_+ - g = P_\Lambda T_K f$ and

$$|f_+|_{r-2\rho,s-2\sigma} \le \left(1 - \frac{|f|_{r,s}}{\alpha\rho\sigma/2q}\right)^{-1} \left(\frac{|g|_* + |f|_{r,s}}{\alpha\rho\sigma/q} + e^{-K\sigma}\right) |f|_{r,s},$$

where

$$|g|_* = \sum_j \frac{1}{2}\left(\frac{\rho}{r_j - r + \rho} + \frac{\sigma}{s_j - s + \sigma}\right) |g|_{r_j,s_j},$$

if $g = \sum_j g_j$ with terms g_j bounded on $V_{r_j,s_j} \supseteq V_{r,s}$. Moreover,

$$|W(\Phi - id)|_{\mathcal{P}}, \; \frac{1}{2}\left|W(D\Phi - I)W^{-1}\right|_{\mathcal{P}} \le \frac{q}{\alpha\rho\sigma}\,|f|_{r,s}$$

uniformly on $V_{r-2\rho,s-2\sigma}$, where $W = \operatorname{diag}\left(\rho^{-1}I_n, \sigma^{-1}I_n\right)$ and I_n denotes the n-dimensional unit matrix.

The proof of the Iterative Lemma is a purely technical matter and is relegated to Appendix A. Here we show how the Normal Form Lemma follows from it.

Proof of the Normal Form Lemma. For the very first step let $\rho_0 = r/8$, $\sigma_0 = s/6$ and $r_1 = r - 2\rho_0 = \frac{3}{4}r$, $s_1 = s - 2\sigma_0 = \frac{2}{3}s$. Since we assume $Ks \geq 6$, we have

$$\epsilon_0 \stackrel{\text{def}}{=} \epsilon \leq \frac{1}{2^7 q} \frac{\alpha r s}{Ks} \leq \frac{\alpha \rho_0 \sigma_0}{16q}.$$

So the Iterative Lemma applies. There exists a coordinate transformation $\Phi_0 \colon V_{r_1,s_1} \to V_{r,s}$ so that $H \circ \Phi_0 = h + g_0 + f_1$ with $g_0 = P_\Lambda T_K f_\epsilon$ and

$$|f_1|_{r_1,s_1} \leq \frac{8}{7} \left(\frac{q\epsilon_0}{\alpha \rho_0 \sigma_0} + e^{-K\sigma_0} \right) \epsilon_0.$$

In general, the first summand is dominant, and assuming that $e^{-K\sigma_0} \leq \dfrac{q\epsilon_0}{\alpha\rho_0\sigma_0}$ we obtain

$$\epsilon_1 \stackrel{\text{def}}{=} |f_1|_{r_1,s_1} \leq \frac{8}{3} \frac{q\epsilon_0^2}{\alpha\rho_0\sigma_0} = \frac{2^7 q \epsilon_0^2}{\alpha r s} \leq \frac{\epsilon_0}{4}.$$

Otherwise we perform the first step with *twice* the step size and already obtain the final result

$$|f_1|_{r_1,s_1} \leq \frac{8}{7} \left(\frac{q\epsilon_0}{4\alpha\rho_0\sigma_0} + e^{-2K\sigma_0} \right) \epsilon_0 \leq \frac{8}{7} \left(\frac{1}{4} + \frac{1}{e} \right) e^{-K\sigma_0} \epsilon_0 \leq e^{Ks/6} \epsilon.$$

Moreover, for Φ_0 we have

$$|W_0(\Phi_0 - id)|_{\mathcal{P}}, \ \frac{1}{2} \left| W_0 (D\Phi_0 - I) W_0^{-1} \right|_{\mathcal{P}} \leq \frac{q\epsilon_0}{\alpha\rho_0\sigma_0}$$

with $W_0 = \operatorname{diag}(\rho_0^{-1} I_n, \sigma_0^{-1} I_n)$ in both cases.

Next we apply the Iterative Lemma N times with fixed parameters

$$\rho = \frac{r}{8N}, \qquad \sigma = \frac{s}{4N},$$

where the integer N is chosen so that $e^{-K\sigma}$ is bounded by $1/8$ but now is as *large* as possible. Thus we choose N so that

$$N \leq \frac{Ks}{12 \ln 2} < N + 1.$$

In particular, $Ks \geq 8N$. We can assume that $N \geq 1$, since otherwise there is nothing more to do.

We will obtain a succession of transformations $\Phi_i\colon V_{i+1} \to V_i = V_{r_i,s_i}$ for $1 \le i \le N$, where $r_i = r_1 - 2(i-1)\rho$ and similarly s_i, so that the product $\Psi = \Phi_0\circ\Phi_1\circ\cdots\circ\Phi_N$ maps $V_{N+1} = V_{r/2,s/6}$ into $V_{r,s}$. Moreover, for the components of the successive hamiltonians $H_i = h + g_{i-1} + f_i$ on V_i with corresponding norms $|\cdot|_i$ we will find the estimates

$$\epsilon_i \overset{\text{def}}{=} |f_i|_i \le 4^{1-i}\epsilon_1, \qquad |g_i - g_{i-1}|_i \le \epsilon_i$$

for $1 \le i \le N+1$ and $1 \le i \le N$, respectively, as we now verify by induction.

Indeed, the smallness condition of the Iterative Lemma is satisfied by H_i at each step, since we have

$$\epsilon_i \le \epsilon_1 = \frac{2^7 q\epsilon_0^2}{\alpha r s} \le \frac{\alpha r s}{2^{13} q N^2} = \frac{\alpha\rho\sigma}{2^8 q},$$

which follows from

$$\epsilon_0 \le \frac{1}{2^7 q}\frac{\alpha r s}{Ks} \le \frac{1}{2^{10} q}\frac{\alpha r s}{N}.$$

To estimate f_{i+1} we write $g_{i-1} = (g_{i-1} - g_{i-2}) + \cdots + (g_1 - g_0) + g_0$ and get

$$\begin{aligned}
|g_{i-1}|_* &\le \frac{1}{2}(\epsilon_{i-1} + \cdots + \epsilon_1) + \frac{1}{4}\left(\frac{\rho}{\rho_0} + \frac{\sigma}{\sigma_0}\right)\epsilon_0 \\
&\le \epsilon_1 + \frac{\epsilon_0}{N} \\
&\le \frac{1}{2^8}\frac{\alpha\rho\sigma}{q} + \frac{1}{2^5}\frac{\alpha\rho\sigma}{q} \\
&\le \frac{1}{2^4}\frac{\alpha\rho\sigma}{q}.
\end{aligned}$$

Moreover, $|f_i|_i \le \dfrac{1}{2^8}\dfrac{\alpha\rho\sigma}{q}$ and $e^{-K\sigma} \le \dfrac{1}{8}$. Hence, we obtain

$$|f_{i+1}|_{i+1} \le \left(1 - \frac{1}{2^7}\right)^{-1}\left(\frac{1}{2^4} + \frac{1}{2^8} + \frac{1}{8}\right)|f_i|_i \le \frac{\epsilon_i}{4} \le 4^{-i}\epsilon_1$$

as we wanted to show. We also have $|g_i - g_{i-1}|_i = |P_\Lambda T_K f_i|_i \le |f_i|_i$ as required, and in addition

$$|W(\Phi_i - id)|_{\mathcal{P}},\ \frac{1}{2}\left|W(D\Phi_i - I)W^{-1}\right|_{\mathcal{P}} \le \frac{q\epsilon_i}{\alpha\rho\sigma}$$

uniformly on V_{i+1} with $W = \operatorname{diag}(\rho^{-1}I_n, \sigma^{-1}I_n)$.

Let $H = h + g + f_*$ be the final hamiltonian after N such iterations. Then we find

$$|f_*|_{r_*,s_*} = |f_{N+1}|_{N+1} \le 4^{-N-1}\epsilon_0 \le 4^{-Ks/12\ln 2}\epsilon_0 = e^{-Ks/6}\epsilon$$

and

$$|g - g_0|_{r_*,s_*} \le \frac{4}{3}\cdot\epsilon_1 \le \frac{2^8 q\epsilon_0^2}{\alpha r s} \le \frac{2^5 qK}{\alpha r}\cdot\epsilon^2.$$

Finally, the estimates for the Φ_i and the usual telescoping arguments yield

$$|W_0(\Psi - id)|_{\mathcal{P}}, \ \frac{1}{2}\left|W_0(D\Psi - I)W^{-1}\right|_{\mathcal{P}} \le \frac{2q\epsilon_0}{\alpha\rho_0\sigma_0} \le \frac{2^5 qK}{\alpha r}\cdot\epsilon$$

uniformly on V_{r_*,s_*}. This in particular contains the estimate

$$\|\Pi_I\Psi - I\| \le \frac{2q\epsilon_0}{\alpha\sigma_0} \le \frac{2qK}{\alpha}\cdot\epsilon.$$

This proves the Normal Form Lemma. ∎

3 Stability Estimates

Bounds on the variation of the action are obtained by analyzing the hamiltonian system in normal form. This analysis is simplest in the completely nonresonant regime. Here, no further assumptions are needed. In particular, no convexity hypothesis is needed.

Proposition 1 (Nonresonant Stability Estimate). *Suppose the domain $D \subseteq P$ is completely α, K-nonresonant. If*

$$\epsilon \le \frac{1}{2^7 q}\,\frac{\alpha r}{K}, \qquad r \le \min\left(\frac{\alpha}{pMK}, r_0\right),$$

then

$$\|I(t) - I_0\| \le r \quad \textit{for} \quad |t| \le \frac{sr}{5\epsilon}\,e^{Ks/6}$$

for every orbit of the perturbed system with initial position in $D \times \mathbb{T}^n$.

Proof. First we assume that $Ks \ge 6$. Let $\tilde r = r/2$, $\tilde s = s/6$. By the Normal Form Lemma there exists a real analytic symplectic transformation $\Psi\colon V_{\tilde r,\tilde s}D \to$

$V_{r,s}D$, $(\tilde{I},\tilde{\theta}) \mapsto (I,\theta)$, which takes the given hamiltonian H into $H\circ\Psi = h+g+f_*$ with $\partial_\theta g = 0$ and $|f_*|_{\tilde{r},\tilde{s}} \le \eta = \epsilon e^{-Ks/6}$. Moreover,

$$\|\tilde{I} - I\| \le \rho = \frac{r}{2^6} = \frac{\tilde{r}}{2^5}$$

uniformly on the smaller domain. Thus, Ψ^{-1} takes any given initial position in $D\times\mathbb{T}^n$ into some initial position in $U_\rho D \times \mathbb{T}^n$.

Fix such an initial position $(\tilde{I}_0,\tilde{\theta}_0)$. The real ball B_0 of radius $\tilde{r}-\rho$ around $\tilde{I}_0$ is contained in $U_{\tilde{r}}D$. As long as the forward evolution of the action $\tilde{I}$ stays inside B_0 we have $d\tilde{I}/dt = -\partial_\theta f_*$ and therefore

$$\|\tilde{I}(t) - \tilde{I}_0\| \le \int_0^t \|\partial_\theta f_*\|\, d\tau \le t \sup_{B_0\times\mathbb{T}^n} \|\partial_\theta f_*\| \le \frac{t\eta}{e\tilde{s}}$$

by the Cauchy estimate of Lemma B.3. It follows that

$$\|\tilde{I}(t) - \tilde{I}_0\| \le \tilde{r} - \rho \quad \text{for} \quad 0 \le t \le \frac{sr}{5\epsilon}\, e^{Ks/6} \le \frac{e\tilde{s}(\tilde{r}-\rho)}{\eta}.$$

Going back to original coordinates we obtain

$$\begin{aligned} \|I(t) - I(0)\| &\le \|I(t) - \tilde{I}(t)\| + \|\tilde{I}(t) - \tilde{I}(0)\| + \|\tilde{I}(0) - I(0)\| \\ &\le \rho + (\tilde{r} - \rho) + \rho \le r \end{aligned}$$

for the same time interval. The estimate for backward orbits is, of course, the same.

For $Ks \le 6$ the same reasoning applies to the hamiltonian in original coordinates, letting $g = 0$ and $f_* = f_\epsilon$. ∎

Next we look at the hamiltonian in the vicinity of resonances. Here, we follow an idea of Gallavotti [7], also used by Lochak [12,13], that due to the convexity of the unperturbed hamiltonian an orbit stays in a *fixed* coordinate neighbourhood of a resonance for an exponentially long time interval. Indeed, on account of the Normal Form Lemma, near a Λ-resonant point I_* the action essentially evolves in the plane Π through I_* spanned by the vectors in Λ. The restriction of h to Π is a convex function with a critical point at I_*, which in addition is nearly invariant because of energy conservation. Thus, the hamiltonian h is almost a Liapunov function for a certain amount of time.

To make this quantitative we introduce the following notion. Let Λ be a nontrivial sublattice of $\mathbb{Z}^n$ characterizing a resonance, and

$$R_\Lambda = \left\{\, \omega \in \mathbb{R}^n : k\cdot\omega = 0 \text{ for all } k \in \Lambda \,\right\}$$

the linear space of *exact* Λ*-resonances*. A domain $D \subseteq P$ is said to be δ*-close to exact* Λ*-resonances,* if

$$\|\omega(I) - R_\Lambda\| = \min_{\upsilon \in R_\Lambda} \|\omega(I) - \upsilon\| \le \delta$$

for all I in D.

We first consider the hamiltonian in normal form coordinates.

Lemma 1. *Suppose $D \subseteq P$ is δ-close to exact Λ-resonances, and the hamiltonian $H = h + g + f_*$ is in Λ-resonant normal form up to f_* on $V_{r,s}D$ with $|g|_{r,s} + |f_*|_{r,s} \le \zeta$. If h is m-convex and*

$$\delta \le \frac{mr}{4}, \qquad \zeta \le \frac{mr^2}{32},$$

then for every orbit with initial action I_0 in $U_\rho D$, $\rho = \delta/4M$, one has

$$\|I(t) - I_0\| \le r - \rho \quad \text{for} \quad |t| \le T_* = \frac{s}{12\|\omega(I_0)\|}\,\frac{mr^2}{\eta},$$

where $\eta = |f_*|_{r,s}$.

Note that the lemma also applies to $\Lambda = \mathbb{Z}^n$. In this case there is no proper normal form to speak of. The assumptions, however, imply that h is convex in D and close to its absolute minimum, which assures stability.

The lemma also makes sense when f_* vanishes yielding eternal stability. Thus, very near resonances in a *convex* integrable system, instability is not caused by the resonant terms of a perturbation but rather by the nonresonant ones which can not be transformed away.

Proof of the Lemma. Fix an initial position (I_0, θ_0) in $U_\rho D \times \mathbb{T}^n$ with $\rho = \delta/4M$ and consider its forward orbit. The real ball B_0 of radius $r - \rho$ around I_0 is contained in $U_r D$, and there is a positive, possibly infinite time T_e of first exit of $I(t)$ from B_0. Let $T = \min(T_e, T_*)$.

Following Lochak [12] we consider $\Delta h = h(I(t))\big|_0^T$ and $\Delta I = I(t)\big|_0^T$. Expanding around I_0 in the ball B_0 we get

$$\Delta h = \langle \omega_0, \Delta I \rangle + \int_0^1 (1 - s)\, \langle Q(I(s))\Delta I, \Delta I \rangle \, ds,$$

where $\omega_0 = \omega(I_0)$, and $I(s) = I_0 + s\Delta I$ describes a straight line in B_0. Hence,

$$\frac{m}{2}\|\Delta I\|^2 \le |\Delta h| + |\langle \omega_0, \Delta I\rangle|$$

by the convexity of Q.

Let P denote the orthogonal projection onto the real space spanned by Λ, and let $Q = I - P$. By the choice of ρ and the mean value theorem, the domain $U_\rho D$ is $\frac{5}{4}\delta$-close to exact Λ-resonances. Therefore,

$$\begin{aligned} |\langle \omega_0, P\Delta I\rangle| &\le \|P\omega_0\|\,\|P\Delta I\| \\ &\le \frac{5}{4}\delta\cdot\|\Delta I\| \le \frac{5}{2m}\delta^2 + \frac{m}{6}\|\Delta I\|^2. \end{aligned}$$

The vector $Q\Delta I$ is given by the integral along the orbit of $\partial_\theta f_*$ only, since g is in Λ-resonant normal form. Thus,

$$\begin{aligned} |\langle \omega_0, Q\Delta I\rangle| &\le \int_0^T |\langle \omega_0, \partial_\theta f_*\rangle|\, d\tau \\ &\le T\|\omega_0\| \sup_{B_0\times\mathbb{T}^n} \|\partial_\theta f_*\| \le T\|\omega_0\|\cdot\frac{1}{es}| f_*|_{r,s} \end{aligned}$$

by the Cauchy estimate of Lemma B.3. Finally, $\Delta H = H(I(t),\theta(t))\big|_0^T$ vanishes by energy conservation and therefore $|\Delta h| \le |\Delta g| + |\Delta f_*| \le 2\zeta$.

Collecting terms and using the notation of the lemma and its hypotheses on δ, ζ and T we thus arrive at

$$\begin{aligned} \frac{m}{3}\|\Delta I\|^2 &\le 2\zeta + \frac{5}{2m}\delta^2 + T\|\omega_0\|\cdot\frac{1}{es}| f_*|_{r,s} \\ &\le \frac{mr^2}{16} + \frac{5mr^2}{32} + \frac{mr^2}{32} \\ &= \frac{mr^2}{4} < \frac{m}{3}(r-\rho)^2, \end{aligned}$$

since $\rho = \delta/4M \le r/16$ by assumption and thus $r \le 16/15\,(r-\rho)$. It follows that $\|\Delta I\| < r - \rho$ and consequently that $T = T_*$. This proves the lemma for forward orbits, but backward orbits are, of course, the same. ∎

As a matter of fact, at any point I the action essentially evolves in the corresponding level set of h and thus in the hyperplane orthogonal to $\omega(I)$. It is only on these hyperplanes where the convexity of h is really needed to make the preceding argument work. This leads to the weaker notion of *quasi-convexity*, which was made quantitative in Section 1.

Lemma 2. *The previous lemma remains valid for an l, m-quasi-convex hamiltonian h, if $r \le 2l/m$, and if $\|\omega(I_0)\|$ is replaced by*

$$\Omega = \sup_{\|I-I_0\| \le r-\rho} \|\omega(I)\|.$$

Proof. We continue with the same set up and the same notation as before, but this time assume that $T_e < T_*$ to reach a contradiction. Thus we assume that $\|\Delta I\| = r-\rho$ at a first time $T = T_e < T_*$. We show that this implies $|\langle \omega(I(s)), \Delta I \rangle| \le l\|\Delta I\|$ for all $0 \le s \le 1$. Then we can again apply the convexity argument, arriving at $\|\Delta I\| < r - \rho$, which yields the desired contradiction.

Indeed, proceeding as before we have

$$\begin{aligned} |\langle \omega(I(s)), P\Delta I \rangle| &\le \frac{5}{4}\delta\|\Delta I\|, \\ |\langle \omega(I(s)), Q\Delta I \rangle| &\le T\Omega \cdot \frac{1}{es}|f_*|_{r,s}, \end{aligned}$$

with $\Omega = \sup_{I\in B_0} \|\omega(I)\|$. In view of $\delta \le mr/4$, the hypotheses on T_* and $r \le 2(r-\rho)$ it follows that

$$|\langle \omega(I(s)), \Delta I \rangle| \le \frac{mr}{3}\|\Delta I\| + \frac{mr}{16}(r-\rho) \le \frac{mr}{2}\|\Delta I\| \le l\|\Delta I\|$$

as we wanted to show. ∎

We now transfer the contents of the last two lemmata to our hamiltonian in original coordinates.

Proposition 2 (Resonant Stability Lemma). *Suppose the domain $D \subseteq P$ is δ-close to exact Λ-resonances, but α, K-nonresonant modulo Λ. If h is l, m-quasi-convex and*

$$\epsilon \le \frac{p}{q}\frac{mr^2}{2^7}, \qquad \delta \le \frac{mr}{8}, \qquad r \le \min\left(\frac{\alpha}{pMK}, \frac{4l}{m}, r_0\right)$$

with $q \ge 3$, then for every orbit with initial position in $D \times \mathbb{T}^n$ the estimate

$$\|I(t) - I_0\| \le r \quad \textit{for} \quad |t| \le \frac{s}{288\Omega}\frac{mr^2}{\epsilon}e^{Ks/6}$$

holds, where $\varOmega = \sup_{\|I-I_0\|\le r} \|\omega(I)\|$. *In particular, in the fully resonant case* $\Lambda = \mathbb{Z}^n$, *this estimate holds for all* t.

Proof. We take the freedom to increase δ so that

$$\delta = \frac{mr}{8}.$$

The hypotheses imply that ϵ satisfies the assumptions of the Normal Form Lemma. Letting $\tilde{r} = r/2$, $\tilde{s} = s/6$, and assuming first that $Ks \ge 6$, there exists a transformation $\Psi\colon V_{\tilde{r},\tilde{s}}D \to V_{r,s}D$, $(\tilde{I}, \tilde{\theta}) \mapsto (I, \theta)$, that takes the original hamiltonian H into the Λ-resonant normal form $H \circ \Psi = h + g + f_*$ up to f_* with

$$|g|_{\tilde{r},\tilde{s}} + |f_*|_{\tilde{r},\tilde{s}} \le \zeta = 2\epsilon, \qquad |f_*|_{\tilde{r},\tilde{s}} \le \eta = e^{-Ks/6}\epsilon.$$

Moreover, with $\alpha \ge pMKr$,

$$\|\tilde{I} - I\| \le \frac{2qK\epsilon}{\alpha} \le \frac{2q\epsilon}{pMr} \le \frac{mr}{2^6 M} = \frac{\delta}{8M}$$

uniformly on the smaller domain by our choice of δ. Hence, Ψ^{-1} takes any initial position in $D \times \mathbb{T}^n$ in the original coordinates to some initial position in $U_\rho D \times \mathbb{T}^n$, $\rho = \delta/4M$, in normal form coordinates.

The two previous lemmata now apply, since also

$$\zeta \le \frac{mr^2}{2^7} = \frac{m\tilde{r}^2}{32}, \qquad \delta = \frac{mr}{8} = \frac{m\tilde{r}}{4},$$

and $\tilde{r} \le 2l/m$. Hence, for every orbit with initial position in $U_\rho D \times \mathbb{T}^n$ we have $\|\tilde{I}(t) - \tilde{I}_0\| \le \tilde{r} - \rho$ for

$$|t| \le \frac{\tilde{s}}{12\tilde{\varOmega}}\,\frac{m\tilde{r}^2}{\eta} = \frac{s}{288\tilde{\varOmega}}\,\frac{mr^2}{\epsilon}\,e^{Ks/6}, \qquad \tilde{\varOmega} = \sup_{\|\tilde{I}-\tilde{I}_0\|\le\tilde{r}-\rho} \|\omega(\tilde{I})\|.$$

This yields the bound $\|I(t) - I_0\| \le r$ as in the proof of Proposition 1 for the stated time interval, since $\tilde{\varOmega} \le \varOmega$.

If, on the other hand, $Ks < 6$, then we simply apply those two lemmata to the given hamiltonian H with $g = 0$ and $f_* = f_\epsilon$. And if $\Lambda = \mathbb{Z}^n$, then we may set $g = f_\epsilon$ and $f_\epsilon = 0$ to obtain the last claim. ∎

Incidentally, the last lemma and its proof also hold in the nonresonant case, with the assumption of δ-closeness simply dropped. But Proposition 1 clearly provides a better estimate in this situation.

4 Geometry of Resonances

In order to apply our stability estimates to *all* orbits with initial positions in $P \times \mathbb{T}^n$ we need to cover all resonances in P by properly chosen neighbourhoods of 'nonresonant resonances'. To this end, an analogous covering is first constructed in frequency space and then pulled back to P via the frequency map.

In principle such a covering is easily constructed. From a given neighbourhood of all resonances of some order d, $1 \le d < n$, a sufficiently *large* neighbourhood of all resonances of order $d+1$ is removed, so that blocks of 'nonresonant resonances of order d' remain. It is the quantitative aspect which is somewhat more subtle.

First we observe that it suffices to consider resonances characterized by *maximal* K*-lattices* Λ in $\mathbb{Z}^n$, the family of which is denoted by $\mathbb{M}_K$. These are lattices that are generated by vectors in $\mathbb{Z}^n_K = \{ k : |k| \le K \}$ and are not properly contained in any other lattice of the *same* dimension. For, if Λ were not a K-lattice, then its effective dimension relevant for the nonresonance conditions is smaller than its true dimension, since integer vectors of norm greater than K are never considered. And if Λ were not maximal, then we could enlarge it to a maximal lattice of the same dimension without affecting the associated resonance space R_Λ nor the nonresonance conditions modulo Λ.

In the following construction a key rôle is played by the *volume* of a nontrivial integer lattice Λ. This quantity, denoted by $|\Lambda|$, is defined as the volume of the parallelepiped, or fundamental domain, spanned by the vectors of any choice of basis for Λ.

This definition makes sense. If A is an $n \times d$-matrix, then its d columns span a parallelepiped in n-space whose d-dimensional volume is $\sqrt{\det A^t A}$. If A and $\tilde{A}$ are two $n \times d$-matrices whose column vectors form two bases for a given d-dimensional lattice Λ, then there is a unimodular $d \times d$-matrix U such that $\tilde{A} = AU$. Thus,

$$\det \tilde{A}^t \tilde{A} = \det U^t A^t A U = \det U \det A^t A \det U = \det A^t A.$$

Hence, there is no ambiguity in defining $|\Lambda| = \sqrt{\det A^t A}$ as *the* volume of a nontrivial lattice Λ, where A is any $n \times d$-matrix whose columns form a basis for Λ. – For the trivial lattice the volume is set equal to 1 for consistency.

Now, fix n positive parameters $\lambda_1 < \lambda_2 < \cdots < \lambda_n$. With each nontrivial maximal K-lattice Λ of dimension d we associate its *resonance zone*

$$Z_\Lambda = \{ \omega : \|\omega - R_\Lambda\| < \delta_\Lambda \}, \qquad \delta_\Lambda = \frac{\lambda_d}{|\Lambda|},$$

which is the open neighbourhood of radius δ_Λ around the $n-d$-dimensional space of exact Λ-resonances $R_\Lambda = \{\omega : k\cdot\omega = 0 \text{ for all } k\in\Lambda\}$. Removing a neighbourhood of all next order resonances we obtain the *resonance blocks*

$$B_\Lambda = Z_\Lambda \backslash Z^*_{d+1},$$

where $Z^*_d = \bigcup_{\dim\Lambda=d} Z_\Lambda$ for $1\le d\le n$ and $Z^*_{n+1}=\varnothing$. We also set

$$Z_\Theta = \mathbb{R}^n, \qquad B_\Theta = Z_\Theta\backslash Z^*_1,$$

where Θ stands for the trivial lattice containing only 0. The block B_Θ comprises the completely nonresonant regime.

Clearly, we have

$$\begin{aligned}\mathbb{R}^n = Z_\Theta &= B_\Theta \cup Z^*_1\\ &= B_\Theta\cup B^*_1\cup Z^*_2\\ &\;\vdots\\ &= B_\Theta\cup B^*_1\cup\cdots\cup B^*_n,\end{aligned}$$

where $B^*_d = \bigcup_{\dim\Lambda=d} B_\Lambda$. So we obtain a covering of frequency space by resonance blocks B_Λ with $\Lambda\in\mathbb{M}_K$.

Geometric Lemma. *Let $E>0$ and $A\ge E+\sqrt{2}$. If*

$$\frac{\lambda_{d+1}}{\lambda_d}\ge AK, \qquad 1\le d<n,$$

then each block B_Λ for a nontrivial Λ is α_Λ, K-nonresonant modulo Λ with

$$\alpha_\Lambda \ge EK\delta_\Lambda,$$

while B_Θ is completely α_Θ, K-nonresonant with $\alpha_\Theta=\lambda_1$.

Proof. In the following it suffices to assume $\|k\|\le K$ rather than the stronger inequality $|k|\le K$.

First, let $\omega\in B_\Theta$. Fix k with $0<\|k\|\le K$, and let Λ be the unique one-dimensional maximal lattice containing k. Then $|\Lambda|\le\|k\|$, hence $\delta_\Lambda\ge\lambda_1/\|k\|$, and so

$$|k\cdot\omega| = \|k\|\cdot\|\omega - R_\Lambda\| \ge \|k\|\,\delta_\Lambda \ge \lambda_1,$$

since $\omega\notin Z_\Lambda$. This proves the lemma for B_Θ.

Now, let Λ a maximal K-lattice of dimension $d \geq 1$ and let $\omega \in B_\Lambda$. We may assume that $d < n$, since there is nothing to prove for $\Lambda = \mathbb{Z}^n$. Fix k not in Λ with $\|k\| \leq K$. Since Λ is maximal, the integer vector k is linearly independent of Λ, and the maximal K-lattice Λ_+ generated by Λ and k together has dimension $d+1 \leq n$.

Let P and P_+ denote the orthogonal projections onto the real spaces spanned by Λ and Λ_+, respectively, and let $Q = I - P$. We then have

$$\begin{aligned} k \cdot \omega &= k \cdot P_+\omega \\ &= Qk \cdot QP_+\omega + Pk \cdot PP_+\omega \\ &= Qk \cdot (I - P)P_+\omega + Pk \cdot P\omega \\ &= Qk \cdot (P_+\omega - P\omega) + Pk \cdot P\omega, \end{aligned}$$

since $PP_+ = P$.

Let V and V_+ denote the volumes of Λ and Λ_+, respectively, and let λ, λ_+ be short for λ_d, λ_{d+1}. Then

$$\|Qk\| \geq \frac{V_+}{V},$$

because the volume of Λ_+ is not bigger than the volume of Λ times the length of the component of k orthogonal to Λ, which is Qk. Moreover,

$$\|P\omega\| \leq \frac{\lambda}{V} \leq \frac{K\lambda}{V_+}, \qquad \|P_+\omega\| \geq \frac{\lambda_+}{V_+} \geq \frac{AK\lambda}{V_+},$$

since $\omega \in Z_\Lambda$, but $\omega \notin Z_{\Lambda_+}$. Thus,

$$\|P_+\omega - P\omega\| = \sqrt{\|P_+\omega\|^2 - \|P\omega\|^2} \geq \frac{K\lambda}{V_+}\sqrt{A^2 - 1}$$

by Pythagoras. Finally, the vectors Qk and $P_+\omega - P\omega$ are *collinear*, because they both lie on the line orthogonal to Λ within the real space spanned by Λ_+. Hence the absolute value of their scalar product equals the product of their lengths, and we obtain

$$\begin{aligned} |k \cdot \omega| &\geq \|Qk\| \cdot \|P_+\omega - P\omega\| - \|Pk\| \cdot \|P\omega\| \\ &\geq \frac{K\lambda}{V}\sqrt{A^2 - 1} - \frac{K\lambda}{V} \geq EK\delta_\Lambda, \end{aligned}$$

since $\sqrt{A^2 - 1} - 1 \geq A - \sqrt{2} \geq E$. ∎

We now pull back this covering of frequency space to a covering of action space. The following formulation of the result includes the parameters m and M only for later convenience. Here they are not needed.

Covering Lemma. *Given positive parameters* p, b, r_0 *and* $K \geq 1$*, fix a constant* $A \geq \dfrac{pM}{bm} + \sqrt{2}$ *and let*

$$r_\Lambda = \frac{r_0}{A^n K^n} \, \frac{A^d K^d}{|\Lambda|}, \qquad d = \dim \Lambda,$$

and

$$\alpha_\Lambda = pMKr_\Lambda, \qquad \delta_\Lambda = bmr_\Lambda.$$

Then there exists a covering of P *by resonance blocks* D_Λ, $\Lambda \in \mathbb{M}_K$, *such that each block* D_Λ *is* α_Λ, K*-nonresonant modulo* Λ, *but* δ_Λ*-close to exact* Λ*-resonances for nontrivial* Λ.

Proof. Define the blocks B_Λ with respect to the parameters

$$\lambda_d = \frac{bmr_0}{A^n K^n} \, A^d K^d, \qquad 1 \leq d \leq n.$$

The Geometric Lemma applies with $E = pM/bm$. For nontrivial Λ we have $\delta_\Lambda = \lambda_d / |\Lambda| = bmr_\Lambda$, and each block B_Λ is α_λ, K-nonresonant modulo Λ with $\alpha_\Lambda = EK\delta_\Lambda = pMKr_\Lambda$. This also applies to the block B_Θ, as one readily verifies. Setting

$$D_\Lambda = \{ I \in P : \omega(I) \in B_\Lambda \}$$

and discarding all sets which turn out to be empty we obtain the postulated covering of P. ∎

With this choice of parameters the covering of frequency space is optimal in the following sense. On one hand, in the completely nonresonant block B_Θ we have

$$|k \cdot \omega| \geq \alpha_\Theta \sim \frac{1}{K^{n-1}}, \qquad 0 \neq |k| \leq K,$$

for all K, which can not be improved on account of Dirichlet's theorem on diophantine approximations. On the other hand, for any bounded region G the Lebesgue measure of $Z_1^* \cap G$ is bounded by

$$\sum_{\dim \Lambda = 1} \delta_\Lambda \sim \frac{mr_0}{K^n} \sum_{1 \leq \|k\| \leq K} \frac{K}{\|k\|} \sim mr_0.$$

So this sum is of the order of 1, which could not be any better.

5 Proof of Theorem 1 and Further Results

Proof of Theorem 1. We now fix the constants $q = 9$, $p = 9/8$ and, in the Covering Lemma, $b = 1/8$ and

$$A = 11\frac{M}{m} \geq \frac{pM}{bm} + \sqrt{2}.$$

Then, for every $r_0 > 0$ and every $K \geq 1$, there exists a covering of P by suitable resonance blocks D_Λ, $\Lambda \in \mathbb{M}_K$, with parameters $\alpha_\Lambda = pMKr_\Lambda$ and $\delta_\Lambda = mr_\Lambda/8$, where

$$r_\Lambda = \frac{r_0}{A^n K^n} \frac{A^d K^d}{|\Lambda|}, \qquad d = \dim \Lambda.$$

In particular, $r_0/A^n K^n = r_\Theta \leq r_\Lambda \leq r_{\mathbb{Z}^n} = r_0$ for all those Λ.

The stability estimates now apply to all blocks simultaneously, if we have $r_0 \leq 4l/m$ and

$$\epsilon \leq \frac{mr_\Theta^2}{2^{10}} = \frac{\epsilon_0}{K^{2n}}, \qquad \epsilon_0 = \frac{mr_0^2}{2^{10}A^{2n}}.$$

Conversely, for $\epsilon \leq \epsilon_0$ we may choose $K \geq 1$ so that equality is achieved and the stability estimates apply to every block with

$$K = \left(\frac{\epsilon_0}{\epsilon}\right)^a, \qquad a = \frac{1}{2n}.$$

For the stability radius we find the upper bounds

$$r_\Lambda \leq \begin{cases} r_0 & \text{for } \Lambda = \mathbb{Z}^n, \\ r_0/AK & \text{otherwise.} \end{cases}$$

The first alternative applies if I_0 falls into the block $D_{\mathbb{Z}^n}$ which is tantamount to $\|\omega(I_0)\| \leq \delta_{\mathbb{Z}^n} = mr_0/8$. Otherwise the second alternative applies, which yields the postulated stability radius R_0/K.

To bound the stability time from below in the resonant regime, we observe that for $d \geq 1$, we have $mr_\Lambda^2/\epsilon \geq 2^{10}A^2$ and thus

$$\frac{s}{288\Omega} \frac{mr_\Lambda^2}{\epsilon} \geq A^2 \frac{s}{\Omega_0}, \qquad \Omega_0 = \sup_{\|I - I_0\| \leq R_0/K} \|\omega(I)\|$$

This yields the postulated stability time. Moreover, this time is infinite in the block $D_{\mathbb{Z}^n}$, that is, for $\|\omega(I_0)\| \leq mr_0/8$.

In the nonresonant regime D_Θ we have

$$\frac{r_\Theta}{\epsilon} = \frac{r_0}{A^n K^n}\,\frac{K^{2n}}{\epsilon_0} = \frac{2^{10} A^n K^n}{m r_0}.$$

Since we can assume that $\Omega_0 \geq \|\omega(I_0)\| \geq m r_0/8$ and $n \geq 2$, we thus have

$$\frac{s r_\Theta}{5\epsilon} \geq \frac{A^n K^n s}{\Omega_0} \geq A^2 \frac{s}{\Omega_0},$$

so at least the same stability time is obtained. ∎

Further Results. The estimates of Theorem 1 apply uniformly to *all* orbits in phase space. Thus they reflect the worst possible cases either for the stability radius or the stability time. In certain regions of phase space, however, much better bounds are available: the stability radius is much smaller in the completely nonresonant domain, while the stability times are much larger in the close vicinity of resonances.

The nonresonant domain is the home of maximal invariant tori of KAM-type, on which orbits enjoy *perpetual* stability, with a bound on the variation of their actions of the order of $\sqrt{\epsilon}$ [1,22]. But although of large measure, the union of all these tori is a *nowhere dense* Cantor set, hence has no interior points. Restricting oneself to a *finite*, but exponentially large time interval, the much simpler Nekhoroshev estimate provides the same stability radius on a relatively *open* subset of phase space of the same large measure. Moreover, since resonances need not be considered, neither the detailed geometric analysis nor the resonant stability estimates are required for this result, and the convexity assumption may also be relaxed if necessary.

From now on we again write s_0 rather than s for the width of the analyticity domain in θ.

Theorem 2. *Suppose h is l, m-quasi-convex and $\epsilon \leq \epsilon_0$. Then*

$$\|I(t) - I_0\| \leq R_\Theta \sqrt{\frac{\epsilon}{\epsilon_0}} \quad \textit{for} \quad |t| \leq T_\Theta \sqrt{\frac{\epsilon_0}{\epsilon}} \exp\left(\frac{s_0}{6}\left(\frac{\epsilon_0}{\epsilon}\right)^a\right)$$

with

$$R_\Theta = \frac{r_0}{A^n}, \qquad T_\Theta = \frac{2^7 A^n s_0}{m r_0}$$

for every orbit with initial position in a relatively open subset of $P \times \mathbb{T}^n$ of relative measure $1 - O(r_0)$.

Proof. Consider the covering of P used in the previous proof. Within the completely nonresonant block D_Θ the stability radius is $r_\Theta = r_0/A^n K^n$, and the stability time is bounded from below by $e^{Ks_0/6}$ times

$$\frac{s_0 r_\Theta}{5\epsilon} \geq \frac{2^7 A^n s_0}{m r_0} \cdot K^n .$$

With $K^n = \sqrt{\epsilon_0/\epsilon}$ the stability estimates follow.

For the measure estimate we may assume that P is bounded. Let G be its bounded image under the frequency map. The complement of the nonresonant block B_Θ is covered by $Z = Z_1^*$, the union of all first order resonance zones, for which we found the measure estimate

$$\mu(Z \cap G) = O(m r_0)$$

at the end of the previous section. The pull back of Z via the frequency map covers the complement of D_Θ in P. Hence, if h is m-convex, then the complement's measure is bounded by

$$\mu(\omega^{-1}(Z) \cap P) \leq m^{-1} \mu(Z \cap G) = O(r_0) ,$$

as we wanted to show.

This argument needs some modification, if h is only l, m-quasi-convex. First of all, by definition h is m-convex on that subdomain of P which is mapped into the ball $\{ \omega : \|\omega\| \leq l \}$. So here the preceding argument applies. In the complementary domain we consider the restriction ω_E of ω to a nonempty energy surface $\{h = E\}$, and in frequency space the retraction $\pi \colon \omega \mapsto \omega / \|\omega\|$ of the domain $X = \{ \omega : \|\omega\| > l \}$ to the unit sphere. The l, m-quasi-convexity implies that the Jacobian of $\pi \circ \omega_E$ is m-convex, as one verifies in suitable coordinates. Moreover, in frequency space,

$$\mu(\pi(Z \cap X)) = O(m r_0) .$$

It follows that the $(n-1)$-dimensional measure of the complement of D_Θ on each energy surface is bounded by

$$\mu(\omega_E^{-1}(Z) \cap P) \leq m^{-1} \mu(\pi(Z \cap X)) = O(r_0) .$$

Taking the union over a finite range of energy values yields an upper bound of the same order for the n-dimensional measure of the complement of D_Θ in P. ∎

Next we look closer at orbits in the vicinity of resonances. Let Λ be a non-trivial integer lattice of dimension d, let R_Λ be the linear space of exact Λ-resonant frequencies of dimension $\nu = n - d$, and

$$S_\Lambda = \{ I \in P : \omega(I) \in R_\Lambda \}$$

the associated resonance surface in action space P. At or very near S_Λ the resonance relations play the rôle of d additional integrals of motion reducing the effective number of degrees of freedom to $\nu = n - d$. In addition, convexity keeps orbits near S_Λ for a long time. The upshot is that the stability estimates depend on the smaller number ν of *free* frequencies rather than the number n of *all* frequencies.

This stabilizing effect of resonances in *convex* systems has been observed by several authors. In the controversial papers by Molchanov [16,17] the stability of the solar system — if it is stable — is attributed to the presence of many almost resonances among the planets rather than the presence of many invariant KAM-tori. Later, this idea was successfully pursued by a group of Italian mathematicians around Galgani. In studying systems with many degrees of freedom they emphasized the rôle of *resonances* and the use of Nekhoroshev estimates in order to control the dynamics in their vicinity by providing bounds which depend essentially only on the smaller number of *free* frequencies.

Recall that Λ is a K-lattice if it is generated by vectors in $\{|k| \le K\}$.

Theorem 3. *Suppose h is l, m-quasi-convex, and S_Λ is a nonempty resonance surface in P associated with a nontrivial K_Λ-lattice Λ of dimension d and codimension $\nu = n - d$. If*

$$\epsilon \le \frac{\epsilon_\Lambda}{K_\Lambda^{2\nu}}, \qquad \epsilon_\Lambda = \frac{m r_0^2}{2^{10} A^{2\nu} |\Lambda|^2},$$

with $r_0 \le 4l/m$, then for every orbit with initial position in $U_\rho S_\Lambda \times \mathbb{T}^n$, $\rho = 4M^{-1}\sqrt{\epsilon m}$, one has

$$\|I(t) - I_0\| \le R_0 \left(\frac{\epsilon}{\epsilon_\Lambda}\right)^{a_\Lambda} \quad \textit{for} \quad |t| \le T_0 \exp\left(\frac{s_0}{6}\left(\frac{\epsilon_\Lambda}{\epsilon}\right)^{a_\Lambda}\right),$$

except when $\|\omega(I_0)\| \le m r_0/8$ in which case $\|I(t) - I_0\| \le r_0$ for all t. The parameters are

$$a_\Lambda = \frac{1}{2\nu}, \qquad R_0 = \frac{r_0}{A}, \qquad T_0 = \frac{s_0}{\Omega_0}$$

with Ω_0 as in Theorem 1. Moreover, with probability $1 - O(r_0)$ within the domain $U_\rho S_\Lambda \times \mathbb{T}^n$ the stability radius is even $32\sqrt{\epsilon/m}$.

This result, as well as Theorem 5 below, may be generalized in the same way as Theorem 1 in Section 1.

Proof. We use the same kind of covering as in the proof of Theorem 1 but this time choose K *larger*. Namely,

$$K = \left(\frac{\epsilon_\Lambda}{\epsilon}\right)^{a_\Lambda} \geq K_\Lambda \quad \Longleftrightarrow \quad \epsilon = \frac{mr_\Lambda^2}{2^{10}} \leq \frac{\epsilon_\Lambda}{K_\Lambda^{2\nu}}$$

in view of the definition of ϵ_Λ and r_Λ. In frequency space the blocks B_L with $\Lambda \subseteq L \in \mathbb{M}_K$ cover the space of exact Λ-resonances R_Λ, and $r_\Lambda \leq r_L \leq r_0$ for all those Λ. Hence their pull backs D_L cover a neighbourhood of S_Λ in P of radius

$$\rho = \frac{\delta_\Lambda}{M} = \frac{mr_\Lambda}{8M} = \frac{4}{M}\sqrt{\epsilon m}.$$

The resonant stability estimate applies to all these blocks simultaneously. The upper bound for the stability radius is the same as before, and the stability time is bounded from below by $e^{Ks_0/6}$ times

$$\frac{s_0}{288\Omega_0}\,\frac{mr_\Lambda^2}{\epsilon} \geq \frac{s_0}{\Omega_0}$$

as claimed. Finally, within the block D_Λ the stability radius is $r_\Lambda = 32\sqrt{\epsilon/m}$, and the relative measure of all sets $D_L \cap U_\rho S_\Lambda$ with $L \supsetneq \Lambda$ is again $1 - O(r_0)$ as one easily verifies. ∎

Incidentally, the proof shows that the same estimates hold in fact on the larger neighbourhood of S_Λ, on which the unperturbed frequencies are $\delta_\Lambda = 4\sqrt{\epsilon m}$-close to exact Λ-resonances.

The estimates of Theorem 3 are the better the smaller the number ν of free frequencies. They are best in small neighbourhoods of unperturbed *periodic* orbits. This special case deserves a special statement, since periodic orbits are better characterized by their minimal period rather than by an integer lattice of resonance relations. More importantly, the pertaining proof does not depend on the Covering Lemma, but reduces directly to the Resonant Stability Estimate. In addition, the construction of the Normal Form only requires an averaging with respect to a single fast variable — the longitude of the periodic orbit — and thus may dispense with Fourier series expansions, small divisors and exponentially weighted norms as was pointed out by Lochak [13]. Here, however, we rely on the Normal Form Lemma as stated.

Theorem 4. *Suppose h is l, m-quasi-convex, and I_* is the action of an unperturbed periodic orbit of minimal period T. If*

$$\epsilon \leq \frac{mr^2}{2^8}, \qquad r \leq \min\left(\frac{\pi}{MT}, \frac{4l}{m}, r_0\right),$$

then for every orbit with $\|I_0 - I_*\| \leq \dfrac{mr}{8M}$ one has

$$\|I(t) - I_0\| \leq r \quad \textit{for} \quad |t| \leq T_* \exp\left(\frac{s_0}{2MTr}\right),$$

where $T_* = \dfrac{s_0}{4\|\omega(I_*)\|}$.

In particular, choosing r *minimal* we obtain

$$\|I(t) - I_0\| \leq 16\sqrt{\frac{\epsilon}{m}} \quad \text{for} \quad |t| \leq T_* \exp\left(\frac{s_0}{32MT}\sqrt{\frac{m}{\epsilon}}\right)$$

for every orbit with $\|I_0 - I_*\| \leq \dfrac{2}{M}\sqrt{\epsilon m}$.

Proof. By assumption there exists a nonzero integer vector k_* in lowest terms such that

$$\omega(I_*) = \alpha_* k_*, \qquad \alpha_* = \frac{2\pi}{T}.$$

It follows that the resonant point I_* itself is α_*, K-nonresonant modulo Λ_* *for all* K, where Λ_* is the $(n-1)$-dimensional integer lattice perpendicular to k_*.

The ball B of radius $\rho = mr/8M$ around I_* is then α, K-nonresonant modulo Λ_* and δ-close to the exact Λ_*-resonance $\omega(I_*)$ with

$$\alpha = \alpha_* - MK\rho, \qquad \delta = M\rho = \frac{mr}{8}.$$

Choosing

$$K = \frac{\pi}{MTr} \geq 1,$$

one easily checks that $\alpha \geq pMKr$ with $p = 3/2$. Thus, all the hypotheses of the Resonant Stability Lemma are satisfied with $p = 3/2$, $q = 3$. As a result, for every orbit starting in $B \times \mathbb{T}^n$ we have the stability radius r and the stability time

$$\frac{s_0}{288\Omega_0}\frac{mr^2}{\epsilon}\,e^{Ks_0/6} \geq \frac{s_0}{2\Omega_0}\exp\left(\frac{s_0}{2MTr}\right),$$

where $\Omega_0 \leq \sup_{\|I-I_*\|\leq r+\rho}\|\omega(I)\|$. Since, for $\|I-I_*\| \leq r+\rho$,

$$\|\omega(I)-\omega(I_*)\| \leq M(r+\rho) \leq 2Mr \leq \frac{2\pi}{T} \leq \|\omega(I_*)\|,$$

we have $\Omega_0 \leq 2\|\omega(I_*)\|$. ∎

In the work of Lochak [12,13] the exponential stability of periodic orbits is the starting point for the more general Nekhoroshev estimates. The idea is to approximate an arbitrary initial position by periodic ones. The one additional ingredient required is Dirichlet's theorem on simultaneous diophantine approximations: for every $\omega \in \mathbb{R}^n$ and natural number $Q \geq 1$,

$$\min_{q=1,2,\ldots,Q}\langle q\omega\rangle \leq \frac{1}{Q^{1/n}},$$

where $\langle x\rangle = \min_{k\in\mathbb{Z}^n}|x-k|$ denotes the distance of x to the integer lattice with respect to the sup-norm $|\cdot|$.

This implies that after fixing any positive $\delta \leq 1$ and any integer $Q \geq 1$, there exists a real T and a nonzero integer vector k for every ω in $|\omega| \geq \delta$ such that

$$\left|\omega - T^{-1}k\right| \leq \frac{1}{TQ^{1/(n-1)}}, \qquad 1 \leq T \leq Q/\delta.$$

To cover all phase space we thus want to apply Theorem 4 simultaneously to all periodic frequencies $T^{-1}k$ with

$$\rho \sim r \sim \frac{1}{TQ^{1/(n-1)}}.$$

The worst smallness condition arises for $T \sim Q$, hence we need

$$\epsilon \precsim r^2 \sim Q^{\frac{2n}{n-1}}.$$

Conversely, we may choose $Q \sim \epsilon^{-\frac{n-1}{2n}}$. As a result, the common stability time is again of the order of the exponential of

$$\frac{1}{Tr} \sim Q^{\frac{1}{n-1}} \sim \epsilon^{-\frac{1}{2n}}.$$

We do not make explicit the constants involved here.

Linear Integrable Systems. We conclude this last section by discussing two simple special cases. First we consider a linear integrable hamiltonian h,

$$h(I) = \omega \cdot I,$$

where ω is a fixed frequency vector satisfying the diophantine conditions

$$|k \cdot \omega| \geq \frac{\gamma_0}{|k|^{\tau-1}}, \qquad 0 \neq k \in \mathbb{Z}^n$$

with some $\gamma_0 > 0$ and $\tau \geq n$. In this case, all of P is in the nonresonant regime, and the Nonresonant Stability Estimate applies to the hamiltonian $H = h + f_\epsilon$. No convexity is required, and no geometry is involved, either. Such hamiltonians were studied, for example, in [4,5,8].

Theorem 5. *Suppose h is linear, ω satisfies the preceding diophantine conditions, and*

$$|f_\epsilon|_{P,r_0,s_0} \leq \epsilon \leq \epsilon_0 = \frac{\gamma_0 r_0}{2^7}.$$

Then for every orbit with initial position in $P \times \mathbb{T}^n$ one has

$$\|I(t) - I_0\| \leq r_0 \quad \textit{for} \quad |t| \leq \frac{T_0}{\epsilon} \exp\left(\frac{s_0}{6}\left(\frac{\epsilon_0}{\epsilon}\right)^{1/\tau}\right),$$

where $T_0 = s_0 r_0 / 5$.

Proof. The domain P is completely α_K, K-nonresonant for every K with $\alpha_K = \gamma_0 / K^{\tau-1}$. Letting $M \to 0$ and $q \to 1$, the Nonresonant Stability Estimate applies with $r = r_0$, if

$$\epsilon \leq \frac{1}{2^7} \frac{\gamma_0 r_0}{K^\tau} = \frac{\epsilon_0}{K^\tau}.$$

That is, for $\epsilon \leq \epsilon_0$ we may choose $K = (\epsilon_0/\epsilon)^{1/\tau} \geq 1$. Then, for every orbit with initial position in $P \times \mathbb{T}^n$ we have the bound $\|I(t) - I_0\| \leq r_0$ for $|t| \leq (s_0 r_0 / 5\epsilon) \cdot e^{K s_0 / 6}$ as claimed. ∎

Kinetic Energy plus Potential. Finally, consider the classical hamiltonian

$$H = \tfrac{1}{2} \langle I, I \rangle + U(\theta)$$

on the phase space $\mathbb{R}^n \times \mathbb{T}^n$. Thus, $h = \frac{1}{2}\langle I, I\rangle$, and $f_\epsilon = U(\theta)$ is independent of I and ϵ. We assume that

$$|U|_{s_0} \le 1$$

by a proper scaling of time and space. Of course, in this autonomous system, all orbits are bounded on account of energy conservation and the convexity of h. It is the non-autonomous case, where U depends analytically on t in a uniform strip, which is of real interest, and which is analyzed in [10]. But it turns out that the same estimates hold in both cases. So we include the autonomous case here as an illustration, while the non-autonomous case will appear elsewhere.

Theorem 6. *Let H be as above. Then for every orbit with initial action I_0 in the annulus $R \le \|I\| \le 2R$, with R sufficiently large, one has*

$$\|I(t) - I_0\| \le c_0 R^{1-1/n} \quad \textit{for} \quad |t| \le \frac{c_1 s_0}{R} \exp\left(\frac{s_0}{c_0 c_1} R^{1/n}\right),$$

where $c_0 = 2^{5/n}$ and $c_1 = 66$.

Proof. We are going to apply Theorem 1 to sufficiently large annuli

$$A_R = \{ I : R \le \|I\| \le 2R \},$$

choosing $r_0 = R$. The integrable hamiltonian $h = \frac{1}{2}\langle I, I\rangle$ is uniformly convex with $\omega(I) = I$ and $Q(I) = I_n$ for all I, whence $m = 1$ and $M = 1$, and l may be chosen arbitrarily large. So we only need to ensure that

$$|U|_{s_0} \le 1 \le \epsilon_0 = \frac{R^2}{2^{10} A^{2n}}, \qquad A = 11.$$

That is, we need $R \ge 2^5 A^n$. We obtain the stability radius

$$R_0 \epsilon_0^{-a} = \frac{R}{A} \cdot \left(\frac{2^5 A^n}{R}\right)^{1/n} = 2^{5/n} R^{1-1/n},$$

since on A_R we have $\|\omega(I)\| \ge R$. The stability time is bounded from below by

$$A^2 \frac{s_0}{\Omega_0} \exp\left(\frac{s_0 \epsilon_0^a}{6}\right) \ge \frac{40 s_0}{R} \exp\left(\frac{s_0}{6}\left(\frac{R}{2^5 A^n}\right)^{1/n}\right),$$

since $\Omega_0 \le \sup_{I \in V_{r_0} A_R} \|\omega(I)\| = 3R$. ∎

A Proof of the Iterative Lemma

The Iterative Lemma is proven in the usual fashion using the Lie-transform method in its simplest form. The canonical transformation Φ to be constructed is written as the time-1-map of the flow X_ϕ^t of a hamiltonian vectorfield X_ϕ:

$$\Phi = X_\phi^t\big|_{t=1}.$$

Given a hamiltonian $H_0 = h + f_0$ we may then use Taylor's formula to write

$$\begin{aligned} H_0 \circ \Phi &= h \circ X_\phi^t\big|_{t=1} + f_0 \circ X_\phi^t\big|_{t=1} \\ &= h + \{h, \phi\} + \int_0^1 (1-t)\,\{\{h,\phi\},\phi\} \circ X_\phi^t \, dt \\ &\quad + f_0 + \int_0^1 \{f_0, \phi\} \circ X_\phi^t \, dt \\ &= h + \{h,\phi\} + f_0 + \int_0^1 \{(1-t)\,\{h,\phi\} + f_0, \phi\} \circ X_\phi^t \, dt. \end{aligned}$$

We apply this expansion with $f_0 = g + T_K f$ and choose ϕ in such a way that $\{h, \phi\} + f_0 = g_+$ is again in resonant normal form with respect to Λ. Equivalently, setting $\chi = g_+ - g$ we solve

$$\{\phi, h\} + \chi = T_K f$$

for ϕ and χ. For the given hamiltonian $H = H_0 + f - T_K f$ we then find $H \circ \Phi = h + g_+ + f_+$ with

$$f_+ = \int_0^1 \{g + f_t, \phi\} \circ X_\phi^t \, dt + (f - T_K f) \circ X_\phi^1,$$

where $f_t = (1-t)\chi + t\, T_K f$.

Formally, a solution of $\{\phi, h\} + \chi = T_K f$ is given by

$$\chi = P_\Lambda T_K f, \qquad \phi = L^{-1}\left(T_K f - P_\Lambda T_K f\right),$$

where L^{-1} denotes the unique inverse of the linear operator $\{\,\cdot\,, h\}$ restricted to the space of functions $\{\, u : T_K u = u,\, P_\Lambda u = 0\,\}$. In terms of Fourier coefficients,

$$\begin{aligned} \chi_k &= f_k && \text{for } k \in \mathbb{Z}_K^n \cap \Lambda \\ \phi_k &= \frac{f_k}{ik \cdot \omega(I)} && \text{for } k \in \mathbb{Z}_K^n \setminus \Lambda, \end{aligned}$$

with all other coefficients being set to zero.

For the estimates let $\delta = |f|_{r,s} < \alpha\rho\sigma/2q$. On the domain D we have $|k\cdot\omega(I)| \ge \alpha$ for $k \in \mathbb{Z}^n_K \setminus \Lambda$, hence

$$|k\cdot\omega(I)| \ge \alpha - MKr \ge \left(1 - \frac{1}{p}\right)\alpha = \frac{\alpha}{q}$$

on $V_r D$ by the mean value theorem and our assumptions on Q and r. In view of the definition of the norm $|\cdot|$ we then have

$$|\chi|_{r,s} \le \delta, \qquad |\phi|_{r,s} \le \frac{q}{\alpha}\delta.$$

The Cauchy estimates of Appendix B then yield

$$|\phi_I|_\infty \le \frac{q\delta}{\alpha\rho} \le \frac{\sigma}{2}, \qquad \|\phi_\theta\| \le |\phi_\theta|_1 \le \frac{q\delta}{\alpha\sigma} \le \frac{\rho}{2}$$

uniformly on $V_{r-\rho,s-\sigma}$. Equivalently, the hamiltonian vectorfield X_ϕ satisfies

$$\left|WX_\phi\right|_{\mathcal{P}} \le \frac{q\delta}{\alpha\rho\sigma} \le \frac{1}{2}, \qquad W = \operatorname{diag}\left(\rho^{-1}I_n, \sigma^{-1}I_n\right),$$

uniformly on $V_{r-\rho,s-\sigma}$. Consequently, its flow X^t_ϕ maps $V_{r-\frac{3}{2}\rho,s-\frac{3}{2}\sigma}$ into $V_{r-\rho,s-\sigma}$ for $0 \le t \le 1$ and satisfies

$$\left|W(X^t_\phi - id)\right|_{\mathcal{P}} \le \frac{q\delta}{\alpha\rho\sigma}, \qquad 0 \le t \le 1,$$

on that smaller domain. The $|W\cdot|_{\mathcal{P}}$-distance of its boundary to its subdomain $V_{r-2\rho,s-2\sigma}$ is 1/2, so that by the generalized Cauchy inequality of Appendix C we also have

$$\left|W(DX^t_\phi - I_{2n})W^{-1}\right|_{\mathcal{P}} \le 2\cdot\frac{q\delta}{\alpha\rho\sigma}, \qquad 0 \le t \le 1,$$

uniformly on $V_{r-2\rho,s-2\sigma}$. This proves our claims for the map $\Phi = X^t_\phi\big|_{t=1}$.

It remains to estimate f_+. Obviously, $|f_t|_{r,s} \le |f|_{r,s}$ and thus

$$|\{f_t,\phi\}|_{r-\rho,s-\sigma} \le \frac{1}{\rho\sigma}|f_t|_{r,s}|\phi|_{r,s} \le \frac{q}{\alpha\rho\sigma}|f|^2_{r,s}$$

by Lemma B.4. Similarly, for each term g_j in g,

$$\left|\{g_j,\phi\}\right|_{r-\rho,s-\sigma} \le \frac{1}{e}\left(\frac{1}{(r_j-r+\rho)\sigma}+\frac{1}{(s_j-s+\sigma)\rho}\right)\left|g_j\right|_{r_j,s_j}|\phi|_{r,s}$$
$$\le \frac{1}{2}\left(\frac{\rho}{r_j-r+\rho}+\frac{\sigma}{s_j-s+\sigma}\right)\left|g_j\right|_{r_j,s_j}\cdot\frac{q}{\alpha\rho\sigma}|f|_{r,s}$$

and hence

$$|\{g,\phi\}|_{r-\rho,s-\sigma} \le \frac{q}{\alpha\rho\sigma}|g|_*|f|_{r,s}$$

in the notation of the Iterative Lemma. Finally, $|f - T_K f|_{r,s-\sigma} \le e^{-K\sigma}|f|_{r,s}$, and we have the general estimate

$$\left|u\circ X_\phi^t\right|_{r-2\rho,s-2\sigma} \le \left(1-\frac{2}{\rho\sigma}|\phi|_{r,s}\right)^{-1}|u|_{r-\rho,s-\sigma}$$

uniformly for $0\le t\le 1$ by Lemma B.5. Putting these pieces together we obtain

$$|f_+|_{r-2\rho,s-2\sigma} \le \left(1-\frac{2q\delta}{\alpha\rho\sigma}\right)^{-1}\left(\frac{|g|_*+|f|_{r,s}}{\alpha\rho\sigma/q}+e^{-K\sigma}\right)|f|_{r,s}$$

as claimed. ∎

B Weighted Norms

We consider functions that are analytic on complex neighbourhoods of a fixed domain $D\times\mathbb{T}^n$. For fixed I we set

$$|u|_{I,s} = \sum_{k\in\mathbb{Z}^n}|u_k(I)|\,e^{|k|s},$$

so that $|u|_{r,s} = |u|_{D,r,s} = \sup_{I\in V_rD}|u|_{I,s}$. We also write V_r for V_rD.

Lemma B.1. *For the sup-norm $|\cdot|_{r,s}$ on $V_{r,s}$ one has, for all $\sigma>0$,*

$$|u|_{r,s} \le |u|_{r,s} \le \coth^n\sigma\ |u|_{r,s+2\sigma}\,.$$

Proof. The first inequality is obvious. The second one follows from the familiar estimate $|u_k(I)| \le e^{-|k|s}|u|_{I,s}$ for the Fourier coefficients of analytic functions and the identity $\sum_k e^{-2|k|\sigma} = \coth^n\sigma$. ∎

Lemma B.2. *For all* $I \in V_r$, $|uv|_{I,s} \le |u|_{I,s}\,|v|_{I,s}$.

Proof. The k-th Fourier coefficient of uv is $(uv)_k = \sum_l u_{k-l} v_l$. Hence,

$$\begin{aligned}
|uv|_{I,s} &= \sum_k |(uv)_k(I)|\, e^{|k|s} \\
&\le \sum_k \sum_l |u_{k-l}(I)|\, e^{|k-l|s} \cdot |v_l(I)|\, e^{|l|s} \\
&= \sum_k |u_k(I)|\, e^{|k|s} \cdot \sum_l |v_l(I)|\, e^{|l|s} \\
&= |u|_{I,s}\, |v|_{I,s}
\end{aligned}$$

as claimed. ∎

Lemma B.3. *For* $0 < \sigma < s$, $0 < \rho < r$ *and* $I \in V_{r-\rho}$,

$$\sum_{1\le i\le n} \left|u_{\theta_i}\right|_{I,s-\sigma} \le \frac{1}{e\sigma} \|u\|_{r,s}, \qquad \max_{1\le i\le n} \left|u_{I_i}\right|_{I,s} \le \frac{1}{\rho} \|u\|_{r,s}.$$

Proof. For the θ-derivatives we have

$$\begin{aligned}
\sum_{1\le i\le n} \left|u_{\theta_i}\right|_{I,s-\sigma} &= \sum_{1\le i\le n} \sum_k |k_i|\, |u_k(I)|\, e^{|k|(s-\sigma)} \\
&= \sum_k |k|\, |u_k(I)|\, e^{|k|(s-\sigma)} \\
&\le \sup_{t\ge 0} t e^{-t\sigma} \cdot |u|_{I,s} \\
&\le \frac{1}{e\sigma}\, |u|_{I,s}
\end{aligned}$$

uniformly for all $I \in V_r$. For the partial derivative with respect to I_i at a point I in $V_{r-\rho}$ we write

$$u_{I_i} = \frac{1}{2\pi i} \int_\gamma \frac{u(I+\zeta,\theta)}{\zeta}\, d\zeta$$

with a circle γ in the I_i-plane around the origin of radius ρ. We obtain

$$\left|u_{I_i}\right|_{I,s} \le \frac{1}{2\pi} \int_\gamma \frac{|u|_{I+\zeta,s}}{|\zeta|}\, |d\zeta| \le \frac{1}{\rho} \|u\|_{r,s}$$

uniformly in i and $I \in V_{r-\rho}$. ∎

Lemma B.4.

$$|\{u, v\}|_{r-\rho,s-\sigma} \le \frac{1}{e}\left(\frac{1}{(r_0-r+\rho)\sigma}+\frac{1}{(s_0-s+\sigma)\rho}\right)|u|_{r_0,s_0}|v|_{r,s}$$

for $0 < r-\rho < r_0$, $0 < s-\sigma < s_0$ *and* $\rho, \sigma > 0$.

Proof. Fix $I \in V_{r-\rho}$. Then, by the preceding lemmata,

$$\begin{aligned} |\langle u_I, v_\theta\rangle|_{I,s-\sigma} &\le \sum_{1\le i\le n} \left|u_{I_i}\right|_{I,s-\sigma} \left|v_{\theta_i}\right|_{I,s-\sigma} \\ &\le \max_{1\le i\le n} \left|u_{I_i}\right|_{I,s-\sigma} \sum_{1\le i\le n} \left|v_{\theta_i}\right|_{I,s-\sigma} \\ &\le \frac{1}{r_0-(r-\rho)} |u|_{r_0,s_0} \cdot \frac{1}{e\sigma} |v|_{r,s}. \end{aligned}$$

Likewise for $|\langle u_\theta, v_I\rangle|_{I,s-\sigma}$. ∎

Lemma B.5. *If* $|\phi|_{r_0,s_0} < \dfrac{\rho\sigma}{2}$*, then*

$$\left|u \circ X_\phi^1\right|_{r-\rho,s-\sigma} \le \left(1-\frac{2}{\rho\sigma}|\phi|_{r_0,s_0}\right)^{-1} |u|_{r,s}$$

for $0 < \rho < r \le r_0-\rho$ *and* $0 < \sigma < s \le s_0-\sigma$.

The smallness condition implies $X_\phi^t\colon V_{r-\rho,s-\sigma} \to V_{r,s}$ for $0 \le t \le 1$. This fact, however, is not used explicitly in the following proof.

Proof. Consider the Lie series expansion

$$u \circ X_\phi^1 = \sum_{h\ge 0} \frac{1}{h!} \operatorname{ad}_\phi^h u,$$

where

$$\operatorname{ad}_\phi^0 u = u, \quad \operatorname{ad}_\phi^h u = \left\{\operatorname{ad}_\phi^{h-1} u, \phi\right\}, \quad h \ge 1.$$

For $h \ge 1$, let $\tilde\rho = \rho/h$, $\tilde\sigma = \sigma/h$. Let $|\cdot|_i = |\cdot|_{r-i\tilde\rho,s-i\tilde\sigma}$ for $1 \le i \le h$. We then have

$$\begin{aligned} \left|\operatorname{ad}_\phi^i u\right|_i &\le \left(\frac{1}{e\tilde\rho(s_0-s+i\tilde\sigma)}+\frac{1}{e\tilde\sigma(r_0-r+i\tilde\rho)}\right)|\phi|_{r_0,s_0}\left|\operatorname{ad}_\phi^{i-1} u\right|_{i-1} \\ &\le \frac{2}{e\tilde\rho\tilde\sigma}\,\frac{1}{h+i}\,|\phi|_{r_0,s_0}\left|\operatorname{ad}_\phi^{i-1} u\right|_{i-1}. \end{aligned}$$

Hence,

$$\left|\mathrm{ad}_\phi^h u\right|_{r-\rho,s-\sigma} \le \left(\frac{2}{e\tilde{\rho}\tilde{\sigma}}\right)^h \frac{h!}{(2h)!} |\phi|_{r_0,s_0}^h |u|_{r,s}.$$

Observing that

$$\left(\frac{2}{e\tilde{\rho}\tilde{\sigma}}\right)^h \frac{1}{(2h)!} \le \left(\frac{2h^2}{e\rho\sigma}\right)^h \frac{e^{2h}}{(2h)^{2h}} \le \left(\frac{2}{\rho\sigma}\right)^h$$

we arrive at

$$\begin{aligned}
\left|u \circ X_\phi^1\right|_{r-\rho,s-\sigma} &\le \sum_{h\ge 0} \frac{1}{h!}\left|\mathrm{ad}_\phi^h u\right|_{r-\rho,s-\sigma} \\
&\le \sum_{h\ge 0} \left(\frac{2}{\rho\sigma}\right)^h |\phi|_{r_0,s_0}^h |u|_{r,s} \\
&= \left(1 - \frac{2}{\rho\sigma}|\phi|_{r_0,s_0}\right)^{-1} |u|_{r,s}. \quad \blacksquare
\end{aligned}$$

C The General Cauchy Inequality

Let A and B be two complex Banach spaces with norms $|\cdot|_A$ and $|\cdot|_B$, and let F be an analytic map from an open subset of A into B. The first derivative $d_v F$ of F at v is a linear map from A into B, whose induced operator norm is

$$|d_v F|_{B,A} = \max_{u\ne 0} \frac{|d_v F(u)|_B}{|u|_A}.$$

The Cauchy inequality can be stated as follows.

Lemma C.1. *Let F be an analytic map from the open ball of radius r around v in A into B such that $|F|_B \le M$ on this ball. Then the inequality*

$$|d_v F|_{B,A} \le \frac{M}{r}$$

holds.

Proof. Let $u \ne 0$ in A. Then $f(z) = F(v + zu)$ is an analytic map from the complex disc $|z| < r/|u|_A$ in $\mathbb{C}$ into B that is uniformly bounded by M. Hence,

$$|d_0 f|_B = |d_v F(u)|_B \le \frac{M}{r} \cdot |u|_A$$

by the usual Cauchy inequality. The above statement follows, since $u \neq 0$ was arbitrary. ∎

References

[1] V.I. Arnold (Ed.), *Dynamical Systems III*, Encyclopaedia of Mathematical Sciences Vol. 3, Springer, Berlin-Heidelberg-New York, 1988.

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

[3] G. Benettin, L. Galgani, A. Giorgilli, *Realization of holonomic constraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory. Part II*, Comm. Math. Phys. **121** (1989), 557–601.

[4] G. Benettin, G. Gallavotti, *Stability of motions near resonances in quasi-integrable Hamiltonian systems*, J. Stat. Phys. **44** (1986), 293–338.

[5] A. Delshams, P. Gutiérrez, *Effective stability for nearly integrable Hamiltonian systems*, Proceedings of the International Conference on Differential Equations, Barcelona 1991 (C. Perelló, C. Simó, J. Solá-Morales, eds.), World Scientific, Singapore, 1993.

[6] F. Fassò, *Lie series method for vector fields and Hamiltonian perturbation theory*, Jour. Appl. Math. Phys. **41** (1990), 843–864.

[7] G. Gallavotti, *Quasi-integrable mechanical systems*, Phénomènes critiques, systèmes aléatoires, théories de jauge, Part II (Les Houches 1984), North-Holland, Amsterdam-New York, 1986, 539–624.

[8] A. Giorgilli, *New insights on the stability problem from recent results in classical perturbation theory*, Preprint, Quaderno n. 35, Università degli studi di Milano, 1990.

[9] A. Giorgilli, L. Galgani, *Rigorous estimates for the series expansions of Hamiltonian perturbation theory*, Celestial Mech. **37** (1985), 95–112.

[10] A. Giorgilli, E. Zehnder, *Exponential stability for time dependent potentials*, Z. Ang. Math. Phys. **43** (1992), 827–855.

[11] Yu.S. Il'yashenko, *A steepness test for analytic functions*, Uspekhi Mat. Nauk **41** (1986), 193–194; Russian Math. Surveys **41** (1986), 229–230.

[12] P. Lochak, *Stabilité en temps exponentiels des systèmes Hamiltoniens proches de systèmes intégrables: résonances et orbites fermés*, Preprint, Laboratoire de Mathématiques de l'Ecole Normale Supérieure, 1990.

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

[14] P. Lochak, C. Meunier, *Multiphase Averaging for Classical Systems*, Applied Mathematical Sciences Vol. 72, Springer, 1988.

[15] P. Lochak, A.I. Neishtadt, *Estimates of stability time for nearly integrable systems with a quasiconvex Hamiltonian*, Chaos **2** (1992), 492–499.

[16] A.M. Molchanov, *The resonant structure of the solar system*, Icarus **8** (1968), 203–215.

[17] A.M. Molchanov, *The reality of resonances in the solar system*, Icarus **11** (1969), 104–110.

[18] A.I. Neishtadt, *The separation of motions in systems with rapidly rotating phase*, J. Appl. Math. Mech. **48** (1984), no. 2, 133–139; Prikl. Mat. Mekh. **48** (1984), no. 2, 197–204.

[19] N.N. Nekhoroshev, *Behaviour of Hamiltonian systems close to integrable*, Functional. Anal. Appl. **5** (1971), 338–339.

[20] N.N. Nekhoroshev, *An exponential estimate of the time of stability of nearly-integrable Hamiltonian systems I*, Uspekhi Mat. Nauk **32** (1977), 5–66; Russian Math. Surveys **32** (1977), 1–65.

[21] N.N. Nekhoroshev, *An exponential estimate of the time of stability of nearly-integrable Hamiltonian systems II*, Trudy Sem. Petrovsk. **5** (1979), 5–50; *Topics in Modern Mathematics, Petrovskii Sem. No. 5*, O.A. Oleinik Ed., Consultant Bureau, New York, 1985.

[22] J. Pöschel, *Integrability of Hamiltonian systems on Cantor sets*, Comm. Pure Appl. Math. **35** (1982), 653–695.