Version 1.2, Oct 6, 1998
In: *Int. Math. Res. Not.* 1999, No. 4, 203–215

On Nekhoroshev's Estimate at an Elliptic Equilibrium

JÜRGEN PÖSCHEL
Universität Stuttgart

Dedicated to Jürgen Moser on the occasion of his 70th birthday

1 Results

Nekhoroshev in his celebrated paper [7] showed that under a perturbation of order ε of a generic integrable Hamiltonian the actions of an arbitrary orbit vary only of the order of ε^b over a time interval of the order of $\exp(\varepsilon^{-a})$. Here, a and b are positive *stability exponents* which depend only on the number of degrees of freedom and the *steepness* of the unperturbed Hamiltonian. This result was proven for Hamiltonians in action-angle-coordinates, but Nekhoroshev conjectured that it should also hold, for example, in full neighbourhoods of elliptic equilibria.

In the eighties a lot of research was devoted to improving the stability exponents for the particular case of convex or quasi-convex Hamiltonians – the steepest Hamiltonians – resulting finally in the probably optimal values

$$a = \frac{1}{2n}, \qquad b = \frac{1}{2n}$$

obtained independently and simultaneously by Lochak & Neishtadt [5] and the author [9]. The latter paper employs Nekhoroshev's "classical" technique of analyzing the system in certain regions around all possible types of resonances and covering the whole phase space by such "blocks". The former paper employs Lochak's novel idea [3] of analyzing just neighbourhoods of periodic orbits of the unperturbed system and approximating all other initial positions by periodic ones.

These results, too, are restricted to Hamiltonians in action-angle-coordinates. To apply them to an elliptic equilibrium, one has to exclude cusp-shaped domains in which these coordinates become singular. Hence it seemed possible that through these cusps orbits could escape at a faster rate.

But recently, Guzzo, Fassò & Benettin [1] and Niederman [8], again simultaneously and independently, showed that Nekhoroshev's conjecture is indeed correct, at least for convex systems. The former paper uses the "classical" technique transfered to cartesian coordinates and has the potential to be generalized to steep systems – see also [2]. The latter paper uses Lochak's technique, which in fact is not tied to any coordinates at all.

These results are important, but the proofs are somewhat technical and not easily accessible to an unprepared reader. In this note we therefore want to present a complete proof, which is an *immediate* extension of Lochak's transparent exposition in [4], without any further complications. In fact, the extension is so straightforward that in this respect there does not seem to be any significant difference between Hamiltonians in rectangular or action-angle coordinates.

Consider a real analytic Hamiltonian near an elliptic equilibrium whose characteristic frequencies are nonresonant up to order $l \geq 4$. Then there exist symplectic coordinates in $\mathbb{R}^n \times \mathbb{R}^n$ such that

$$H = \langle \alpha, I \rangle + \tfrac{1}{2} \langle AI, I \rangle + B(I) + P(x, y),$$

where

$$I = (I_1, \dots, I_n), \qquad I_j = \tfrac{1}{2}(x_j^2 + y_j^2),$$

the term B is at least of order 3 in I and absent for $4 \leq l \leq 5$, and $P = O_{l+1}(x, y)$ is of order $l + 1$ in x and y. The coefficients of A and B are uniquely determined, they are the so called Birkhoff invariants of H.

Throughout this note we consider I as a function of x and y. That is,

$$I\colon \ \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n, \quad (x, y) \mapsto I = \big(\dots, \tfrac{1}{2}(x_j^2 + y_j^2), \dots\big).$$

It is legitimate to *think* of them as action coordinates, but they are *never introduced* as such. So, for example, with a solution $(x(t), y(t))$ of our Hamiltonian system we associate the curve

$$I(t) = \big(\dots, \tfrac{1}{2}(x_j^2(t) + y_j^2(t)), \dots\big)$$

in $\mathbb{R}^n$.

As usual, we consider the case where the integrable Hamiltonian

$$H_0 = \langle \alpha, I \rangle + \tfrac{1}{2} \langle AI, I \rangle + B(I)$$

is convex in I for sufficiently small I. This is equivalent to A being positive definite. – Let $|I| = |I_1| + \cdots + |I_n|$.

Theorem 1.1 ([2, 8]) *Suppose A is positive definite. Then for every orbit of the Hamiltonian H with $|I(0)| < \delta^2$ sufficiently small one has*

$$|I(t) - I(0)| < c\delta^{2+\lambda a} \quad \textit{for} \quad |t| < \frac{1}{|\alpha|} \exp\left(d\delta^{-\lambda a}\right),$$

with

$$a = \frac{1}{2n}, \qquad \lambda = l - 3,$$

where the constants c and d depend only on A and the dimension n.

Remark 1. Note that δ is, up to a factor, the euclidean distance of the initial position to the origin.

Remark 2. If the components of α are all of the same sign, then H is definite, and the perpetual stability of the equilibrium is not an issue. Still, also in this case the theorem provides additional, nontrivial information about the variations of the actions over exponentially long time intervals.

Remark 3. If α is non-resonant, then $l = \infty$, and some *hyperexponential* estimate will hold, extending the result of Morbidelli & Giorgilli [6] to elliptic equilibria. In fact, estimates can be obtained for arbitrary non-resonant α in terms of the function

$$\omega(m) = \min_{0<|k|\le m} |\langle k, \alpha \rangle|,$$

not just for the special case of diophantine α. We did not pursue this further.

The estimates of Theorem 1.1 are given in coordinates in which the Hamiltonian is in Birkhoff normal form up to order l included. Hence they are not “invariant”. Passing to more general coordinates, however, one does not much affect the stability bounds, while the stability time remains the same.

Corollary 1.2 *Suppose the Hamiltonian*

$$H = \langle \alpha, I \rangle + \dots = \frac{1}{2} \sum_{j=1}^{n} \alpha_j (x_j^2 + y_j^2) + \dots$$

is in Birkhoff normal form up to order $m \geq 2$, but could be brought into a convex Birkhoff normal form up to order $l \geq 4$. Then for every orbit with $|I(0)| < \delta^2$ sufficiently small one has

$$|I(t) - I(0)| < c \max\left(\delta^{2+\lambda a}, \delta^{m+1}\right) \quad \textit{for} \quad |t| < \frac{1}{|\alpha|} \exp\left(d\delta^{-\lambda a}\right),$$

where the constants are the same as in Theorem 1.1.

We reduce Theorem 1.1 to a more standard situation by performing the scaling

$$x = \delta\tilde{x}, \quad y = \delta\tilde{y},$$

which maps the ball $|\tilde{I}| < 1$ one-to-one onto the ball $|I| < \delta^2$. Dividing the resulting Hamiltonian by δ^4 we obtain

$$\tilde{H} = \langle \tilde{\alpha}, \tilde{I} \rangle + \tfrac{1}{2} \langle A\tilde{I}, \tilde{I} \rangle + \delta^{-4} B(\delta^2 \tilde{I}) + \varepsilon F(\tilde{x}, \tilde{y}, \varepsilon),$$

where

$$\tilde{\alpha} = \frac{\alpha}{\delta^2}, \qquad \varepsilon = \delta^{l-3} = \delta^{\lambda}, \qquad \delta = \varepsilon^{1/\lambda},$$

and F is real analytic and uniformly bounded on a fixed ball around the origin. For this Hamiltonian we prove the following result.

Theorem 1.3 *Suppose A is positive definite. If ε is sufficiently small, then for every orbit of the Hamiltonian $\tilde{H}$ with $|\tilde{I}(0)| < 1$ one has*

$$|\tilde{I}(t) - \tilde{I}(0)| < c\varepsilon^a \quad \textit{for} \quad |t| < \frac{1}{|\tilde{\alpha}|} \exp\left(d\varepsilon^{-a}\right)$$

with $a = 1/2n$, where the constants c and d depend only on A and the dimension n, but not on $\tilde{\alpha}$.

We observe that Theorem 1.3 gives precisely the same stability exponents as the usual Nekhoroshev estimates for convex Hamiltonian systems in action-angle-coordinates, such as in [4] or [9]. There is no restriction on $\tilde{\alpha}$, so the singular dependence of $\tilde{\alpha}$ on ε does not affect this result except for a minor influence on the bound for $|t|$.

Theorem 1.1 follows from Theorem 1.3, the definitions of $\tilde{\alpha}$ and ε, and the scaling of the Hamiltonian by δ^2. So it suffices to prove the latter.

2 Proof of Theorem 1.3

Set up. We drop the tilde from the notation, and consider the Hamiltonian

$$H = \langle \alpha, I \rangle + \tfrac{1}{2} \langle AI, I \rangle + G(I) + \varepsilon F(z, \varepsilon),$$

where now $z = (x, y)$ and

$$I = I(z) = \bigl(\ldots, \tfrac{1}{2}(z_j^2 + z_{n+j}^2), \ldots\bigr).$$

In view of the scaling performed in the previous section to obtain this Hamiltonian we can make the following assumptions.

Assumptions. There exist a positive constant M such that

$$\frac{1}{M} \le A \le M.$$

The integrable Hamiltonian G is real analytic on the fixed complex ball B_4 with

$$|G|_4 \le \frac{1}{4M},$$

where $B_r = \{z : \|z\| < r\} \subset \mathbb{C}^n \times \mathbb{C}^n$, and $|\cdot|_r$ denotes the sup-norm over B_r. The perturbing Hamiltonian F is also real analytic on B_4 with

$$|F|_4 \le 1,$$

uniformly for all small ε.

Here and later, *real analytic* means *analytic* in each complex variable and *real* for real arguments. We then want to prove that for every orbit of this Hamiltonian with $|I(0)| < 1$ one has

$$|I(t) - I(0)| < c\varepsilon^a \quad \text{for} \quad |t| < \frac{1}{|\alpha|} \exp\bigl(d\varepsilon^{-a}\bigr).$$

This is done in three steps.

Normal Form. Consider the integrable Hamiltonian

$$H_0 = \langle \alpha, I \rangle + \tfrac{1}{2} \langle AI, I \rangle + G(I).$$

Along a real solution with initial position z^o we have $I(t) = I(0) = I(z^o) =: I^o$, and the motion is quasi-periodic with frequencies

$$\omega^o = \omega(I^o) = \frac{\partial H_0}{\partial I}(I^o).$$

Consider the case of a T*-periodic* frequency vector:

$$T\omega^o \in 2\pi\mathbb{Z}^n, \qquad T\omega^o \neq 0,$$

for some real T, giving rise to a non-trivial T-periodic solution. The set of all real initial positions z with these frequencies ω^o forms a torus

$$\mathcal{T}(I^o) = \left\{ z : I(z) = I^o \right\} \subset \mathbb{R}^n \times \mathbb{R}^n,$$

whose dimension equals the number of positive components of I^o. In the following, $|I^o| < 1$.

To obtain a normal form for H in a neighbourhood of $\mathcal{T}(I^o)$, let

$$J = I - I^o, \quad \text{or} \quad I = I^o + J,$$

keeping in mind that these are just short hand notations for real analytic expressions in z. Up to an irrelevant additive constant, we then can write

$$H_0 = \left\langle \omega^o, I \right\rangle + g(I)$$

with

$$g(I) = \tfrac{1}{2} \langle AJ, J \rangle + G(I^o + J) - G(I^o) - dG_{I^o}(J).$$

By our assumptions,

$$\frac{1}{2M} \left| I - I^o \right|^2 \leq |g(I)| \leq 2M \left| I - I^o \right|^2 \tag{1}$$

on B_3. The total Hamiltonian H is then

$$H = h(I) + g(I) + f(z)$$

with $h(I) = \langle \omega^{\text{o}}, I \rangle$, g as above, and $f(z) = \varepsilon F(z,\varepsilon)$, where we do not indicate the dependence of f on ε for brevity. The Hamiltonians h and g are integrable and hence in involution, the flow of h is T-periodic, and $|f|_3 \le \varepsilon$ is small.

We study this Hamiltonian H on the complex domains

$$\begin{aligned} D_{r,s} &= \left\{ z : \left| I(z) - I^{\text{o}} \right| < r, \quad \|z\| < s \right\} \\ &= \{ z : |J(z)| < r, \quad \|z\| < s \} \subset \mathbb{C}^n \times \mathbb{C}^n. \end{aligned}$$

For $|I^{\text{o}}| < 1$ and $r > 0$, $s > 1$, these are nonempty neighbourhoods of $\mathcal{T}(I^{\text{o}})$.

In the following we do not attempt to obtain estimates with particularly sharp constants. Indeed, we suppress all constants, and use the notations $u <\cdot\, v$ and $u \cdot\!< v$ to indicate that $u < cv$ and $cu < v$, respectively, with some constant $c \ge 1$, which depends only on M and the dimension n. Similarly for $u =\cdot\, v$ later on.

Let $|\cdot|_{r,s}$ denote the sup-norm over $D_{r,s}$, and X_h the vector field of a Hamiltonian h. The following lemma is proven in the next section.

Lemma 2.1 (Normal Form Lemma) *Consider $H = h+g+f$ as described above. If*

$$mT\varepsilon \cdot\!< r, \qquad mTr \cdot\!< 1$$

with an integer $m \ge 1$ and $0 < r \cdot\!< 1$, then there exists a real analytic, symplectic transformation $\Psi\colon D_{2r,2} \to D_{3r,3}$ with $|\Psi - \mathrm{id}|_{2r,2} <\cdot\, T\varepsilon$, such that

$$H \circ \Psi = h + g + \hat{g} + \hat{f}$$

with $\{h, \hat{g}\} = 0$ and

$$|\hat{g}|_{2r,2} <\cdot\, \varepsilon, \qquad |\hat{f}|_{2r,2} <\cdot\, 2^{-m}\varepsilon, \qquad \left| X_{\hat{f}} \right|_{2r,2} < 2^{-m}\varepsilon. \tag{2}$$

Here, the implicit constants are in fact independent of M and n.

Local Stability. We use the normal form to prove a stability result near periodic solutions which is completely analogous to Theorem 1 in [4].

Lemma 2.2 (Local Stability Estimate) *Consider the Hamiltonian of the set up, and let z^{o} be an initial position in $\|z^{\text{o}}\| < 1$ with a T-periodic frequency vector $\omega^{\text{o}} = \omega(I^{\text{o}})$. If*

$$\varepsilon \cdot\!< r^2, \qquad mTr \cdot\!< 1$$

with an integer $m \geq 1$ *and* $0 < r \cdot\!< 1$, *then for every initial position with amplitudes* $I(0)$ *satisfying* $|I(0) - I^{\mathrm{o}}| \cdot\!< r$ *one has*

$$\left|I(t) - I^{\mathrm{o}}\right| < r \quad \textit{for} \quad |t| < \frac{2^m}{|\omega^{\mathrm{o}}|}.$$

Proof. By hypothesis, the normalizing transformation of Lemma 2.1 satisfies

$$|\Psi - \mathrm{id}|_{2r,2} <\!\cdot\, T\varepsilon \cdot\!< Tr^2 \cdot\!< r,$$

so the image of $D_{2r,2}$ under Ψ covers $D_{r,1}$. It therefore suffices to prove the claim for the Hamiltonian in normal form. So let $H = h + g + \hat{g} + \hat{f}$ with $\{h, \hat{g}\} = 0$ satisfying (2).

Along an orbit, consider $h(t) = h(z(t))$. Then

$$\dot{h} = \frac{d}{dt}h = \{h, H\} = \{h, \hat{f}\} = dh(X_{\hat{f}}).$$

As long as the orbit stays inside the domain of the normal form we thus have

$$\left|\dot{h}\right| \leq \left|\omega^{\mathrm{o}}\right| \left|X_{\hat{f}}\right|_{2r,2} < \left|\omega^{\mathrm{o}}\right| 2^{-m}\varepsilon.$$

By energy conservation, $H(t) = H(0)$ for all t, hence

$$|g(t)| \leq |g(0)| + |h(t) - h(0)| + |\hat{g}(t) - \hat{g}(0)| + |\hat{f}(t) - \hat{f}(0)|.$$

By the estimates for $\dot{h}$ as well as for $\hat{g}$ and $\hat{f}$, we obtain

$$|g(t)| <\!\cdot\, |g(0)| + \varepsilon \quad \text{for} \quad |t| < \frac{2^m}{|\omega^{\mathrm{o}}|},$$

and with the convexity (1) of g,

$$\frac{1}{2M}\,|J(t)|^2 <\!\cdot\, 2M\,|J(0)|^2 + \varepsilon \quad \text{for} \quad |t| < \frac{2^m}{|\omega^{\mathrm{o}}|}.$$

Together with $|J(0)| \cdot\!< r$ and $\varepsilon \cdot\!< r^2$ we conclude that

$$|J(t)| < r \quad \text{for} \quad |t| < \frac{2^m}{|\omega^{\mathrm{o}}|}.$$

In particular, for that time interval the orbit stays inside the domain of the normal form. ∎

Lemma 2.2 can be interpreted in a variety of ways – see [4]. Here we move on directly to global estimates.

Global Stability. Here, we follow almost literally [4, p 890–1]. – The idea is to approximate an arbitrary initial position by periodic orbits of the integrable reference system and to apply the local stability estimates around these periodic orbits. The basic ingredient of this approach is Dirichlet's theorem on simultaneous approximations, which states that for every $\omega \in \mathbb{R}^n$ and every integer $Q \geq 1$,

$$\min_{1 \leq q \leq Q, q \in \mathbb{Z}} \min_{p \in \mathbb{Z}^n} |q\omega - p|_\infty \leq \frac{1}{Q^{1/n}}, \tag{3}$$

where $|\cdot|_\infty$ denotes the maximum norm.

Consider an arbitrary initial position with amplitudes I in $|I| < 1$, and let

$$\omega = \omega(I) = \frac{\partial H_0}{\partial I}(I)$$

be its frequencies with respect to the integrable reference Hamiltonian H_0. Without loss of generality we may assume that $|\omega|_\infty \geq 1$. Given $\varepsilon > 0$, choose

$$Q = \varepsilon^{-a(n-1)}$$

with an exponent a yet to be determined. Scaling down ω slightly until one of its components hits an integer, approximating the remaining $n - 1$ components by a rational frequency vector according to (3) and then scaling back, we find that there exists a $2\pi T$-periodic frequency vector ω^{o} such that

$$\left|\omega - \omega^{\mathrm{o}}\right|_\infty \leq \frac{1}{T Q^{1/(n-1)}} = \frac{\varepsilon^a}{T}, \qquad \tfrac{1}{2} \leq T \leq Q.$$

This frequency vector in turn corresponds to a unique amplitude vector I^{o} due to the convexity of the frequency map of H_0. That is, we have $\omega^{\mathrm{o}} = \omega(I^{\mathrm{o}})$ with

$$\left|I - I^{\mathrm{o}}\right| <\cdot \frac{\varepsilon^a}{T}$$

in view of the assumptions of the set up. To apply the local stability estimate around $\mathcal{T}(I^{\mathrm{o}})$ in such a way that the initial amplitude I is covered by it, we set

$$r =\cdot \frac{\varepsilon^a}{T}$$

with a sufficiently *large* implicit constant. Then we have $|I - I^{\mathrm{o}}| \cdot\!< r$. Moreover, in view of $T \le Q$, the hypothesis $\varepsilon \cdot\!< r^2 =\!\cdot\, \varepsilon^{2a}/T^2$ is satisfied, if

$$\varepsilon <\!\cdot\, \frac{\varepsilon^{2a}}{Q^2} = \varepsilon^{2na},$$

which is possible with $a = 1/2n$.

It remains to satisfy the second assumption of Lemma 2.2, which amounts to a bound on the integer m. We need $mTr =\!\cdot\, m\varepsilon^a \cdot\!< 1$, so we may choose

$$m \cdot\!= \varepsilon^{-a}.$$

Given an initial amplitude $I(0)$ and ε sufficiently small, we then apply Lemma 2.2 around a suitably chosen periodic orbit with amplitudes I^{o} with $|I(0) - I^{\mathrm{o}}| \cdot\!< r$. We conclude that

$$\left|I(t) - I^{\mathrm{o}}\right| < r \quad \text{for} \quad |t| < \frac{1}{|\omega|} \exp(d\varepsilon^{-a}).$$

A fortiori, we then also have $|I(t) - I(0)| < r <\!\cdot\, \varepsilon^a$ over the same time interval. This proves Theorem 1.3.

3 Proof of Lemma 2.1

Instead of Lemma 2.1 proper we prove the following variant of it which is formulated in terms of vector field norms. It clearly applies as well and gives the same conclusions.

Lemma 3.1 *Consider a real analytic Hamiltonian $H = h + g + f$ on $D_{3r,3}$, where h and g are integrable, the flow of $h = \langle \omega, I \rangle$ is T-periodic, and*

$$\left|X_g\right|_{3r,3} < \delta, \qquad \left|X_f\right|_{3r,3} < \varepsilon.$$

If

$$\delta <\!\cdot\, r, \qquad mT\varepsilon \cdot\!< r, \qquad mTr \cdot\!< 1 \tag{4}$$

with some integer $m \ge 1$ and $0 < r \cdot\!< 1$, then there exists a real analytic, symplectic transformation $\Psi\colon D_{2r,2} \to D_{3r,3}$ with $|\Psi - \mathrm{id}|_{2r,2} <\!\cdot\, T\varepsilon$, such that $H \circ \Psi =$

$h + \tilde{g} + \hat{f}$ with

$$\left|X_{\tilde{g}} - X_g\right|_{2r,2} < 2\varepsilon, \qquad \left|X_{\hat{f}}\right| < 2^{-m}\varepsilon,$$

and $\{h, \tilde{g}\} = 0$.

As usual, this lemma is proven by iterating an averaging transformation a finite number of times – namely, m times. The general step of this procedure is described in the next lemma.

Lemma 3.2 (Iterative Lemma) *Suppose the Hamiltonian $H = h + g + f$ is real analytic on $D_{r,s}$, where the flow of $h = \langle \omega, I \rangle$ is T-periodic and*

$$\left|X_g - X_{g_0}\right|_{r,s} < \gamma, \qquad \left|X_{g_0}\right|_{r,s} < \delta, \qquad \left|X_f\right|_{r,s} < \varepsilon$$

with an integrable Hamiltonian g_0. If

$$T\varepsilon \cdot\!< \rho \cdot\!< \sigma$$

with $0 < \rho < r$ and $0 < \sigma < s$, then there exists a real analytic, symplectic transformation $\Phi\colon D_{r-\rho,s-\sigma} \to D_{r,s}$, such that $H \circ \Phi = h + g_+ + f_+$ with

$$\left|X_{g_+} - X_g\right|_{r,s} < \varepsilon, \qquad \left|X_{f_+}\right|_{r-\rho,s-\sigma} <\!\cdot\, T\left(\frac{\delta}{\sigma} + \frac{\gamma + \varepsilon}{\rho}\right)\varepsilon,$$

and $\{h, g_+ - g\} = 0$. Moreover, $|\Phi - \mathrm{id}|_{r-\rho,s-\sigma} < T\varepsilon$.

Proof. As usual, Φ is written as the time-1-map X_ϕ^1 of a Hamiltonian vector field X_ϕ, where ϕ solves the homological equation

$$\{\phi, h\} = f - \bar{f}.$$

Standard calculations then lead to $H \circ \Phi = h + g_+ + f_+$ with $g_+ = g + \bar{f}$ and

$$f_+ = \int_0^1 \{g + f_t, \phi\} \circ X_\phi^t \,\mathrm{d}t, \qquad f_t = tf + (1-t)\bar{f}.$$

The operator $\{\cdot, h\}$ is semi-simple, so $\bar{f}$ is determined as the projection of f onto the null space of this operator. Whence $\{g_+ - g, h\} = \{\bar{f}, h\} = 0$. Then ϕ is the preimage of $f - \bar{f}$ in its range. Since the flow of h is T-periodic, these solutions are given by

$$\bar{f} = \frac{1}{T}\int_0^T f \circ X_h^t \,\mathrm{d}t, \qquad \phi = \frac{1}{T}\int_0^T t(f - \bar{f}) \circ X_h^t \,\mathrm{d}t.$$

The expressions for the corresponding vector fields are

$$X_{\bar{f}} = \frac{1}{T}\int_0^T (X_h^t)^* X_f \,\mathrm{d}t, \qquad X_\phi = \frac{1}{T}\int_0^T t(X_h^t)^*(X_f - X_{\bar{f}}) \,\mathrm{d}t.$$

From these representations we immediately get $\left|X_{\bar{f}}\right|_{r,s} < \varepsilon$ and $\left|X_\phi\right|_{r,s} < T\varepsilon$, hence also the first of the postulated estimates. From $T\varepsilon \cdot\!< \rho \cdot\!< \sigma$ by hypothesis we find that $X_\phi^t \colon D_{r-\rho,s-\sigma} \to D_{r,s}$ for $-1 \le t \le 1$, because $\left|\dot{I}\right| \le \left|X_\phi\right| |z|$. Also,

$$\left|X_\phi^t - \mathrm{id}\right|_{r-\rho,s-\sigma} <\!\cdot\; T\varepsilon, \qquad \left|DX_\phi^t - I\right|_{r-\rho,s-\sigma} <\!\cdot\; \frac{T\varepsilon}{\rho}$$

for $-1 \le t \le 1$.

It remains to estimate f_+, or rather

$$X_{f_+} = \int_0^1 (X_\phi^t)^* \left[X_g + X_{f_t}, X_\phi\right] \mathrm{d}t.$$

It suffices to estimate the Lie bracket itself, hence $\left[X_{f_t}, X_\phi\right]$, $\left[X_g - X_{g_0}, X_\phi\right]$ and $\left[X_{g_0}, X_\phi\right]$. As to the first bracket, we write

$$\left[X_{f_s}, X_\phi\right] = \frac{\mathrm{d}}{\mathrm{d}t}(X_\phi^t)^* X_{f_s}\Big|_{t=0},$$

and observe that the flow of X_ϕ^t starting in $D_{r-\rho,s-\sigma}$ exists within $D_{r,s}$ for *complex* t with

$$|t| \cdot\!< \frac{\rho}{\left|X_\phi\right|_{r,s}}.$$

Hence, with Cauchy's estimate we obtain

$$\left|\left[X_{f_s}, X_\phi\right]\right|_{r-\rho,s-\sigma} <\!\cdot\; \frac{1}{\rho}\left|X_{f_s}\right|_{r,s}\left|X_\phi\right|_{r,s} <\!\cdot\; \frac{T\varepsilon}{\rho} \cdot \varepsilon.$$

The estimate of $\left|\left[X_g - X_{g_0}, X_\phi\right]\right|_{r-\rho,s-\sigma}$ is similar, using $\left|X_g - X_{g_0}\right|_{r,s} < \gamma$.

The bracket $[X_{g_0}, X_\phi]$, however, is estimated slightly differently – and this is the only point in the whole scheme where one should not apply the most obvious estimate. We write

$$[X_{g_0}, X_\phi] = -\frac{\mathrm{d}}{\mathrm{d}t}(X_{g_0}^t)^* X_\phi \Big|_{t=0}.$$

Since g_0 is integrable, the flow $X_{g_0}^t$ leaves I invariant. So the existence of this flow is only restricted by the condition that $\|X_{g_0}^t(z)\| < s$ for $\|z\| < s - \sigma$. Whence this flow exists for complex t with

$$|t| \cdot\!< \frac{\sigma}{|X_{g_0}|_{r,s}},$$

and by Cauchy's estimate we obtain

$$\big|[X_{g_0}, X_\phi]\big|_{r-\rho, s-\sigma} <\!\cdot \frac{1}{\sigma}\, |X_{g_0}|_{r,s}\, |X_\phi|_{r,s} <\!\cdot \frac{T\delta}{\sigma} \cdot \varepsilon.$$

This proves the lemma. ∎

Proof of Lemma 3.1. We apply the Iterative Lemma m times, so that

$$H_i \circ \Phi_i = H_{i+1} = h + g_{i+1} + f_{i+1}, \qquad 0 \le i \le m-1,$$

starting with $H_0 = h + g_0 + f_0 = h + g + f = H$, the given Hamiltonian, and choose uniformly

$$\rho = \frac{r}{m}, \qquad \sigma = \frac{1}{m}$$

for each step. We claim that at each step we obtain $\{h, g_i\} = 0$ and

$$|X_{g_i} - X_{g_0}|_{r-i\rho, s-i\sigma} < 2(1-2^{-i})\varepsilon, \qquad |X_{f_i}|_{r-i\rho, s-i\sigma} < 2^{-i}\varepsilon.$$

This clearly holds for $i = 0$, so we can proceed by induction.

To apply the Iterative Lemma for some $i \ge 0$ we observe that

$$T\varepsilon = mT\varepsilon \cdot \frac{1}{m} \cdot\!< \frac{r}{m} = \rho \cdot\!< \sigma$$

by the hypothesis (4) and $r \cdot\!< 1$. So the lemma applies. The estimate for $X_{g_{i+1}}$ is immediate. To estimate $X_{f_{i+1}}$ we note that with $\delta <\cdot r$, $\gamma < 2\varepsilon$ we have

$$T\frac{\delta}{\sigma} <\cdot mTr, \qquad T\frac{\gamma+\varepsilon}{\rho} <\cdot \frac{mT\varepsilon}{r},$$

so with hypothesis (4) we can arrange that

$$T\left(\frac{\delta}{\sigma}+\frac{\gamma+\varepsilon}{\rho}\right) \cdot< \frac{1}{2}.$$

This gives the estimate for $X_{f_{i+1}}$.

Finally we note that

$$|\Phi_i - \mathrm{id}|_{r-(i+1)\rho, s-(i+1)\sigma} < T\varepsilon_i = 2^{-i}T\varepsilon,$$

from which the estimate for $\Psi = \Phi_0 \circ \cdots \circ \Phi_{m-1}$ follows by routine arguments. – This completes the proof of Lemma 3.1. ∎

References

[1] M. Guzzo, F. Fassò & G. Benettin, On the stability of elliptic equilibria. *Math. Phys. Electr. J.* **4** (1998), paper 1, 16 pages.

[2] F. Fassò, M. Guzzo & G. Benettin, Nekhoroshev-stability of elliptic equilibria of Hamiltonian systems. *Comm. Math. Phys.* (to appear).

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

[4] P. Lochak, Hamiltonian perturbation theory: periodic orbits, resonances and intermittency. *Nonlinearity* **6** (1993), 885–904.

[5] P. Lochak & A. Neishtadt, Estimates of stability time for nearly integrable systems with a quasiconvex Hamiltonian. *Chaos* **2** (1992), 495–499.

[6] A. Morbidelli & A. Giorgilli, Superexponential stability of KAM tori. *J. Stat. Phys.* **78** (1995), 1607–1617.

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

[8] L. Niederman, Nonlinear stability around an elliptic equilibrium point in a Hamiltonian system. *Nonlinearity* (to appear).

[9] J. Pöschel, Nekhoroshev estimates for quasi-convex hamiltonian systems. *Math. Z.* **213** (1993), 187–216.

Mathematisches Institut A, Universität Stuttgart, Pfaffenwaldring 57, D-70569 Stuttgart
poschel@mathematik.uni-stuttgart.de