Nekhoroshev type stability results for Hamiltonian systems with an additional transversal component

Markus Kunze[1] & David M.A. Stuart[2]

[1] Mathematisches Institut, Universität Köln, Immermannstr. 49-51, D - 50931 Köln, Germany

[2] Centre for Mathematical Sciences, Wilberforce Road, Cambridge, CB3 OWA, England

Key words: Hamiltonian systems, Nekhoroshev stability, normal forms

Abstract

We prove exponential stability theorems of Nekhoroshev type for motion in the neighbourhood of an elliptic fixed point in Hamiltonian systems having an additional transverse component of arbitrary dimension, under certain conditions on this transverse component. We consider both the cases (i) of a strongly constrained motion, and (ii) of a weak perturbation.

1 Introduction and statement of main results

An integrable Hamiltonian, written in action angle variables $(I,\phi)=(I_1,\dots I_n,\phi_1,\dots\phi_n)\in\mathbb{R}^n\times(\mathbb{R}/\mathbb{Z})^n$, takes the form $H_{\rm int}(I_1,\dots I_n)$, and the corresponding equations of motion imply that the action variables I_j are constant while the angle variables ϕ_j evolve at the uniform rate $\frac{\partial H_{\rm int}}{\partial I_j}$. For a nonintegrable perturbation of such a system, described by a smooth Hamiltonian of the form

$$H_{\rm int}(I_1,\dots I_n)+\varepsilon H_{\rm pert}(I_1,\dots I_n,\phi_1,\dots\phi_n)\,,$$

Nekhoroshev proved the following *exponential stability* estimate in [7]: let $H_{\rm int}$ satisfy a condition known as steepness, then there exist positive numbers $R_0,T_0,\varepsilon_0,a,b$ such that for all small ε

$$|I(t)-I(0)|\le R_0\varepsilon^b\quad\text{for}\quad |t|\le T_0e^{(\frac{\varepsilon_0}{\varepsilon})^a}\,. \tag{1.1}$$

This says that for small ε the action variables are almost, or effectively, constant since they vary little over exponentially long time scales. In fact the main theorem in [7, §4.4] proves exponential stability bounds for slightly more general perturbations

$$H_{\rm pert}=H_{\rm pert}(I_1,\dots I_n,\phi_1,\dots\phi_n,\xi_1,\dots\xi_N,\eta_1,\dots\eta_N)$$

in which there is dependence upon an additional set of N (Darboux) conjugate pairs (ξ_j,η_j); we shall refer to these extra variables as the *transverse component.*

It was also conjectured in [7] that, under appropriate conditions, such exponential stability should hold in a sufficiently small neighbourhood of an elliptic equilibrium point. Following a preliminary result in [5, §IV.2, Theorem 4] this was proved in [3, 9] and then in [10] under convexity hypotheses which can be described as follows: let

- $\{(x_j,y_j)\}_{j=1}^n$ be Darboux coordinates on $\mathbb{R}^{2n}$, and define $I_j=(x_j^2+y_j^2)/2$,
- $\alpha\in\mathbb{R}^n$, and let A be a *strictly positive* $n\times n$ matrix,
- f be a real analytic function vanishing to fifth order at the origin,

then the dynamics in a neighbourhood of the origin in $\mathbb{R}^{2n}$ for the Hamiltonian

$$H_0=\langle\alpha,I\rangle+\frac{1}{2}\langle AI,I\rangle+f$$

satisfies exponential stability estimates; see theorem 5.5 for a precise statement.

In view of the above it is to be expected that exponential stability for $\{I_j\}_{j=1}^n$ might continue to hold in a neighbourhood of an elliptic fixed point, under perturbations depending also on an additional transverse component. In this paper we study this situation in detail. We shall consider perturbations in which the additional transverse variable $\zeta = (\xi, \eta) \in \mathbb{R}^{2N}$, while the original phase space $\mathbb{R}^{2n} \ni z = (x, y)$ is a symplectic subspace of the new enlarged phase space $\mathbb{R}^{2n} \times \mathbb{R}^{2N}$. We shall deal with Hamiltonians coupling z and ζ of the form[1]

$$H = H_0 + \sigma\Lambda, \quad \Lambda = O(|\zeta|^2) \quad \text{as} \quad \zeta \to 0 \tag{1.2}$$

and ask the question: under which conditions does Nekhoroshev exponential stability hold for $z \in \mathbb{R}^{2n}$ in a neighbourhood of the origin? We shall consider (1.2) in both limits $\sigma \nearrow +\infty$ and $\sigma \searrow 0$.

Counterintuitively perhaps, the case $\sigma \nearrow +\infty$ can sometimes be regarded as a perturbation of a Hamiltonian flow on $\mathbb{R}^{2n}$, as is discussed in section 3. To be precise, this is the case when Λ is such as to force the flow onto the $\mathbb{R}^{2n} \times \{0\}$ subspace for large σ, on which subspace the dynamics is governed by the restricted Hamiltonian $H_0(z) = H(z, 0)$, that is, motion in a constraining potential. (To ensure this, it will be required that $\Lambda = 0$ if and only if $\zeta = 0$; see section 3 for the precise conditions). In this case we have the main theorem 3.4, which can be stated heuristically as:

> *Exponential stability estimates like* (1.1) *continue to hold for the* $\mathbb{R}^{2n}$ *projection of the flow in* $\mathbb{R}^{2n} \times \mathbb{R}^{2N}$ *determined by the Hamiltonian* $H = H_0 + \sigma\Lambda$*, with* Λ *a constraining potential, when* σ *is sufficiently large (independent of* N*).*

Such results in general come only with the assurance that they hold for sufficiently large σ, but without precise quantitative information on their domain of validity (at least in the absence of more special assumptions). In fact theorem 3.4 is obtained by combining general compactness results based on the Arzelà-Ascoli theorem with the standard Nekhoroshev estimate for $H_0(z)$. We provide details, (i) to explain the method in a simple case, (ii) to clarify how quantitative information on the domains can be derived in special cases (theorem 3.9) and (iii) to emphasize the difference with the case $\sigma \searrow 0$ which is treated in section 4. We remark that stronger results could be proved with the assumption that Λ vanishes faster than quadratically as $\zeta \to 0$, but we are not aware of any likely applications in this case. The convergence of the strongly constrained motion to the formal limiting motion has previously been studied in [11, 14]. However, in addition to deriving the exponential estimates in this limit, we give a new short proof of the convergence itself which emphasizes that it is an immediate consequence of "asymptotic non-transfer of energy in the limit" - see (3.6) and its role in the proof of lemma 3.2. Perturbative methods of Nekhoroshev type were applied before in [1, 2] to obtain exponential estimates on the energy exchange between the constraining forces and this limiting equation.

The more obvious perturbative problem in (1.2) arises by considering small σ; here our second main theorem 4.2 can be stated heuristically as:

> *There exists a neighbourhood* $\mathcal{N}$ *of the origin in* $\mathbb{R}^{2n} \times \mathbb{R}^{2N}$ *and* σ_0 *such that for* $0 < \sigma < \sigma_0$ *and initial data in* $\mathcal{N}$ *exponential stability estimates like* (1.1) *hold for the flow projected onto* $\mathbb{R}^{2n}$*. All of the neighbourhoods and estimates can be bounded explicitly and uniformly in* N*.*

(See also theorems 4.1 and 4.6 for alternative formulations.) The proof of theorem 4.2 relies on a normal form lemma 5.4 which involves applying the method of averaging in a way which couples z and ζ. The method of proof for this auxiliary lemma is known in principle, but is nevertheless included for convenience. Some results in a similar direction were obtained before in [8, p. 1713].

Apart from the difficult limit $N \to \infty$ (infinite-dimensional case), another possibility for future work and generalizing our results would consist of trying to relax the differentiability assumptions, as in [4].

2 Some notation

In general we will be concerned with real analytic Hamiltonians $H = H(z, \zeta)$ depending on the variables $(z, \zeta) \in \mathbb{R}^{2n} \times \mathbb{R}^{2N}$ or $(z, \zeta) \in \mathbb{C}^{2n} \times \mathbb{C}^{2N}$. (By real analytic mapping from a complex domain into another complex vector space, we mean a complex analytic mapping which maps real vectors into real vectors.) Denoting $z = (z_1, \dots, z_n)$

[1]It is possible to put the coupling between z and ζ into either f or Λ, and we make different choices depending upon which is most convenient.

for $z_j = (x_j, y_j) \in \mathbb{C}^2$ for $1 \le j \le n$ we write $I_j = (x_j^2 + y_j^2)/2 \in \mathbb{C}$, and also $\zeta_j = (\xi_j, \eta_j) \in \mathbb{C}^2$ for $1 \le j \le N$. Define the domains

$$\mathcal{D}_{a,\,b,\,c} = \{(z,\zeta) \in \mathbb{C}^{2n} \times \mathbb{C}^{2N} : |I - I^0| < a, |z| < b, |\zeta| < c\}$$

for $a, b, c > 0$, where $I^0 \in \mathbb{R}^n$ is given and

$$|I - I^0| = \sum_{j=1}^{n} |I_j - I_j^0|, \quad |z|^2 = \sum_{j=1}^{n} (|x_j|^2 + |y_j|^2) \quad \text{and} \quad |\zeta|^2 = \sum_{j=1}^{N} (|\xi_j|^2 + |\eta_j|^2).$$

The norm of a matrix $A \in \mathbb{R}^{n\times n}$ is the operator norm w.r.t. the l_1-norm $|I| = \sum_{j=1}^n |I_j|$. We always view I as a function of z and note the estimate, with $\tilde z_j = (\tilde x_j, \tilde y_j)$:

$$\begin{aligned} |I(\tilde z) - I(z)| &\le \frac{1}{2}\Big(\sum_{j=1}^{n} |\tilde x_j - x_j|^2 + |\tilde y_j - y_j|^2\Big)^{\frac12}\Big(\sum_{j=1}^{n} |\tilde x_j + x_j|^2 + |\tilde y_j + y_j|^2\Big)^{\frac12} \\ &\le \frac12 |\tilde z - z|\,(|\tilde z|^2 + 2|\tilde z||z| + |z|^2)^{\frac12} = \frac12 |\tilde z - z|\,(|\tilde z| + |z|). \end{aligned} \tag{2.1}$$

The Hamiltonian vector field generated by a function $f = f(z,\zeta)$ is written as X_f, and the associated flow as X_f^t. We shall refer to integral curves of X_f also as integral curves of f when no confusion seems likely. The supremum norm of functions or vector fields on $\mathcal{D}_{a,\,b,\,c}$ is denoted by $|\cdot|_{a,\,b,\,c}$. For $\underline{r} = (r_1, r_2, r_3)$ we will write $\mathcal{D}_{\underline{r}} = \mathcal{D}_{r_1,\,r_2,\,r_3}$ and $|\cdot|_{\underline{r}} = |\cdot|_{r_1,\,r_2,\,r_3}$.

Let Π_1 (resp. Π_2) be the orthogonal projection operator onto the $\mathbb{C}^{2n}$ (resp. $\mathbb{C}^{2N}$) factor of $\mathbb{C}^{2n} \times \mathbb{C}^{2N}$. We will refer to $\zeta = \Pi_2(z,\zeta)$ as the *transverse* component. The symbols $C, C_1, C_2, \dots$ are reserved for constants which are allowed to depend only on n.

In practice, the domains $\mathcal{D}_{\underline{r}}$ will be used when $r_1 \ll r_2$ and when $I_j^0 = I_j(z^0)$ for some $z^0 \in \mathbb{C}^{2n}$ with $|z^0| < r_2$. In this situation the $\mathcal{D}_{\underline{r}}$ are non-empty open neighbourhoods of the sets $\bigcap_{j=1}^n \{z : I_j(z) = I_j^0\}$ on which the "action" variables I_j take particular values. Ignoring the transverse component, these domains may be pictured as very small islands (open neighbourhoods) around $\bigcap_{j=1}^n \{z : I_j = I_j^0\}$, all contained within an overall small (but less small!) neighbourhood $\{z : |z| < r_2\}$.

3 Constrained motion: the case of large σ

In this section, we consider the case in which there is a transverse variable, $\zeta \in \mathbb{R}^{2N}$, which is subject to a strong constraining potential. Precisely, we consider real analytic Hamiltonians of the form

$$H(z,\zeta) = H_0(z) + \sigma\Lambda(z,\zeta)\,, \tag{3.1}$$

in the limit $\sigma \to +\infty$, assuming that

$$\Lambda \ge 0 \qquad \text{and} \qquad \Lambda(z,\zeta) = 0 \ \text{ if and only if } \zeta = 0. \tag{3.2}$$

The idea is that in this limit Λ forces the motion onto the set $\Lambda = 0$, thus dynamically enforcing the constraint $\zeta = 0$. We will work under the assumption that there exist positive numbers c_0, c_1, p such that for real (z,ζ)

$$H(z,\zeta) \ge c_0(|z|^p + \sigma|\zeta|^p) - c_1\,, \tag{3.3}$$

and also that for all $R > 0$ there exists $c_2(R) > 0$ such that

$$\left|\frac{\partial\Lambda}{\partial z}(z,\zeta)\right| \le c_2(R)\Lambda(z,\zeta) \quad \text{for} \quad |z| \le R\,. \tag{3.4}$$

Remark 3.1 *The numbers c_0, c_1, c_2, p, and hence the bound* (3.4)*, are assumed to be independent of σ.*

The first result does not require analyticity:

Lemma 3.2 *Assume that H is a $C^{1,1}$ function (i.e. C^1 with Lipschitz derivative) of the form* (3.1)*, also verifying* (3.2)*-*(3.4)*. Let there be given real initial data $(z^\sigma(0), \zeta^\sigma(0))$ for $\sigma \ge \sigma_0$, such that*

(i) $z^\sigma(0) \to z(0)$ *as* $\sigma \to +\infty$*;*

(ii) $\sigma\Lambda(z^\sigma(0), \zeta^\sigma(0)) \to 0$ *as* $\sigma \to +\infty$*;*

(iii) $\sup_{\sigma \geq \sigma_0} H(z^\sigma(0), \zeta^\sigma(0)) = E < \infty$.

Then there exist, for each $\sigma \geq \sigma_0$*, global integral curves* $(z^\sigma(t), \zeta^\sigma(t))$ *of* H *which have the property that*

$$\lim_{\sigma \to +\infty} \max_{|t| \leq T} \big(|z^\sigma(t) - z(t)| + |\zeta^\sigma(t)| \big) = 0 \tag{3.5}$$

for any $T > 0$*, where* $z(t)$ *is an integral curve of the Hamiltonian* $H(z, 0) = H_0(z)$*. Furthermore*

$$\lim_{\sigma \to +\infty} \max_{|t| \leq T} \Big(|H_0(z^\sigma(t)) - H_0(z^\sigma(0))| + \sigma\, \Lambda(z^\sigma(t), \zeta^\sigma(t)) \Big) = 0\,. \tag{3.6}$$

Proof of lemma 3.2 The co-ercivity in (3.3) together with energy conservation implies the bound

$$\sigma |\zeta^\sigma(t)|^p + |z^\sigma(t)|^p \leq \frac{c_1 + E}{c_0}\,, \tag{3.7}$$

which is uniform in σ, and shows that $\zeta^\sigma(t) = O(\sigma^{-\frac{1}{p}})$, uniformly in t. To obtain compactness for $z^\sigma(t)$ we use the z component of the differential equation, i.e.

$$\frac{d}{dt}\, z^\sigma = \Pi_1\, X_H(z^\sigma, \zeta^\sigma)\,,$$

conservation of energy, (3.4) and (3.7) to deduce that $\dot z^\sigma(t)$ is bounded, uniformly in t and $\sigma \geq \sigma_0$. It follows from the Arzelà-Ascoli theorem that there exists a subsequence converging uniformly on bounded intervals $[-T, T]$ to a continuous limit $z = z(t)$. To prove that this limit is an integral curve of H_0 we consider the integrated form of the equation:

$$z^\sigma(t) = z^\sigma(0) + \int_0^t \Pi_1\, X_H(z^\sigma(s), \zeta^\sigma(s))\, ds = \int_0^t \Big[X_{H_0}(z^\sigma(s)) + \sigma\, \Pi_1\, X_\Lambda(z^\sigma(s), \zeta^\sigma(s)) \Big]\, ds\,. \tag{3.8}$$

Notice first that it is possible to take the limit of this equation once we know (3.6) holds, on account of (3.4). So we first prove (3.6). Energy conservation $H_0(z^\sigma(t)) + \sigma\Lambda(z^\sigma(t), \zeta^\sigma(t)) = H_0(z^\sigma(0)) + \sigma\Lambda(z^\sigma(0), \zeta^\sigma(0))$ and the assumptions (i), (ii) imply that subsequentially $\lim_{\sigma \to +\infty} \sup_{|t| \leq T} \sigma\Lambda(z^\sigma(t), \zeta^\sigma(t))$ exists for all $T > 0$, and:

$$\begin{aligned} Q(T) \;&:=\; \lim_{\sigma \to +\infty} \sup_{|t| \leq T} \sigma\Lambda(z^\sigma(t), \zeta^\sigma(t)) \\ &=\; \lim_{\sigma \to +\infty} \sup_{|t| \leq T} \Big[H_0(z^\sigma(0)) - H_0(z^\sigma(t)) \Big] = \sup_{|t| \leq T} \Big[H_0(z(0)) - H_0(z(t)) \Big]\,. \end{aligned}$$

On the other hand, the equation of motion and (3.4) imply that

$$\begin{aligned} \Big| H_0(z^\sigma(t)) - H_0(z^\sigma(0)) \Big| \;&=\; \Big| \int_0^t \frac{d}{ds} \left[H_0(z^\sigma(s)) \right] ds \Big| = \Big| \int_0^t \langle DH_0(z^\sigma(s)), \dot z^\sigma(s) \rangle\, ds \Big| \\ &=\; \Big| \sigma \int_0^t \langle DH_0(z^\sigma(s)), \Pi_1\, X_\Lambda(z^\sigma(s), \zeta^\sigma(s)) \rangle\, ds \Big| \\ &\leq\; c_* \int_0^t \sup_{|s'| \leq s} \sigma\Lambda(z^\sigma(s'), \zeta^\sigma(s'))\, ds\,, \end{aligned}$$

with c_* uniform as $\sigma \to +\infty$. Now, defining $Q^\sigma(T) := \sup_{|t| \leq T} \sigma\Lambda(z^\sigma(t), \zeta^\sigma(t))$, we deduce from this and energy conservation that

$$Q^\sigma(t) \leq Q^\sigma(0) + c_* \int_0^t Q^\sigma(s)\, ds\,. \tag{3.9}$$

It follows from the Gronwall inequality, that $Q^\sigma(T) \leq Q^\sigma(0) e^{c_* T}$. Taking the limit $\sigma \to +\infty$ we obtain $Q(T) \leq c_* \int_0^T Q(s)\, ds$, and, using $Q(0) = 0$, we deduce that $Q(T) = 0$ for all $T \geq 0$, which implies the validity of (3.6). It then follows from (3.4), (3.8) and assumption (i) that

$$z(t) = z(0) + \int_0^t \Pi_1 X_{H_0}(z(s), 0)\, ds, \tag{3.10}$$

i.e. the curve $t \mapsto z(t)$ is the integral curve of the Hamiltonian $H(z,0) = H_0(z)$ starting at $z(0)$, which is unique since H_0 defines a Lipschitz continuous Hamiltonian vector field by assumption. It follows from the uniqueness of this limit curve that all subsequences have a subsequence which converges to the same limit, and hence that $(z^\sigma(t), \zeta^\sigma(t))$ converges to $(z(t), 0)$ without recourse to subsequences, as asserted in the lemma. $\square$

Remarks 3.3 *(a) The conclusion* (3.5) *says in words that in the limit the curve is constrained to lie on the* $\zeta = 0$ *subspace, while* (3.6) *says in words that in the limit all the energy is in the* z *variable, and this variable evolves in a way that conserves* $H_0(z)$ *- this evolution is in fact the Hamiltonian evolution determined by* $H_0(z)$*; see (3.10).*

(b) Clearly the conditions on H, H_0, Λ *only need to hold on some open set containing the region defined in* (3.7). *Also, in* (3.3) *the function* $c_0|\cdot|^p$ *could be replaced by any function tending to* $+\infty$ *at* ∞.

We now consider the dynamics for (3.1) under the situation described in lemma 3.2 and assuming in addition that the initial values $z(0) = (x(0), y(0)) \in \mathbb{R}^{2n}$ are close to the equilibrium point $(0,0)$ for the real analytic Hamiltonian. In that case it is a consequence of lemma 3.2 that for sufficiently large σ properties of the integral curves of H_0 will be inherited by those of H on finite time intervals. Theorem 3.4 expresses this fact for the particular case of the Nekhoroshev exponential stability bound (which is recalled for H_0 in theorem 5.5).

Theorem 3.4 *Let* H *be a real analytic function of the form*

$$H(z,\zeta) = H_0(z) + \sigma\Lambda(z,\zeta), \quad \textit{with} \quad H_0(z) = \langle \alpha, I(z) \rangle + \frac{1}{2}\langle AI(z), I(z) \rangle + f(z), \quad \alpha \in \mathbb{R}^n \setminus \{0\}, \tag{3.11}$$

such that $\langle AI, I \rangle \geq \frac{1}{M}|I|^2$ *and* f *is real analytic so that* $f(z) = \mathcal{O}(z^5)$ *as* $|z| \to 0$*, and also verifying* (3.2)-(3.4). *Fix* $a \in]0, \frac{1}{1+3n}[$. *Then there exist positive numbers* K, k *(depending on* a, n, α, M *and* $\|A\|$*) and* θ_0 *(depending on* a, n, α, M, $\|A\|$ *and* f*) with the following properties. If* $t \mapsto (z^\sigma(t), \zeta^\sigma(t))$ *is an integral curve of* H *which verifies the conditions in lemma 3.2 for large* $\sigma > 0$*, and if* $\mathbf{I}(0) = I(z^\sigma(0))$ *is such that* $|\mathbf{I}(0)| = \theta^2$ *for some* $0 < \theta \leq \theta_0$*, then* $\mathbf{I}(t) = I(z^\sigma(t))$ *satisfies*

$$|\mathbf{I}(t) - \mathbf{I}(0)| \leq K\theta^{2+a} \quad \textit{for} \quad |t| \leq e^{\frac{k}{\theta^a}}, \tag{3.12}$$

for sufficiently large $\sigma \geq \sigma_0$ *(depending on the initial conditions and* θ*).*

Remark 3.5 *The proof here depends upon the compactness method used to prove lemma 3.2, and does not provide very explicit or optimal information on the number* σ_0 *above which a bound like (3.12) holds - rather it proves that the bound will hold eventually, along any sequence of integral curves satisfying the conditions in lemma 3.2. In this regard, compare with the statements of theorem 4.2 and 4.6.*

However, more explicit quantitative information on domains of validity for Nekhoroshev bounds can be extracted under additional structural hypotheses as explained in theorem 3.9. It is to derive this, as well as for expository purposes, that we now present a proof of the bound (3.12) *as a large* σ *stability estimate for the integral curves of* H*, rather than attempt to read it off from the* $\sigma = \infty$ *case.*

Remark 3.6 *To match the notation in lemma 3.2, the integral curve for* H *is written as* $t \mapsto (z^\sigma(t), \zeta^\sigma(t))$*, but in the proof we drop the additional* σ *superscript to simplify the notation.*

Beginning of proof of theorem 3.4 Following [5] this will be deduced from three facts:

(a) periodic orbits are dense in a neighbourhood of the fixed point,

(b) motion in a neighbourhood of a periodic orbit satisfies long-time stability estimates, on account of the normal form lemma 5.4, and

(c) a priori control of the effect of the transverse component ζ is provided by (3.6).

To begin with, since $z(0) \in \mathbb{R}^{2n}$ has real components, note that $|z(0)|^2 = \sum_{j=1}^n |z_j(0)|^2 = 2\sum_{j=1}^n |\mathbf{I}_j(0)| = 2|\mathbf{I}(0)| = 2\theta^2$. In what follows the parameter θ will be used as a book-keeping device, i.e. all quantities which need to be controlled will be controlled in terms of θ. We are going to apply the normalization lemma 5.4 to $H_0 = H_0(z)$, i.e. averaging will be performed in the z variable only. Therefore we make the following modification of the notation defined in section 2:

Throughout this proof only, we write $\mathcal{D}_{\underline{r}} = \mathcal{D}_{r_1,r_2}$ and $\mathcal{D}_{r_1,r_2} = \{z \in \mathbb{C}^{2n} : |I - I^0| < r_1, |z| < r_2\}$ and drop the third component from the definition of the corresponding norms $|\cdot|_{\underline{r}}$.

First we apply corollary 5.8 with I replaced by $\mathbf{I}(0)$ and $g = 0$ in (5.26) below. Then $\Omega(I) = \alpha + AI$ and there exist $K_1 > 0$ (depending on α and A) and $\theta_1 > 0$ (depending on α, A, a and n) such that the following holds. If $|\mathbf{I}(0)| = \theta^2$ for some $0 < \theta \leq \theta_1$, then there are $I^0 \in \mathbb{R}^n$ and $\tau > 0$ satisfying

(i) $|\mathbf{I}(0) - I^0|_\infty \leq K_1 \frac{\theta^{2+a}}{\tau}$, and

(ii) $\pi \leq \tau \leq 4\pi\theta^{-a(n-1)}$

and such that $\omega^0 = \alpha + AI^0$ is τ/θ^2-periodic, i.e. $T\omega^0 \in 2\pi\mathbb{Z}^n$ for $T = \tau/\theta^2$. We will call this orbit the *approximating periodic orbit*. Up to a constant, which does not affect the flow, we rewrite H_0 from (3.11) as

$$H_0(z) = \langle \omega^0, I \rangle + \frac{1}{2}\langle A(I - I^0), I - I^0 \rangle + f(z). \tag{3.13}$$

We will now apply the following result on stability in a neighbourhood of periodic orbits:

Lemma 3.7 (Local Stability) *Consider the Hamiltonian H from (3.11). Assume also that H_0 is written as in* (3.13)*, with f real analytic on an open neighbourhood of $\overline{\mathcal{D}}_{3\underline{r}}$ with $|f|_{3\underline{r}} \leq \varepsilon$ and $r_1, r_2 > 0$ such that*

$$r_1 < \frac{1}{4} r_2^2, \quad \varepsilon M < \frac{r_1^2}{2200}, \quad \textit{and} \quad |I^0| < \frac{r_2^2}{16}. \tag{3.14}$$

Assume further that $\omega^0 \in \mathbb{R}^n$ is such that $T\omega^0 \in 2\pi\mathbb{Z}^n$, and that for some $m \in \mathbb{N}$ and $l_2 > 0$

$$54m\|A\|r_1 T \leq \frac{1}{4}, \quad m^2 \varepsilon T < \frac{l_2 r_1^2}{r_2^2}. \tag{3.15}$$

Let $t \mapsto (z(t), \zeta(t))$ be an integral curve for H whose initial data $(z(0), \zeta(0))$ are such that

$$\sigma\,\Lambda(z(0), \zeta(0)) \leq \frac{r_1^2}{360M}, \quad \textit{and} \quad |\mathbf{I}(0) - I^0| \leq l_1 r_1. \tag{3.16}$$

Then, for l_1, l_2 sufficiently small, there holds

$$|\mathbf{I}(t) - I^0| < r_1 \quad \textit{for} \quad |t| \leq t_*, \tag{3.17}$$

where $t_* > 0$ is any time such that

$$t_* \leq \frac{3 \cdot 2^m r_1}{50|\omega^0| r_2^2} \quad \textit{and} \quad t_* \sigma \max_{|t| \leq t_*} \Lambda(z(t), \zeta(t)) \leq \frac{8\varepsilon}{5 r_2 c_2(r_2) |\omega^0|}, \tag{3.18}$$

with c_2 from (3.4). To be precise, the following choices for l_1, l_2 will suffice:

$$l_1 = \min\left\{\frac{1}{4}, \frac{1}{5\sqrt{M\|A\|}}\right\}, \quad l_2 = \min\left\{\frac{1}{2592}, \frac{1}{120\sqrt{M\|A\|}}\right\}. \tag{3.19}$$

Proof of lemma 3.7 As already stated we apply the normal form lemma 5.4, specialized to the case that there is no ζ dependence, to $H_0 = H_0(z)$ so that all the conditions involving r_3 or σ are to be disregarded, and also $g = 0$ and $\delta = 0$. The conditions in (iv)–(vi) of that lemma are then easily seen to be satisfied as a consequence of (3.14), (3.15) and (3.19). Hence there exists a real analytic symplectic transformation $\Psi : \mathcal{D}_{2\underline{r}} \to \mathcal{D}_{3\underline{r}}$ such that, on $\mathcal{D}_{2\underline{r}}$,

$$\tilde{H}_0(\tilde{z}) := H_0 \circ \Psi(\tilde{z}) = \langle \omega^0, I(\tilde{z}) \rangle + \frac{1}{2}\langle A(I(\tilde{z}) - I^0), I(\tilde{z}) - I^0 \rangle + \hat{g}(\tilde{z}) + \hat{f}(\tilde{z})$$

and with the properties:

(a) $\quad |\Psi - \mathrm{id}|_{2\underline{r}} \leq \dfrac{18\, m r_2}{r_1}\, \varepsilon T$,

(b) $\quad |\hat{g}|_{2\underline{r}} \leq 2\varepsilon$ and $\{\hat{g}, h\} = 0$,

(c) $\quad |\hat{f}|_{2\underline{r}} \leq 2^{-m}\varepsilon.$

The total Hamiltonian is now $\tilde{H}(\tilde{z},\zeta) = \tilde{H}_0(\tilde{z}) + \sigma\tilde{\Lambda}(\tilde{z},\zeta)$ where $\Lambda(\Psi(\tilde{z}),\zeta) = \tilde{\Lambda}(\tilde{z},\zeta)$ defines $\tilde{\Lambda}$. Note that in order to distinguish integral curves of the normal form Hamiltonian $\tilde{H}$ from those of the original Hamiltonian H, its z-variables are marked by a tilde; the relation is $z = \Psi(\tilde{z})$. We first obtain bounds for $\tilde{\mathbf{I}}(t) - I^0 = I(\tilde{z}(t)) - I^0$ for the flow of $\tilde{H}$. We will then show that these imply (3.17) for $\mathbf{I}(t) - I^0 = I(z(t)) - I^0$ with $z(t) = \Psi(\tilde{z}(t))$, using lemma 3.8 below to ensure that $z(t) \in \mathcal{D}_{\underline{r}}$ can indeed be written thus. But for the moment we assume this and consider an integral curve $t \mapsto (\tilde{z}(t),\zeta(t))$ of $X_{\tilde{H}}$ such that $t \mapsto \tilde{z}(t) \in \mathcal{D}_{5\underline{r}/3}$. Since in general $\{G(\tilde{I}),F(\tilde{I})\} = 0$, using (b) we obtain for $\tilde{h}(t) = h(\tilde{z}(t)) = \langle\omega^0, I(\tilde{z}(t))\rangle = \langle\omega^0,\tilde{\mathbf{I}}(t)\rangle$ the relation

$$\frac{d\tilde{h}}{dt} = \langle Dh, X_{\tilde{H}}\rangle = \{h,\tilde{H}\} = \{h,\hat{f}+\sigma\tilde{\Lambda}\} = \langle Dh, X_{\hat{f}}\rangle + \sigma\langle Dh, \Pi_1 X_{\tilde{\Lambda}}\rangle\,.$$

Next observe that

$$\tilde{z} \in \mathcal{D}_{5\underline{r}/3} \quad\text{and}\quad |\tilde{w}-\tilde{z}| \leq \frac{r_1}{10\,r_2} \quad\Longrightarrow\quad \tilde{w} \in \mathcal{D}_{2\underline{r}}, \tag{3.20}$$

since $|\tilde{w}| \leq |\tilde{w}-\tilde{z}| + |\tilde{z}| < r_1/10r_2 + 5r_2/3 < 2r_2$ by the condition $r_1 < r_2^2/4$ in (3.14); using in addition (2.1) we obtain

$$|I(\tilde{w}) - I^0| \leq |I(\tilde{w}) - I(\tilde{z})| + |I(\tilde{z}) - I^0| \leq \frac{1}{2}\left(|\tilde{w}-\tilde{z}| + 2|\tilde{z}|\right)|\tilde{w}-\tilde{z}| + \frac{5r_1}{3} < 2r_1.$$

Thus we can bound by means of Cauchy's estimate, cf. (5.7) below:

$$|X_{\hat{f}}|_{5\underline{r}/3} \leq \frac{10\,r_2|\hat{f}|_{2\underline{r}}}{r_1} \leq \frac{10\,r_2 2^{-m}\varepsilon}{r_1}\,. \tag{3.21}$$

Also by Cauchy's estimate, (3.20) and (a) we obtain

$$|D\Psi - 1|_{5\underline{r}/3} \leq \frac{10\,r_2}{r_1}\,|\Psi - \mathrm{id}|_{2\underline{r}} \leq \frac{180\,mr_2^2}{r_1^2}\,\varepsilon T \leq \frac{180\,m^2r_2^2}{r_1^2}\,\varepsilon T < 180\,l_2 < \frac{1}{2}\,, \tag{3.22}$$

from which we derive the pointwise estimate $|\Pi_1 X_{\tilde{\Lambda}}| \leq |\frac{\partial\Lambda}{\partial z}||D\Psi| \leq 3c_2(r_2)\Lambda/2$, using also (3.4). It follows that, for as long as $\tilde{z}(t)$ remains in $\mathcal{D}_{5\underline{r}/3}$ and $z(t) = \Psi(\tilde{z}(t)) \in \mathcal{D}_{\underline{r}}$,

$$\left|\frac{d\tilde{h}}{dt}(t)\right| = |\langle D\tilde{h}, X_{\hat{f}} + \sigma\Pi_1 X_{\tilde{\Lambda}}\rangle| \leq \frac{5}{3}\,|\omega^0|\,r_2\left(\frac{10\,r_2 2^{-m}\varepsilon}{r_1} + \frac{3\sigma}{2}\,c_2(r_2)\,|\Lambda(z(t),\zeta(t))|\right). \tag{3.23}$$

From the definition of t_* we deduce that, for $|t| \leq t_*$ as in (3.17)-(3.18),

$$|\tilde{h}(t) - \tilde{h}(0)| \leq \varepsilon + \frac{5\sigma}{2}\,r_2c_2(r_2)|\omega^0|\,|t|\max_{|t'|\leq|t|}\Lambda(z(t'),\zeta(t')) \leq 5\varepsilon\,.$$

Energy conservation $\tilde{H}(\tilde{z}(t),\zeta(t)) = \tilde{H}(\tilde{z}(0),\zeta(0))$ together with $\Lambda \geq 0$ from (3.2) and the convexity assumption (strict positivity of the matrix A) then give:

$$\begin{aligned}\frac{1}{2M}\,|\tilde{\mathbf{I}}(t) - I^0|^2 &\leq \frac{1}{2}\,\|A\|\,|\tilde{\mathbf{I}}(0) - I^0|^2 + |\tilde{h}(t) - \tilde{h}(0)| + 2|\hat{g}|_{5\underline{r}/3} + 2|\hat{f}|_{5\underline{r}/3} + \sigma\,\Lambda(z(0),\zeta(0))\\ &\leq \frac{1}{2}\,\|A\|\,|\tilde{\mathbf{I}}(0) - I^0|^2 + 5\varepsilon + 6\varepsilon + \sigma\Lambda(z(0),\zeta(0)),\end{aligned} \tag{3.24}$$

so that due to (3.14)-(3.19) and for $|t| \leq t_*$:

$$\begin{aligned}|\tilde{\mathbf{I}}(t) - I^0|^2 &\leq M\|A\|\,|\tilde{\mathbf{I}}(0) - I^0|^2 + 22\,\varepsilon M + 2M\sigma\Lambda(z(0),\zeta(0))\\ &\leq M\|A\|\,|\tilde{\mathbf{I}}(0) - I^0|^2 + \frac{r_1^2}{100} + 2M\sigma\Lambda(z(0),\zeta(0))\end{aligned} \tag{3.25}$$

for as long as $\tilde{z}(t) \in \mathcal{D}_{5\underline{r}/3}$ and $z(t) \in \mathcal{D}_{\underline{r}}$.

Now to deduce (3.17) it is necessary both to show that (3.25) implies the inequality in (3.17), and also to justify the assumption that $\tilde{z}(t) \in \mathcal{D}_{5\underline{r}/3}$ and $z(t) \in \mathcal{D}_{\underline{r}}$ made above in deriving (3.25). To this end suppose that

$|\mathbf{I}(0) - I^0| \le l_1 r_1$ for an integral curve $t \mapsto (z(t), \zeta(t))$ of the original Hamiltonian vector field X_H. Since we are considering real-valued solutions of the Hamiltonian equations,

$$\begin{aligned} |z(0)|^2 &= 2|\mathbf{I}(0)| \le 2(|\mathbf{I}(0) - I^0| + |I^0|) \\ &\le 2l_1 r_1 + \frac{1}{8} r_2^2 < \Big(\frac{l_1}{2} + \frac{1}{8}\Big) r_2^2 \le \frac{r_2^2}{4}. \end{aligned}$$

Therefore we have $z(0) \in \mathcal{D}_{\underline{r}/2}$. Denote by $t_0 > 0$ the longest time such that $z(t) \in \mathcal{D}_{\underline{r}}$ for all $|t| \le t_0$.

The point of the following lemma 3.8 is to show that a sufficiently large neighbourhood of the approximating periodic orbit is covered by the transformation Ψ (as a consequence of (a) and the various assumptions on the parameters used). This ensures that stability information just derived for integral curves of the transformed Hamiltonian $\tilde{H}$ will imply stability information for the integral curves of H on a sufficiently large neighbourhood of this periodic orbit. Here we write $\mathcal{D}_{\underline{r}}^{(\mathrm{real})} = \mathcal{D}_{r_1, r_2}^{(\mathrm{real})}$ where

$$\mathcal{D}_{a,\, b}^{(\mathrm{real})} = \{z \in \mathbb{R}^{2n} : |I(z) - I^0| < a, |z| < b\},$$

and similarly we denote

$$B_r^{(\mathrm{real})}(w) = \{z \in \mathbb{R}^{2n} : |z - w| < r\}$$

for $r > 0$ and $w \in \mathbb{R}^{2n}$.

Lemma 3.8 *Under the hypotheses of lemma 3.7, Ψ satisfies $\Psi(\mathcal{D}_{5\underline{r}/3}^{(\mathrm{real})}) \supset \mathcal{D}_{\underline{r}}^{(\mathrm{real})}$.*

Proof of lemma 3.8 According to (a) and (3.22) we have $|\Psi - \mathrm{id}|_{2\underline{r}} \le \frac{18\, m r_2}{r_1}\, \varepsilon T =: \mu$ and $|D\Psi - 1|_{5\underline{r}/3} < 1/2$. Hence $D\Psi(z)$ is invertible for every $z \in \mathcal{D}_{5\underline{r}/3}$, and accordingly $\Psi : \mathcal{D}_{5\underline{r}/3} \to \Psi(\mathcal{D}_{5\underline{r}/3}) =: \mathcal{W}$ is a real-analytic diffeomorphism such that $\|D\Psi^{-1}(w)\| \le 2$ for $w \in \mathcal{W}$. Now fix $w \in \mathcal{D}_{\underline{r}}^{(\mathrm{real})}$. Then $B_\delta^{(\mathrm{real})}(w) \subset \mathcal{D}_{3\underline{r}/2}^{(\mathrm{real})}$ for $\delta = \frac{r_1}{4\, r_2}$, as can be shown using $r_1 < r_2^2/4$ and (2.1), analogously to (3.20). Furthermore, for $w \in \mathcal{D}_{\underline{r}}^{(\mathrm{real})}$,

$$|w - \Psi(w)| \le \mu < \frac{\delta}{2}$$

due to $\frac{18\, m r_2}{r_1}\, \varepsilon T \le \frac{4 \cdot 18\, m^2 r_2^2}{r_1^2}\, \varepsilon T \times \frac{r_1}{4 r_2} \le 72\, l_2 \delta < \frac{\delta}{2}$. In other words, we have $w \in B_{\delta/2}^{(\mathrm{real})}(\Psi(w))$. Next we apply lemma 5.9 below and use the fact that Ψ is real on real vectors, to deduce that

$$\Psi\Big(\overline{B_\delta^{(\mathrm{real})}(w)}\Big) \supset \overline{B_{\delta/2}^{(\mathrm{real})}(\Psi(w))}.$$

To summarize, for fixed $w \in \mathcal{D}_{\underline{r}}^{(\mathrm{real})}$ we obtain

$$w \in B_{\delta/2}^{(\mathrm{real})}(\Psi(w)) \subset \Psi\Big(\overline{B_\delta^{(\mathrm{real})}(w)}\Big) \subset \Psi\Big(\overline{\mathcal{D}_{3\underline{r}/2}^{(\mathrm{real})}}\Big) \subset \Psi(\mathcal{D}_{5\underline{r}/3}^{(\mathrm{real})}),$$

and this concludes the proof of lemma 3.8. □

Continuation of the proof of lemma 3.7 Due to lemma 3.8 we may write $z(t) = \Psi(\tilde{z}(t))$ for $|t| \le t_0$ with an integral curve $t \mapsto (\tilde{z}(t), \zeta(t))$ of $X_{\tilde{H}}$ such that $t \mapsto \tilde{z}(t) \in \mathcal{D}_{5\underline{r}/3}^{(\mathrm{real})}$. Then by (a) and (3.14)-(3.19),

$$\begin{aligned} |\tilde{\mathbf{I}}(0) - I^0| &\le |\tilde{\mathbf{I}}(0) - \mathbf{I}(0)| + |\mathbf{I}(0) - I^0| \le \frac{1}{2}\Big(|\tilde{z}(0)| + |z(0)|\Big)\, |\tilde{z}(0) - \Psi(\tilde{z}(0))| + l_1 r_1 \\ &\le \frac{1}{2}\Big(\frac{5r_2}{3} + r_2\Big)\, \frac{18\, m r_2}{r_1}\, \varepsilon T + l_1 r_1 \le \frac{24\, m^2 r_2^2}{r_1}\, \varepsilon T + l_1 r_1 \le (24\, l_2 + l_1)\, r_1. \end{aligned} \tag{3.26}$$

Then we can apply (3.25) and use $t_* \le T$ together with (3.16) to obtain

$$\begin{aligned} |\tilde{\mathbf{I}}(t) - I^0|^2 &\le M\|A\|\, |\tilde{\mathbf{I}}(0) - I^0|^2 + \frac{r_1^2}{100} + 2M\sigma\, \Lambda(z(0), \zeta(0)) \\ &\le \Big(M\|A\|\, (24\, l_2 + l_1)^2 + \frac{1}{100}\Big)\, r_1^2 + \frac{r_1^2}{180} \\ &\le \Big(\frac{4}{100} + \frac{1}{100}\Big)\, r_1^2 + \frac{r_1^2}{180} \\ &= \frac{r_1^2}{20} + \frac{r_1^2}{180} < \Big(\frac{r_1}{4}\Big)^2 \end{aligned} \tag{3.27}$$

for $|t| \le \min\{t_*, t_0\}$, since l_2, l_1 are such that $M\|A\|\,(24\,l_2 + l_1)^2 \le \frac{4}{100} = (\frac{1}{5})^2$. In the same manner as for (3.26) this in turn leads to

$$
\begin{aligned}
|\mathbf{I}(t) - I^0| &\le |\mathbf{I}(t) - \tilde{\mathbf{I}}(t)| + |\tilde{\mathbf{I}}(t) - I^0| \\
&\le \frac{1}{2}\Big(|\tilde{z}(t)| + |z(t)|\Big)\,|\tilde{z}(t) - \Psi(\tilde{z}(t))| + \frac{r_1}{4} \\
&\le \Big(24\,l_2 + \frac{1}{4}\Big) r_1 < \frac{r_1}{2}
\end{aligned} \tag{3.28}
$$

for $|t| \le \min\{t_*, t_0\}$, due to $24 l_2 < \frac{1}{4}$. Since also $r_1 < r_2^2/4$, this implies that for such times

$$
|z(t)|^2 = 2|\mathbf{I}(t)| \le 2\Big(|\mathbf{I}(t) - I^0| + |I^0|\Big) \le 2\Big[\frac{r_2^2}{8} + \frac{r_2^2}{16}\Big] < r_2^2\,.
$$

Hence we see that $\min\{t_*, t_0\} < t_0$, or in other words $\min\{t_*, t_0\} = t_*$. Thus (3.17) is a consequence of (3.28). □

Completion of proof of theorem 3.4 We now aim to show that the stability bound (3.17), applied in the neighbourhood of the approximating periodic orbit obtained prior to lemma 3.7, implies (3.12). Since f vanishes to fifth order we take $r_2 = 8\theta$ and $\varepsilon = C_1\theta^5$ to ensure that $|f|_{3\underline{r}} \le \sup\{|f(z)| : |z| \le 3r_2\} \le C_0(3r_2)^5 \le \varepsilon$, where $C_1 = 24^5\,C_0$ has to be chosen large enough (depending on f). In addition, let

$$
m = [\delta\,\theta^{-a}] \in \mathbb{N} \quad \text{and} \quad r_1 = \frac{L\theta^{2+a}}{\tau}, \tag{3.29}
$$

where $\delta, L > 0$ will be fixed below; recall that the period of the approximating periodic orbit is $T = \tau/\theta^2$. We will now verify that having fixed l_1, l_2 satisfying (3.19), the conditions (3.14)-(3.15) can be made to hold by making θ sufficiently small and choosing δ, L appropriately. To start with

$$
\frac{r_1}{r_2^2} = \frac{L\theta^a}{64\tau} \le \frac{L\theta^a}{64\pi}
$$

by (ii), and hence the first condition of (3.14) holds if θ is small enough. In addition,

$$
mr_1T = [\delta\,\theta^{-a}]\,\frac{L\theta^{2+a}}{\tau}\,\frac{\tau}{\theta^2} \le \delta L
$$

whence we need to have

$$
\delta L \le \frac{1}{216\,\|A\|} \tag{3.30}
$$

to validate the first condition of (3.15). Next,

$$
\frac{\varepsilon}{r_1^2} = \frac{C_1\theta^5\tau^2}{L^2\theta^{4+2a}} \le \frac{16\pi^2 C_1}{L^2}\,\theta^{1-2an}
$$

by (ii) shows that we can fulfill the second condition of (3.14) for θ sufficiently small, due to $a < \frac{1}{2n}$. Concerning the condition on $|I^0|$ in (3.14), here

$$
|I^0| \le |\mathbf{I}(0) - I^0| + |\mathbf{I}(0)| \le nK_1\,\frac{\theta^{2+a}}{\tau} + \theta^2 \le \Big(nK_1\,\frac{\theta^a}{\pi} + 1\Big)\theta^2 \le 2\theta^2
$$

by (i) and (ii) for θ small enough. Hence

$$
\frac{|I^0|}{r_2^2} \le \frac{2\theta^2}{64\theta^2} < \frac{1}{16},
$$

and thus all of (3.14) is verified, provided that (3.30) can be ensured. To establish the second condition of (3.16), note that

$$
\frac{|\mathbf{I}(0) - I^0|}{r_1} \le nK_1\,\frac{\theta^{2+a}}{\tau}\,\frac{\tau}{L\theta^{2+a}} = \frac{nK_1}{L}
$$

by (i). Accordingly, we need to have

$$
\frac{nK_1}{L} \le l_1 \tag{3.31}
$$

for l_1 from (3.19). For the last condition of (3.15) finally

$$\frac{m^2\varepsilon T r_2^2}{r_1^2} \leq \frac{\delta^2\theta^{-2a}C_1\theta^5\tau^3\cdot 64\theta^2}{\theta^2 L^2\theta^{4+2a}} = \frac{64\, C_1\delta^2}{L^2}\,\theta^{1-4a}\,\tau^3 \leq \frac{4096\,\pi^3 C_1\delta^2}{L^2}\,\theta^{1-a(1+3n)}$$

by (ii). Since $a < \frac{1}{1+3n}$, the right-hand side is smaller than l_2 from (3.19), if θ is sufficiently small. Altogether, (3.14) and (3.15) will be satisfied, provided that (3.30) and (3.31) hold. This can be achieved by explicitly taking

$$L = \frac{nK_1}{l_1} \quad \text{and} \quad \delta = \frac{1}{216\,\|A\|L}.$$

We thus have shown so far that there is $\theta_0 > 0$ (depending on the quantities as stated in the theorem) such that for $0 < \theta \leq \theta_0$ the assumptions (3.14) and (3.15) from lemma 3.7 hold, as does the second condition of (3.16). Now fix $0 < \theta \leq \theta_0$ and put $t_* = \frac{3\cdot 2^m r_1}{50|\omega^0|r_2^2}$ (depending on θ). Then (3.6) from lemma 3.2 ensures that

$$\lim_{\sigma\to+\infty}\max_{|t|\leq t_*} \sigma\,\Lambda(z(t),\zeta(t)) = 0\,;$$

(recall that $(z(t),\zeta(t)) = (z^\sigma(t),\zeta^\sigma(t))$ in the notation of lemma 3.2). To be more explicit, the Gronwall bound from (3.9) implies that:

$$t_*\sigma\max_{|t|\leq t_*}\Lambda(z^\sigma(t),\zeta^\sigma(t)) \leq t_*\,e^{c_* t_*}\,\sigma\Lambda(z^\sigma(0),\zeta^\sigma(0))\,. \tag{3.32}$$

In particular, referring to the hypotheses of lemma 3.2, the first condition of (3.16) and the second condition from (3.18) will be satisfied, if $\sigma \geq \sigma_0$ for an appropriate $\sigma_0 = \sigma_0(\theta) > 0$ depending on the initial data and θ. Therefore lemma 3.7 applies and we deduce from (3.17) that $|\mathbf{I}(t) - I^0| < r_1$ for $|t| \leq t_*$. Now combine this with (i) to bound, for $|t| \leq t_*$,

$$|\mathbf{I}(t) - \mathbf{I}(0)| \leq |\mathbf{I}(t) - I^0| + |I^0 - \mathbf{I}(0)| \leq r_1 + nK_1\,\frac{\theta^{2+a}}{\tau} \leq \frac{2L}{\pi}\,\theta^{2+a} = K\theta^{2+a}\,,$$

where we have defined $K = \frac{2L}{\pi}$. (The penultimate inequality holds since $r_1 = \frac{L\theta^{2+a}}{\tau}$ and $L \geq nK_1$).

It remains to observe that by (ii),

$$t_* = \frac{3\cdot 2^m L\theta^a}{3200\,|\omega^0|\tau} \geq \frac{3\cdot 2^m L\theta^{an}}{12800\,\pi|\omega^0|},$$

so that with $B = \frac{3L}{12800\,\pi|\omega^0|}$ and m as in (3.29) and for θ small enough (reducing θ_0 further if necessary)

$$\ln t_* \geq \ln B + an\ln\theta + (\ln 2)\,[\delta\,\theta^{-a}] \geq k\theta^{-a}$$

for any $k < (\ln 2)\delta$, and in particular for $k = \frac{\ln 2}{2}\,\delta$, completing the proof of (3.12) and the theorem. □

As already remarked, theorem 3.4 does not provide very explicit quantitative information on the domains on which the bound (3.12) holds, only the assurance that it holds for sufficiently large σ along a sequence of solutions verifying the conditions in lemma 3.2 - this situation arises because of the implicit determination of σ_0 via the bound (3.32). However, the situation can improve in particular situations, when the nonlinear interaction has a special structure which allows extraction of more precise information on the domains, as we now explain. We don't have in mind any specific application, but just want to point out an example of how variants of theorem 3.4 might look, and how they are proved. We assume that there are additional smooth functions J_k, $k = 1,\ldots l$, of $\zeta \in \mathbb{R}^{2N}$ which all Poisson commute with Λ:

$$\{J_k,\Lambda\} = 0 \quad \text{for} \quad k = 1,\ldots,l, \tag{3.33a}$$

and with the additional property that for all $R > 0$ there exists $c_3(R) > 0$ such that:

$$\left|\frac{\partial\Lambda}{\partial z}(z,\zeta)\right| \leq c_3(R)\sum_{k=1}^{l}|J_k(\zeta)| \quad \text{for} \quad |z| \leq R\,. \tag{3.33b}$$

Alternatively, we may work drop (3.2) and work under the hypothesis:

$$|\Lambda(z,\zeta)| + \left|\frac{\partial\Lambda}{\partial z}(z,\zeta)\right| \leq c_3(R)\sum_{k=1}^{l}|J_k(\zeta)| \quad \text{for} \quad |z| \leq R\,. \tag{3.33c}$$

Then we have the following quantitative version of theorem 3.4:

Theorem 3.9 *Let H be a real analytic function of the form* (3.11) *such that $\langle AI, I\rangle \geq \frac{1}{M}|I|^2$ and f is real analytic so that $f(z) = \mathcal{O}(z^5)$ as $|z| \to 0$, and also verifying* (3.2)-(3.3) *and* (3.33a)-(3.33b). *Fix $a \in]0, \frac{1}{1+3n}[$. Then there exist positive numbers K, k (depending on a, n, α, M and $\|A\|$) and θ_0 (depending on a, n, α, M, $\|A\|$ and f) with the following properties. If $t \mapsto (z(t), \zeta(t))$ is an integral curve of H and $\mathbf{I}(0) = I(z(0))$ is such that $|\mathbf{I}(0)| = \theta^2$ for some $0 < \theta \leq \theta_0$, then $\mathbf{I}(t) = I(z(t))$ satisfies*

$$|\mathbf{I}(t) - \mathbf{I}(0)| \leq K\theta^{2+a} \quad \textit{for} \quad |t| \leq e^{\frac{k}{\theta^a}}$$

for initial data such that:

$$\sigma \sum_{k=1}^{l} |J_k(\zeta(0))| \leq \frac{\theta^4 e^{-\frac{k}{\theta^a}}}{K_2} \quad \textit{and} \quad \sigma\Lambda(z(0), \zeta(0)) \leq \frac{L^2\theta^{4+2an}}{(4\pi)^2 360M}, \tag{3.34}$$

with K_2 such that:

$$K_2 \geq 5C_*|\omega^0|/C_1, \tag{3.35}$$

where $c_3(8\theta) \leq C_*$ for $\theta \leq 1$ and C_1 is as defined just prior to (3.29).

Alternatively, under the same assumptions except for the replacement of (3.2) *and* (3.33b) *by* (3.33c), *and with the condition* (3.34) *on the initial data replaced by:*

$$\sigma \sum_{k=1}^{l} |J_k(\zeta(0))| \leq \min\left\{\frac{\theta^4 e^{-\frac{k}{\theta^a}}}{K_2}, \frac{L^2\theta^{4+2an}}{2(4\pi)^2 360MC_*}\right\}, \tag{3.36}$$

the same conclusion holds.

Proof The Hamiltonian is the same as in theorem 3.4, and so the proof of theorem 3.9 proceeds almost identically to the proof of theorem 3.4 and lemmas 3.7 and 3.8. There are two points at which the argument if modified:

- firstly, the condition (3.16) required to apply lemma 3.7 is an explicit consequence of (3.29) and the second inequality in (3.34); and
- secondly, to bound $\frac{d\tilde{h}}{dt}$ the estimate (3.23) is now replaced by

  $$\left|\frac{d\tilde{h}}{dt}(t)\right| = |\langle D\tilde{h}, X_{\hat{f}} + \sigma\Pi_1 X_{\tilde{\Lambda}}\rangle| \leq \frac{5}{3}|\omega^0|\, r_2 \left(\frac{10\, r_2 2^{-m}\varepsilon}{r_1} + \frac{3\sigma}{2} c_3(r_2) \sum_{k=1}^{l} |J_k(\zeta(0))|\right).$$

  (The fact that this holds with the J_k evaluated at $\zeta(0)$ is a consequence of the assumption that they Poisson commute with Λ and so are constants of motion.)

Now continuing the proof, to ensure that $|\tilde{h}(t) - \tilde{h}(0)| \leq 5\varepsilon$ for $|t| \leq e^{\frac{k}{\theta^a}}$ we require

$$e^{\frac{k}{\theta^a}} \frac{5\sigma}{2} r_2 c_3(r_2)|\omega^0| \sum_{l=1}^{k} |J_k(\zeta(0))| \leq 4\varepsilon$$

which, using the definitions $r_2 = 8\theta$ and $\varepsilon = C_1\theta^5$ from the paragraph preceding (3.29), is a consequence of (3.34) for sufficiently large K_2 such that (3.35) holds. Now that the inequality $|\tilde{h}(t) - \tilde{h}(0)| \leq 5\varepsilon$ is derived, the remainder of the proof is exactly as that of theorem 3.4

For the proof of the alternative version asserted in the last sentence, notice that (3.33c) and (3.36) imply that

$$\sigma \max_{|t'| \leq |t|} \Lambda(z(t'), \zeta(t')) \leq \frac{L^2\theta^{4+2an}}{2(4\pi)^2 360M},$$

by the conservation of the $\{J_k\}$. From this the same bound (3.23) for $\frac{dh}{dt}$ follows, and from then on the proof differs only at the stage (3.24), at which point the use of $\Lambda \geq 0$ from (3.2) is to be avoided, and replaced by

$$\frac{1}{2M}\left|\tilde{\mathbf{I}}(t) - I^0\right|^2 \leq \frac{1}{2}\|A\|\left|\tilde{\mathbf{I}}(0) - I^0\right|^2 + 5\varepsilon + 6\varepsilon + 2\sigma \max_{|t'| \leq |t|} \Lambda(z(t'), \zeta(t')). \tag{3.37}$$

With the defintions in (3.29) we have $4M\sigma \max_{|t'| \leq |t|} \Lambda(z(t'), \zeta(t')) \leq \frac{r_1^2}{180}$ which implies $|\tilde{\mathbf{I}}(t) - I^0|^2 \leq \left(\frac{r_1}{4}\right)^2$ as in (3.27), and from there on the remainder of the proof is identical to that of theorem 3.4. □

Remark 3.10 *This last quantitative result holds whenever the stated assumptions in* (3.34) *hold, which is in principle more general than the assumption* $\sigma \to +\infty$ *made in theorem 3.4 (the strongly constrained case).*

An example for theorem 3.9: Examples for this theorem arise naturally in systems with a continuous symmetry group with associated conserved quantities $\{J_k\}$, the "Noether charges". A well known example of such a conserved quantity is the total probability in quantum mechanics, $\int |\psi(x)|^2\,dx$, which is the Noether charge associated to global invariance under phase rotation of the wave function, $\psi \mapsto e^{i\theta}\psi$. Coupling z to such a wave function would correspond to an infinite dimensional transverse variable (here ζ is the wave function $\psi(x)$), taking us beyond the scope of the present article. In order to obtain a simple finite dimensional model problem, we can replace the spatial domain by a finite periodic one dimensional lattice: points are labelled by $j = 1, \ldots, N$ and a discrete wave function $\zeta_j = (\xi_j, \eta_j) \in \mathbb{R}^2$ or $\zeta_j = \xi_j + i\eta_j \in \mathbb{C}$ is given as a function on the lattice with periodic boundary conditions $\zeta_{j+N} = \zeta_j$. In this context a natural possible lattice discretization of the quantum mechanical energy is of the form:

$$\Lambda(\zeta) = \sum_j V_j\,|\zeta_j|^2 + \sum_{|j-j'|=1} |\zeta_j - \zeta_{j'}|^2 \,. \tag{3.38}$$

The first term in this expression is the equivalent of the potential energy $\int V(x)|\psi(x)|^2\,dx$ in quantum mechanics, while the second is a discrete kinetic energy (analogous to $\int |\nabla\psi(x)|^2\,dx$ in quantum mechanics). Now consider, for example, the Hamiltonian (3.11) with H_0 as described in the forgoing theorem, and with Λ as in (3.38), but now with $V_j = V_j(z)$. There is an overall global phase invariance $\zeta_j = \xi_j + i\eta_j \mapsto e^{i\theta}(\xi_j + i\eta_j)$ and the associated Noether charge is

$$J = \sum_j |\zeta_j|^2 = \sum_j (\xi_j^2 + \eta_j^2)$$

which is a conserved quantity for the Hamiltonian dynamics determined by $H = H_0 + \sigma\Lambda$. The assumptions (3.33b) and (3.33c) hold with $c_3 \sim \mathit{const.}\,(1 + \max_j |\partial_z V_j|)$ and theorem 3.9 thus applies. The Hamiltonian described in this example might be regarded as a very simple caricature for the description of the interaction of a very large molecule, described by the Hamiltonian H_0, with electrons on a lattice, described by the discrete wave function ζ_j in a tight binding approximation. However we do not claim that this model is realistic: for this to be justified physically a more complicated model would be needed, including at a minimum prefactors in the above expression with dependence on N in order to ensure that the correspondence with the quantum mechanical energy holds in the continuum large N limit. A more complicated analysis would be required for such physically realistic systems, but we included this model just to indicate how the structural feature involving Noether charges could be exploited in principle.

4 Nekhoroshev stability in the case of small σ

In this section, the function $t \mapsto \big(z(t), \zeta(t)\big) \in \mathbb{R}^{2n} \times \mathbb{R}^{2N}$ is an integral curve of the real-analytic Hamiltonian

$$H(z,\zeta) = \langle \alpha, I(z)\rangle + \frac{1}{2}\,\langle AI(z), I(z)\rangle + f_\sigma(z,\zeta) + \sigma\Lambda(\zeta)\,, \quad \alpha \in \mathbb{R}^n \setminus \{0\}\,. \tag{4.1}$$

We will consider the case that f_σ is allowed to depend on σ and satisfies

$$|f_\sigma(z,\zeta)| \le C_0\big(|z|^5 + |\zeta|^2|z|^4 + \sigma|\zeta|^2|z|\big) \tag{4.2}$$

in a sufficiently large neighbourhood of the origin. In addition we will always assume that

$$\langle AI, I\rangle \ge \frac{1}{M}\,|I|^2 \quad \text{and} \quad \Lambda(\zeta) \ge \frac{|\zeta|^2}{2}\,, \tag{4.3}$$

and

$$\Lambda(\zeta) \le \frac{C_\Lambda}{2}\,|\zeta|^2 \quad \text{and} \quad |D\Lambda(\zeta)| \le C_\Lambda|\zeta|. \tag{4.4}$$

(These conditions are all understood to hold on some open set in $\mathbb{R}^{2n} \times \mathbb{R}^{2N}$ in which the integral curve lies. In fact the quadratic growth assumptions on Λ in (4.3)-(4.4) could be replaced by more general assumptions, but in order to avoid the introduction of yet another set of parameters we will stick to the case of quadratic bounds, which is both the most natural and the most important for applications.)

We will prove that exponential stability estimates like (3.12) hold for the projected motion in the z-plane, together with long time bounds for $\zeta(t)$, as long as σ is sufficiently small.

Theorem 4.1 *Let $t \mapsto \big(z(t), \zeta(t)\big)$ be an integral curve of the real-analytic Hamiltonian H verifying (4.1)-(4.4). There exist constants $\sigma_0, k > 0$ and $0 < p_1 < p_2 < 1$, $1 < 2p_2 < q_1$, with the following properties. If $0 < \sigma \le \sigma_0$ and if the initial data are such that $\mathbf{I}(0) = I(z(0))$ and $\mathbf{\Lambda}(0) = \Lambda(\zeta(0))$ satisfy*

$$|\mathbf{I}(0)| = O(\sigma^{p_1}), \quad \sigma\mathbf{\Lambda}(0) = O(\sigma^{q_1}),$$

then the quantities $\mathbf{I}(t) = I(z(t))$ and $\mathbf{\Lambda}(t) = \Lambda(\zeta(t))$ satisfy

$$|\mathbf{I}(t) - \mathbf{I}(0)| = O(\sigma^{p_2}) \quad \text{and} \quad \sigma\mathbf{\Lambda}(t) = O(\sigma^{2p_2}) \quad \text{for} \quad |t| \le e^{\frac{k}{\sigma^{q_2}}},$$

where $q_2 = p_2 - p_1$. All of the exponents and implicit constants are independent of N.

This theorem is a consequence of:

Theorem 4.2 *Let $t \mapsto \big(z(t), \zeta(t)\big)$ be an integral curve of the real-analytic Hamiltonian H verifying (4.1)-(4.4). Fix $a \in]0, \min\{\frac{1}{4(n-1)}, \frac{1}{1+3n}\}[$. Then there exist positive numbers C_E, $\theta_0 < 1$, K, k with the following properties. If the initial data are such that $\mathbf{I}(0) = I(z(0))$ and $\mathbf{\Lambda}(0) = \Lambda(\zeta(0))$ satisfy*

$$|\mathbf{I}(0)| \le \theta^2, \quad \sigma\mathbf{\Lambda}(0) \le C_E \theta^{4+2an} \quad \text{and} \quad \sigma = \theta^{2+2a(2n-1)}, \tag{4.5}$$

for $0 < \theta \le \theta_0$, then $\mathbf{I}(t) = I(z(t))$ and $\mathbf{\Lambda}(t) = \Lambda(\zeta(t))$ satisfy

$$|\mathbf{I}(t) - \mathbf{I}(0)| \le K\theta^{2+a} \quad \text{and} \quad \sigma\mathbf{\Lambda}(t) \le K\theta^{4+2a} \quad \text{for} \quad |t| \le e^{\frac{k}{\theta^a}}. \tag{4.6}$$

The numbers C_E, θ_0, K, k depend on a, n, $\|A\|$, C_0, M, C_Λ, but not on N.

Remarks 4.3 *(a) Theorem 4.1 is a direct consequence of theorem 4.2: it suffices to take $a \in]0, \min\{\frac{1}{4(n-1)}, \frac{1}{1+3n}\}[$ and define:*

$$p_1 = \frac{2}{2 + 2a(2n-1)}, \quad q_1 = \frac{4 + 2an}{2 + 2a(2n-1)},$$
$$p_2 = \frac{2 + a}{2 + 2a(2n-1)}, \quad q_2 = \frac{a}{2 + 2a(2n-1)}.$$

Notice that $a < \frac{1}{1+3n} < \frac{1}{2(n-1)}$ implies that p_2 thus defined satisfies $2p_2 > 1$.

(b) The idea of the proof is to combine Nekhoroshev estimates for the dynamics of z with a priori estimates for ζ, which derive from the lower bound for Λ in (4.3). The crucial point is to understand the flow of energy, and show that it can be controlled on very long time scales. This is achieved via the use of a normal form for the Hamiltonian (lemma 5.4), which is valid in the neighbourhood of a (dense) set of approximating periodic orbits.

(c) Regarding the allowed conditions on f, the crucial thing is that on an appropriate neighbourhood f is bounded by a number ε satisfying the conditions in (4.8) and (4.9) and satisfying the scaling relations in the paragraph "Completion of proof of theorem 4.2", which ensure applicability of the normal form lemma in §5.1.2; validation of these conditions relies upon assumptions on f like (4.2). This latter condition may not be the most general, but we have not been able to include certain terms, in particular those linear in ζ (except in as much as they can be bounded by the terms on the right hand side of (4.2) by e.g. Cauchy-Schwarz). However, such terms can sometimes be removed beforehand, e.g. terms of the form "$z\,\zeta$" can potentially be eliminated by symplectic diagonalization when setting up the problem.

Beginning of proof of theorem 4.2 We follow the same basic strategy as in the proof of theorem 3.4, and start in identical fashion by introducing an approximating periodic orbit by corollary 5.8 (with I replaced by $\mathbf{I}(0)$ and $g = 0$). This provides a frequency vector $\omega^0 = \alpha + AI^0$ which is τ/θ^2-periodic, i.e. $T\omega^0 \in 2\pi\mathbb{Z}^n$ for $T = \tau/\theta^2$, such that:

(i) $\max_{1 \le j \le n} |\mathbf{I}_j(0) - I_j^0| \le C_A \frac{\theta^{2+a}}{\tau}$, and

(ii) $\pi \le \tau \le 4\pi\theta^{-a(n-1)}$.

This is a periodic orbit for the unperturbed z part of the motion. The number C_A is just a bound for the inverse of the map $I \mapsto \alpha + AI$ and depends on M, n. Up to a constant, which does not affect the flow, we rewrite H as:

$$H(z,\zeta) = \langle \omega^0, I\rangle + \frac{1}{2}\langle A(I-I^0), I-I^0\rangle + f_\sigma(z,\zeta) + \sigma\Lambda(\zeta)\,. \tag{4.7}$$

We will now apply the following result on stability in a neighbourhood of periodic orbits, which is the analogue of lemma 3.7:

Lemma 4.4 (Stability in a neighbourhood of a periodic orbit) *Assume that $\omega^0 \in \mathbb{R}^n$ is such that $T\omega^0 \in 2\pi\mathbb{Z}^n$. Consider a Hamiltonian of the form (4.7), verifying (4.3) and (4.4), which is real analytic on an open neighborhood of $\overline{\mathcal{D}}_{3\underline{r}}$ so that $|f_\sigma|_{3\underline{r}} \le \varepsilon$ and with $r_1, r_2, r_3 > 0$ such that*

$$r_1 < \min\left\{\frac{1}{4}\,r_2^2\,,\,2r_2r_3\right\},\quad \varepsilon M < l_0 r_1^2\,,\quad |I^0| < \frac{r_2^2}{16}\,,\quad \textit{and}\quad r_1^2 \le 4\sigma M r_3^2\,, \tag{4.8}$$

for some positive l_0. Assume further m is a positive integer such that

$$54m\|A\|r_1T \le \frac{1}{6}\,,\quad m^2\varepsilon T < \frac{l_2 r_1^2}{r_2^2}\,,\quad \textit{and}\quad C_\Lambda\sigma m T \le \frac{r_1}{54 r_2 r_3}\,. \tag{4.9}$$

Then for initial data satisfying

$$|\mathbf{I}(0) - I^0| \le l_1 r_1,\quad \sigma\mathbf{\Lambda}(0) \le \frac{r_1^2}{200M} \tag{4.10}$$

and with $l_0, l_1, l_2 > 0$ sufficiently small (depending only on M, $\|A\|$)

$$|\mathbf{I}(t) - I^0| \le r_1,\quad \sigma\mathbf{\Lambda}(t) \le \frac{r_1^2}{16M},\quad \textit{and}\quad |\zeta(t)| \le r_3 \quad \textit{for}\quad |t| \le t_* = \frac{3\cdot 2^m r_1}{10|\omega^0| r_2^2}. \tag{4.11}$$

To be specific the following choices for l_0, l_1, l_2 will suffice:

$$l_0 = \frac{1}{2200}\,,\quad l_1 = \min\left\{\frac{1}{4}, \frac{1}{20\sqrt{M\|A\|}}\right\},\quad l_2 = \min\left\{\frac{1}{3888}, \frac{1}{480\sqrt{M\|A\|}}\right\}. \tag{4.12}$$

Proof of lemma 4.4 We apply the normal form lemma 5.4 with g and δ set to zero: the conditions in (iv)–(vi) of that lemma are then easily seen to be satisfied as a consequence of (4.8)-(4.9), with l_2 as in (4.12). Hence there exists a real analytic symplectic transformation $\Psi : \mathcal{D}_{2\underline{r}} \to \mathcal{D}_{3\underline{r}}$, such that on $\mathcal{D}_{2\underline{r}}$,

$$\tilde{H} := H\circ\Psi = \langle \omega^0, I(\tilde{z})\rangle + \frac{1}{2}\langle A(I(\tilde{z}) - I^0), I(\tilde{z}) - I^0\rangle + \hat{g}(\tilde{z},\tilde{\zeta}) + \hat{f}_\sigma(\tilde{z},\tilde{\zeta}) + \sigma\Lambda(\tilde{\zeta})$$

and with the properties:

(a) $\quad |\Psi - \mathrm{id}|_{2\underline{r}} \le \dfrac{18\,m r_2}{r_1}\,\varepsilon T =: \mu,$

(b) $\quad |\hat{g}|_{2\underline{r}} \le 2\varepsilon$ and $\{\hat{g}, h\} = 0$ for $h = \langle \omega^0, I\rangle$,

(c) $\quad |\hat{f}_\sigma|_{2\underline{r}} \le 2^{-m}\varepsilon.$

(Notice that σ is fixed in lemma 5.4, so that lemma can be applied to f_σ depending on σ and yields a new $\hat{f}_\sigma$, also depending on σ, obeying the bound in (c)). The variables in the normal form Hamiltonian $\tilde{H}$ are distinguished by a tilde, and are related to the original variables by $(z,\zeta) = \Psi(\tilde{z},\tilde{\zeta})$. The crucial point is the small rate of change of $\tilde{h}(t) = h(\tilde{z}(t)) = \langle \omega^0, \tilde{\mathbf{I}}(t)\rangle$, where $t \mapsto \big(\tilde{z}(t), \tilde{\zeta}(t)\big) \in \mathcal{D}_{2\underline{r}}$ is an integral curve for $X_{\tilde{H}}$ and we write $\tilde{\mathbf{I}}(t)$ and $\tilde{\mathbf{\Lambda}}(t)$ in place of $I(\tilde{z}(t))$ and $\Lambda(\tilde{\zeta}(t))$, respectively. Calculating the derivative, using (b) and the fact that Λ depends only on the transverse variable $\tilde{\zeta}$, we find:

$$\frac{d\tilde{h}}{dt} = \langle Dh, X_{\tilde{H}}\rangle = \{h, \tilde{H}\} = \{h, \hat{f}_\sigma\} = \langle Dh, X_{\hat{f}_\sigma}\rangle.$$

Since $h(\tilde{z}) = \langle \omega^0, I(\tilde{z})\rangle$ depends only on $\tilde{z}$ this can be estimated using only a bound for $\Pi_1 X_{\hat{f}_\sigma}$ which can be obtained in the same way as (3.20)-(3.21):

$$|\Pi_1 X_{\hat{f}_\sigma}|_{5\underline{r}/3} \le \frac{10\, r_2 |\hat{f}_\sigma|_{2\underline{r}}}{r_1} \le \frac{10\, r_2 2^{-m}\varepsilon}{r_1}$$

for as long as the solution remains in $\mathcal{D}_{5\underline{r}/3}$, during which time:

$$\left|\frac{d\tilde{h}}{dt}\right| = |\langle Dh, X_{\hat{f}_\sigma}\rangle| \le \frac{5}{3}\,|\omega^0|\, r_2\, |X_{\hat{f}_\sigma}|_{5\underline{r}/3} \le \frac{50\,|\omega^0| r_2^2\, 2^{-m}\varepsilon}{3r_1}. \tag{4.13}$$

Energy conservation, the convexity assumption (strict positivity of the matrix A) and the coercivity assumption (4.3) on Λ then imply:

$$\begin{aligned}\frac{1}{2M}\,|\tilde{\mathbf{I}}(t) - I^0|^2 + \frac{\sigma}{2}\,|\tilde{\zeta}(t)|^2 &\le \frac{1}{2}\,\|A\|\,|\tilde{\mathbf{I}}(0) - I^0|^2 + |\tilde{h}(t) - \tilde{h}(0)| + 2|\hat{g}|_{5\underline{r}/3} + 2|\hat{f}_\sigma|_{5\underline{r}/3} + \sigma\tilde{\mathbf{\Lambda}}(0)\\ &\le \frac{1}{2}\,\|A\|\,|\tilde{\mathbf{I}}(0) - I^0|^2 + \frac{50\,|\omega^0| r_2^2\, 2^{-m}\varepsilon}{3r_1}\,|t| + 6\varepsilon + \sigma\tilde{\mathbf{\Lambda}}(0). \end{aligned} \tag{4.14}$$

To go further we must relate the initial data in the original and tilde variables. By (4.10) we know $|\mathbf{I}(0) - I^0| \le l_1 r_1 < r_1/2$, and since we are considering real-valued solutions of the Hamiltonian equations,

$$\begin{aligned}|z(0)|^2 &= 2|\mathbf{I}(0)| \le 2(|\mathbf{I}(0) - I^0| + |I^0|)\\ &\le 2l_1 r_1 + \frac{1}{8}\,r_2^2 < \Big(\frac{l_1}{2} + \frac{1}{8}\Big) r_2^2 \le \frac{r_2^2}{4}\,,\end{aligned}$$

and also $|\zeta(0)|^2 \le 2\mathbf{\Lambda}(0) < (\frac{r_3}{2})^2$ by the final conditions in (4.8) and (4.10) of the lemma and (4.3). Therefore we have

$$(z(0), \zeta(0)) \in \mathcal{D}_{\underline{r}/2}.$$

But (a) and (4.8)-(4.9) then imply that

$$|\tilde{z}(0)| \le |z(0)| + \frac{18 m r_2}{r_1}\,\varepsilon T \le |z(0)| + \frac{18 l_2 r_1}{r_2} \le |z(0)| + \frac{9}{2}\,l_2 r_2,$$

so that for l_2 as in (4.12) we get $|\tilde{z}(0)| \le \frac{5r_2}{3}$. But then, using (a) again,

$$\begin{aligned}|\tilde{\mathbf{I}}(0) - I^0| &\le |\tilde{\mathbf{I}}(0) - \mathbf{I}(0)| + |\mathbf{I}(0) - I^0| \le \frac{1}{2}\,\Big(|\tilde{z}(0)| + |z(0)|\Big)\,|\tilde{z}(0) - \Psi(\tilde{z}(0))| + l_1 r_1\\ &\le \frac{1}{2}\Big(\frac{5r_2}{3} + r_2\Big)\,\frac{18\, m r_2}{r_1}\,\varepsilon T + l_1 r_1 \le \frac{24\, m^2 r_2^2}{r_1}\,\varepsilon T + l_1 r_1 \le (24\, l_2 + l_1)\, r_1\end{aligned} \tag{4.15}$$

by (4.9) and (4.10). Thus restricting $|t|$ as in (4.11) we obtain from (4.14):

$$|\tilde{\mathbf{I}}(t) - I^0|^2 \le M\|A\|\,(24\, l_2 + l_1)^2 r_1^2 + 2M(11\,\varepsilon + \sigma\tilde{\mathbf{\Lambda}}(0)). \tag{4.16}$$

It remains to consider $\tilde{\mathbf{\Lambda}}(0)$. By the fundamental theorem of calculus and the assumption (4.4) on $|D\Lambda|$ we have

$$|\Lambda(\zeta) - \Lambda(\tilde{\zeta})| \le C_\Lambda(|\zeta|\mu + \mu^2),$$

since $|\tilde{\zeta} - \zeta| \le \mu$ by (a). From (4.8) and (4.9) it follows that $\mu < r_3$ and thus

$$\sigma|\mathbf{\Lambda}(0) - \tilde{\mathbf{\Lambda}}(0)| \le 2\sigma C_\Lambda r_3 \mu = \frac{36\sigma C_\Lambda m r_2 r_3 \varepsilon T}{r_1} \le \frac{36\sigma C_\Lambda m r_1 r_2 r_3 l_0 T}{M} \le \frac{36 l_0 r_1^2}{54M} \le \frac{r_1^2}{200M} \tag{4.17}$$

due to $l_0 \le 3/400$. Hence, using also the final condition in (4.10), $\sigma\tilde{\mathbf{\Lambda}}(0) \le \frac{r_1^2}{100M}$ and so by (4.16), since the conditions in (4.12) ensure that $(24\, l_2 + l_1)\sqrt{M\|A\|} \le \frac{1}{10}$,

$$|\tilde{\mathbf{I}}(t) - I^0|^2 \le \Big(\frac{1}{100} + 22\, l_0 + \frac{1}{50}\Big) r_1^2 < \Big(\frac{r_1}{4}\Big)^2. \tag{4.18}$$

Using the final condition in (4.8) we have similarly from (4.14):

$$|\tilde{\zeta}(t)|^2 \leq 2\tilde{\mathbf{\Lambda}}(t) \leq \frac{\|A\|}{\sigma}(24\,l_2+l_1)^2 r_1^2 + \frac{22\,\varepsilon}{\sigma} + 2\tilde{\mathbf{\Lambda}}(0) < \frac{r_1^2}{25\sigma M} < \Big(\frac{r_3}{2}\Big)^2, \tag{4.19}$$

for as long as $\big(\tilde{z}(t),\tilde{\zeta}(t)\big) \in \mathcal{D}_{5\underline{r}/3}$ and with $|t|$ restricted as in (4.11).

Now to deduce (4.11) it is necessary to transfer the information in (4.18)–(4.19) back to bounds on the original variables z, ζ, Λ, and $I - I^0$. So let $t \mapsto (z(t),\zeta(t)) \in \mathbb{R}^{2n}\times\mathbb{R}^{2N}$ be the integral curve of the original Hamiltonian vector field X_H. Since $(z(0),\zeta(0)) \in \mathcal{D}_{\underline{r}/2}$, we can define $t_0>0$ to be the longest time such that $(z(t),\zeta(t)) \in \mathcal{D}_{\underline{r}}$ for all $|t| \leq t_0$, since such a $t_0>0$ exists by continuity. The point of the following lemma 4.5 is to show that a sufficiently large neighbourhood (to be precise $\mathcal{D}_{\underline{r}}$) of the approximating periodic orbit determined by I^0 is covered by the transformation Ψ, as a consequence of (a) and the various assumptions on the parameters used. This ensures that stability information derived for integral curves of the transformed Hamiltonian $\tilde{H}$ does indeed imply stability information for the integral curves of H on a sufficiently large neighbourhood of this periodic orbit. In what follows we write $\mathcal{D}_{\underline{r}}^{(\mathrm{real})} = \mathcal{D}_{\underline{r}} \cap (\mathbb{R}^{2n}\times\mathbb{R}^{2N})$, and similarly we denote

$$B_\delta^{(\mathrm{real})}(w,\eta) = \{(z,\zeta)\in\mathbb{R}^{2n}\times\mathbb{R}^{2N} : |z-w|+|\zeta-\eta|<\delta\}$$

for $\delta>0$ and $(w,\eta)\in\mathbb{R}^{2n}\times\mathbb{R}^{2N}$.

Lemma 4.5 *Under the hypotheses of lemma 4.4, Ψ satisfies $\Psi(\mathcal{D}_{5\underline{r}/3}^{(\mathrm{real})}) \supset \mathcal{D}_{\underline{r}}^{(\mathrm{real})}$.*

Proof of lemma 4.5 According to (a) we have $|\Psi - \mathrm{id}|_{2\underline{r}} \leq \mu = \frac{18\,mr_2}{r_1}\varepsilon T$. Thus from (3.20) in conjunction with Cauchy's estimate and (4.8)–(4.12) we obtain:

$$\begin{aligned}|D\Psi - 1|_{5\underline{r}/3} &\leq \max\Big\{\frac{3}{r_3},\frac{10\,r_2}{r_1}\Big\}|\Psi-\mathrm{id}|_{2\underline{r}}\\ &\leq \frac{180\,mr_2^2}{r_1^2}\varepsilon T \leq \frac{180\,m^2r_2^2}{r_1^2}\varepsilon T < 180\,l_2 < \frac{1}{2}.\end{aligned}$$

Hence $D\Psi(z)$ is invertible for every $z\in\mathcal{D}_{5\underline{r}/3}$, and accordingly $\Psi:\mathcal{D}_{5\underline{r}/3}\to\Psi(\mathcal{D}_{5\underline{r}/3}) =: \mathcal{W}$ is a real-analytic diffeomorphism such that $\|D\Psi^{-1}(w,\eta)\|\leq 2$ for $(w,\eta)\in\mathcal{W}$. Now fix $(w,\eta)\in\mathcal{D}_{\underline{r}}^{(\mathrm{real})}$. Then $B_\delta^{(\mathrm{real})}(w,\eta)\subset\mathcal{D}_{3\underline{r}/2}^{(\mathrm{real})}$ for $\delta=\frac{r_1}{4\,r_2}<\frac{r_3}{2}$, as can be shown using the first condition in (4.8) and (2.1), analogously to (3.20). Furthermore, for $(w,\eta)\in\mathcal{D}_{\underline{r}}^{(\mathrm{real})}$,

$$|(w,\eta)-\Psi(w,\eta)| \leq \mu < \frac{\delta}{2}$$

due to $\frac{18\,mr_2}{r_1}\varepsilon T \leq \frac{4\cdot 18\,m^2r_2^2}{r_1^2}\varepsilon T \times \frac{r_1}{4r_2} \leq 72\,l_2\delta < \frac{\delta}{2}$. In other words, we have $(w,\eta)\in B_{\delta/2}^{(\mathrm{real})}\big(\Psi(w,\eta)\big)$. Next we apply lemma 5.9 below, and use the fact that Ψ is real on real vectors, to deduce that

$$\Psi\Big(\overline{B_\delta^{(\mathrm{real})}(w,\eta)}\Big) \supset \overline{B_{\delta/2}^{(\mathrm{real})}\Big(\Psi(w,\eta)\Big)}.$$

To summarize, for fixed $(w,\eta)\in\mathcal{D}_{\underline{r}}^{(\mathrm{real})}$ we obtain

$$(w,\eta)\in B_{\delta/2}^{(\mathrm{real})}\Big(\Psi(w,\eta)\Big) \subset \Psi\Big(\overline{B_\delta^{(\mathrm{real})}(w,\eta)}\Big) \subset \Psi\Big(\overline{\mathcal{D}_{3\underline{r}/2}^{(\mathrm{real})}}\Big) \subset \Psi(\mathcal{D}_{5\underline{r}/3}^{(\mathrm{real})}),$$

and this concludes the proof of lemma 4.5. □

Continuation of the proof of lemma 4.4 Due to lemma 4.5, and referring to the definition of t_0, we may write $(z(t),\zeta(t)) = \Psi(\tilde{z}(t),\tilde{\zeta}(t))$ for $|t|\leq t_0$ with an integral curve

$$t\mapsto\big(\tilde{z}(t),\tilde{\zeta}(t)\big)\in\mathcal{D}_{5\underline{r}/3}^{(\mathrm{real})}$$

of $X_{\tilde{H}}$. Then we can apply (4.18)-(4.19) to obtain $|\tilde{\mathbf{I}}(t)-I^0|<r_1/4$ and $|\tilde{\zeta}(t)|<r_3/2$ for $|t|\leq\min\{t_*,t_0\}$, where $t_* = \frac{3\cdot 2^m r_1}{10|\omega^0|r_2^2}$. In the same manner as for (4.15) this in turn leads to

$$\begin{aligned}|\mathbf{I}(t)-I^0| &\leq |\mathbf{I}(t)-\tilde{\mathbf{I}}(t)|+|\tilde{\mathbf{I}}(t)-I^0|\\ &\leq \frac{1}{2}\Big(|\tilde{z}(t)|+|z(t)|\Big)\,|\tilde{z}(t)-\Psi(\tilde{z}(t))| + \frac{r_1}{4}\\ &\leq \Big(24\,l_2+\frac{1}{4}\Big)r_1 < \frac{r_1}{2}\end{aligned} \tag{4.20}$$

for $|t| \leq \min\{t_*, t_0\}$ and as $24l_2 < \frac{1}{4}$. Since also $r_1 < r_2^2/4$, this implies that for such times

$$|z(t)|^2 = 2|\mathbf{I}(t)| \leq 2\Big(|\mathbf{I}(t) - I^0| + |I^0|\Big) \leq 2\Big(\frac{r_2^2}{8} + \frac{r_2^2}{16}\Big) < r_2^2. \tag{4.21}$$

Also, as in the derivation of (4.17) and by (4.19), we have for $|t| \leq \min\{t_*, t_0\}$

$$\sigma\mathbf{\Lambda}(t) \leq \sigma|\mathbf{\Lambda}(t) - \tilde{\mathbf{\Lambda}}(t)| + \sigma\tilde{\mathbf{\Lambda}}(t) \leq 2C_\Lambda r_3\sigma\mu + \sigma\tilde{\mathbf{\Lambda}}(t) \leq \frac{r_1^2}{200M} + \frac{r_1^2}{50M} < \frac{r_1^2}{16M}, \tag{4.22}$$

and furthermore by (a),

$$|\zeta(t)| \leq |\tilde{\zeta}(t) - \zeta(t)| + |\tilde{\zeta}(t)| \leq \mu + \frac{r_3}{2} < r_3, \tag{4.23}$$

the latter since $\mu \leq 18l_2r_1/r_2 < 36l_2r_3 < r_3/2$. Altogether from (4.21) and (4.23) we conclude that $\min\{t_*, t_0\} < t_0$, or in other words $\min\{t_*, t_0\} = t_*$, and so the assertions in (4.11) follow as a consequence of (4.20), (4.22), and (4.23). □

Completion of proof of theorem 4.2 We now aim to show that the stability bound (4.11), applied in the neighbourhood of the approximating periodic orbit obtained prior to lemma 4.4, implies (4.6). Recall that the period of the approximating periodic orbit is $T = \tau/\theta^2$, and define $r_2 = 8\theta$ and

$$m = [\delta\,\theta^{-a}], \quad r_1 = \frac{L\theta^{2+a}}{\tau}, \quad \text{and} \quad r_3 = P\theta^{1+2a(1-n)}, \tag{4.24}$$

where δ, L, $P > 0$ will be fixed below. To ensure that $|f_\sigma|_{3\underline{r}} \leq \varepsilon$ define $\varepsilon = C_1\theta^5$, so that

$$|f_\sigma|_{3\underline{r}} \leq \sup\{|f_\sigma(z,\zeta)| : |z| \leq 3r_2, |\zeta| \leq r_3\} \leq C_0\big[(3r_2)^5 + r_3^2(3r_2)^4 + \sigma r_3^2(3r_2)\big] \leq \varepsilon\,,$$

with the choice $C_1 = (24^5 + 24^4P^2 + 24P^2)C_0$; here we have used $\theta < 1$, $\sigma \leq \theta^{2+4a(n-1)}$ due to (4.5), and the restriction $a < \frac{1}{4(n-1)}$ from the beginning of the theorem statement. We will now verify that having defined $l_0, l_1, l_2 > 0$ by (4.12), the conditions (4.8)–(4.10) can be made to hold by making θ sufficiently small and choosing δ, L, P appropriately.

The conditions in (4.8). To start with

$$\frac{r_1}{r_2^2} = \frac{L\theta^a}{64\tau} \leq \frac{L\theta^a}{64\pi} \quad \text{and} \quad \frac{r_1}{r_2r_3} \leq \frac{L\theta^{a(2n-1)}}{8\pi P}$$

by (ii), and hence the first condition of (4.8) holds if θ is small enough (depending upon L, P, a, n). Next,

$$\frac{\varepsilon}{r_1^2} = \frac{C_1\theta^5\tau^2}{L^2\theta^{4+2a}} \leq \frac{16\pi^2C_1}{L^2}\,\theta^{1-2an}$$

by (ii) shows that we can fulfil the second condition of (4.8) for θ sufficiently small (depending upon L, C_0, P, a, n), due to $a < \frac{1}{2n}$. Concerning the condition on $|I^0|$ in (4.8), here

$$|I^0| \leq |\mathbf{I}(0) - I^0| + |\mathbf{I}(0)| \leq nC_A\,\frac{\theta^{2+a}}{\tau} + \theta^2 \leq \Big(nC_A\,\frac{\theta^a}{\pi} + 1\Big)\theta^2 \leq 2\theta^2$$

by (i) and (ii) for θ small enough (depending upon C_A, a, n). Hence

$$\frac{|I^0|}{r_2^2} \leq \frac{2\theta^2}{64\theta^2} < \frac{1}{16}.$$

The final condition in (4.8) reads as

$$1 \geq \frac{r_1^2}{4\sigma Mr_3^2} = \frac{L^2}{4MP^2\tau^2}$$

recall (4.5). Since $\tau \geq \pi$, this follows from

$$\frac{L^2}{P^2} \leq 4\pi^2M. \tag{4.25}$$

The conditions in (4.9). Next

$$mr_1T = [\delta\,\theta^{-a}]\,\frac{L\theta^{2+a}}{\tau}\,\frac{\tau}{\theta^2} \le \delta L$$

whence the restriction

$$\delta L \le \frac{1}{324\,\|A\|} \tag{4.26}$$

is sufficient to validate the first condition of (4.9). For the second condition of (4.9) calculate

$$\frac{m^2\varepsilon T r_2^2}{r_1^2} \le \frac{\delta^2\theta^{-2a}C_1\theta^5\tau^3\times 64\theta^2}{\theta^2L^2\theta^{4+2a}} = \frac{64\,C_1\delta^2}{L^2}\,\theta^{1-4a}\,\tau^3 \le \frac{4096\,\pi^3C_1\delta^2}{L^2}\,\theta^{1-a(1+3n)}$$

by (ii). Since $a < \frac{1}{1+3n}$, the right-hand side is smaller than l_2 from (4.12), if θ is sufficiently small (depending upon δ, L, C_0, P, a, n). The final condition in (4.9) is

$$1 \le \frac{r_1}{54C_\Lambda\sigma mTr_2r_3} = \frac{L\theta^{a(1-2n)}}{432C_\Lambda P\tau^2\,[\delta\theta^{-a}]},$$

which due to (ii) is a consequence of $432C_\Lambda P(4\pi)^2[\delta\theta^{-a}] \le L\theta^{-a}$. This in turn holds if

$$\delta P \le \frac{L}{6912\pi^2C_\Lambda}. \tag{4.27}$$

The conditions in (4.10). The first one holds because

$$\frac{|\mathbf{I}(0)-I^0|}{r_1} \le nC_A\,\frac{\theta^{2+a}}{\tau}\,\frac{\tau}{L\theta^{2+a}} = \frac{nC_A}{L}$$

by (i). Accordingly, we need to have

$$\frac{nC_A}{L} \le l_1 \tag{4.28}$$

for l_1 from (4.12). The second condition holds because $r_1 = \frac{L\theta^{2+a}}{\tau} \ge \frac{L\theta^{2+an}}{4\pi}$ by (ii), so that due to (4.5)

$$\sigma\mathbf{\Lambda}(0) \le C_E\theta^{4+2an} \le \frac{C_E(4\pi)^2r_1^2}{L^2} = \frac{r_1^2}{200M}, \quad \text{taking} \quad C_E = \frac{L^2}{(4\pi)^2200M}.$$

Altogether, the conditions necessary to apply lemma 4.4 will be satisfied provided that the restrictions in (4.25), (4.26), (4.27) and (4.28) hold. The latter can be achieved by explicitly taking

$$L = \frac{nC_A}{l_1}, \quad P = \frac{L}{2\pi\sqrt{M}} \quad \text{and} \quad \delta = \min\left\{\frac{1}{324\,\|A\|L}\,,\,\frac{L}{6912\pi^2C_\Lambda P}\right\}.$$

Therefore lemma 4.4 can be used, and we deduce from (4.11) that

$$|\mathbf{I}(t)-I^0| \le r_1, \quad \sigma\mathbf{\Lambda}(t) \le \frac{r_1^2}{16M}, \quad \text{and} \quad |\zeta(t)| \le r_3 \quad \text{for} \quad |t| \le \frac{3\times 2^m r_1}{10|\omega^0|r_2^2} =: t_*.$$

Now combine the former with (i) to bound

$$|\mathbf{I}(t)-\mathbf{I}(0)| \le |\mathbf{I}(t)-I^0| + |I^0-\mathbf{I}(0)| \le r_1 + nC_A\frac{\theta^{2+a}}{\tau} \le \frac{2L}{\pi}\,\theta^{2+a}$$

(since clearly $L > nC_A$) for $|t| \le t_*$. Thus, recalling that $\tau \ge \pi$ by (ii), we can achieve the bounds in (4.6) with

$$K = \max\left\{\frac{2L}{\pi}, \frac{L^2}{16M\pi^2}\right\},$$

and

$$t_* = \frac{3\times 2^m L\theta^a}{640\,|\omega^0|\tau} \ge \frac{3\times 2^m L\theta^{an}}{2560\,\pi|\omega^0|}\,.$$

It follows that with $B = \frac{3L}{2560\,\pi|\omega^0|}$ and m as in (4.24),

$$\ln t_* = \ln B + an\ln\theta + (\ln 2)\,[\delta\,\theta^{-a}] \geq k\,\theta^{-a}\,,$$

for θ small enough and any $k < (\ln 2)\delta$, thus completing the proof of (4.6). □

In some applications it may be desirable to prove that the stability estimates hold for sufficiently small σ in an open set in $\mathbb{R}^{2n}$ which is essentially determined by the unperturbed z motion. In these circumstances the following variant of theorem 4.2 is natural:

Theorem 4.6 *Let* $t \mapsto \big(z(t), \zeta(t)\big)$ *be an integral curve of the real-analytic Hamiltonian* H *verifying* (4.1), (4.3)-(4.4) *and*

$$|f_\sigma(z,\zeta)| \leq C_0\big(|z|^5 + \sigma|\zeta|^2|z|\big)$$

(in place of (4.2)*). Fix* $a \in]0, \min\{\frac{1}{4(n-1)}, \frac{1}{1+3n}\}[$. *Then there exist positive numbers* C_E, $\theta_0 < 1$, K, k *with the following properties. If the initial data are such that* $\mathbf{I}(0) = I(z(0))$ *and* $\mathbf{\Lambda}(0) = \Lambda(\zeta(0))$ *satisfy*

$$|\mathbf{I}(0)| \leq \theta^2\,, \quad \sigma\mathbf{\Lambda}(0) \leq C_E\theta^{4+2an} \quad \textit{and} \quad 0 < \sigma \leq \theta^{2+2a(2n-1)}\,,$$

for $0 < \theta \leq \theta_0$, *then* $\mathbf{I}(t) = I(z(t))$ *and* $\mathbf{\Lambda}(t) = \Lambda(\zeta(t))$ *satisfy the stability estimate (4.6) as in theorem 4.2.*

Proof of theorem 4.6 Only a small modification of the proof of theorem 4.2 is needed. We start by using periodic approximation and lemma 4.4 in identical fashion, but then in (4.24) replace the definition of r_3 by

$$r_3 = \frac{P\theta^{2+a}}{\sqrt{\sigma}} \tag{4.29}$$

(leaving the definitions of r_1, r_2, m unchanged). When σ is equal to its maximum allowed value, $\theta^{2+2a(2n-1)}$, this reproduces the value of r_3 in (4.24), but as σ gets smaller r_3 defined in (4.29) increases in such a way that σr_3^2 is unchanged. With this understood, it is easy to see that the conditions (4.9)-(4.10) in lemma 4.4 continue to hold with the new definition of r_3, and

$$|f_\sigma|_{3\underline{r}} \leq \sup\{|f_\sigma(z,\zeta)| : |z| \leq 3r_2, |\zeta| \leq r_3\} \leq C_0\big[(3r_2)^5 + \sigma r_3^2(3r_2)\big] \leq \varepsilon\,,$$

so we may set $\varepsilon = C_2\theta^5$, with the choice $C_2 = (24^5 + 24P^2)C_0$. From this, the estimate (4.6) follows as a consequence of lemma 4.4, exactly as in the completion of the proof of theorem 4.2. □

5 Some auxiliary results

In this section we first prove the normal form results used in the main text, then as an aside we recall in section 5.2 the classical Nekhoroshev stability estimate (which corresponds to $N = 0$, i.e. no transversal components), and finally some technical results are collected in section 5.3.

5.1 Transformation to normal form

In order to make this paper self-contained, and for convenience of the reader, we include a full proof of a result concerning a normal form transformation which in principle is known and variants of which could also be found in other works [3, 9, 10]. However, since we needed a precise statement on how certain quantities depend on the assumptions, we feel it is appropriate to include the proof in full. In section 5.1.1 we do a single step of improvement of the interaction term, while the full normal form transformation in section 5.1.2 is then obtained by iteratively applying this single step.

The actual application of the normal form lemma 5.4 to prove stability theorems requires validation of the hypotheses (i)-(vi), which requires assumptions on f such as (4.2). However, the results of this section hold for general analytic f satisfying the hypotheses of the lemmas which follow.

5.1.1 One step improvement of the interaction term

We start with an integrable Hamiltonian $\langle \omega^0, I\rangle + \frac{1}{2}\langle A(I-I^0), I-I^0\rangle$ and a further Hamiltonian $\Lambda(\zeta)$ on $\mathbb{R}^{2N}$. We will introduce a coupling and use the following lemma iteratively to successively reduce the interaction. In the proof we will sometimes abbreviate:

$$h(z) = \langle \omega^0, I(z)\rangle \quad \text{and} \quad g_0(z) = \frac{1}{2}\langle A(I(z)-I^0), I(z)-I^0\rangle. \tag{5.1}$$

Lemma 5.1 (Iteration step) *Consider the Hamiltonian*

$$H(z,\zeta) = \langle \omega^0, I(z)\rangle + \frac{1}{2}\langle A(I-I^0), I-I^0\rangle + g(z,\zeta) + f(z,\zeta) + \sigma\Lambda(\zeta),$$

where $\omega^0, I^0 \in \mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$ is a symmetric matrix and $T, \sigma > 0$ are fixed such that $T\omega^0 \in 2\pi\mathbb{Z}^n$ holds. The functions g and f are assumed to be real analytic on an open set containing $\overline{\mathcal{D}}_{\underline{r}}$ for $\underline{r} = (r_1, r_2, r_3)$ with $r_1, r_2, r_3 > 0$, whereas Λ is assumed to be real analytic on an open set containing $\{|\zeta| \le r_3\}$. We suppose that for some $\delta, \varepsilon > 0$ and some constant $C_\Lambda > 0$,

(i) $|g|_{\underline{r}} \le \delta$ *and* $\{g, h\} = 0$,

(ii) $|f|_{\underline{r}} \le \varepsilon$,

(iii) $|D\Lambda(\zeta)| \le C_\Lambda |\zeta|$ *for* $|\zeta| \le r_3$.

If $\rho_1 \in]0, r_1[$, $\rho_2 \in]0, r_2[$, $\rho_3 \in]0, r_3[$ are such that

$$\varepsilon T < \frac{1}{9}\Big(\min\Big\{\frac{\rho_1}{r_2}, \rho_2, \rho_3\Big\}\Big)^2, \tag{5.2}$$

then there exists a real analytic symplectic transformation

$$\Phi:\ \mathcal{D}_{\underline{r}-\underline{\rho}} \to \mathcal{D}_{\underline{r}}$$

such that, on $\mathcal{D}_{\underline{r}-\underline{\rho}}$,

$$H\circ\Phi = \langle \omega^0, I(z)\rangle + \frac{1}{2}\langle A(I-I^0), I-I^0\rangle + g_+(z,\zeta) + f_+(z,\zeta) + \sigma\Lambda(\zeta) \tag{5.3}$$

and with the properties:

(a) $\quad |\Phi - \mathrm{id}|_{\underline{r}-\underline{\rho}} \le \dfrac{3\,\varepsilon T}{\min\{\frac{\rho_1}{r_2}, \rho_2, \rho_3\}}$,

(b) $\quad |g_+|_{\underline{r}} \le \delta + \varepsilon$ *and* $\{g_+, h\} = 0$,

(c) $\quad |f_+|_{\underline{r}-\underline{\rho}} \le \left[\dfrac{6\|A\| r_1 r_2}{\rho_2} + \dfrac{36\,(\delta+\varepsilon)}{(\min\{\frac{\rho_1}{r_2}, \rho_2, \rho_3\})^2} + \dfrac{3\sigma C_\Lambda r_3}{2\rho_3}\right]\varepsilon T.$

Proof of lemma 5.1 We start by averaging over the flow generated by h: let

$$\bar{f}(z,\zeta) = \frac{1}{T}\int_0^T (f\circ X_h^t)(z,\zeta)\,dt. \tag{5.4}$$

Explicitly,

$$\begin{aligned} X_h^t(z,\zeta) &= (z_1(t), \dots, z_n(t), \zeta), \\ z_j(t) &= R_j(t) z_j, \quad z_j = (x_j, y_j), \\ R_j(t) &= \begin{pmatrix} \cos(\omega_j^0 t) & \sin(\omega_j^0 t) \\ -\sin(\omega_j^0 t) & \cos(\omega_j^0 t) \end{pmatrix}. \end{aligned}$$

Since $T\omega^0 \in 2\pi\mathbb{Z}^n$ we get $R_j(t+T) = R_j(t)$ and the flow X_h^t is T-periodic. In addition the matrices are real and $R_j^T R_j = \mathrm{id}$, so that $|z(t)| = |z(0)|$ and $\mathbf{I}_j(t) = (x_j(t)^2 + y_j(t)^2)/2 = \mathbf{I}_j(0)$. Then X_h^t leaves invariant every domain $\mathcal{D}_{\underline{r}}$; in particular, $\bar{f}$ is well defined on $\mathcal{D}_{\underline{r}}$, the domain of f, and $|\bar{f}|_{\underline{r}} \leq |f|_{\underline{r}} \leq \varepsilon$. Now define

$$\varphi(z,\zeta) = \frac{1}{T}\int_0^T t\,((f-\bar{f})\circ X_h^t)(z,\zeta)\,dt \tag{5.5}$$

which is well defined on $\mathcal{D}_{\underline{r}}$ and satisfies

$$\{\varphi, h\} = f - \bar{f} \quad \text{and} \quad |\varphi|_{\underline{r}} \leq T|f|_{\underline{r}} \leq \varepsilon T. \tag{5.6}$$

[To establish (5.6), we use

$$\begin{aligned}
\frac{d}{ds}\left[s\,(f-\bar{f})\circ X_h^{t+s}\right] &= s\,\frac{d}{ds}\left[(f-\bar{f})\circ X_h^{t+s}\right] + (f-\bar{f})\circ X_h^{t+s} \\
&= s\,\frac{d}{dt}\left[(f-\bar{f})\circ X_h^{t+s}\right] + (f-\bar{f})\circ X_h^{t+s},
\end{aligned}$$

which upon integration $\int_0^T ds$ yields

$$T(f-\bar{f})\circ X_h^{t+T} = \frac{d}{dt}\int_0^T s\,(f-\bar{f})\circ X_h^{t+s}\,ds + \int_0^T (f-\bar{f})\circ X_h^{t+s}\,ds.$$

Therefore

$$\begin{aligned}
\{\varphi, h\} &= \{\varphi, h\}\circ X_h^t\Big|_{t=0} = \frac{d}{dt}\Big(\varphi\circ X_h^t\Big)\Big|_{t=0} = \frac{1}{T}\frac{d}{dt}\Big(\int_0^T s\,(f-\bar{f})\circ X_h^{t+s}\,ds\Big)\Big|_{t=0} \\
&= f - \bar{f} - \frac{1}{T}\int_0^T (f-\bar{f})\circ X_h^s\,ds \\
&= f - \bar{f},
\end{aligned}$$

(as a consequence of $X_h^T = \mathrm{id}$ and the fact that $\bar{f}\circ X_h^s$ is independent of s since

$$\frac{d}{ds}\Big(\bar{f}\circ X_h^s\Big) = \frac{d}{ds}\left(\frac{1}{T}\int_0^T f\circ X_h^{s+t}\,dt\right) = \frac{1}{T}\int_0^T \frac{d}{dt}\,(f\circ X_h^{s+t})\,dt = 0,$$

so that the integral in the penultimate line is zero.) The formula for φ can be estimated in the obvious way given the remarks already made on the action of X_h^t, completing the proof of (5.6).]

Estimates for the derivatives of φ follow from Cauchy's theorem:

$$\left|\frac{\partial\varphi}{\partial z}\right|_{\underline{r}-\underline{\rho}/3} \leq \frac{3|\varphi|_{\underline{r}}}{\min\{\frac{\rho_1}{r_2}, \rho_2\}}, \qquad \left|\frac{\partial\varphi}{\partial\zeta}\right|_{\underline{r}-\underline{\rho}/3} \leq \frac{3|\varphi|_{\underline{r}}}{\rho_3}, \tag{5.7}$$

since $(z,\zeta)\in\mathcal{D}_{\underline{r}-\underline{\rho}/3}$ and $|z-w| \leq \frac{1}{3}\min\{\frac{\rho_1}{r_2}, \rho_2\}$ implies $(w,\zeta)\in\mathcal{D}_{\underline{r}}$. In fact, by (2.1),

$$\begin{aligned}
|I(w) - I^0| &\leq |I(w) - I(z)| + |I(z) - I^0| \leq \frac{1}{2}\,|w-z|(|w-z| + 2|z|) + r_1 - \rho_1/3 \\
&< \frac{\rho_1}{6r_2}\Big(\rho_2/3 + 2(r_2 - \rho_2/3)\Big) + r_1 - \rho_1/3 < r_1.
\end{aligned}$$

This implies bounds for the corresponding Hamiltonian vector field X_φ:

$$\left|\Pi_1 X_\varphi\right|_{\underline{r}-\underline{\rho}/3} \leq \frac{3\,\varepsilon T}{\min\{\frac{\rho_1}{r_2}, \rho_2\}}, \qquad \left|\Pi_2 X_\varphi\right|_{\underline{r}-\underline{\rho}/3} \leq \frac{3\,\varepsilon T}{\rho_3}. \tag{5.8}$$

Remark 5.2 *These Hamiltonian vector fields have, respectively, $2n$ and $2N$ components and the bounds (5.8) hold using the Euclidean norm with respect to these components; see [5, Lemma 1] or [6, Prop. 3 in §6] for an abstract treatment for maps between Banach spaces.*

We now introduce

$$\begin{aligned}
\Phi &= X_\varphi^1, \qquad\qquad \text{(the time one map of the flow of } \varphi) \\
g_+ &= g + \bar{f}, & (5.9) \\
f_+ &= \int_0^1 \{g_0 + g + f_t, \varphi\} \circ X_\varphi^t \, dt \;+\; \sigma\,(\Lambda \circ \Phi - \Lambda), & (5.10)
\end{aligned}$$

where g_0 is as in (5.1) and $f_t = tf + (1-t)\bar{f}$ for $t \in [0,1]$. To verify that (5.3) holds with the properties asserted, observe that

$$\frac{d}{dt}\Big[(g_0 + g + f_t) \circ X_\varphi^t\Big] = \{g_0 + g + f_t, \varphi\} \circ X_\varphi^t + (f - \bar{f}) \circ X_\varphi^t,$$

and consequently

$$(g_0 + g + f) \circ \Phi - (g_0 + g + \bar{f}) = \int_0^1 \{g_0 + g + f_t, \varphi\} \circ X_\varphi^t \, dt + \int_0^1 (f - \bar{f}) \circ X_\varphi^t \, dt. \qquad (5.11)$$

Since

$$\frac{d}{dt}\Big(h \circ X_\varphi^t\Big) = -\,\{\varphi, h\} \circ X_\varphi^t = -\,(f - \bar{f}) \circ X_\varphi^t$$

by (5.6), it follows from (5.11) that

$$(g_0 + g + f + h) \circ \Phi - (g_0 + g + \bar{f} + h) = \int_0^1 \{g_0 + g + f_t, \varphi\} \circ X_\varphi^t \, dt.$$

Thus

$$\begin{aligned}
H \circ \Phi &= (h + g_0 + g + f) \circ \Phi + \sigma\Lambda \circ \Phi \\
&= h + g_0 + g + \bar{f} + \sigma\Lambda + \sigma\,(\Lambda \circ \Phi - \Lambda) \\
&\quad + \int_0^1 \{g_0 + g + f_t, \varphi\} \circ X_\varphi^t \, dt \\
&= h + g_0 + g_+ + \sigma\Lambda + f_+
\end{aligned}$$

which is the form of $H \circ \Phi$ asserted in (5.3), with the functions g_+ and f_+ being defined in (5.9) and (5.10), respectively. To check the estimate (c) in the lemma we split up f_+ as follows:

$$\begin{aligned}
f_+ &= \int_0^1 \{g_0, \varphi\} \circ X_\varphi^t \, dt + \int_0^1 \{g + f_t, \varphi\} \circ X_\varphi^t \, dt + \sigma\,(\Lambda \circ \Phi - \Lambda) \\
&= f_{+,1} + f_{+,2} + f_{+,3}. & (5.12)
\end{aligned}$$

In order to derive the bounds for the $f_{+,j}$ quantities and to justify the preceding calculation we summarize some mapping properties of the flows in the following proposition, thus also establishing statement (a) in the lemma since $\Phi = X_\varphi^1$.

Proposition 5.3 (Mapping properties for the flows X_φ^t and $X_{g_0}^t$) *Under the assumptions of lemma 5.1 the Hamiltonian flows generated by φ and g_0 have the following properties:*

(i) *For real times $|t| \le 1$, the flow X_φ^t satisfies*

$$X_\varphi^t : \mathcal{D}_{\underline{r}-\underline{\rho}} \to \mathcal{D}_{\underline{r}-2\underline{\rho}/3}, \qquad (5.13)$$

$$X_\varphi^t : \mathcal{D}_{\underline{r}-2\underline{\rho}/3} \to \mathcal{D}_{\underline{r}-\underline{\rho}/3} \quad \textit{and} \qquad (5.14)$$

$$|X_\varphi^t - \mathrm{id}|_{\underline{r}-2\underline{\rho}/3} \le |X_\varphi|_{\underline{r}-\underline{\rho}/3}\,|t| \le \frac{3\,\varepsilon T}{\min\{\frac{\rho_1}{r_2}, \rho_2, \rho_3\}}, \qquad (5.15)$$

and for complex times t such that

$$|t| < \lambda \quad \textit{for} \quad \lambda = \frac{1}{18\,\varepsilon T}\left(\min\left\{\frac{\rho_1}{r_2}, \rho_2, \rho_3\right\}\right)^2 \qquad (5.16)$$

the flow X_φ^t is analytic on $\mathcal{D}_{\underline{r}-\underline{\rho}/6}$ and satisfies

$$X_\varphi^t : \mathcal{D}_{\underline{r}-2\underline{\rho}/3} \to \mathcal{D}_{\underline{r}-\underline{\rho}/2} \subset \mathcal{D}_{\underline{r}-\underline{\rho}/3}. \qquad (5.17)$$

(ii) *For complex times t such that*

$$|t| < \tau \quad \text{for} \quad \tau = \frac{\rho_2}{3\|A\|\, r_1 r_2} \tag{5.18}$$

the flow $X^t_{g_0}$ is analytic on $\mathcal{D}_{\underline{r}-\underline{\rho}/3}$ and satisfies

$$X^t_{g_0} : \mathcal{D}_{\underline{r}-2\underline{\rho}/3} \to \mathcal{D}_{\underline{r}-\underline{\rho}/3}, \tag{5.19}$$

$$X^t_{g_0} : \mathcal{D}_{\underline{r}-\underline{\rho}/3} \to \mathcal{D}_{\underline{r}}\,. \tag{5.20}$$

Proof of proposition 5.3 (i) To begin with, the equation $(d/dt)X^t_\varphi = X_\varphi(X^t_\varphi)$ reads

$$(\dot z(t), \dot\zeta(t)) = (\Pi_1 X_\varphi, \Pi_2 X_\varphi)(z(t), \zeta(t))$$

where $(z(t), \zeta(t)) = X^t_\varphi(z(0), \zeta(0))$ for some fixed $(z(0), \zeta(0)) \in \mathcal{D}_{\underline{r}-\underline{\rho}}$. This implies, by (5.8), that

$$|\dot z(t)| \le \frac{3\,\varepsilon T}{\min\{\frac{\rho_1}{r_2}, \rho_2\}} \quad \text{and} \quad |\dot\zeta(t)| \le \frac{3\,\varepsilon T}{\rho_3}, \tag{5.21}$$

at least as long as the solution stays in $\mathcal{D}_{\underline{r}-\underline{\rho}/3}$, during which time

$$|\dot I| = \Big|\sum_{j=1}^n (x_j \dot x_j + y_j \dot y_j)\Big| \le |z(t)||\dot z(t)| \le \frac{3\,r_2 \varepsilon T}{\min\{\frac{\rho_1}{r_2}, \rho_2\}}.$$

Writing $\mathbf{I}(t) = I(z(t))$ with $z(0) \in \mathcal{D}_{\underline{r}-\underline{\rho}}$ we deduce from (5.2) that for $|t| \le 1$

$$|\mathbf{I}(t) - I^0| \le |\mathbf{I}(t) - \mathbf{I}(0)| + |\mathbf{I}(0) - I^0| \le \frac{3\,r_2 |t| \varepsilon T}{\min\{\frac{\rho_1}{r_2}, \rho_2\}} + r_1 - \rho_1 < r_1 - \frac{2}{3}\,\rho_1.$$

Furthermore, using (5.2) again,

$$|z(t)| \le |z(0)| + \frac{3\,|t|\varepsilon T}{\min\{\frac{\rho_1}{r_2}, \rho_2\}} \le r_2 - \rho_2 + \frac{3\,|t|\varepsilon T}{\min\{\frac{\rho_1}{r_2}, \rho_2\}} < r_2 - \frac{2}{3}\,\rho_2$$

and

$$|\zeta(t)| \le |\zeta(0)| + \frac{3\,|t|\varepsilon T}{\rho_3} \le r_3 - \rho_3 + \frac{3\,|t|\varepsilon T}{\rho_3} < r_3 - \frac{2}{3}\,\rho_3.$$

This argument shows in particular that if the ρ_j are chosen in accordance with (5.2), then the solution starting in $\mathcal{D}_{\underline{r}-\underline{\rho}}$ will remain in $\mathcal{D}_{\underline{r}-2\underline{\rho}/3}$ for all times $|t| \le 1$. This proves (5.13), and verification of (5.14) is analogous. Moreover, (5.15) follows from (5.21).

For the complex case, since X_φ is analytic, the flow (X^t_φ) is defined locally and is locally analytic on $\mathbb{C}^{2n} \times \mathbb{C}^{2N}$ and for complex t. To find for which $t \in \mathbb{C}$ and between which domains this is true, we just repeat the argument that led to (5.13), and it is found that for $|t| < \lambda$ with λ as in (5.16) the flow is well defined, analytic and satisfies (5.17).

(ii) Again, since X_{g_0} is analytic, the flow $(X^t_{g_0})$ is defined locally, and is locally analytic, on $\mathbb{C}^{2n} \times \mathbb{C}^{2N}$ for complex t. Observe that

$$\frac{d}{dt}\Big(I \circ X^t_{g_0}\Big) = \{I, g_0\} \circ X^t_{g_0} = 0$$

for the function $I = I(z)$, since $g_0 = g_0(I)$ only depends on z through $I = (I_1, \dots I_n)$. In addition, since g_0 is independent of $\zeta = \Pi_2(z, \zeta)$ we have

$$\frac{d}{dt}\Big(\zeta \circ X^t_{g_0}\Big) = \{\zeta, g_0\} \circ X^t_{g_0} = 0.$$

In other words, both I and ζ are preserved by the flow, so that restrictions on the time which ensure (5.19)-(5.20) arise only from the condition on z. To prove (5.20) for instance, write $(z(t), \zeta(t)) = X^t_{g_0}(z(0), \zeta(0))$ and $\mathbf{I}(t) = I(z(t))$. Then by the foregoing observation:

$$\begin{aligned} |\mathbf{I}(t) - I^0| &= |\mathbf{I}(0) - I^0| < r_1 - \rho_1/3 < r_1, \\ |\zeta(t)| &= |\zeta(0)| < r_3 - \rho_3/3 < r_3, \end{aligned}$$

for all times, provided that initially $(z(0), \zeta(0)) \in \mathcal{D}_{\underline{r}-\underline{\rho}/3}$. Furthermore,

$$|\dot{z}(t)| \leq \left|\frac{d}{dt} X^t_{g_0}\right| = |X_{g_0}(X^t_{g_0})| \leq |X_{g_0}|_{\underline{r}}$$

as long as the flow stays in $\mathcal{D}_{\underline{r}}$. Using the definition of g_0 in (5.1), for any unit $2n$ vector $(\mathbf{a}, \mathbf{b}) = (a_1, \ldots, a_n, b_1, \ldots b_n)$ we can estimate

$$\begin{aligned} |(\mathbf{a} \cdot \nabla_x + \mathbf{b} \cdot \nabla_y) g_0| &= |\sum_{i,j=1}^{n} A_{ij}(I - I^0)_i (a_j x_j + b_j y_j)| \\ &\leq \|A\| \, |I - I^0| \max_{1 \leq j \leq n} |a_j x_j + b_j y_j| \leq \|A\| \, |I - I^0| \, |z|. \end{aligned}$$

So, maximizing over the unit vector, we can bound the Euclidean norm for X_{g_0} as

$$|X_{g_0}(z, \zeta)| \leq \|A\| \, |I - I^0| \, |z|$$

(using the l^1 operator norm on A). It follows that

$$|X_{g_0}|_{\underline{r}} \leq \|A\| \, r_1 r_2 \quad \text{for any} \quad \underline{r}.$$

Hence the desired bound $|z(t)| < r_2$ is obtained by inserting (5.18) into the estimate:

$$|z(t)| \leq |z(0)| + |t| \, |X_{g_0}|_{\underline{r}} < r_2 - \frac{1}{3}\rho_2 + \|A\| \, r_1 r_2 \, |t| < r_2.$$

To summarize, it has been shown that (5.20) is verified for $|t| < \tau$, and (5.19) follows in the same way. □

Continuation of proof of lemma 5.1 So far the statements (5.3) and (a) of the lemma are proved. Next, notice that the first assertion in (b) follows immediately from the definition of g_+ in (5.9), and the assumption $|f|_{\underline{r}} \leq \varepsilon$. To establish the second assertion in (b) we need to prove that $\{\bar{f}, h\} = 0$ (in view of $g_+ = g + \bar{f}$ and $\{g, h\} = 0$) which follows directly from the definition (5.4):

$$\{\bar{f}, h\} = \frac{1}{T}\int_0^T \{f \circ X^t_h, h\} \, dt = \frac{1}{T}\int_0^T \frac{d}{dt}(f \circ X^t_h) \, dt = 0.$$

To complete the proof of the lemma it remains to verify (c), which is now done by estimating each of the three terms in (5.12).

Estimation of $f_{+,1}$*:* As a consequence of (ii) in the previous proposition, the function

$$F(t) = \varphi \circ X^t_{g_0}(z, \zeta)$$

is analytic for complex times t as in (5.18) and for $(z, \zeta) \in \mathcal{D}_{\underline{r}-\underline{\rho}/3}$, since φ is defined on $\mathcal{D}_{\underline{r}}$. Then by Cauchy's estimate

$$|\{g_0, \varphi\}(z, \zeta)| = |F'(0)| \leq \frac{2}{\tau} \sup_{|t| = \tau/2} |F(t)|$$

for every $(z, \zeta) \in \mathcal{D}_{\underline{r}-2\underline{\rho}/3}$. To bound $F(t) = \varphi \circ X^t_{g_0}$ we just observe that by (5.19) and (5.6),

$$|\varphi \circ X^t_{g_0}|_{\underline{r}-2\underline{\rho}/3} \leq |\varphi|_{\underline{r}-\underline{\rho}/3} \leq |\varphi|_{\underline{r}} \leq \varepsilon T,$$

which leads to the estimate $|\{g_0, \varphi\}|_{\underline{r}-2\underline{\rho}/3} \leq 2\varepsilon T/\tau$. Hence, by (5.13) in the previous proposition,

$$|f_{+,1}|_{\underline{r}-\underline{\rho}} = \left| \int_0^1 \{g_0, \varphi\} \circ X^t_\varphi \, dt \right|_{\underline{r}-\underline{\rho}} \leq |\{g_0, \varphi\}|_{\underline{r}-2\underline{\rho}/3} \leq \frac{6\|A\| \, r_1 r_2 \varepsilon T}{\rho_2}.$$

Estimation of $f_{+,2}$*:* Next, to bound $f_{+,2} = \int_0^1 \{g + f_t, \varphi\} \circ X^t_\varphi \, dt$ we proceed in a similar fashion, but using the flow X^t_φ in place of $X^t_{g_0}$. To treat the first term in the integral define

$$G(t) = g \circ X^t_\varphi(z, \zeta)$$

where $(z,\zeta)\in\mathcal{D}_{\underline{r}-2\underline{\rho}/3}$ is fixed. By (i) in the previous proposition this is analytic for complex times t as in (5.16), so that Cauchy's estimate gives

$$|\{g,\varphi\}|_{\underline{r}-2\underline{\rho}/3} = |G'(0)| \le \frac{2}{\lambda}\sup_{|t|=\lambda/2}|G(t)|_{\underline{r}-2\underline{\rho}/3}.$$

By (5.17), G is bounded as $|G(t)|_{\underline{r}-2\underline{\rho}/3}\le |g|_{\underline{r}-\underline{\rho}/3}\le\delta$ for these t, leading to the overall bound

$$\Big|\int_0^1\{g,\varphi\}\circ X_\varphi^t\,dt\Big|_{\underline{r}-\underline{\rho}} \le |\{g,\varphi\}|_{\underline{r}-2\underline{\rho}/3}\le 2\delta/\lambda.$$

The second term in the integral defining $f_{+,2}$ is handled in exactly the same way, leading to the same bound with δ replaced by ε, since $|f_t|_{\underline{r}}\le\varepsilon$ for $t\in[0,1]$. Therefore altogether

$$|f_{+,2}|_{\underline{r}-\underline{\rho}}\le\frac{2(\delta+\varepsilon)}{\lambda}=\frac{36\,\varepsilon T(\delta+\varepsilon)}{\left(\min\{\frac{\rho_1}{r_2},\rho_2,\rho_3\}\right)^2}.$$

Estimation of $f_{+,3}$: The last contribution to f_+ arises from $f_{+,3}=\sigma\,(\Lambda\circ\Phi-\Lambda)$. By definition of $\Phi=X_\varphi^1$ this can be rewritten as

$$f_{+,3}=\sigma\,(\Lambda\circ X_\varphi^1-\Lambda\circ X_\varphi^0)=\sigma\int_0^1\frac{d}{dt}\,(\Lambda\circ X_\varphi^t)\,dt=\sigma\int_0^1\{\Lambda,\varphi\}\circ X_\varphi^t\,dt, \tag{5.22}$$

so that, using (5.13) and $\{\Lambda,\varphi\}=\langle D\Lambda,\Pi_2X_\varphi\rangle$ (the latter due to $\Lambda=\Lambda(\zeta)$), we deduce

$$|f_{+,3}|_{\underline{r}-\underline{\rho}}\le\sigma\,|\langle D\Lambda,\Pi_2X_\varphi\rangle|_{\underline{r}-2\underline{\rho}/3}\le\sigma\,C_\Lambda r_3\,|\Pi_2X_\varphi|_{\underline{r}-2\underline{\rho}/3}.$$

Since only the $\zeta_j=(\xi_j,\eta_j)$ derivatives of φ contribute to Π_2X_φ, this can be combined with Cauchy's estimate as

$$|f_{+,3}|_{\underline{r}-\underline{\rho}}\le\frac{3\sigma C_\Lambda r_3}{2\rho_3}\,|\varphi|_{\underline{r}}\le\frac{3\sigma C_\Lambda r_3}{2\rho_3}\,\varepsilon T$$

by (5.6). If we add together these bounds on $|f_{+,j}|_{\underline{r}-\underline{\rho}}$, then (c) is obtained. □

5.1.2 Transformation to normal form

We iterate lemma 5.1 m times to prove the following result.

Lemma 5.4 (Normal form) *Consider the Hamiltonian*

$$H(z,\zeta)=\langle\omega^0,I\rangle+\frac{1}{2}\,\langle A(I-I^0),I-I^0\rangle+g(z,\zeta)+f(z,\zeta)+\sigma\Lambda(\zeta),$$

where $\omega^0,I^0\in\mathbb{R}^n$, $A\in\mathbb{R}^{n\times n}$ is a symmetric matrix and $T,\sigma>0$ are fixed such that $T\omega^0\in 2\pi\mathbb{Z}^n$ holds. The functions g and f are assumed to be real analytic on an open neighbourhood of $\overline{\mathcal{D}}_{3\underline{r}}$, and Λ is assumed to be real analytic on an open neighbourhood of $\{|\zeta|\le 3r_3\}$. We suppose that

(i) $|g|_{3\underline{r}}\le\delta$ *and* $\{g,h\}=0$,

(ii) $|f|_{3\underline{r}}\le\varepsilon$,

(iii) $|D\Lambda(\zeta)|\le C_\Lambda|\zeta|$ *for* $|\zeta|\le 3r_3$,

(iv) $r_1<2r_2^2$ *and* $r_1<2r_2r_3$,

(v) $m^2\varepsilon T<\dfrac{r_1^2}{81\,r_2^2}$,

(vi) $54m\|A\|r_1T+\dfrac{324\,(\delta+2\varepsilon)m^2r_2^2T}{r_1^2}+\dfrac{9\sigma C_\Lambda mT}{2}\le\dfrac{1}{2}$

for some $\delta, \varepsilon > 0$ and $C_\Lambda > 0$. Then there exists a real analytic symplectic transformation

$$\Psi : \mathcal{D}_{2\underline{r}} \to \mathcal{D}_{3\underline{r}}$$

such that, on $\mathcal{D}_{2\underline{r}}$,

$$H \circ \Psi = \langle \omega^0, I \rangle + \frac{1}{2} \langle A(I - I^0), I - I^0 \rangle + \hat{g}(z, \zeta) + \hat{f}(z, \zeta) + \sigma\Lambda(\zeta)$$

and with the properties:

(a) $\quad |\Psi - \mathrm{id}|_{2\underline{r}} \leq \dfrac{18\, m r_2}{r_1}\, \varepsilon T,$

(b) $\quad |\hat{g}|_{2\underline{r}} \leq \delta + 2\varepsilon$ *and* $\{\hat{g}, h\} = 0,$

(c) $\quad |\hat{f}|_{2\underline{r}} \leq 2^{-m} \varepsilon.$

Proof We apply the iterative lemma (lemma 5.1) m times, where at the j^{th} stage $\underline{r}$ is taken to be $3\underline{r} - j\underline{r}/m$ and $\underline{\rho} = \underline{r}/m$ with $j = 0, \ldots, m-1$. For $j = 0$ we need to check (5.2), which reads as

$$\varepsilon T < \frac{1}{9\, m^2} \Big(\min \Big\{ \frac{r_1}{3 r_2}, r_2, r_3 \Big\} \Big)^2.$$

According to (iv) we have $\min\{\frac{r_1}{3r_2}, r_2, r_3\} = \frac{r_1}{3r_2}$ and the condition becomes

$$\varepsilon T < \frac{1}{81\, m^2} \frac{r_1^2}{r_2^2}$$

which is verified by (v). Thus lemma 5.1 yields a real analytic symplectic transformation $\Phi_1 : \mathcal{D}_{3\underline{r} - \underline{r}/m} \to \mathcal{D}_{3\underline{r}}$ such that, on $\mathcal{D}_{3\underline{r} - \underline{r}/m}$,

$$H \circ \Phi_1 = \langle \omega^0, I \rangle + \frac{1}{2} \langle A(I - I^0), I - I^0 \rangle + g_1(z, \zeta) + f_1(z, \zeta) + \sigma\Lambda(\zeta)$$

and moreover:

- $|\Phi_1 - \mathrm{id}|_{3\underline{r} - \underline{r}/m} \leq \dfrac{9\, m r_2}{r_1}\, \varepsilon T,$
- $|g_1|_{3\underline{r}} \leq \delta + \varepsilon$ and $\{g_1, h\} = 0,$
- $|f_1|_{3\underline{r} - \underline{r}/m} \leq \left[54 m \|A\| r_1 + \dfrac{324\, (\delta + \varepsilon) m^2 r_2^2}{r_1^2} + \dfrac{9 \sigma C_\Lambda m}{2} \right] \varepsilon T \leq \dfrac{\varepsilon}{2},$

the latter in view of (vi). Put $\Psi_1 = \Phi_1$. For the induction step assume that we have constructed a real analytic symplectic transformation Ψ_j such that, on $\mathcal{D}_{3\underline{r} - j\underline{r}/m}$,

$$H \circ \Psi_j = \langle \omega^0, I \rangle + \frac{1}{2} \langle A(I - I^0), I - I^0 \rangle + g_j(z, \zeta) + f_j(z, \zeta) + \sigma\Lambda(\zeta)$$

with

- $|\Psi_j - \mathrm{id}|_{3\underline{r} - j\underline{r}/m} \leq \dfrac{9\, m r_2}{r_1}\, \varepsilon T \displaystyle\sum_{i=0}^{j-1} 2^{-i},$
- $|g_j|_{3\underline{r} - (j-1)\underline{r}/m} \leq \delta + \varepsilon \displaystyle\sum_{i=0}^{j-1} 2^{-i}$ and $\{g_j, h\} = 0,$
- $|f_j|_{3\underline{r} - j\underline{r}/m} \leq 2^{-j} \varepsilon.$

In order to apply the iterative lemma to this Hamiltonian (and with ε replaced by $2^{-j}\varepsilon$ and δ replaced by $\delta + \varepsilon\sum_{i=0}^{j-1} 2^{-i}$), we have to see that (5.2) holds, which reads as

$$2^{-j}\varepsilon T < \frac{1}{9\,m^2}\Big(\min\Big\{\frac{r_1}{r_2(3-j/m)}, r_2, r_3\Big\}\Big)^2. \tag{5.23}$$

Since $\frac{r_1}{r_2(3-j/m)} \le \frac{r_1}{2r_2} \le \min\{r_2, r_3\}$ by (iv), (5.23) reduces to

$$2^{-j}\varepsilon T < \frac{1}{9\,m^2}\Big(\frac{r_1}{r_2(3-j/m)}\Big)^2,$$

which is a consequence of (v). Therefore lemma 5.1 applies, yielding a real analytic symplectic transformation

$$\Phi_{j+1} : \mathcal{D}_{3\underline{r}-(j+1)\underline{r}/m} \to \mathcal{D}_{3\underline{r}-j\underline{r}/m}$$

such that, on $\mathcal{D}_{3\underline{r}-(j+1)\underline{r}/m}$,

$$H \circ \Psi_j \circ \Phi_{j+1} = \langle \omega^0, I\rangle + \frac{1}{2}\langle A(I-I^0), I-I^0\rangle + g_{j+1}(z,\zeta) + f_{j+1}(z,\zeta) + \sigma\Lambda(\zeta)$$

and furthermore by the hypotheses:

- $$|\Phi_{j+1} - \mathrm{id}|_{3\underline{r}-(j+1)\underline{r}/m} \le \frac{3\cdot 2^{-j}\varepsilon T m}{\big(\frac{r_1}{r_2(3-j/m)}\big)} \le \frac{9\,mr_2}{r_1}\,2^{-j}\varepsilon T,$$
- $$|g_{j+1}|_{3\underline{r}-j\underline{r}/m} \le \delta + \varepsilon\sum_{i=0}^{j-1} 2^{-i} + 2^{-j}\varepsilon = \delta + \varepsilon\sum_{i=0}^{j} 2^{-i} \text{ and } \{g_{j+1}, h\} = 0,$$
- $$\begin{aligned}|f_{j+1}|_{3\underline{r}-(j+1)\underline{r}/m} &\le \Big[6m\|A\|r_1(3-j/m)^2 + \frac{36\,(\delta+\varepsilon\sum_{i=0}^{j} 2^{-i})m^2}{(\frac{r_1}{r_2(3-j/m)})^2} + \frac{3\sigma C_\Lambda(3-j/m)m}{2}\Big]\,2^{-j}\varepsilon T\\ &\le 2^{-(j+1)}\varepsilon.\end{aligned}$$

Now define $\Psi_{j+1} = \Psi_j \circ \Phi_{j+1}$ and estimate

$$\begin{aligned}|\Psi_{j+1} - \mathrm{id}|_{3\underline{r}-(j+1)\underline{r}/m} &\le |(\Psi_j - \mathrm{id})\circ\Phi_{j+1}|_{3\underline{r}-(j+1)\underline{r}/m} + |\Phi_{j+1} - \mathrm{id}|_{3\underline{r}-(j+1)\underline{r}/m}\\ &\le |\Psi_j - \mathrm{id}|_{3\underline{r}-j\underline{r}/m} + \frac{9\,mr_2}{r_1}\,2^{-j}\varepsilon T \le \frac{9\,mr_2}{r_1}\,\varepsilon T\sum_{i=0}^{j} 2^{-i}\end{aligned}$$

to deduce that the inductive assumptions hold also at this step. The process terminates at $j = m-1$ and we can define $\hat{g} = g_{m-1}$, $\hat{f} = f_{m-1}$, and $\Psi = \Psi_{m-1}$. □

5.2 Nekhoroshev stability in the case $N = 0$

We recall the statement of Nekhoroshev stability in the case $N = 0$, so that only the z component appears. We assume that the initial values $z(0) = (x(0), y(0)) \in \mathbb{R}^{2n}$ are close to the equilibrium point $(0,0)$ for the real analytic Hamiltonian:

$$H(z) = \langle \alpha, I(z)\rangle + \frac{1}{2}\langle AI(z), I(z)\rangle + f(z), \quad \text{with} \quad \langle AI, I\rangle \ge \frac{1}{M}|I|^2 = \frac{1}{M}\Big(\sum_{j=1}^{n}|I_j|\Big)^2$$

and $f(z) = \mathcal{O}(z^5)$ for $|z| \to 0$. In that case we have the following theorem:

Theorem 5.5 *There exist positive numbers K, k, a (depending on n, α, M and $\|A\|$) and θ_0 (depending on n, α, M, $\|A\|$ and f) with the following properties. If $\mathbf{I}(0) = I(z(0))$ is such that $|\mathbf{I}(0)| = \theta^2$ for some $0 < \theta \le \theta_0$, then $\mathbf{I}(t) = I(z(t))$ satisfies*

$$|\mathbf{I}(t) - \mathbf{I}(0)| \le K\theta^{2+a} \quad \textit{for} \quad |t| \le e^{\frac{k}{\theta^a}}.$$

Proof This is the classical Nekhoroshev bound of [7] for the case of an elliptic equilibrium, see [3, 5, 9, 10]. The proof can also be extracted from the proof of our theorem 3.4 above, although not in its putative sharpest form (with $a = \frac{1}{2n}$). □

5.3 Some further lemmas

In this section $|\cdot|_\infty$ denotes the maximum norm on $\mathbb{R}^n$, i.e. $|x|_\infty = \max_{1\le j\le n}|x_j|$.

Lemma 5.6 (Dirichlet) *For every $Q\in\mathbb{N}$ and $\omega\in\mathbb{R}^n$,*

$$\min_{q\in\{1,\dots,Q\}}\ \min_{p\in\mathbb{Z}^n}|q\omega-p|_\infty \le \frac{1}{Q^{1/n}}.$$

Proof See [12, Thm. 1B, p. 34]. □

Corollary 5.7 *For every $Q\in\mathbb{N}$ and $\omega\in\mathbb{R}^n$ such that $|\omega|_\infty>1$ there exists $\omega^0\in\mathbb{R}^n$ and $T\in[2\pi(1-|\omega|_\infty^{-1})\,,2\pi Q]$ and such that $\omega^0T\in 2\pi\mathbb{Z}^n$ and*

$$|\omega-\omega^0|_\infty \le \frac{2\pi}{TQ^{1/(n-1)}}. \tag{5.24}$$

Proof We may assume that $\omega_n=|\omega|_\infty>1$, where $\omega=(\omega_1,\dots,\omega_n)$. Then write

$$\omega=\frac{\omega_n}{[\omega_n]}\Big(\frac{[\omega_n]}{\omega_n}\,\omega_1,\dots,\frac{[\omega_n]}{\omega_n}\,\omega_{n-1},[\omega_n]\Big)$$

and apply lemma 5.6 to find $q\in\{1,\dots,Q\}$ and $p\in\mathbb{Z}^{n-1}$ such that

$$\Big|q\,\frac{[\omega_n]}{\omega_n}\,\omega_j-p_j\Big|\le\frac{1}{Q^{1/(n-1)}},\quad 1\le j\le n-1. \tag{5.25}$$

Defining $T=2\pi q\,[\omega_n]/\omega_n$ and

$$\omega^0=\frac{\omega_n}{[\omega_n]}\Big(\frac{p_1}{q},\dots,\frac{p_{n-1}}{q},[\omega_n]\Big),$$

it follows that $T\omega^0=2\pi(p_1,\dots,p_{n-1},q[\omega_n])\in2\pi\mathbb{Z}^n$. Furthermore, $1-1/\omega_n\le[\omega_n]/\omega_n\le1$ and $1\le q\le Q$ yield the bound on T. Finally, (5.24) is a direct consequence of (5.25). □

As a further corollary we obtain the density of periodic orbits in sufficiently small neighbourhoods of elliptic equilibria for convex integrable Hamiltonians. The frequency ω^0 is called $T-$*periodic* if $\omega^0T\in2\pi\mathbb{Z}^n$. Now to be precise consider a real analytic Hamiltonian on $\mathbb{R}^{2n}$ of the form

$$\langle\alpha,I\rangle+\frac12\langle AI,I\rangle+g(I) \tag{5.26}$$

where A is a strictly positive matrix and $g(I)=O(|I|^3)$, and the notation is as in the introduction. The function

$$I\mapsto\Omega(I)=\alpha+AI+Dg(I)=\alpha+AI+O(|I|^2) \tag{5.27}$$

is invertible in a neighbourhood of the origin in $\mathbb{R}^n$ by the inverse function theorem, since $D\Omega(I)=A+O(|I|)$ is invertible for $|I|$ small enough. The smooth inverse Ω^{-1} is defined on a neighbourhood of $\alpha=\Omega(0)$.

Corollary 5.8 *Given a function $\Omega:\mathbb{R}^n\to\mathbb{R}^n$ as in (5.27), with $\alpha\in\mathbb{R}^n\setminus\{0\}$, and a number $a>0$, there exist $C>0$ (depending upon Ω) and $\theta_0>0$ (depending upon Ω,α,a,n) such that the following holds: if $I\in\mathbb{R}^n$ and $|I|=\theta^2$ for some $0<\theta\le\theta_0$, then there exist $I^0\in\mathbb{R}^n$ and $\tau>0$ such that*

(i) $|I-I^0|_\infty\le C\,\frac{\theta^{2+a}}{\tau}$,

(ii) $\pi\le2\pi(1-\theta^2|\Omega(I)|_\infty^{-1})\le\tau\le4\pi\theta^{-a(n-1)}$, *and*

(iii) $\omega^0=\Omega(I^0)$ *is τ/θ^2-periodic, i.e.* $\omega^0\frac{\tau}{\theta^2}\in2\pi\mathbb{Z}^n$.

Proof By the above remarks there is $\varepsilon > 0$ such that $\Omega : B_\varepsilon(0) \to \Omega(B_\varepsilon(0)) =: U$ is smoothly invertible and $C = 2\pi\|D\Omega^{-1}\|_{L^\infty(\overline{U})} < \infty$. Choose δ obeying $0 < \delta < |\alpha|_\infty$ such that $B_\delta(\alpha) \subset U$. Next fix $\theta_0 > 0$ sufficiently small that for $0 < \theta \le \theta_0$ and $|I| = \theta^2$ there holds:

$$|\Omega(I) - \alpha|_\infty < \delta/2, \quad \theta^2 < \min\{\varepsilon, |\alpha|_\infty/4\}, \quad \theta^{a(n-1)} < 1, \quad 2\theta^{2+a} < \delta/2;$$

hence θ_0 depends on Ω, α, a and n. Now if $|I| = \theta^2$ for some $0 < \theta \le \theta_0$, then $|I| < \varepsilon$ ensures that $\omega = \Omega(I) \in U$ is well-defined, and $|\omega|_\infty > |\alpha|_\infty/2$ since $\delta < |\alpha|_\infty$. Putting $Q = [\theta^{-a(n-1)}] + 1$ and $\tilde{\omega} = \theta^{-2}\omega$, we have $|\tilde{\omega}|_\infty \ge \theta^{-2}|\alpha|_\infty/2 > 2 > 1$. Therefore corollary 5.7 applies and yields the existence of $\tau > 0$ and $\tilde{\omega}^0 \in \mathbb{R}^n$ such that $\tilde{\omega}^0\tau \in 2\pi\mathbb{Z}^n$, $2\pi(1 - |\tilde{\omega}|_\infty^{-1}) \le \tau \le 2\pi Q$ and

$$|\tilde{\omega} - \tilde{\omega}^0|_\infty \le \frac{2\pi}{\tau Q^{1/(n-1)}}. \tag{5.28}$$

Also, $\tau \ge 2\pi(1 - |\tilde{\omega}|_\infty^{-1}) \ge \pi$ follows since $|\tilde{\omega}|_\infty > 2$. Defining $\omega^0 = \theta^2\tilde{\omega}^0$, we get $\omega^0 \frac{\tau}{\theta^2} \in 2\pi\mathbb{Z}^n$. Furthermore,

$$|\omega - \omega^0|_\infty = \theta^2|\tilde{\omega} - \tilde{\omega}^0|_\infty \le \frac{2\pi\theta^2}{\tau Q^{1/(n-1)}} \le \frac{2\theta^2}{\theta^{-a}} = 2\theta^{2+a} < \delta/2$$

implies that $|\omega^0 - \alpha|_\infty \le |\omega^0 - \omega|_\infty + |\Omega(I) - \alpha|_\infty < \delta/2 + \delta/2 = \delta$, and consequently $\omega^0 \in U$ so that $I^0 = \Omega^{-1}(\omega^0)$ is well defined. Then (ii) follows from $2\pi Q \le 2\pi(\theta^{-a(n-1)} + 1) \le 4\pi\theta^{-a(n-1)}$. Finally, concerning (i) it suffices to note that since both ω and ω^0 lie in the ball $B_\delta(\alpha) \subset U$

$$|I - I^0|_\infty = |\Omega^{-1}(\omega) - \Omega^{-1}(\omega^0)|_\infty \le \|D\Omega^{-1}\|_{L^\infty(\overline{U})}\,|\omega - \omega^0|_\infty \le \|D\Omega^{-1}\|_{L^\infty(\overline{U})}\,\frac{2\pi\theta^2}{\tau\theta^{-a}} = C\,\frac{\theta^{2+a}}{\tau}$$

by (5.28), completing the proof. □

We also need the following quantitative version of the inverse mapping theorem.

Lemma 5.9 *Let X, Y be Banach spaces and suppose that $U \subset X$ is open. If $\Psi : U \to \Psi(U) \subset Y$ is a homeomorphism, Ψ^{-1} is Lipschitz continuous with constant* $\mathrm{Lip}(\Psi^{-1}) < \lambda$, *and $\overline{B_r(x)} \subset U$, then*

$$\Psi(\overline{B_r(x)}) \supset \overline{B_{r/\lambda}(\Psi(x))}.$$

Proof See [13, Prop. I.3, p. 50]. □

References

[1] Benettin G., Galgani L. & Giorgilli A.: Realization of holonomic constraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory. Part I, *Comm. Math. Phys.* **113**, 87-103 (1987)

[2] Benettin G., Galgani L. & Giorgilli A.: 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**, 557-601 (1989)

[3] Fassò F., Guzzo M. & Benettin G.: Nekhoroshev-stability of elliptic equilibria of Hamiltonian systems, *Comm. Math. Phys.* **197**, 347-360 (1998)

[4] Bounemoura A.: Nekhoroshev estimates for finitely differentiable quasi-convex Hamiltonians, *J. Differential Equations* **249**, 2905-2920 (2010)

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

[6] Nachbin L.: *Topology on Spaces of Holomorphic Mappings*, Springer, Berlin-New York 1969

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

[8] Niederman L.: Stability over exponentially long times in the planetary problem, *Nonlinearity* **9**, 1703-1751 (1996)

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

[10] Pöschel J.: On Nekhoroshev's estimate at an elliptic equilibrium, *Internat. Math. Res. Notices*, 203-215 (1999)

[11] Rubin H. & Ungar P.: Motion under a strong confining force, *Comm. Pure Appl. Math.* **10**, 65-87 (1957)

[12] Schmidt W.M.: *Diophantine Approximations and Diophantine Equations*, LNM 1467, Springer, Berlin-New York 1991

[13] Shub M.: *Global Stability of Dynamical Systems*, Springer, Berlin-New York 1987

[14] Takens F.: Motion under the influence of a strong constraining force, pp. 425-445 in *Global Theory of Dynamical Systems*, Proc. Internat. Conf., Northwestern Univ., Evanston, Ill., 1979, LNM 819, Springer, Berlin-New York 1980