ANNALI DELLA
SCUOLA NORMALE SUPERIORE DI PISA
Classe di Scienze

JÜRGEN PÖSCHEL

A KAM-theorem for some nonlinear partial differential equations

Annali della Scuola Normale Superiore di Pisa, Classe di Scienze 4e série, tome 23, nº 1 (1996), p. 119-148

<http://www.numdam.org/item?id=ASNSP_1996_4_23_1_119_0>

A KAM-Theorem for some Nonlinear Partial Differential Equations

JÜRGEN PÖSCHEL

Introduction

In this paper a KAM-theorem about the existence of quasi-periodic motions in some infinite dimensional hamiltonian systems is proven. In [5] and [8] this theorem is applied to some nonlinear Schrödinger and wave equation on the interval $[0, \pi]$, respectively, and we refer to these sources for motivation and background. Here we concern ourselves with the basic KAM-theorem, which is the very foundation of these applications.

The first theorem of this kind is due to Eliasson [2], who proved the existence of invariant tori of less than maximal dimension in nearly integrable hamiltonian systems of finite degrees of freedom. Thereafter, the result was extended to infinite degrees of freedom systems by Wayne [10], the author [7] and, independently of Eliasson's work, by Kuksin – see [4] and the references therein. We refer to [4, 7] for more historical remarks, and to [4] for further applications. The relations of the present paper to [4] and [7] will be discussed in the last section.

1. - Statement of Results

We consider small perturbations of an infinite dimensional hamiltonian in the parameter dependent normal form

$$N = \sum_{1 \leq j \leq n} \omega_j(\xi) y_j + \frac{1}{2} \sum_{j \geq 1} \Omega_j(\xi)(u_j^2 + v_j^2)$$

Pervenuto alla Redazione il 28 Ottobre 1994 e in forma definitiva il 18 Febbrario 1995.

on a phase space

$$\mathcal{P}^{a,p} = \mathbb{T}^n \times \mathbb{R}^n \times \ell^{a,p} \times \ell^{a,p} \ni (x, y, u, v),$$

where $\mathbb{T}^n$ is the usual n-torus with $1 \leq n < \infty$, and $\ell^{a,p}$ is the Hilbert space of all real (later complex) sequences $w = (w_1, w_2, \ldots)$ with

$$\|w\|_{a,p}^2 = \sum_{j\geq 1} |w_j|^2 j^{2p} e^{2aj} < \infty,$$

where $a \geq 0$ and $p \geq 0$. The frequencies $\omega = (\omega_1, \ldots, \omega_n)$ and $\Omega = (\Omega_1, \Omega_2, \ldots)$ depend on n parameters $\xi \in \Pi \subset \mathbb{R}^n$, Π a closed bounded set of positive Lebesgue measure, in a way described below.

The hamiltonian equations of motion of N are

$$\dot{x} = \omega(\xi), \quad \dot{y} = 0, \quad \dot{u} = \Omega(\xi)v, \quad \dot{v} = -\Omega(\xi)u,$$

where $(\Omega u)_j = \Omega_j u_j$. Hence, for each $\xi \in \Pi$, there is an invariant n-dimensional torus $\mathcal{T}_0^n = \mathbb{T}^n \times \{0,0,0\}$ with frequencies $\omega(\xi)$, which has an elliptic fixed point in its attached uv-space with frequencies $\Omega(\xi)$. Hence $\mathcal{T}_0^n$ is linearly stable. The aim is to prove the persistence of a large portion of this family of linearly stable rotational tori under small perturbations $H = N + P$ of the hamiltonian N. To this end the following assumptions are made.

Assumption A: Nondegeneracy. The map $\xi \mapsto \omega(\xi)$ is a lipeomorphism between Π and its image, that is, a homemorphism which is Lipschitz continuous in both directions. Moreover, for all integer vectors $(k,l) \in \mathbb{Z}^n \times \mathbb{Z}^\infty$ with $1 \leq |l| \leq 2$,

$$|\{\xi : \langle k, \omega(\xi)\rangle + \langle l, \Omega(\xi)\rangle = 0\}| = 0$$

and

$$\langle l, \Omega(\xi)\rangle \not\equiv 0 \text{ on } \Pi,$$

where $|\cdot|$ denotes Lebesgue measure for sets, $|l| = \sum_j |l_j|$ for integer vectors, and $\langle \cdot, \cdot \rangle$ is the usual scalar product.

Assumption B: Spectral Asymptotics. There exist $d \geq 1$ and $\delta < d - 1$ such that

$$\Omega_j(\xi) = j^d + \cdots + O(j^\delta),$$

where the dots stand for fixed lower order terms in j, allowing also negative exponents. More precisely, there exists a fixed, parameter-independent sequence $\bar{\Omega}$ with $\bar{\Omega}_j = j^d + \cdots$ such that the tails $\tilde{\Omega}_j = \Omega_j - \bar{\Omega}_j$ give rise to a Lipschitz map

$$\tilde{\Omega} : \Pi \to \ell_\infty^{-\delta},$$

where ℓ_∞^p is the space of all real sequences with finite norm $|w|_p = \sup_j |w_j| j^p$. – Note that the coefficient of j^d can always be normalized to one by rescaling the time. So there is no loss of generality by this assumption. Also, there is no restriction on finite numbers of frequencies.

Assumption C: Regularity. The perturbation P is real analytic in the space coordinates and Lipschitz in the parameters, and for each $\xi \in \Pi$ its hamiltonian vector space field $X_P = (P_y, -P_x, P_v, -P_u)^T$ defines near $\mathcal{T}_0^n$ a real analytic map

$$X_P : \mathcal{P}^{a,p} \to \mathcal{P}^{a,\bar{p}}, \quad \begin{cases} \bar{p} \geq p \text{ for } d > 1, \\ \bar{p} > p \text{ for } d = 1. \end{cases}$$

We may also assume that $p - \bar{p} \leq \delta < d - 1$ by increasing δ, if necessary.

To make this quantitative we introduce complex $\mathcal{T}_0^n$-neighbourhoods

$$D(s,r) : |\mathrm{Im}\, x| < s, \ |y| < r^2, \ \|u\|_{a,p} + \|v\|_{a,p} < r,$$

where $|\cdot|$ denotes the sup-norm for complex vectors, and weighted phase space norms

$$\text{(1)} \qquad |\!|\!| W |\!|\!|_r = |\!|\!| W |\!|\!|_{\bar{p},r} = |X| + \frac{1}{r^2}|Y| + \frac{1}{r}\|U\|_{a,\bar{p}} + \frac{1}{r}\|V\|_{a,\bar{p}}$$

for $W = (X, Y, U, V)$. Then we assume that X_P is real analytic in $D(s,r)$ for some positive s, r uniformly in ξ with finite norm $|\!|\!| X_P |\!|\!|_{r,D(s,r)} = \sup_{D(s,r)} \|X_P\|_r$, and that the same holds for its Lipschitz semi-norm

$$|\!|\!| X_P |\!|\!|_r^{\mathcal{L}} = \sup_{\xi \neq \varsigma} \frac{|\!|\!| \Delta_{\xi\varsigma} X_P |\!|\!|_r}{|\xi - \varsigma|},$$

where $\Delta_{\xi\varsigma} X_P = X_P(\cdot, \xi) - X_P(\cdot, \varsigma)$, and where the supremum is taken over Π.

The main result decomposes into two parts, an analytic and a geometric one, formulated as Theorem A and B, respectively. In the former the existence of invariant tori is stated under the assumption that a certain set of diophantine frequencies is not empty. The latter assures that this is indeed the case.

To state the main results we assume that

$$|\omega|_{\Pi}^{\mathcal{L}} + |\!|\!| \Omega |\!|\!|_{-\delta,\Pi}^{\mathcal{L}} \leq M < \infty, \quad |\omega^{-1}|_{\omega(\Pi)}^{\mathcal{L}} \leq L < \infty,$$

where the Lipschitz semi-norms are defined analogously to $|\!|\!| X_P |\!|\!|_r^{\mathcal{L}}$. Moreover, we introduce the notations

$$\langle l \rangle_d = \max\left(1, \left|\sum j^d l_j\right|\right), \quad A_k = 1 + |k|^\tau,$$

where $\tau \geq n + 1$ is fixed later. Finally, let $Z = \{(k,l) \neq 0, |l| \leq 2\} \subset \mathbb{Z}^n \times \mathbb{Z}^\infty$.

THEOREM A. *Suppose $H = N + P$ satisfies assumptions A, B and C, and*

$$\epsilon = |\!|\!| X_P |\!|\!|_{r,D(s,r)} + \frac{\alpha}{M} |\!|\!| X_P |\!|\!|_{r,D(s,r)}^{\mathcal{L}} \leq \gamma\alpha,$$

where $0 < \alpha \le 1$ is another parameter, and γ depends on n, τ and s. Then there exists a Cantor set $\Pi_\alpha \subset \Pi$, a Lipschitz continuous family of torus embeddings $\Phi : \mathbb{T}^n \times \Pi_\alpha \to \mathcal{P}^{a,\bar{p}}$, and a Lipschitz continuous map $\omega_* : \Pi_\alpha \to \mathbb{R}^n$, such that for each ξ in Π_α the map Φ restricted to $\mathbb{T}^n \times \{\xi\}$ is a real analytic embedding of a rotational torus with frequencies $\omega_*(\xi)$ for the hamiltonian H at ξ.

Each embedding is real analytic on $|\operatorname{Im} x| < \frac{s}{2}$, and

$$\| \Phi - \Phi_0 \|_r + \frac{\alpha}{M} \| \Phi - \Phi_0 \|_r^{\mathcal{L}} \le c\epsilon/\alpha,$$

$$|\omega_* - \omega| + \frac{\alpha}{M} |\omega_* - \omega|^{\mathcal{L}} \le c\epsilon,$$

uniformly on that domain and Π_α, where Φ_0 is the trivial embedding $\mathbb{T}^n \times \Pi \to \mathcal{T}_0^n$, and $c \le \gamma^{-1}$ depends on the same parameters as γ.

Moreover, there exist Lipschitz maps ω_ν and Ω_ν on Π for $\nu \ge 0$ satisfying $\omega_0 = \omega$, $\Omega_0 = \Omega$ and

$$|\omega_\nu - \omega| + \frac{\alpha}{M} |\omega_\nu - \omega|^{\mathcal{L}} \le c\epsilon,$$

$$|\Omega_\nu - \Omega|_{-\delta} + \frac{\alpha}{M} |\Omega_\nu - \Omega|_{-\delta}^{\mathcal{L}} \le c\epsilon,$$

such that $\Pi \backslash \Pi_\alpha \subset \bigcup \mathcal{R}_{kl}^\nu(\alpha)$, where

$$\mathcal{R}_{kl}^\nu(\alpha) = \left\{ \xi \in \Pi : |\langle k, \omega_\nu(\xi) \rangle + \langle l, \Omega_\nu(\xi) \rangle| < \alpha \frac{\langle l \rangle_d}{A_k} \right\},$$

and the union is taken over all $\nu \ge 0$ and $(k,l) \in Z$ such that $|k| > K_0 2^{\nu-1}$ for $\nu \ge 1$ with a constant $K_0 \ge 1$ depending only on n and τ.

REMARK 1. We will see at the end of Section 4 that around each torus there exists another normal form of the hamiltonian having an elliptic fixed point in the uv-space. Thus all the tori are linearly stable. Moreover, their frequencies are diophantine.

REMARK 2. The rôle of the parameter α is the following. In applications the size of the perturbation usually depends on a small parameter, for example the size of the neighbourhood around an elliptic fixed point. One then wants to choose α as another function of this parameter in order to obtain useful estimates for $|\Pi \backslash \Pi_\alpha|$. See [5, 8] for examples.

REMARK 3. Theorem A only requires the frequency map $\xi \mapsto \omega(\xi)$ to be Lipschitz continuous, but not to be a homeomorphism or lipeomorphism. This only matters for Theorem B.

We now verify that the Cantor set Π_α is not empty, and that indeed $|\Pi \backslash \Pi_\alpha| \to 0$ as α tends to zero. In the case $d = 1$, let κ be a positive number such that the unperturbed frequencies satisfy

$$\frac{\Omega_i - \Omega_j}{i-j} = 1 + O(j^{-\kappa}), \quad i > j,$$

uniformly on Π. Without loss of generality, we can assume that $-\delta \le \kappa$ by increasing δ, if necessary.

THEOREM B. *Let ω_ν and Ω_ν for $\nu \ge 0$ be Lipschitz maps on Π satisfying*

$$|\omega_\nu - \omega|,\ |\Omega_\nu - \Omega|_{-\delta} \le \alpha,\ |\omega_\nu - \omega|^{\mathcal{L}},\ |\Omega_\nu - \Omega|^{\mathcal{L}}_{-\delta} \le \frac{1}{2L},$$

and define the sets $\mathcal{R}^\nu_{kl}(\alpha)$ as in Theorem A *choosing τ as in* (22). *Then there exists a finite subset $\mathcal{X} \subset Z$ and a constant $\tilde{c}$ such that*

$$\left| \bigcup_{\substack{(k,l)\notin\mathcal{X} \\ \nu}} \mathcal{R}^\nu_{kl}(\alpha) \right| \le \tilde{c}\rho^{n-1}\alpha^\mu, \quad \mu = \begin{cases} 1 & \text{for } d > 1, \\ \dfrac{\kappa}{\kappa+1} & \text{for } d = 1, \end{cases}$$

for all sufficiently small α, where $\rho = \operatorname{diam}\Pi$. The constant $\tilde{c}$ and the index set $\mathcal{X}$ are monotone functions of the domain Π: they do not increase for closed subsets of Π. In particular, if $\delta \le 0$, then $\mathcal{X} \subset \{(k,l) : 0 < |k| \le 16LM\}$.

By slightly sharpening the smallness condition the frequency maps of Theorem A satisfy the assumptions of Theorem B, and we may conclude that the measure of *all* sets $\mathcal{R}^\nu_{kl}(\alpha)$ tends to zero.

COROLLARY C. *If in Theorem* A, *the constant γ is replaced by a smaller constant $\tilde{\gamma} \le \gamma/2LM$ depending on the set $\mathcal{X}$, then*

$$|\Pi \backslash \Pi_\alpha| \le \left| \bigcup \mathcal{R}^\nu_{kl}(\alpha) \right| \to 0 \quad \textit{as } \alpha \to 0. \tag{2}$$

In particular, if $\delta \le 0$, then one may take $\tilde{\gamma} = \dfrac{\gamma}{2(LM)^{\tau+1}}$.

The point of choosing $\tilde{\gamma}$ is to make sure that $\max_{(k,l)\in\mathcal{X}} |k| \le K_0$, so that for $(k,l) \in \mathcal{X}$ we only need to consider the sets $\mathcal{R}^0_{kl}(\alpha)$, which are defined in terms of the *unperturbed* frequencies. Then $|\mathcal{R}^0_{kl}(\alpha)| \to 0$ as $\alpha \to 0$ by Assumption A.

In the applications [5, 8] the unperturbed frequencies are in fact *affine* functions of the parameters. In the case $d > 1$, as it happens in the nonlinear Schrödinger equation, we then immediately obtain $|\Pi\backslash\Pi_\alpha| \le \tilde{c}\rho^{n-1}\alpha$. In the case $d = 1$, however, α appears with the exponent $\mu < 1$, and it happens that for the nonlinear wave equation the present estimate is not sufficient to conclude that the set of bad frequencies is smaller than the set of all frequencies (which also depends on a small parameter). The following better estimate is required, which we only formulate for the case needed.

THEOREM D. *Suppose that in Theorem* A *the unperturbed frequencies are affine functions of the parameters. Then*

$$|\Pi\backslash\Pi_\alpha| \le \tilde{c}\rho^{n-1}\alpha^{\tilde{\mu}}, \quad \tilde{\mu} = \begin{cases} 1 & \text{for } d > 1, \\ \dfrac{\kappa}{\kappa + 1 - \pi/4} & \text{for } d = 1, \end{cases}$$

for all sufficiently small α, where π is any number in $0 \le \pi < \min(\bar{p} - p, 1)$. In this case the constant $\tilde{c}$ also depends on π and $\bar{p} - p$.

The rest of the paper consists almost entirely of the proofs of the preceding results, which employs the usual Newton type iteration procedure to handle small divisor problems. In Section 2 the relevant linearized equation is considered, and in Section 3 one step of the iterative scheme is described. The iteration itself takes place in Section 4, and Section 5 provides the estimates of the measure of the excluded set of parameters. In Section 6 some refinement of these measure estimates is undertaken, and in Section 7 we finally observe that the results imply that a certain class of normal forms is structurally stable. The paper concludes with a few remarks relating this paper to previous work, in particular [4] and [7].

2. - The linearized equation

The KAM-theorem is proven by the usual Newton-type iteration procedure, which involves an infinite sequence of coordinate changes and is described in some detail for example in [7]. Each coordinate change Φ is obtained as the time-1-map $X_F^t|_{t=1}$ of a hamiltonian vectorfield X_F. Its generating hamiltonian F as well as some correction $\hat{N}$ to the given normal form N are a solution of the linearized equation

$$\{F, N\} + \hat{N} = R,$$

which is the subject of this section. One then finds that Φ takes the hamiltonian $H = N + R$ into $H \circ \Phi = N_+ + R_+$, where $N_+ = N + \hat{N}$ is the new normal form and $R_+ = \int_0^1 \{(1-t)\hat{N} + tR, F\} \circ X_F^t \mathrm{d}t$ the new error term.

We suppose that in complex coordinates $z = \frac{1}{\sqrt{2}}(u - \mathrm{i}v)$ and $\bar{z} = \frac{1}{\sqrt{2}}(u + \mathrm{i}v)$ we have $N = \langle \omega(\xi), y \rangle + \langle \Omega(\xi), z\bar{z} \rangle$ and

$$R = \sum_{2|m|+|q+\bar{q}| \le 2} \sum_k R_{kmq\bar{q}} e^{\mathrm{i}\langle k,x \rangle} y^m z^q \bar{z}^{\bar{q}}, \tag{3}$$

with coefficients depending on $\xi \in \Pi$, such that $X_R : \mathcal{P}^{a,p} \to \mathcal{P}^{a,\bar{p}}$ is real analytic and Lipschitz in ξ. The mean value of such a hamiltonian is defined as

$$[R] = \sum_{|m|+|q|=1} R_{omqq} y^m z^q \bar{z}^q$$

and is of the same form as N.

LEMMA 1. *Suppose that uniformly on* Π,

$$|\langle k, \omega(\xi)\rangle + \langle l, \Omega(\xi)\rangle| \geq \alpha \frac{\langle l\rangle_d}{A_k}, \quad (k,l) \in Z,$$

where $\alpha > 0$ *and* $A_k \geq 1$. *Then the linearized equation* $\{F, N\} + \hat{N} = R$ *has a solution* F, $\hat{N}$ *that is normalized by* $[F] = 0$, $[\hat{N}] = \hat{N}$, *and satisfies*

$$\|X_{\hat{N}}\|_r \leq \|X_R\|_r, \quad \|X_F\|_{r,-\sigma} \leq \frac{2^6 B_\sigma}{\alpha} \|X_R\|_r,$$
$$\|X_{\hat{N}}\|_r^{\mathcal{L}} \leq \|X_R\|_r^{\mathcal{L}}, \quad \|X_F\|_{r,-\sigma}^{\mathcal{L}} \leq \frac{2^8 B_\sigma}{\alpha}\left(\|X_R\|_r^{\mathcal{L}} + \frac{M}{\alpha}\|X_R\|_r\right),$$

for $0 < \sigma \leq s$, *where* $M = |\omega|^{\mathcal{L}} + |\Omega|_{-\delta}^{\mathcal{L}}$ *and* $B_\sigma^2 = 2^n \sum_k (1+|k|)^2 A_k^4 e^{-2|k|\sigma}$, *and the short hands* $\|\cdot\|_r = \|\cdot\|_{r,D(s,r)}$ *and* $\|\cdot\|_{r,-\sigma} = \|\cdot\|_{r,D(s-\sigma,r)}$ *are used.*

The estimates hold in fact with $|\Omega|_{1-d}^{\mathcal{L}}$ in place of $|\Omega|_{-\delta}^{\mathcal{L}}$, but this slightly better result is not needed later. Concerning the dependence on σ the above estimates are very crude but sufficient for our purposes. Much better estimates have been obtained by Rüssmann – see for example [9].

PROOF. Writing expansions for F and $\hat{N}$ analogous to that for R and using the nonresonance assumptions one finds by comparison of coefficients that $\hat{N} = [R]$ and

$$\mathrm{i} F_{kmq\bar{q}} = \begin{cases} \dfrac{R_{kmq\bar{q}}}{\langle k,\omega\rangle + \langle q-\bar{q}, \Omega\rangle} & \text{for } |k| + |q-\bar{q}| \neq 0, \\ 0 & \text{otherwise,} \end{cases}$$

for all ξ, which is not indicated. With the chosen normalization this solution is also unique.

For the estimates we decompose $R = R^0 + R^1 + R^2$, where R^j comprises all terms with $|q+\bar{q}| = j$, and furthermore

$$R^0 = R^{00},$$
$$R^1 = \langle R^{10}, z\rangle + \langle R^{01}, \bar{z}\rangle,$$
$$R^2 = \langle R^{20} z, z\rangle + \langle R^{11} z, \bar{z}\rangle + \langle R^{02}\bar{z}, \bar{z}\rangle,$$

where the R^{ij} depend on x, ξ, and R^{00} depends in addition on y. With a similar decomposition of F and $\hat{N}$, the linearized equation decomposes into

$$\{F^{ij}, N\} = R^{ij} - [R^{ij}], \quad \hat{N}^{ij} = [R^{ij}],$$

and it suffices to discuss each term individually. In the following we do this for $\dot{R} = R^{10}$ and $\ddot{R} = R^{11}$. To shorten notation, let $\|\cdot\| = \|\cdot\|_{a,\bar{p}}$.

Consider the term $\dot{F} = F^{10}$. We have $\dot{R} = R_z|_{z,\bar{z}=0}$ and thus

$$\|\dot{R}\|_{D(s)} \leq r \,|\!|\!| X_R |\!|\!|_r,$$

where $D(s) = \{|\mathrm{Im}\, x| < s\}$. This is an analytic map into $\ell^{a,\bar{p}}$ with a Fourier series expansion whose coefficients $\dot{R}_k$ satisfy the usual L^2-bound

$$\sum_k \|\dot{R}_k\|^2 e^{2|k|s} \leq 2^n \|\dot{R}\|^2_{D(s)}.$$

Each coefficient is a Lipschitz map $\Pi \to \ell^{a,\bar{p}}$, and the corresponding coefficient of $\dot{F}$ is given by

$$\mathrm{i}\dot{F}_{k,j} = \frac{\dot{R}_{k,j}}{\langle k, \omega \rangle + \Omega_j}, \quad j \geq 1.$$

By the small divisor assumptions we have $|\langle k, \omega\rangle + \Omega_j| \geq \alpha / A_k$ and thus $\|\dot{F}_k\| \leq (A_k/\alpha)\|\dot{R}_k\|$ uniformly on Π. It follows that

$$\begin{aligned}\|\dot{F}\|_{D(s-\sigma)} &\leq \sum_k \|\dot{F}_k\| e^{|k|(s-\sigma)} \\ &\leq \frac{1}{\alpha}\sqrt{\sum_k A_k^2 e^{-2|k|\sigma}}\,\sqrt{\sum_k \|\dot{R}_k\|^2 e^{2|k|s}} \leq \frac{B_\sigma}{\alpha}\|\dot{R}\|_{D(s)},\end{aligned}$$

or $\frac{1}{r}\|\dot{F}\|_{D(s-\sigma)} \leq \frac{B_\sigma}{\alpha} |\!|\!| X_R |\!|\!|_r$.

To control the Lipschitz semi-norm of $\dot{F}$, let $\delta_{k,j} = \langle k, \omega\rangle + \Omega_j$ and $\Delta = \Delta_{\xi\varsigma}$ for ξ, $\varsigma \in \Pi$. Then we have $\mathrm{i}\Delta\dot{F}_{k,j} = \Delta(\delta_{k,j}^{-1}\dot{R}_{k,j}) = \delta_{k,j}^{-1}(\xi)\Delta\dot{R}_{k,j} + \dot{R}_{k,j}(\varsigma)\Delta\delta_{k,j}^{-1}$ and

$$-\Delta\delta_{k,j}^{-1} = \frac{\Delta\delta_{k,j}}{\delta_{k,j}(\xi)\delta_{k,j}(\varsigma)} = \frac{\langle k, \Delta\omega\rangle + \Delta\Omega_j}{\delta_{k,j}(\xi)\delta_{k,j}(\varsigma)}.$$

The small divisor assumptions give $|\delta_{k,j}| \geq \alpha j^d / A_k$. Therefore,

$$|\Delta\delta_{k,j}^{-1}| \leq \frac{A_k^2}{\alpha^2}(|k|\,|\Delta\omega| + |\Delta\Omega_j|_\delta j^{-d})$$

and hence

$$\|\Delta\dot{F}_k\| \leq \frac{A_k}{\alpha}\|\Delta\dot{R}_k\| + \frac{A_k^2}{\alpha^2}(|k|\,|\Delta\omega| + |\!|\Delta\Omega|\!|_{-\delta})\|\dot{R}_k\|.$$

Summing up the Fourier series as before we obtain

$$\|\Delta\dot{F}\|_{D(s-\sigma)} \leq \frac{B_\sigma}{\alpha}\|\Delta\dot{R}\|_{D(s)} + \frac{B_\sigma}{\alpha^2}(|\Delta\omega| + |\!|\Delta\Omega|\!|_{-\delta})\|\dot{R}\|_{D(s)}.$$

Dividing by $|\xi - \varsigma|$ and taking the supremum over $\xi \neq \varsigma$ in Π we arrive at

$$\frac{1}{r} \|\dot{F}\|^{\mathcal{L}}_{D(s-\sigma)} \leq \frac{B_\sigma}{\alpha} \left(\| X_R \|^{\mathcal{L}}_r + \frac{M}{\alpha} \| X_R \|_r \right).$$

Consider now the term $\ddot{F} = F^{11}$. We have $\ddot{R} = \partial_{\bar{z}} \partial_z R$, hence by the generalized Cauchy inequality of Lemma A.3,

$$\|\ddot{R}\|_{D(s)} \leq \frac{1}{r} \|R_z\|_{D(s,r)} \leq \| X_R \|_r$$

in the operator norm for bounded linear operators $\ell^{a,p} \longrightarrow \ell^{a,\bar{p}}$. This is equivalent to the statement that $\tilde{R} = (v_i \ddot{R}_{ij} w_j)$ is a bounded linear operator of ℓ^2 into itself with operator norm $|||\tilde{R}|||_{D(s)} = \|\ddot{R}\|_{D(s)}$, where v_i, w_j are certain weights whose explicit form does not matter here.

Expanding $\tilde{R}$ into its Fourier series with operator valued coefficient $\tilde{R}_k$ we have, as before, $\sum_k |||\tilde{R}_k|||^2 e^{2|k|s} \leq 2^n |||\tilde{R}|||_{D(s)}$. The corresponding coefficient of $\tilde{F}_k = (\tilde{F}_{k,ij})$ is given by

$$\mathrm{i}\tilde{F}_{k,ij} = \frac{\tilde{R}_{k,ij}}{\langle k, \omega \rangle + \Omega_i - \Omega_j}, \qquad |k| + |i - j| \neq 0,$$

while $\tilde{F}_{0,jj} = 0$, and the coefficients $\tilde{R}_{0,jj}$ are absorbed by $\hat{N}$. The small divisor assumptions imply that $|\langle k, \omega \rangle + \Omega_i - \Omega_j| \geq \alpha(1 + |i - j|)/A_k$, since $d \geq 1$. Hence, by Lemma A.1 we obtain $|||\tilde{F}_k||| \leq 3(A_k/\alpha)|||\tilde{R}_k|||$ uniformly in Π, and summing up as before, $|||\tilde{F}|||_{D(s-\sigma)} \leq 3(B_\sigma/\alpha)|||\tilde{R}|||_{D(s)}$. Going back to the operator norm $\|\cdot\|$ and multiplying by z we arrive at

$$\frac{1}{r} \|\ddot{F} z\|_{D(s-\sigma,r)} \leq \frac{3B_\sigma}{\alpha} \| X_R \|_r.$$

The Lipschitz estimate follows the same lines as the one for $\dot{F}$. So we consider $\mathrm{i}\Delta\tilde{F}_{k,ij} = \delta^{-1}_{k,ij} \Delta\tilde{R}_{k,ij} + \tilde{R}_{k,ij} \Delta\delta^{-1}_{k,ij}$ with $\delta_{k,ij} = \langle k, \omega \rangle + \Omega_i - \Omega_j$. Since $2|i^d - j^d| \geq |i - j|(i^{d-1} + j^{d-1})$, the small divisor assumptions imply

$$|\delta_{k,ij}| \geq \frac{\alpha}{2A_k} (i^\delta + j^\delta)|i - j|, \quad i \neq j.$$

We thus obtain

$$|\Delta\delta^{-1}_{k,ij}| \leq \frac{4A_k^2}{\alpha^2} (|k|\,|\Delta\omega| + 2\|\Delta\Omega\|_{-\delta}) \frac{1}{|i - j|}$$

and

$$|||\Delta\tilde{F}_k||| \leq \frac{A_k}{\alpha} |||\Delta\tilde{R}_k||| + \frac{2^5 A_k^2}{\alpha^2} (|k|\,|\Delta\omega| + \|\Delta\Omega\|_{-\delta}) |||\tilde{R}_k|||.$$

This leads to

$$\|\ddot{F}\|^{\mathcal{L}}_{D(s-\sigma)} \leq \frac{B_\sigma}{\alpha} \|\ddot{R}\|^{\mathcal{L}}_{D(s)} + \frac{2^5 B_\sigma}{\alpha^2} M \|\ddot{R}\|_{D(s)}$$

and

$$\frac{1}{r} \|\ddot{F}z\|^{\mathcal{L}}_{D(s-\sigma,r)} \leq \frac{2^5 B_\sigma}{\alpha} \left(\|X_R\|^{\mathcal{L}}_r + \frac{M}{\alpha} \|X_R\|_r \right).$$

The terms F^{10} and F^{11} exhibit all the difficulties involved with infinitely many degrees of freedom. All the other components F^{ij} admit the same estimates, or even better ones. To each component of the hamiltonian vectorfield X_F, at most eight such terms are contributing. The estimates of X_F thus follow.

The estimates of $X_{\hat{N}}$ follow from the observation that $\hat{N}_y$ is the $\mathbb{T}^n$-mean value of R_y, and $\hat{N}_{z\bar{z}}$ is the diagonal of the $\mathbb{T}^n$-mean value of $R_{z\bar{z}}$. ∎

For our purposes the estimates of Lemma 1 may be condensed as follows. For $\lambda \geq 0$, define

$$\|X\|^\lambda_r = \|X\|_r + \lambda \|X\|^{\mathcal{L}}_r.$$

Since we will always use the symbol 'λ' in this rôle, there should be no confusion with exponentiation. Also, $\|\cdot\|^*_r$ stands for either $\|\cdot\|_r$ or $\|\cdot\|^{\mathcal{L}}_r$.

LEMMA 2. *The estimates of Lemma* 1 *imply that*

$$\|X_{\hat{N}}\|^*_{r,D(s,r)} \leq \|X_R\|^*_{r,D(s,r)}, \quad \|X_F\|^\lambda_{r,D(s-\sigma,r)} \leq \frac{aB_\sigma}{\alpha} \|X_R\|^\lambda_{r,D(s,r)}$$

for $0 < \sigma \leq s$ *and* $0 \leq \lambda \leq \alpha/M$ *with some absolute constant* a. *Moreover, if* $A_k = 1 + |k|^\tau$, *then*

$$B_\sigma \leq \frac{b}{\sigma^{2\tau+n+1}} \tag{4}$$

with some constant $b \geq 1$ *depending on* n *and* τ.

3. - The KAM Step

At the general ν-th step of the iteration scheme we are given a hamiltonian $H_\nu = N_\nu + P_\nu$, where $N_\nu = \langle \omega_\nu(\xi), y\rangle + \langle \Omega_\nu(\xi), z\bar{z}\rangle$ is a normal form and P_ν is a perturbation that is real analytic on $D(s_\nu, r_\nu)$. Both are Lipschitz in ξ, which varies over a closed set Π_ν, on which $|\omega_\nu|^{\mathcal{L}} + |\Omega_\nu|^{\mathcal{L}}_{-\delta} \leq M_\nu$ and

$$|\langle k, \omega_\nu(\xi)\rangle + \langle l, \Omega_\nu(\xi)\rangle| \geq \alpha_\nu \frac{\langle l\rangle_d}{A_k}, \quad (k,l) \in Z. \tag{5}$$

For the duration of this section we now drop the index ν and write '+' for '$\nu + 1$' to simplify notation. Thus, $P = P_\nu$, $P_+ = P_{\nu+1}$, and so on. Also, we write $<$ in estimates in order to suppress various multiplicative constants, which depend only on n and τ and could be made explicit, but need not be. Indeed, the only dependence on τ enters through the constant b in (4).

To perform the next step of the iteration we assume that the perturbation is so small that we can choose $0 < \eta < \frac{1}{8}$ and $0 < \sigma < s$, $\sigma \leq 1$, such that

$$\| X_P \|_{r,D(s,r)} + \frac{\alpha}{M} \| X_P \|^{\mathcal{L}}_{r,D(s,r)} \leq \frac{\alpha \sigma^t \eta^2}{c_0}, \tag{6}$$

where $t = 2\tau + n + 2$ and c_0 is some sufficiently large constant depending only on n and τ. On the other hand, for the KAM step we need not assume that the frequency map ω is a homeomorphism or lipeomorphism.

Approximating P. We approximate P by its Taylor polynomial R in y, z, $\bar{z}$ of the form (3). This amounts to corresponding approximations of the partials P_x, P_y, P_z, $P_{\bar{z}}$ which constitute the vectorfield X_P. Since P is analytic, all these approximations are given by certain Cauchy integrals, and the estimates are the same as in a finite dimensional setting. We obtain

$$\| X_R \|^*_{r,D(s,r)} < \| X_P \|^*_r, \quad \| X_R - X_P \|^*_{\eta r, D(s,4\eta r)} < \eta \| X_P \|^*_r. \tag{7}$$

Solution of the linearized equation. Since the small divisor estimates (5) are supposed to hold, we can solve the linearized equation $\{F, N\} + \hat{N} = R$ with the help of Lemmata 1 and 2. Together with the preceding estimate of X_R we obtain

$$\| X_{\hat{N}} \|^*_{r,D(s,r)} < \| X_P \|^*_r, \quad \| X_F \|^\lambda_{r,D(s-\sigma,r)} < \frac{1}{\alpha \sigma^{t-1}} \| X_P \|^\lambda_r, \tag{8}$$

for $0 \leq \lambda \leq \alpha/M$. Furthermore we have $\| DX_F \|_{r,r,D(s-2\sigma,r/2)} < \sigma^{-1} \| X_F \|_{r,D(s-\sigma,r)}$, where on the left we use the operator norm

$$\| L \|_{r,s} = \sup_{W \neq 0} \frac{\| LW \|_{\bar{p},r}}{\| W \|_{p,s}},$$

with $\| \cdot \|_{\bar{p},r}$ defined in (1), and $\| \cdot \|_{p,r}$ defined analogously. This follows by the generalized Cauchy estimate of Lemma A.3 and the observation that every point in $D(s-2\sigma, r/2)$ has at least $\| \cdot \|_{p,r}$-distance $\sigma/2$ to the boundary of $D(s-\sigma, r)$.

Coordinate transformation. The preceding estimates and assumption (6) imply that

$$\frac{1}{\sigma} \| X_F \|_{r,D(s-\sigma,r)}, \quad \| DX_F \|_{r,r,D(s-2\sigma,r/2)} < \frac{\eta^2}{c_0} \tag{9}$$

is small. Hence the flow X_F^t exists on $D(s-3\sigma, r/4)$ for $-1 \le t \le 1$ and takes this domain into $D(s-2\sigma, r/2)$, and by Lemma A.4 we have

$$\| X_F^t - id \|^*_{r,D(s-3\sigma,r/4)} \lessdot \| X_F \|^*_{r,D(s-\sigma,r)} \tag{10}$$

for $-1 \le t \le 1$. Furthermore, by the generalized Cauchy estimate,

$$\| DX_F^t - I \|^*_{r,r,D(s-4\sigma,r/8)} \lessdot \frac{1}{\sigma} \| X_F \|^*_{r,D(s-\sigma,r)}, \tag{11}$$

since any point in $D(s-4\sigma, r/8)$ has $\| \cdot \|_r$-distance greater than $\sigma/32$ to the boundary of $D(s-3\sigma, r/4)$.

The new error term. Subjecting $H = N+P$ to the symplectic transformation $\Phi = X_F^t|_{t=1}$ we obtain the new hamiltonian $H \circ \Phi = N_+ + P_+$ on $D(s-5\sigma, \eta r)$, where $N_+ = N + \hat{N}$ and

$$P_+ = (P - R) \circ X_F^1 + \int_0^1 \{R(t), F\} \circ X_F^t \mathrm{d}t$$

with $R(t) = (1-t)\hat{N} + tR$. Hence, the new perturbing vectorfield is

$$X_{P_+} = (X_F^1)^*(X_P - X_R) + \int_0^1 (X_F^t)^*[X_{R(t)}, X_F] \mathrm{d}t.$$

We will show at the end of this section that for $0 \le t \le 1$,

$$\| (X_F^t)^* Y \|^\lambda_{\eta r, D(s-5\sigma,\eta r)} \lessdot \| Y \|^\lambda_{\eta r, D(s-2\sigma, 4\eta r)}. \tag{12}$$

We already estimated $X_P - X_R$, so it remains to consider the commutator $[X_{R(t)}, X_F]$. To shorten notation we write R for $R(t)$.

On the domain $D(s-2\sigma, r/2)$ we have, using $\bar{p} \ge p$,

$$\| [X_R, X_F] \|_r \le \| DX_R \cdot X_F \|_r + \| DX_F \cdot X_R \|_r \le \| DX_R \|_{r,r} \| X_F \|_r + \| DX_F \|_{r,r} \| X_R \|_r.$$

Using the generalized Cauchy estimate and (7) we get

$$\| [X_R, X_F] \|_{r,D(s-2\sigma,r/2)} \lessdot \sigma^{-1} \| X_P \|_r \| X_F \|_{r,D(s-\sigma,r)}.$$

Similarly, on the same domain,

$$\begin{aligned} \| [X_R, X_F] \|_r^{\mathcal{L}} \le\ & \| DX_R \|^{\mathcal{L}}_{r,r} \| X_F \|_r + \| DX_R \|_{r,r} \| X_F \|_r^{\mathcal{L}} \\ & + \| DX_F \|^{\mathcal{L}}_{r,r} \| X_R \|_r + \| DX_F \|_{r,r} \| X_R \|^{\mathcal{L}}_r \\ \lessdot\ & \sigma^{-1} \| X_P \|^{\mathcal{L}}_r \| X_F \|_{r,D(s-\sigma,r)} + \sigma^{-1} \| X_P \|_r \| X_F \|^{\mathcal{L}}_{r,D(s-\sigma,r)}. \end{aligned}$$

Finally, we have $\|Y\|^\lambda_{\eta r} \le \eta^{-2}\|Y\|^\lambda_r$ for any vectorfield Y. So altogether we obtain

$$\|[X_R, X_F]\|^\lambda_{\eta r, D(s-2\sigma, r/2)} \le \frac{1}{\eta^2}\|[X_R, X_F]\|^\lambda_{r, D(s-2\sigma, r/2)}$$

$$\lessdot \frac{1}{\sigma\eta^2}\|X_P\|^\lambda_r \|X_F\|^\lambda_{r, D(s-\sigma, r)} \lessdot \frac{1}{\alpha\sigma^t\eta^2}(\|X_P\|^\lambda_r)^2$$

for $0 \le \lambda \le \alpha/M$. Collecting all terms we then arrive at the estimate

$$\|X_{P_+}\|^\lambda_{\eta r, D(s-5\sigma, \eta r)} \lessdot \frac{1}{\alpha\sigma^t\eta^2}(\|X_P\|^\lambda_r)^2 + \eta\|X_P\|^\lambda_r, \tag{13}$$

$0 \le \lambda \le \alpha/M$, for the new error term.

The new normal form. This is $N_+ = N + \hat{N}$ with $\|X_{\hat{N}}\|^*_r \lessdot \|X_P\|^*_r$. This implies $|\hat{\omega}| \lessdot \|X_P\|_r$ and $\|\hat{\Omega} z\|_{\bar{p}} \lessdot r\|X_P\|_r$ on $D(s, r)$, hence $|\hat{\Omega}|_{\bar{p}-p} \lessdot \|X_P\|_r$ on Π. The same holds for their Lipschitz semi-norms. With $-\delta \le \bar{p} - p$ we get

$$|\hat{\omega}| + |\hat{\Omega}|_{-\delta} \lessdot \|X_P\|_r, \quad |\hat{\omega}|^{\mathcal{L}} + |\hat{\Omega}|^{\mathcal{L}}_{-\delta} \lessdot \|X_P\|^{\mathcal{L}}_r. \tag{14}$$

In order to bound the small divisors for the new frequencies $\omega_+ = \omega + \hat{\omega}$ and $\Omega_+ = \Omega + \hat{\Omega}$ for $|k| \le K$, K to be chosen later, we observe that $|l|_\delta \le |l|_{d-1} \le 2\langle l\rangle_d$, hence

$$|\langle k, \hat{\omega}\rangle + \langle l, \hat{\Omega}\rangle| \le |k|\,|\hat{\omega}| + |l|_\delta |\hat{\Omega}|_{-\delta} \lessdot |k|\,\|X_P\|_r \langle l\rangle_d \le \hat{\alpha}\,\frac{\langle l\rangle_d}{A_k},$$

with some $\hat{\alpha} \gtrdot A'_K\|X_P\|_r$, where $A'_K = K \max_{|k|\le K} A_k$ and the dot represents some constant. Using the bound for the old divisors, the new ones then satisfy

$$|\langle k, \omega_+(\xi)\rangle + \langle l, \Omega_+(\xi)\rangle| \ge \alpha_+\frac{\langle l\rangle_d}{A_k}, \quad |k| \le K, \tag{15}$$

on Π with $\alpha_+ = \alpha - \hat{\alpha}$. In the next section we will make sure that α_+ is positive.

Proof of estimate (12). Fix $\Phi = X^t_F$ and consider $\Phi^* Y = D\Phi^{-1} Y \circ \Phi$. Then Φ maps $U = D(s - 5\sigma, \eta r)$ into $V = D(s - 4\sigma, 2\eta r)$ by the estimate (9). Hence, $\|\Phi^* Y\|_{\eta r, U} \le \|D\Phi^{-1}\|_{\eta r, \eta r, V}\|Y\|_{\eta r, V}$, and

$$\|D\Phi^{-1}\|_{\eta r, \eta r} \le 1 + \|D\Phi^{-1} - I\|_{\eta r, \eta r} \le 1 + \eta^{-2}\|D\Phi^{-1} - I\|_{r,r} \lessdot 1$$

by (11) and (9). So we have $\|\Phi^* Y\|_{\eta r, U} \lessdot \|Y\|_{\eta r, V}$.

As to the Lipschitz semi-norm we observe that both Φ and Y depend on parameters. Therefore,

$$\|\Delta\Phi^* Y\|_{\eta r, U} \le \|\Delta D\Phi^{-1}\|_{\eta r, \eta r, V}\|Y \circ \Phi\|_{\eta r, U} + \|D\Phi^{-1}\|_{\eta r, \eta r, V}\|\Delta(Y \circ \Phi)\|_{\eta r, U}$$

$$\lessdot \|\Delta D\Phi^{-1}\|_{\eta r, \eta r, V}\|Y\|_{\eta r, V} + \|\Delta Y\|_{\eta r, V} + \|DY\|_{\eta r, r, V}\|\Delta\Phi\|_{r, U}.$$

It follows that

$$\left\| \Phi^* Y \right\|_{\eta r, U}^{\mathcal{L}} \lessdot \left\| D\Phi^{-1} - I \right\|_{\eta r, \eta r, V}^{\mathcal{L}} \left\| Y \right\|_{\eta r, V} + \left\| Y \right\|_{\eta r, V}^{\mathcal{L}} + \left\| DY \right\|_{\eta r, r, V} \left\| \Phi - id \right\|_{r, U}^{\mathcal{L}}$$
$$\lessdot \left\| Y \right\|_{\eta r, V}^{\mathcal{L}} + \frac{1}{\sigma \eta^2} \left\| Y \right\|_{\eta r, W} \left\| X_F \right\|_{r, V}^{\mathcal{L}}$$

with $W = D(s - 2\sigma, 4\eta r)$, using the generalized Cauchy estimate and (10), (11). Since $\lambda \left\| X_F \right\|_{r,V}^{\mathcal{L}} \lessdot \sigma \eta^2$ by (8) and (9), we obtain

$$\left\| \Phi^* Y \right\|_{\eta r, U} + \lambda \left\| \Phi^* Y \right\|_{\eta r, U}^{\mathcal{L}} \lessdot \left\| Y \right\|_{\eta r, W} + \lambda \left\| Y \right\|_{\eta r, W}^{\mathcal{L}},$$

as we wanted to show.

4. - Iteration and Proof of Theorem A

To iterate the KAM step infinitely often we now choose sequences for the pertinent parameters. The guiding principle is to choose a geometric sequence for σ, to minimize the error estimate by choice of η, and to keep α and M essentially constant.

Let c_1 be twice the maximum of all implicit constants obtained during the KAM step and depending only on n and τ. For $\nu \geq 0$ set

$$\alpha_\nu = \frac{\alpha_0}{2}(1 + 2^{-\nu}), \quad M_\nu = M_0(2 - 2^{-\nu}), \quad \lambda_\nu = \frac{\alpha_\nu}{M_\nu},$$

and

$$\epsilon_{\nu+1} = \frac{c_1 \epsilon_\nu^\kappa}{(\alpha_\nu \sigma_\nu^t)^{\kappa - 1}}, \quad \sigma_{\nu+1} = \frac{\sigma_\nu}{2}, \quad \eta_\nu^3 = \frac{\epsilon_\nu}{\alpha_\nu \sigma_\nu^t},$$

where $\kappa = \dfrac{4}{3}$. Furthermore, $s_{\nu+1} = s_\nu - 5\sigma_\nu$, $r_{\nu+1} = \eta_\nu r_\nu$, and $D_\nu = D(s_\nu, r_\nu)$. As initial value fix $\sigma_0 = s_0/40 \leq 1/4$ so that $s_0 > s_1 > \cdots \geq s_0/2$, and assume

$$\epsilon_0 \leq \gamma_0 \alpha_0 \sigma_0^t, \quad \gamma_0 \leq (c_0 + 2^{t+3} c_1)^{-3}, \tag{17}$$

where c_0 appears in (6). Finally, let $K_\nu = K_0 2^\nu$ with $K_0^{\tau+1} = \dfrac{1}{c_1 \gamma_0}$.

Iterative Lemma. *Suppose $H_\nu = N_\nu + P_\nu$ is given on $D_\nu \times \Pi_\nu$, where $N_\nu = \langle \omega_\nu(\xi), y \rangle + \langle \Omega_\nu(\xi), z\bar{z} \rangle$ is a normal form satisfying $|\omega_\nu|^{\mathcal{L}} + \left\| \Omega_\nu \right\|_{-\delta}^{\mathcal{L}} \leq M_\nu$,*

$$|\langle k, \omega_\nu(\xi) \rangle + \langle l, \Omega_\nu(\xi) \rangle| \geq \alpha_\nu \frac{\langle l \rangle_d}{A_k}, \quad (k, l) \in Z,$$

on Π_ν, and

$$\left\| X_{P_\nu} \right\|_{r_\nu, D_\nu}^{\lambda_\nu} \leq \epsilon_\nu.$$

Then there exists a Lipschitz family of real analytic symplectic coordinate transformations $\Phi_{\nu+1} : D_{\nu+1} \times \Pi_\nu \to D_\nu$ *and a closed subset*

$$\Pi_{\nu+1} = \Pi_\nu \setminus \bigcup_{|k|>K_\nu} \mathcal{R}_{kl}^{\nu+1}(\alpha_{\nu+1})$$

of Π_ν, *where*

$$\mathcal{R}_{kl}^{\nu+1}(\alpha_{\nu+1}) = \left\{ \xi \in \Pi_\nu : |\langle k, \omega_{\nu+1}\rangle + \langle l, \Omega_{\nu+1}\rangle| < \alpha_{\nu+1} \frac{\langle l \rangle_d}{A_k} \right\},$$

such that for $H_{\nu+1} = H_\nu \circ \Phi_{\nu+1} = N_{\nu+1} + P_{\nu+1}$ *the same assumptions are satisfied with* $\nu + 1$ *in place of* ν.

PROOF. By induction one verifies that $\epsilon_\nu \leq \gamma_0 \alpha_\nu \sigma_\nu^t / 2^\nu$ for all $\nu \geq 0$. With the definition of η_ν this implies $\epsilon_\nu \leq \alpha_\nu \sigma_\nu^t \eta_\nu^2 / c_0$. So the smallness condition (6) of the KAM step is satisfied, and there exists a transformation $\Phi_{\nu+1} : D_{\nu+1} \times \Pi_\nu \to D_\nu$ taking H_ν into $H_{\nu+1} = N_{\nu+1} + P_{\nu+1}$. The new error satisfies the estimate

$$\begin{aligned} \| X_{P_{\nu+1}} \|_{r_{\nu+1}, D_{\nu+1}}^{\lambda_{\nu+1}} &\leq \frac{c_1}{2} \left(\frac{\epsilon_\nu^2}{\alpha_\nu \sigma_\nu^t \eta_\nu^2} + \eta_\nu \epsilon_\nu \right) \\ &\leq \frac{c_1}{2} \left(\frac{\epsilon_\nu^\kappa}{(\alpha_\nu \sigma_\nu^t)^{\kappa-1}} + \frac{\epsilon_\nu^\kappa}{(\alpha_\nu \sigma_\nu^t)^{\kappa-1}} \right) \leq \epsilon_{\nu+1}. \end{aligned}$$

In view of (14) the Lipschitz semi-norm of the new frequencies is bounded by

$$M_\nu + c_1 \| X_{P_\nu} \|_{r_\nu}^{\mathcal{L}} \leq M_\nu + \frac{c_1 \epsilon_\nu}{\alpha_\nu} M_\nu \leq M_\nu (1 + 2^{-\nu-2}) \leq M_{\nu+1}$$

as required. Finally, one verifies that $c_1 \epsilon_\nu \leq c_1 \gamma_0 \alpha_\nu \sigma_\nu^t \leq (\alpha_\nu - \alpha_{\nu+1}) / A'_{K_\nu}$, hence

$$c_1 A'_{K_\nu} \| X_{P_\nu} \|_{r_\nu} \leq \alpha_\nu - \alpha_{\nu+1}.$$

So by (15) the small divisor estimates hold for the new frequencies with parameter $\alpha_{\nu+1}$ up to $|k| \leq K_\nu$. Removing from Π_ν the union of the resonance zones $\mathcal{R}_{kl}^{\nu+1}(\alpha_{\nu+1})$ for $|k| > K_\nu$ we obtain the parameter domain $\Pi_{\nu+1} \subset \Pi_\nu$ with the required properties. ∎

With (10), (11) and (14) we also obtain the following estimates.

ESTIMATES. *For* $\nu \geq 0$,

$$\begin{aligned} \frac{1}{\sigma_\nu} \| \Phi_{\nu+1} - id \|_{r_\nu, D_{\nu+1}}^{\lambda_\nu}, \ \| D\Phi_{\nu+1} - I \|_{r_\nu, r_\nu, D_{\nu+1}}^{\lambda_\nu} &\leq \frac{c_1 \epsilon_\nu}{\alpha_\nu \sigma_\nu^t}, \\ |\omega_{\nu+1} - \omega_\nu|_{\Pi_\nu}^{\lambda_\nu}, \ |\Omega_{\nu+1} - \Omega_\nu|_{-\delta, \Pi_\nu}^{\lambda_\nu} &\leq c_1 \epsilon_\nu. \end{aligned} \tag{18}$$

PROOF OF THEOREM A. Suppose the assumptions of Theorem A are satisfied. To apply the Iterative Lemma with $\nu = 0$, set $s_0 = s$, $r_0 = r, \ldots, N_0 = N$, $P_0 = P$ and $\gamma = \gamma_0\sigma_0^t$. The smallness condition is satisfied, because

$$\lVert X_{P_0}\rVert_{r_0,D_0}^{\lambda_0} = \lVert X_P\rVert_{r,D(s,r)}^{\lambda} \leq \gamma\alpha = \gamma_0\alpha_0\sigma_0^t = \epsilon_0.$$

The small divisor conditions are satisfied by setting $\Pi_0 = \Pi \setminus \bigcup_{k,l} \mathcal{R}_{kl}^0(\alpha_0)$. Then the Iterative Lemma applies, and we obtain a decreasing sequence of domains $D_\nu \times \Pi_\nu$ and transformations $\Phi^\nu = \Phi_1 \circ \ldots \circ \Phi_\nu$: $D_\nu \times \Pi_{\nu-1} \to D_{\nu-1}$ for $\nu \geq 1$, such that $H \circ \Phi^\nu = H_\nu + P_\nu$. Moreover, the estimate (18) hold.

To prove convergence of the Φ^ν we note that the operator norm $\lVert \cdot \rVert_{r,s}$ satisfies $\lVert AB\rVert_{r,s} \leq \lVert A\rVert_{r,r}\lVert B\rVert_{s,s}$ for $r \geq s$. We thus obtain

$$\lVert \Phi^{\nu+1} - \Phi^\nu \rVert_{r_0,D_{\nu+1}} \leq \lVert D\Phi^\nu\rVert_{r_0,r_\nu,D_\nu} \lVert \Phi_{\nu+1} - id\rVert_{r_\nu,D_{\nu+1}},$$

and

$$\lVert D\Phi^\nu\rVert_{r_0,r_\nu,D_\nu} \leq \prod_{\mu=0}^{\nu} \lVert D\Phi^\mu\rVert_{r_\mu,r_\mu,D_\mu} \leq \prod_{\mu\geq 0}(1+2^{-\mu-2}) \leq 2$$

for all $\nu \geq 0$. Also,

$$\begin{aligned}\lVert \Phi^{\nu+1} - \Phi^\nu \rVert_{r_0,D_{\nu+1}}^{\mathcal{L}} \leq \lVert D\Phi^\nu\rVert_{r_0,r_\nu,D_\nu}^{\mathcal{L}} \lVert \Phi_{\nu+1} - id\rVert_{r_\nu,D_{\nu+1}} \\ + \lVert D\Phi^\nu\rVert_{r_0,r_\nu,D_\nu} \lVert \Phi_{\nu+1} - id\rVert_{r_\nu,D_{\nu+1}}^{\mathcal{L}},\end{aligned}$$

where the first factor is uniformly bounded in a similar fashion. It follows that

$$\lVert \Phi^{\nu+1} - \Phi^\nu \rVert_{r_0,D_{\nu+1}}^{\lambda_0} \lessdot \lVert \Phi_{\nu+1} - id\rVert_{r_\nu,D_{\nu+1}}^{\lambda_\nu}.$$

So the Φ^ν converge uniformly on $\bigcap D_\nu \times \Pi_\nu = D(s/2) \times \Pi_\alpha$ to a Lipschitz continuous family of real analytic torus embeddings $\Phi : \mathbb{T}^n \times \Pi_\alpha \to \mathcal{P}^{a,\bar{p}}$, for which the estimates of Theorem A hold. Similarly, the frequencies ω_ν and Ω_ν converge uniformly on Π_α to Lipschitz continuous limits ω_* and Ω_* with estimates as in Theorem A. The embedded tori are invariant rotational tori, because

$$\lVert X_H \circ \Phi^\nu - D\Phi^\nu \cdot X_{N_\nu}\rVert_{r_\nu,D_\nu} \lessdot \lVert (\Phi^\nu)^* X_H - X_{N_\nu}\rVert_{r_\nu,D_\nu} \leq \lVert X_{P_\nu}\rVert_{r_\nu,D_\nu},$$

whence in the limit, $X_H \circ \Phi = D\Phi \cdot X_{\omega_*}$ for each $\xi \in \Pi_\alpha$, where X_{ω_*} is the constant vectorfield ω_* on $\mathbb{T}^n$.

It remains to prove the characterization of the set Π_α. By construction, $\Pi \setminus \Pi_\alpha$ is the union of the inductively defined resonance zones $\mathcal{R}_{kl}^\nu(\alpha_\nu)$ for $\nu \geq 0$ and $|k| > K_{\nu-1}$, where the involved frequencies ω_ν, Ω_ν are Lipschitz on $\Pi_{\nu-1}$, and $K_{-1} = 0$, $\Pi_{-1} = \Pi$. By Lemma A.2, each coordinate function of $\omega_\nu - \omega$ on Π_ν has a Lipschitz continuous extension to Π preserving minimum, maximum and Lipschitz semi-norm. Since we are using the sup-norm for ω, doing this for each component we obtain an extension $\breve{\omega}_\nu : \Pi \to \mathbb{R}^n$ of ω_ν with $|\breve{\omega}_\nu - \omega|_\Pi^\lambda = |\omega_\nu - \omega|_{\Pi_\nu}^\lambda$. The same applies to Ω_ν. It follows that

$$\mathcal{R}_{kl}^\nu(\alpha_\nu) \subset \left\{ \xi \in \Pi : |\langle k, \breve{\omega}_\nu \rangle + \langle l, \breve{\Omega}_\nu \rangle| < \alpha_0 \frac{\langle l \rangle_d}{A_k} \right\}.$$

The latter are the resonance zones described in Theorem A, if we drop the $\breve{}$. This completes the proof of Theorem A. ∎

Actually, more information may be extracted from the preceding construction. On the domain $D_* \times \Pi_\alpha$, $D_* = D(s/2, r/2)$, the normal forms N_ν converge to $N_* = \langle \omega_*(\xi), y \rangle + \langle \Omega_*(\xi), z\bar{z} \rangle$ with frequencies satisfying

$$|\langle k, \omega_*(\xi) \rangle + \langle l, \Omega_*(\xi) \rangle| \geq \frac{\alpha}{2} \frac{\langle l \rangle_d}{A_k}, \quad (k, l) \in \mathcal{Z},$$

on Π_α. Also, the transformations Φ^ν converge to a Lipschitz family of real analytic, symplectic coordinate transformations

$$\Phi : D_* \times \Pi_\alpha \to D_0,$$

because each Φ^ν is of first order in y and second order in z, $\bar{z}$ only, and the corresponding jets can be shown to converge uniformly on $D(s/2) \times \Pi_\alpha$ with appropriate estimates – see [7]. The limit jet then defines Φ. Finally, one checks that $\Phi^* X_H = X_{N_*} + X_{R_*}$, where R_* is of order 3 at $\mathcal{T}_0^n$. That is, the Taylor series expansion of R_* only contains monomials $y^k z^q \bar{z}^{\bar{q}}$ with $2|k| + |q + \bar{q}| \geq 3$. Thus, the perturbed normal form is transformed back into another normal form up to terms of higher order. In particular, the preserved invariant tori are all linearly stable.

5. - Measure Estimates and Proof of Theorem B

In estimating the measure of the resonance zones it is not necessary to distinguish between the various perturbations ω_ν and Ω_ν of the frequencies, since only the size of the perturbation matters. Therefore, we now write ω' and Ω' for all of them, and we have

$$(19) \qquad |\omega' - \omega|, \ |\Omega' - \Omega|_{-\delta} \leq \alpha, \ |\omega' - \omega|^{\mathcal{L}}, \ |\Omega' - \Omega|_{-\delta}^{\mathcal{L}} \leq \frac{1}{2L}.$$

Similarly, we write $\mathcal{R}'_{kl}$ rather than $\mathcal{R}^\nu_{kl}$ for the various resonance zones.

Let $\Lambda = \{l : 1 \leq |l| \leq 2\}$. We can fix $\sigma > 0$ and a constant $D \geq 1$ such that

$$(20) \qquad \langle l \rangle_d \geq D^{-1} |l|_\sigma |l|_\delta$$

for $l \in \Lambda$ where $|l|_\delta = \sum |l_j| j^\delta$. For example, one may take $\sigma = \min(d, d-1-\delta)$ and $D = \frac{9}{2}$, but such specific choices are not important here.

The proof of Theorem B requires a couple of lemmata.

LEMMA 3. *There exists a positive constant β depending on Ω such that*

$$|\langle l, \Omega'(\xi) \rangle| \geq 2\beta \langle l \rangle_d$$

on Π for all $l \in \Lambda$, provided $|\Omega' - \Omega|_{-\delta} \leq \alpha \leq \beta$.

PROOF. Consider the case $\langle l, \Omega' \rangle = \Omega'_i - \Omega'_j$, which is the subtlest. As to the unperturbed frequencies, $\langle l, \Omega \rangle \neq 0$ on Π by assumption A, and

$$\frac{\langle l, \Omega \rangle}{\langle l \rangle_d} \to 1$$

uniformly in ξ by assumption B. Hence there exists a $\beta > 0$ such that $|\langle l, \Omega \rangle| \geq 3D\beta \langle l \rangle_d$ on Π for all $l \in \Lambda$. The result for the perturbed frequencies then follows with $|\langle l, \Omega - \Omega' \rangle| \leq |l|_\delta |\Omega - \Omega'|_{-\delta} \leq D\beta |l|_\sigma^{-1} \langle l \rangle_d \leq D\beta \langle l \rangle_d$. ∎

LEMMA 4. *If $\mathcal{R}'_{kl}(\alpha) \neq \emptyset$ and $\alpha \leq \beta$, then*

$$|k| \geq \vartheta \langle l \rangle_d$$

with $\vartheta = \dfrac{\beta}{|\omega|_\Pi + 1}$.

PROOF. If $\mathcal{R}'_{kl}(\alpha)$ is not empty, then $|\langle k, \omega'(\xi) \rangle + \langle l, \Omega'(\xi) \rangle| < \alpha \langle l \rangle_d$ at some point ξ in Π, and thus $|k|\,|\omega'(\xi)| \geq |\langle l, \Omega'(\xi) \rangle| - \alpha \langle l \rangle_d \geq 2\beta \langle l \rangle_d - \alpha \langle l \rangle_d \geq \beta \langle l \rangle_d$ by Lemma 3. ∎

LEMMA 5. *If $|k| \geq 8LM|l|_\delta$, then*

$$|\mathcal{R}'_{kl}(\alpha)| < c_3 \frac{\alpha}{A_k},$$

with $c_3 = \vartheta^{-1} L^n M^{n-1} \rho^{n-1}$ and $\rho = \operatorname{diam} \Pi$.

PROOF. We introduce the unperturbed frequencies $\varsigma = \omega(\xi)$ as parameters over the domain $\Delta = \omega(\Pi)$ and consider the resonance zones $\mathcal{R}^\Delta_{kl} = \omega(\mathcal{R}'_{kl})$ in Δ. Keeping the old notation for the frequencies we then have $\omega = id$,

$$|\omega' - id|^{\mathcal{L}} \leq \frac{1}{2}, \quad |\Omega'|^{\mathcal{L}}_{-\delta} \leq 2LM$$

for the perturbed frequencies as functions of ς by (19) and $LM \geq 1$.

Now consider $\mathcal{R}^\Delta_{kl}(\alpha)$. Let $\phi(\varsigma) = \langle k, \omega'(\varsigma) \rangle + \langle l, \Omega'(\varsigma) \rangle$. Choose a vector $v \in \{-1, 1\}^n$ such that $\langle k, v \rangle = |k|$ and write $\varsigma = rv + w$ with $r \in \mathbb{R}$, $w \in v^\perp$. As a function of r, we then have, for $t > s$,

$$\langle k, \omega'(\varsigma)\rangle|_s^t = \langle k, \varsigma\rangle|_s^t + \langle k, \omega'(\varsigma) - \varsigma\rangle|_s^t \geq |k|(t-s) - \frac{1}{2}|k|(t-s) = \frac{1}{2}|k|(t-s)$$

and

$$|\langle l, \Omega'(\varsigma)\rangle|_s^t| \leq |l|_\delta |\Omega'|_{-\delta}^{\mathcal{L}}(t-s) \leq 2LM|l|_\delta(t-s) \leq \frac{1}{4}|k|(t-s).$$

Hence, $\phi(rv+w)|_s^t \geq \frac{1}{4}|k|(t-s)$ uniformly in w. It follows that

$$\{r : rv + w \in \Delta, |\phi(rv+w)| \leq \delta\} \subseteq \{r : |r - r_0(w)| \leq 4\delta|k|^{-1}\}$$

with r_0 depending miserably on w, and hence

$$|\mathcal{R}_{kl}^{\Delta}(\alpha)| \leq 4(\operatorname{diam}\Delta)^{n-1}\alpha \cdot \frac{\langle l\rangle_d}{A_k|k|}$$

by Fubini's theorem. Going back to the original parameter domain Π by the inverse frequency map ω^{-1} and observing that $\operatorname{diam}\Delta \leq 2M \operatorname{diam}\Pi$ and $\langle l\rangle_d \leq \vartheta^{-1}|k|$, the final estimate follows. ∎

Now let

$$L_* = \frac{8DLM}{\vartheta}, \quad K_* = 8LM \max_{|l|_\sigma \leq L_*} |l|_\delta,$$

where ϑ and σ are defined in Lemma 4 and (20), respectively. Assume $\alpha \leq \beta$ from now on. The preceding three lemmata then lead to the following conclusion.

LEMMA 6. *If* $|k| \geq K_*$ *or* $|l|_\sigma \geq L_*$, *then*

$$|\mathcal{R}_{kl}'(\alpha)| \lessdot c_3 \frac{\alpha}{A_k}.$$

The same holds for $k \neq 0$, $l = 0$.

PROOF. If $\mathcal{R}_{kl}'(\alpha)$ is not empty and $|l|_\sigma \geq L_*$, then

$$|k| \geq \vartheta\langle l\rangle_d \geq \vartheta D^{-1}|l|_\sigma|l|_\delta \geq 8LM|l|_\delta.$$

But if $|l|_\sigma \leq L_*$, then $|k| \geq K_*$ also implies $|k| \geq 8LM|l|_\delta$. So in both cases, Lemma 5 applies. The case $l = 0$ follows directly from Lemma 5. ∎

Next we consider the "resonance classes"

$$\mathcal{R}_k'(\alpha) = \bigcup_{l\in\Lambda}^{*} \mathcal{R}_{kl}'(\alpha),$$

where the star indicates that we exclude the finitely many resonance zones with $0 < |k| < K_*$ and $0 < |l|_\sigma < L_*$. Note that $\mathcal{R}_{kl}'(\alpha)$ is empty for $k = 0$ and $\alpha \leq \beta$ by Lemma 3.

LEMMA 7. *If $d > 1$, then*

$$|\mathcal{R}'_k(\alpha)| \lessdot c_4 \alpha \frac{|k|^s}{A_k}$$

with $s = \dfrac{2}{d-1}$ *and* $c_4 = \dfrac{c_3}{\vartheta^s}$.

PROOF. By Lemma 4 we may restrict the star-union to $\langle l \rangle_d \le \vartheta^{-1}|k|$, and since $2\langle l \rangle_d \ge |l|_{d-1}$,

$$\operatorname{card}\{l : \langle l \rangle_d \le \vartheta^{-1}|k|\} \le \operatorname{card}\{l : |l|_{d-1} \le 2\vartheta^{-1}|k|\} \lessdot \frac{|k|^s}{\vartheta^s}.$$

The result now follows with Lemma 6. ∎

Recall that for $d = 1$ we have a $\kappa > 0$ and a constant $a \ge 1$ such that

$$\left| \frac{\Omega_i - \Omega_j}{i - j} - 1 \right| \le \frac{a}{j^\kappa}, \quad i > j.$$

LEMMA 8. *If $d = 1$, then*

$$|\mathcal{R}'_k(\alpha)| \lessdot c_5 \alpha^{\kappa^*} \frac{|k|^2}{A_k^{\delta^*}},$$

with $c_5 = \dfrac{ac_3}{\vartheta^2}$, *where* $x^* = \dfrac{|x|}{1+|x|}$ *for real* x.

PROOF. Write $\Lambda = \Lambda^+ \cup \Lambda^-$, where Λ^- contains those $l \in \Lambda$ with two non-zero components of opposite sign, and Λ^+ contains the rest. For $l \in \Lambda^+$ we have $\langle l \rangle_d = |l|$, hence $\operatorname{card}\{l \in \Lambda^+ : \langle l \rangle_d \le \vartheta^{-1}|k|\} \lessdot \vartheta^{-2}|k|^2$ and

$$\left| \bigcup_{l \in \Lambda^+}^{*} \mathcal{R}_{kl}(\alpha) \right| \lessdot c_5 \alpha \frac{|k|^2}{A_k}$$

as in the previous proof.

The minus-case, however, requires more consideration. For $l \in \Lambda^-$ we have $\langle l, \Omega' \rangle = \Omega'_i - \Omega'_j$ and $\langle l \rangle_d = |i - j|$, and up to an irrelevant sign, l is uniquely determined by the two integers $i \ne j$. We may suppose that $i - j = m > 0$. Then $|\langle l, \Omega' - \Omega \rangle| \le |l|_\delta |\Omega' - \Omega|_{-\delta} \le \alpha(i^\delta + j^\delta)$ and $|\langle l, \Omega \rangle - m| \le amj^{-\kappa}$. Therefore

$$\begin{aligned} \mathcal{R}'_{kij}(\alpha) &= \left\{ \xi : |\langle k, \omega' \rangle + \langle l, \Omega' \rangle| < \frac{\alpha m}{A_k} \right\} \\ &\subseteq Q_{kmj} \overset{\text{def}}{=} \left\{ \xi : |\langle k, \omega' \rangle + m| < \frac{\alpha m}{A_k} + \frac{2\alpha}{j^{-\delta}} + \frac{am}{j^\kappa} \right\}. \end{aligned}$$

Moreover, $Q_{kmj} \subseteq Q_{kmj_0}$ for $j \geq j_0$. For fixed $m \leq \vartheta^{-1}|k|$, we then obtain

$$\left| \bigcup_{i-j=m}^{*} \mathcal{R}'_{kij}(\alpha) \right| \leq \sum_{j<j_0} |\mathcal{R}'_{kij}(\alpha)| + |Q_{kmj_0}| < c_3 \left(\frac{j_0 \alpha}{A_k} + \frac{\alpha}{j_0^{-\delta}} + \frac{a}{j_0^{\kappa}} \right). \tag{21}$$

By choosing either $j_0^{1-\delta} = A_k$ or $\alpha j_0^{1+\kappa} = A_k$, whichever gives the better estimate, and using the assumption $-\delta \leq \kappa$ we arrive at

$$\left| \bigcup_{i-j=m}^{*} \mathcal{R}'_{kij}(\alpha) \right| < a c_3 \left(\frac{\alpha}{A_k^{\delta^*}} + \frac{\alpha^{\kappa^*}}{A_k^{\kappa^*}} \right) < a c_3 \frac{\alpha^{\kappa^*}}{A_k^{\delta^*}}.$$

Summing over m,

$$\left| \bigcup_{l \in \Lambda^-}^{*} \mathcal{R}'_{kl}(\alpha) \right| < \sum_{|m| \leq \vartheta^{-1}|k|} \left| \bigcup_{i=j=m}^{*} \mathcal{R}'_{kij}(\alpha) \right| < c_5 |k| \frac{\alpha^{\kappa^*}}{A_k^{\delta^*}}.$$

The two cases together give the final estimate. ∎

PROOF OF THEOREM B. We can choose τ so that

$$\sum_{|k| \geq K} \frac{|k|^s}{A_k}, \sum_{|k| \geq K} \frac{|k|^2}{A_k^{\delta^*}} \leq \sum_{|k| \geq K} \frac{1}{1 + |k|^{n+1}} < \frac{1}{1+K}.$$

For example,

$$\tau \geq \begin{cases} n + 1 + \dfrac{2}{d-1} & \text{for } d > 1, \\ (n+3) \dfrac{\delta - 1}{\delta} & \text{for } d = 1. \end{cases} \tag{22}$$

Letting $\mathcal{X} = \{(k,l) : 0 < |k| < K_*, 0 < |l|_\sigma < L_*\}$ we then obtain

$$\left| \bigcup_{(k,l) \notin \mathcal{X}} \mathcal{R}^{\nu}_{kl}(\alpha) \right| \leq \sum_{\nu \geq 0} \sum_{|k| > K_{\nu-1}} |\mathcal{R}^{\nu}_{k}(\alpha)| < \sum_{\nu \geq 0} \frac{c_6 \alpha^{\mu}}{1 + K_{\nu-1}} < c_6 \alpha^{\mu}$$

by the definition of the resonance classes $\mathcal{R}^{\nu}_{k}(\alpha)$ with μ as in Theorem B and a constant c_6 of the form $\tilde{c}(\operatorname{diam}\Pi)^{n-1}$, where $\tilde{c}$ does not increase when the parameter domain Π decreases. This gives the required estimate. Finally, if $\delta \leq 0$, then $|l|_\delta \leq 2$ for all l and hence $K_* \leq 16LM$. This proves Theorem B. ∎

PROOF OF COROLLARY C. By choosing $\tilde{\gamma} \leq \gamma/2LM$ the frequencies ω_ν and Ω_ν satisfy the assumptions of Theorem B, and thus

$$\left| \bigcup_{(k,l) \notin \mathcal{X}} \mathcal{R}^{\nu}_{kl}(\alpha) \right| \to 0 \quad \text{as } \alpha \to 0.$$

Choosing, in the definition $\gamma = \gamma_0\sigma_0^t$, also $\gamma_0 \leq \dfrac{1}{c_3 K_*^{\tau+1}}$ in addition to (17), then

$$K_0^{\tau+1} = \frac{1}{c_3\gamma_0} \geq K_*^{\tau+1},$$

so the remaining resonance zones are all defined in terms of the unperturbed frequencies. Hence, by Assumption A, the monotonicity of $\mathcal{R}_{kl}^0(\alpha)$ in α and the boundedness of Π, we have $|\mathcal{R}_{kl}^0(\alpha)| \to 0$ as $\alpha \to 0$ for each $(k,l) \in \mathcal{X}$. Since $\mathcal{X}$ is finite, also

$$\left| \bigcup_{\mathcal{X}} \mathcal{R}_{kl}^0(\alpha) \right| \to 0 \quad \text{as } \alpha \to 0,$$

which gives the claim. Finally, if $\delta \leq 0$, then $K_* \leq 16LM$. ∎

6. - Proof of Theorem D

To prove Theorem D we precede the KAM iteration by *one* modified KAM step. For this preparatory step the small divisor estimates (5) are used with a parameter

$$\hat{\alpha} = \alpha^{1-3w} > \alpha,$$

where $w > 0$ is chosen later. Moreover, for $\langle l, \Omega\rangle = \Omega_i - \Omega_j$, $i \neq j$, we use the modified estimate

$$|\langle k, \omega\rangle + \Omega_i - \Omega_j| \geq \frac{\hat{\alpha}}{A_k} \cdot \frac{|i-j|}{\min(i^\pi, j^\pi)} \tag{23}$$

with positive $\pi < \bar{p} - p$. The upshot is that the measure estimates are improved at the expense of deteriorating the regularity of the vectorfield.

Using the modified small divisor estimates in the solution of the linearized equation we obtain

$$\|X_F\|_{\check{p},r,D(s-\sigma,r)}^{\lambda} \leq \frac{aB_\sigma}{\hat{\alpha}} \|X_R\|_{\bar{p},rD(s,r)}^{\lambda}, \quad \check{p} = \bar{p} - \pi > p.$$

Since $\hat{\alpha} > \alpha$, the KAM step applies under the same assumptions as before, but now the estimates of X_F are to be understood in terms of the weaker norm $\|\cdot\|_{\check{p},r}$. Accordingly, the vectorfield of the next perturbation P_0 – the starting point for the iteration – is also bounded in this norm only. Using the notation of Section 4 we obtain

$$\|X_{P_0}\|_{\check{p},\eta r}^{\lambda} \lessdot \frac{1}{\hat{\alpha}\sigma^t\eta^2} (\|X_P\|_{\bar{p},r}^{\lambda})^2 + \eta\|X_P\|_{\bar{p},r}^{\lambda} \lessdot \frac{1}{(\hat{\alpha}\sigma^t)^{\kappa-1}} (\|X_P\|_{\bar{p},r}^{\lambda})^{\kappa}$$

by choosing $\eta^3 = \hat{\alpha}^{-1}\sigma^{-t}\|X_P\|_{\bar{p},r}^{\lambda}$. With the assumption $\|X_P\|_{\bar{p},r}^{\lambda} \leq \gamma\alpha$, the choices $\sigma = \sigma_0$, $\gamma \leq \gamma_0\sigma_0^t$ (as for the first step of the iteration) and $\hat{\alpha} = \alpha^{1-3w}$ we obtain

$$\|X_{P_0}\|_{\check{p},\eta r}^{\lambda} \leq \gamma\check{\alpha}, \quad \check{\alpha} = \alpha^{1+w} < \alpha.$$

For the frequencies ω_0, Ω_0 of the new normal form N_0 the usual estimates (14) hold with $-\delta \le \bar{p} - p$. It is not necessary, however, to keep track of the small divisor estimates for the new frequencies, since the KAM scheme now starts from scratch, with parameters $\check{\alpha}$ and $\check{p}$ instead of α and $\bar{p}$, respectively.

We estimate the measure of the resonance zones eliminated in the first and the subsequent steps. To this end fix τ as in (22) assuming $-\delta \le \check{p} - p$. For brevity, the notation '$\lessdot$' now includes also constants that depend on π and are of the same form as the constants $c_3, \ldots$ in Section 5.

Let $\Xi_{\hat{\alpha}} = \bigcup_{\mathcal{Z}} \mathcal{R}^s_{kl}(\hat{\alpha})$ be the union of the resonance zones eliminated in the preparatory step and defined in terms of the modified small divisor estimates.

LEMMA 9.

$$|\Xi_{\hat{\alpha}}| \lessdot \hat{\alpha}^{\lambda}, \quad \lambda = \begin{cases} 1 & \textit{for } \pi > 1, \\ \dfrac{\kappa}{\kappa + 1 - \pi} & \textit{for } \pi < 1. \end{cases}$$

PROOF. We first show that the estimate of Lemma 8 changes to

$$|\mathcal{R}^s_k(\hat{\alpha})| \lessdot \hat{\alpha}^{\lambda} \frac{|k|^2}{A_k^{\lambda}}. \tag{24}$$

The estimate for $l \in \Lambda^+$ is the same as before, giving a contribution of the size $\hat{\alpha}\,\dfrac{|k|^2}{A_k}$. For $l \in \Lambda^-$ and $\pi > 1$ we have $|\mathcal{R}^s_{kij}(\hat{\alpha}) \lessdot \dfrac{\hat{\alpha}}{A_k j^{\pi}}$, and the sum over all j converges to a similar contribution. For $l \in \Lambda^-$ and $\pi < 1$, however, the modified small divisor estimate (23) gives

$$\begin{aligned} \mathcal{R}^s_{kij}(\hat{\alpha}) &= \left\{ \xi : |\langle k, \omega \rangle + \langle l, \Omega \rangle| < \frac{\hat{\alpha} m}{A_k j^{\pi}} \right\} \\ &\subseteq Q_{kmj} \stackrel{\text{def}}{=} \left\{ \xi : |\langle k, \omega \rangle + m| < \frac{\hat{\alpha} m}{A_k j^{\pi}} + \frac{a m}{j^{\kappa}} \right\}. \end{aligned}$$

There is no contribution from $\Omega' - \Omega$ here, since we are dealing with the unperturbed frequencies. For fixed m we then obtain

$$\left| \bigcup_{i-j=m}^{*} \mathcal{R}^s_{kij}(\hat{\alpha}) \right| \lessdot \frac{\hat{\alpha}}{A_k} \sum_{j \le j_0} \frac{1}{j^{\pi}} + \frac{a}{j_0^{\kappa}} \lessdot \frac{\hat{\alpha}}{A_k j_0^{\pi - 1}} + \frac{a}{j_0^{\kappa}} \lessdot \frac{\hat{\alpha}^{\lambda}}{A_k^{\lambda}}$$

by choosing $j_0^{\kappa+1-\pi} = A_k / \hat{\alpha}$. Then (24) follows by symming over m.

Summing (24) over k we obtain one contribution to the estimate of $|\Xi_{\hat{\alpha}}|$. The other contribution is due to the finitely many resonance zones $\mathcal{R}^s_{kl}(\hat{\alpha})$ with $(k, l) \in \mathcal{X}$. In each of them, $\langle k, \omega \rangle + \langle l, \Omega \rangle$ is a nontrivial affine function of ξ, so one has $|\mathcal{R}^s_{kl}(\hat{\alpha})| \lessdot \hat{\alpha}$. This proves the lemma. ∎

The KAM iteration now starts with the parameter set $\Pi_0 = \Pi \backslash \Xi_{\hat{\alpha}}$ and parameter $\check{\alpha} = \alpha^{1+w}$.

LEMMA 10. *For sufficiently small* α,

$$|\Pi_0 \backslash \Pi_{\check{\alpha}}| \lessdot \max(\alpha, \check{\alpha}^{\mu}), \quad \mu = \frac{\kappa}{\kappa+1} = \kappa^*.$$

PROOF. We show that now the estimate of Lemma 8 changes to

$$|\mathcal{R}'_k(\check{\alpha})| \lessdot \max(\alpha, \check{\alpha}^{\kappa^*}) \frac{|k|^2}{A_k^{\delta^*}}.$$

Again, the estimate for $l \in \Lambda^+$ is the same. For $l \in \Lambda^-$, there is a contribution of order α to the estimate of $\Omega' - \Omega$ from the preparatory step. So instead of (21) we have

$$\left| \bigcup_{i-j=m}^{*} \mathcal{R}'_{kij}(\check{\alpha}) \right| \lessdot \frac{j_0 \check{\alpha}}{A_k} + \frac{\alpha}{j_0^{-\delta}} + \frac{a}{j_0^{\kappa}}.$$

By proper choice of j_0 this gives the bound $\max(\alpha, \check{\alpha}^{\kappa^*}) A_k^{-\delta^*}$ and hence the estimate of $|\mathcal{R}'_k(\check{\alpha})|$. The rest of the proof is analogous to the preceding one. Just note that the functions $\langle k, \omega' \rangle + \langle l, \Omega' \rangle$ are Lipschitz close of order α to nontrivial affine functions of ξ. ∎

The proof of Theorem D is now almost complete. The two lemmata combined give

$$|\Pi \backslash \Pi_\alpha| \lessdot \hat{\alpha}^{\lambda} + \check{\alpha}^{\mu}.$$

For $\pi < \bar{p} - p \leq 1$ the right hand side is minimized by choosing $w = \dfrac{\pi}{4\kappa + 4 - \pi}$, so that

$$\frac{1-3w}{\kappa+1-\pi} = \frac{1+w}{\kappa+1},$$

hence $\hat{\alpha}^{\lambda} = \check{\alpha}^{\mu} = \alpha^{\tilde{\mu}}$ with $\tilde{\mu} = \dfrac{\kappa}{\kappa + 1 - \pi/4}$. This proves Theorem D.

7. - Structural Stability

The results may be used to show that a certain class of hamiltonians is structurally stable. Let

$$N = \langle \omega(\xi), y \rangle + \frac{1}{2} \langle \Omega(\xi), u^2 + v^2 \rangle + \tilde{N}$$

be a hamiltonian on some phase space $\mathcal{P}^{a,p}$ depending on parameters $\xi \in \Pi \subset \mathbb{R}^n$, Π a closed bounded set of positive Lebesgue measure. Let us say that H is a *regular normal form* if the following three conditions are satisfied, with notations as in Section 1.

Condition A^*: Nondegeneracy and Nonresonance. The map $\xi \mapsto \omega(\xi)$ is a lipeomorphism between Π and its image. Moreover, there exist positive constants α_0 and τ_0 such that

$$|\langle k, \omega(\xi)\rangle + \langle l, \Omega(\xi)\rangle| \geq \frac{\alpha_0 \langle l \rangle_d}{1 + |k|^{\tau_0}} \tag{25}$$

for all $(k, l) \in Z$ and $\xi \in \Pi$, where d is defined in condition B^*.

Condition B^*: Spectral Asymptotics. There exist $d \geq 1$, $\delta < d - 1$ and a fixed sequence $\bar{\Omega}$ with $\bar{\Omega}_j = j^d + \ldots$, such that $\Omega_j = \bar{\Omega}_j + \tilde{\Omega}_j$, where the tails $\tilde{\Omega}_j$ define a Lipschitz continuous map $\tilde{\Omega} : \Pi \to \ell_\infty^{-\delta}$.

Condition C^*: Regularity. For each $\xi \in \Pi$ the hamiltonian vectorfield $X_{\tilde{N}}$ defines near $\mathcal{T}_0^n$ a real analytic map

$$X_{\tilde{N}} : \mathcal{P}^{a,p} \to \mathcal{P}^{a,\bar{p}}, \qquad \begin{cases} \bar{p} \geq p \text{ for } d > 1, \\ \bar{p} > p \text{ for } d = 1, \end{cases}$$

which is Lipschitz in ξ and where $\tilde{N}$ is of order 3 at $\mathcal{T}_0^n$ as defined at the end of Section 4.

THEOREM E. *A regular normal form N is structurally stable under sufficiently small perturbations of the same regularity as $\tilde{N}$. That is, for every such perturbation H of N, there exists another Cantor set $\Pi_* \subset \Pi$ of positive Lebesgue measure and a Lipschitz family of real analytic, symplectic coordinate transformations Φ near $\mathcal{T}_0^n$, such that $\Phi^* X_H = X_{N_*}$ with another regular normal form N_* with respect to Π_*.*

PROOF. Let N be a regular normal form. Then assumptions A and B are satisfied, and the parameters L, M, τ and κ are fixed. Theorem B implies that for the union of resonance zones $\mathcal{R}(\alpha) = \bigcup \mathcal{R}_{kl}^{\nu}(\alpha) \subset \Pi$ defined in terms of arbitrary but sufficiently small perturbations of the frequencies ω and Ω as in Theorem A, we have $|\mathcal{R}(\alpha)| \to 0$ as $\alpha \to 0$. Hence, the measure of any of the sets Π_α in Theorem A converges uniformly to the measure of Π as α tends to zero.

Now fix α small enough to make this measure positive. By condition C^*,

$$\| X_{\tilde{N}} \|_{r,D(s,r)} + \frac{\alpha}{M} \| X_{\tilde{N}} \|_{r,D(s,r)}^{\mathcal{L}} < \frac{1}{2} \tilde{\gamma} \alpha \tag{26}$$

for all small positive r and s. Then Theorems A and B apply to every perturbation $H = N + P$ of N, where P is of the same regularity as $\tilde{N}$ and satisfies the same estimate (26) for some positive r and s. We obtain a Cantor set $\Pi_* \subset \Pi$ of positive measure and, by the remark at the end of Section 4, a family of real analytic, symplectic coordinate transformations Φ near $\mathcal{T}_0^n$ such that $\Phi^* X_H = X_{N_+} + X_{R_+}$, where R_+ is of order 3 at $\mathcal{T}_0^n$. Moreover, the frequencies of N_+ satisfy the diophantine conditions (25) with parameters $\alpha/2$ and τ. It follows that the hamiltonian $N_* = N_+ + R_+$ is a regular normal form. ∎

8. - Concluding Remarks

REMARK 1. The regularity condition may be written in the form

$$p - \bar{p} < d - 1, \quad p - \bar{p} \leq 0.$$

In the framework of differential operators on Sobolev spaces, d and $p - \bar{p}$ may be identified with the orders of the linear and nonlinear part of the associated differential operator L, respectively. Thus, L has to be *quasi-linear* by the first condition, and its nonlinear part has to be *bounded* by the second condition.

The first assumption is rather natural. Nonlinearities of the same order as the linearity may cause the blow up of every nontrivial solution [6], so quasi-periodic solutions may not exist at all. The second condition, however, is not necessary, but makes the proof and the result more transparent. It happens to be satisfied by the nonlinear Schrödinger and wave equations in [5, 8]. It may be removed for $d > 1$ at the expense of a more convoluted proof, so that the theorem also applies for example to perturbations of the KdV equation. See [3] as well as a forthcomming publication by S. Kuksin for more details.

REMARK 2. The results of this paper improve on the results obtained in [7] in many ways: – the phase space can be chosen appropriately to suit applications to nonlinear partial differential equations; – the nondegeneracy condition is weaker; – the dependence on the parameters ξ need only be Lipschitz; – the frequencies Ω may only grow linearly, thus violating the finiteness condition in [7].

Moreover, a flaw in the proof of Lemma 8.1 in [7] is fixed, that was pointed out to the author by H. Rüssmann. There not only the t-derivative of the function $\Phi(w + tv)$, but also its Lipschitz semi-norm needs to be controlled in order to obtain the desired measure estimate. Such an estimate is provided here.

Due to the weaker nondegeneracy assumption the result above gives no control over the rate of convergence in the measure estimate (2). However, with more information about the unperturbed frequencies such control is easily obtained. For example, suppose that ω and Ω are differentiable on Π, that $\omega : \Pi \to \Delta$ is a diffeomorphism, and that for some $\alpha_0 > 0$, $\mathcal{R}^0_{kl}(\alpha_0) = \emptyset$ for each $(k, l) \in \mathcal{X}$ for which k lies in the closed convex hull of the set of gradients $\{\partial_\varsigma \langle l, \Omega \circ \omega^{-1}(\varsigma) \rangle : \varsigma \in \Delta\}$. Then the arguments of Lemma 5 and Lemma 8.1 in [7] show that

$$\left| \bigcup_{\mathcal{X}} \mathcal{R}^0_{kl}(\alpha) \right| \leq \tilde{c} \rho^{n-1} \alpha,$$

recovering the result of [7].

REMARK 3. We finally compare our results with those of Kuksin in [4]. By and large, the basic KAM theorems are the same, with the same range of applications to partial differential equations. There are, however, some differences: – the nondegeneracy condition of Assumption A is weaker, as a certain collection of exact resonances is only required to be of measure zero; – we can allow for Sobolev spaces $\ell^{a,p}$ of *exponentially* decreasing sequences by letting $a > 0$, which avoids *a posteriori* arguments about the analyticity of the solutions obtained by the KAM theorem; – the dependence of the measure estimates on the asymptotic properties of the eigenvalues λ_j in the case $d = 1$ is made explicit in terms of the exponent μ in Theorem B. Indeed, in [4] this point was overlooked, and the estimates for this case such as (4.11) on page 77 are not correct. This was later corrected in *An Erratum* available from Sergej Kuksin; see also Appendix 2 in [1].

Another difference is in the proofs. Here, in Theorem A as well as in its proof, the unperturbed hamiltonian N describes a *linear* system of equations, and higher order integrable terms are simply considered as perturbations as well. In Kuksin's set up, the unperturbed system also may contain *nonlinear* terms. This considerably complicates the handling of the linearized equation, and many more careful estimates are required. On the other hand, it provides some greater flexibility in applying the results.

This, however, seems to be of advantage only in the subtle case of small amplitude solutions u of the nonlinear wave equation

$$u_{tt} = u_{xx} - mu - au^3 + O(u^5), \quad m > 0, \quad a \neq 0,$$

on $[0, \pi]$. Here, one has $d = 1$, and the problem is to find sets of nonresonant frequencies of positive measure in the presence of a "small twist". Still, the results of Bobenko and Kuksin [1] for this equation are not better than in [8], because on the other hand, they had to cope with worse asymptotic properties of the frequencies, namely $\kappa = 1$ instead of $\kappa = 2$ as in [8]. – Combining both approaches, one could also handle $O(u^4)$-terms. But such a small improvement requires quite a big effort.

Acknowledgement. This paper was written while the author was visiting the Forschungsinstitut für Mathematik at the ETH Zürich. I like to thank the institute for its stimulating working atmosphere, excellent working conditions and very helpful staff, and the Deutsche Forschungsgemeinschaft for their financial support through a Heisenberg grant. In particular, it is a pleasure to thank Jürgen Moser and Sergej Kuksin for many fruitful discussions on the subject.

A. - Utilities

LEMMA A.1. *If $A = (A_{ij})$ is a bounded linear operator on ℓ^2, then also*

$B = (B_{ij})$ with

$$B_{ij} = \frac{|A_{ij}|}{|i-j|}, \quad i \neq j,$$

and $B_{ii} = 0$ is a bounded linear operator on ℓ^2, and $\|B\| \leq \frac{\pi}{\sqrt{3}} \|A\|$.

PROOF. By the Schwarz inequality, we have

$$\sum_{j \geq 1} |B_{ij}|, \ \sum_{i \geq 1} |B_{ij}| \leq \|A\| \sqrt{\sum_{k \neq 0} \frac{1}{k^2}} = \lambda \|A\|, \quad \lambda = \frac{\pi}{\sqrt{3}},$$

for all i and j. Hence, again by Schwarz,

$$\begin{aligned} \|Bv\|^2 &\leq \sum_i \Big(\sum_j |B_{ij}|\,|v_j| \Big)^2 \leq \sum_i \Big(\sum_j |B_{ij}| \Big) \Big(\sum_j |B_{ij}|\,|v_j|^2 \Big) \\ &\leq \lambda \|A\| \sum_i \sum_j |B_{ij}|\,|v_j|^2 \leq \lambda^2 \|A\|^2 \sum_j |v_j|^2 = \lambda^2 \|A\|^2 \|v\|^2. \ \blacksquare \end{aligned}$$

LEMMA A.2. *Let $F \subset \mathbb{R}^n$ be closed and $u : F \to \mathbb{R}$ a bounded Lipschitz continuous function. Then there exists an extension $U : \mathbb{R}^n \to \mathbb{R}$ of u, which preserves minimum, maximum and Lipschitz semi-norm of u.*

PROOF. Let $\lambda = |u|_F^{\mathcal{L}}$, and define

$$\tilde{u}(x) = \sup_{\xi \in F} (u(\xi) - \lambda |x - \xi|)$$

for $x \in \mathbb{R}^n$. This is an extension of u to all of $\mathbb{R}^n$. By the triangle inequality, $\tilde{u}(x) \geq u(\xi)\lambda|x' - \xi| - \lambda|x - x'|$ for all $\xi \in F$ and hence $\tilde{u}(x) \geq \tilde{u}(x') - \lambda|x - x'|$. Interchanging x and x', we get

$$\frac{|\tilde{u}(x) - \tilde{u}(x')|}{|x - x'|} \leq \lambda.$$

It follows that $|\tilde{u}|_{\mathbb{R}^n}^{\mathcal{L}} = |u|_F^{\mathcal{L}}$. Replacing $\tilde{u}$ above $\max_F u$ by $\max_F u$ does not change its Lipschitz semi-norm, and similarly below $\min_F u$. The resulting function U has all the required properties. ■

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

$$\|d_v G\|_{F,E} = \max_{u \neq 0} \frac{\|d_v G(u)\|_F}{\|u\|_E}.$$

The Cauchy inequality can be stated as follows.

LEMMA A.3. *Let G be an analytic map from the open ball of radius r around v in E into F such that $\|G\|_F \leq M$ on this ball. Then*

$$\|d_v G\|_{F,E} \leq \frac{M}{r}.$$

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

$$\|d_0 f\|_F = \|d_v G(u)\|_F \leq \frac{M}{r} \cdot \|u\|_E$$

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

Let V be an open domain in a real Banach space E with norm $\|\cdot\|$, Π a subset of another real Banach space, and $X : V \times \Pi \to E$ a parameter dependent vectorfield on V, which is C^1 on V and Lipschitz on Π. Let ϕ^t be its flow. Suppose there is a subdomain $U \subset V$ such that $\phi^t : U \times \Pi \to V$ for $-1 \leq t \leq 1$.

LEMMA A.4. *Under the preceding assumptions,*

$$\|\phi^t - id\|_U \leq \|X\|_V,$$
$$\|\phi^t - id\|_U^{\mathcal{L}} \leq \exp(\|DX\|_V)\|X\|_V^{\mathcal{L}},$$

for $-1 \leq t \leq 1$, where all norms are understood to be taken also over Π.

PROOF. Let $0 \leq t \leq 1$. We have $\phi^t - id = \int_0^t X \circ \phi^s \mathrm{d}s$, so the first estimate is clear. To prove the second one, let $\Delta\phi^t = \phi^t(\cdot, \xi) - \phi^t(\cdot, \varsigma)$ for $\xi, \varsigma \in \Pi$. Then

$$\Delta(\phi^t - id) = \int_0^t \Delta(X \circ \phi^s)\mathrm{d}s = \int_0^t \Delta X \circ \phi^s \mathrm{d}s + \int_0^t (X \circ \phi^s(\cdot, \xi) - X \circ \phi^s(\cdot, \varsigma))\mathrm{d}s,$$

hence

$$\|\Delta(\phi^t - id)\|_U \leq \int_0^t \|\Delta X\|_V \mathrm{d}s + \int_0^t \|DX\|_V \|\Delta\phi^s\|_U \mathrm{d}s$$

$$\leq \|\Delta X\|_V + \|DX\|_V \int_0^t \|\Delta(\phi^s - id)\|_U \mathrm{d}s.$$

With Gronwall's inequality it follows that $\|\Delta(\phi^t - id)\|_U \leq \|\Delta X\|_V \exp(t\|DX\|_V)$. Dividing by the norm of $\xi - \varsigma$ and taking the supremum over $\xi \neq \varsigma$ in Π the Lipschitz estimate follows. ∎

REFERENCES

[1] A.I. Bobenko - S.B. Kuksin, *The nonlinear Klein-Gordon equation on an interval as a perturbed sine-Gordon equation*, Comment. Math. Helv. **70** (1995), 63-112.

[2] L.H. Eliasson, *Perturbations of stable invariant tori for Hamiltonian systems*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. **15** (1988), 115-147.

[3] S.B. Kuksin, *Perturbation theory of conditionally periodic solutions of infinite dimensional Hamiltonian systems and its application to the Korteweg-de Vries equation*, Mat. Sb. **136** (1988); English transl. in Math. USSR-Sb. **64** (1989), 397-413.

[4] S.B. Kuksin, Nearly integrable infinite-dimensional Hamiltonian systems, Lecture Notes in Mathematics 1556, Springer, Berlin, 1993.

[5] S.B. Kuksin - J. Pöschel, *Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation*, Ann. of Math. (to appear).

[6] P.D. Lax, *Developement of singularities of solutions of nonlinear hyperbolic partial differential equation*, J. Math. Phys. **5** (1964), 611-613.

[7] J. Pöschel, *On elliptic lower dimensional tori in Hamiltonian systems*, Math. Z. **202** (1989), 559-608.

[8] J. Pöschel, *Quasi-periodic solutions for a nonlinear wave equation*, Comment. Math. Helv. (to appear).

[9] H. Rüssmann, *On an inequality for trigonometric polynomials in several variables*, Analysis, et cetera (P.H. Rabinowitz and E. Zehnder, eds.), Academic Press, Boston, pp. 545-562.

[10] C.E. Wayne, *Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory*, Comm. Math. Phys. **127** (1990), 479-528.

Mathematisches Institut A
Universität Stuttgart
Pfaffenwaldring 57
D-70569 Stuttgart
Germany