Commun. Math. Phys. 118, 119–161 (1988)

Communications in
Mathematical
Physics

Construction of Analytic KAM Surfaces and Effective Stability Bounds

Alessandra Celletti* and Luigi Chierchia**

Forschungsinstitut für Mathematik, ETH-Zentrum, CH-8092 Zürich, Switzerland

To our friend and colleague Paola Calderoni

Abstract. A class of analytic (possibly) time-dependent Hamiltonian systems with d degrees of freedom and the "corresponding" class of area-preserving, twist diffeomorphisms of the plane are considered. Implementing a recent scheme due to Moser, Salamon and Zehnder, we provide a method that allows us to construct "explicitly" KAM surfaces and, hence, to give lower bounds on their breakdown thresholds. We, then, apply this method to the Hamiltonian $H \equiv y^2/2 + \varepsilon(\cos x + \cos(x-t))$ and to the map $(y, x) \to (y + \varepsilon \sin x, x + y + \varepsilon \sin x)$ obtaining, with the aid of computer-assisted estimations, explicit approximations (within an error of $\sim 10^{-5}$) of the golden-mean KAM surfaces for complex values of ε with $|\varepsilon|$ less or equal than, respectively, 0.015 and 0.65. (The experimental numerical values at which such surfaces are expected to disappear are about, respectively, 0.027 and 0.97.) A possible connection between break-down thresholds and singularities in the complex ε-plane is pointed out.

Table of Contents

---

* Supported by Consiglio Nazionale delle Ricerche, Italy
** Permanent address: Dipartimento di Matematica, II Università di Roma, I-00173 Roma, Italy

1. Introduction

a) Problem and Results

As it is well known from Kolmogorov-Arnold-Moser (KAM) theory [21, 1, 27] most of the invariant surfaces of an integrable system do not disappear under the effect of a small perturbation but give rise to invariant tori on which the motion is quasi-periodic with (highly) incommensurate frequencies ("KAM surfaces"). For models with few degrees of freedom, numerical investigations (see, e.g., [18, 17]) and some rigorous results [26, 24] have shown that, if the strength of the perturbation is large enough, KAM surfaces break down.

The existence of these surfaces is particularly relevant for stability theory. In fact, for systems with no more than two degrees of freedom, KAM tori separate the phase-space into disjoint invariant sets making thus possible confinement of motions. Also in higher dimension, where confinement is no longer possible [2], the existence and location of invariant sets might be relevant for practical purposes in view of the slow rate of diffusion allowed by Nekhoroshev's theorem [31, 5, 4, 37].

At a more phenomenological level, the breakdown of KAM tori seems also to be closely related to the "onset of chaos" [18, 14, 20, 12]. In particular, it is believed that, as the perturbative parameter is increased, there is, in a suitable sense, a "last" KAM surface to disappear [17, 23, 12].

In this paper we shall address the problem of providing a rigorous and constructive method able to yield, in concrete cases, "good" lower bounds on the breakdown threshold.

The model that we will mainly consider is a class of (possibly) time dependent Hamiltonian systems with d degrees of freedom with real-analytic Hamiltonian given, in standard canonical coordinates, by

$$H(y,x,t;\varepsilon)\equiv\frac{y^2}{2}+f(x,t;\varepsilon),\qquad y^2\equiv y\cdot y\equiv\sum_{i=1}^{d}\frac{y_i^2}{2},\tag{H}$$

where f has period 2π in each variable $x_1,\ldots,x_d,t$ and depends analytically on the parameter ε.

Several physical systems are represented by such Hamiltonians. An example borrowed from statistical mechanics, describing a system of d rotators with short range interaction, is given by (H) with

$$f(x,t;\varepsilon)\equiv f(x;\varepsilon)\equiv\varepsilon\sum_{i=1}^{d-1}\cos(x_{i+1}-x_i).$$

(For a KAM and Nekhoroshev analysis of these systems see [36, 37].) A low-dimensional example, which will be of particular interest to us, is given by

$$H=\frac{y^2}{2}+\varepsilon[\cos x+\cos(x-t)],\qquad(d=1).\tag{H1}$$

This Hamiltonian, which is the central object of the renormalization theory of [14] (see also [12, 23]), governs the motion of a particle of charge ε, subject to the potential of two longitudinal (electrostatic) waves. The study of (H1) is also relevant in celestial mechanics: After minor modifications and under suitable assumptions, (H1) is a good description of a rigid body on an elliptic orbit with spin-axis parallel to the largest principal moment of inertia and perpendicular to the orbit plane [38].

Often, in applications with few degrees of freedom, it is preferred to work with area-preserving mappings, obtained as Poincaré sections, rather than directly with Hamiltonian systems, and it might be useful to have available a method that can be applied to maps too. Therefore, we will also consider the following class of area-preserving twist maps [which might be considered a formal analogous of (H)]:

$$(y,x)\xrightarrow{\phi}(y_1,x_1)\equiv(y-f_x(x;\varepsilon),x+y-f_x(x;\varepsilon)), \tag{M}$$

f being a real-analytic function periodic in x. For $f=\varepsilon\cos x$, ϕ is the well known Chirikov-Greene standard map.

We will use the models (H) and (M) to illustrate a new KAM technique that allows us to construct analytic KAM surfaces and to have a careful control of the quantities involved.

We will then apply this technique to the Hamiltonian (H1) and to the standard map, proving the existence of the "golden-mean KAM surface" for (complex) values of ε with

$$|\varepsilon|\leqq 0.015 \text{ (H1)}, \qquad |\varepsilon|\leqq 0.65 \text{ (standard map)}.$$

Furthermore, "explicit" approximations to such surfaces will be provided with an error of order 10^{-5}.

To compare these results, we first report the numerical (non-rigorous) expectations. The breakdown threshold for the standard map is believed to be ~ 0.971 [9, 17]. Less settled is a numerical determination of a reliable value for the threshold in the Hamiltonian case; however Escande, using a (non-rigorous) method based on renormalization theory, indicates a value of ~ 0.0276 ([13], see also [14]) and Greene's residue criterion [17], applied to a Poincaré section for (H1) (the so-called "leap-frog integrator with large step size") yields a value of ~ 0.02758 [15].

As for known rigorous results, we recall that there are no homotopically non-trivial invariant curves for the standard map for values of $\varepsilon\geqq 0.985$ [24]. A lower bound on the existence of the golden-mean KAM surface, given by Herman [19], yields a ratio with the numerical expectation of 1/33 for the standard map, while a ratio of 1/40 is obtained in [8] in the Hamiltonian case.

Finally, we mention that numerical extrapolations of our methods give results in good agreement with the above numerical expectations.

(b) KAM *Method*

Let us first consider the Hamiltonian case. We recall that a KAM surface with given frequency (or "rotation") vector $(\omega_1,\dots,\omega_d)\equiv\omega\in\mathbf{R}^d$ for (H) is a $(d+1)$-

dimensional torus described parametrically by

$$\left.\begin{aligned}&(\theta_1,\ldots,\theta_d,\mathrm{t})=(\theta,\mathrm{t})\in\mathbf{T}^{d+1}\to(\theta+u(\theta,t),t)\in\mathbf{T}^{d+1}, \qquad \mathbf{T}\equiv\mathbf{R}/2\pi\mathbf{Z},\\ &\det(I+u_\theta)\neq 0, \qquad (I+u_\theta)_{ij}=\delta_{ij}+\frac{\partial u_i}{\partial\theta_j}, \qquad (1\leqq i,j\leqq d),\end{aligned}\right\} \tag{T}$$

where u is a (vector-valued) function, depending on the parameters ε and ω, with the property that the flow induced by H in the (θ,t)-coordinates linearizes in

$$(\theta_0,t_0)\to(\theta_0+\omega t,t_0+t).$$

This definition, together with Hamilton's equation, is equivalent to require that u satisfies

$$D^2u_i+\frac{\partial f}{\partial x_i}(\theta+u,t;\varepsilon)=0, \qquad i=1,\ldots,d$$

or, more compactly

$$D^2u+f_x(\theta+u,t;\varepsilon)=0, \tag{E}$$

where D is the constant vector field on $\mathbf{T}^{d+1}$ given by

$$D\equiv\omega\cdot\partial_\theta+\partial_t\equiv\sum_{i=1}^{d}\omega_i\frac{\partial}{\partial\theta_i}+\frac{\partial}{\partial t}.$$

Remark 1. Equation (E) plays an important role in Percival's analysis [32] and (with $d=1$) in a particular case of Moser's generalization [29] of Aubry-Mather's theory [3, 25]. For a numerical treatment of (E) (with some special f) see [6].

Remark 2. Notice that if $u(\theta,t;\varepsilon)$ is a solution of (E) so is $(\theta,t)\to c+u(c+\theta,t;\varepsilon)$ for any constant vector c. In the following we will only consider solutions u with vanishing average on $\mathbf{T}^{d+1}$.

Following [30] and [35], one can solve, under suitable hypotheses, Eq. (E) using a Newton iteration procedure. Namely, one starts with an approximate solution of (E), i.e., with a function v for which the error term

$$e=D^2v+f_x(\theta+v,t;\varepsilon) \tag{AE}$$

is small, and constructs, solving a linearization of (AE) a new approximation, v', for which the relative error term satisfies $|e'|\sim O(|e|)^2$. In order to carry out such a procedure, we require that

$$\det\mathcal{M}\neq 0, \qquad \mathcal{M}\equiv I+v_\theta, \tag{C}$$

[which, in view of (T), seems quite natural] and that ω satisfies the standard strong-irrationality assumptions: We assume that exists a number $\tau\geqq d$ such that

$$\gamma\equiv\sup_{\substack{0\neq n\in Z^d\\ m\in Z}}(|\omega\cdot n+m|\,|n|^\tau)^{-1}<\infty, \tag{DC}$$

where

$$\omega\cdot n\equiv\sum_{i=1}^{d}\omega_i n_i, \qquad |n|\equiv\left(\sum_{i=1}^{d}n_i^2\right)^{1/2}.$$

The set of such "Diophantine" vectors will be denoted by $\mathscr{D}_d$ and, from now on, we will attach to any $\omega \in \mathscr{D}_d$ some τ (possibly the smallest) for which (DC) holds together with the relative constant γ.

Thus, starting from some approximate solution $v^{(0)}$ and applying iteratively the Newton step, one obtains a sequence of new approximants $v^{(j)}$, provided (C) holds for any j. One needs, then, to have a quantitative control of the functions involved in the procedure.

For this purpose, we will construct an algorithm (which will refer to as "KAM algorithm") that given upper bounds on norms relative to the approximants $v^{(j)}$ and $e^{(j)}$ provides upper bounds on the corresponding norms of the next approximants $v^{(j+1)}$ and $e^{(j+1)}$. Here, the norms refer to a suitably chosen scale of Banach spaces to be described later (compare Sect. 5 below). We then say that such KAM algorithm converges if

$$\tilde{M}^{(j)} < \infty\,, \qquad \forall j \geqq 0\,, \qquad \lim_{j\to\infty} E^{(j)} = 0\,, \tag{K}$$

where $\tilde{M}^{(j)}$ and $E^{(j)}$ are upper bounds on the norms of $(\mathscr{M}^{(j)})^{-1}$ and $e^{(j)}$. If (K) holds one obtains a solution of (E) as (uniform) limit of the $v^{(j)}$'s; if, for some $j \geqq 0$, $\tilde{M}^{(j)}$ becomes infinite, we say that the algorithm diverges.

It is quite remarkable that, with a finite amount of computations, one can usually decide with reasonable precision whether, for a given initial approximation $v^{(0)}$, the KAM algorithm converges or not.

To give an example, consider the system with Hamiltonian (H1) and let ω be the golden-mean $(\sqrt{5}-1)/2$, for which $\tau = 1$ and $\gamma = (\sqrt{5}+3)/2$. The KAM-torus equation takes the form

$$D^2 u = \varepsilon[\sin(\theta + u) + \sin(\theta + u - t)]\,, \qquad D \equiv \omega \frac{\partial}{\partial\theta} + \frac{\partial}{\partial t}\,. \tag{E1}$$

An obvious (but rather bad) initial choice is $v^{(0)} \equiv 0$, for which

$$e^{(0)} \equiv \varepsilon[\sin\theta + \sin(\theta - t)]\,.$$

In this situation the KAM algorithm presented below converges for $|\varepsilon| \leqq 0.000028$ but diverges at $j = 7$ for $|\varepsilon| = 0.000029$.

To explain this fact we observe that one way to prove the convergence of the KAM algorithm, which we will actually follow, is to find a simple explicit condition ($\equiv$"KAM condition") that if satisfied, for some j_0, by $M^{(j_0)}$ $(\geqq |\mathscr{M}^{(j_0)}|)$, $\tilde{M}^{(j_0)}$ and $E^{(j_0)}$ yields (K). Now, roughly speaking, if the algorithm converges then the KAM condition will eventually be satisfied and the fast rate of convergence of the scheme makes usually possible to check the condition after only few steps (typically $10 \sim 20$).

We discuss now briefly the mapping case. Analogously to the Hamiltonian case, a KAM curve with frequency (or rotation) number ω is a circle represented parametrically by

$$\theta \in \mathbf{T} \to \theta + u(\theta) \in \mathbf{T}\,, \qquad 1 + u_\theta \neq 0\,,$$

so that in the θ-coordinates, ϕ^n corresponds to

$$\theta_0 \to \theta_0 + n\omega, \qquad n \in \mathbf{Z}.$$

The KAM-curve equation for (M) is the readily seen to be

$$D^2 u + f_x(\theta + u; \varepsilon) = 0, \tag{EM}$$

where D denotes here the linear finite-difference operator

$$(Du)(\theta) = u\left(\theta + \frac{\omega}{2}\right) - u\left(\theta - \frac{\omega}{2}\right). \tag{DM}$$

At this point, the above discussion for the Hamiltonian systems holds word-by-word in the present situation with the only exception of the Diophantine condition (DC), which becomes now $\omega/2\pi \in \mathscr{D}_1$, necessary to control the inverse of D.

(c) The Initial Guess

The efficiency of a Newton algorithm is of course related to the initial guess.

The choice of $v^{(0)}$, which in conjunction with the above KAM algorithm, will yield the mentioned results for (H1) and the standard map, is related to the analyticity properties of the KAM surfaces *in the parameter* ε.

As already pointed out in [28], KAM surfaces [of systems like (H) with $f(x,t;0)=0$] are analytic in ε near the origin (for a new proof avoiding the use of a Newton method, see [11]). In fact, a trivial byproduct of the above Newton scheme will be that if one starts with an approximant v, which is analytic in ε in some domain $B \subset \mathbf{C}$ and if (C) is satisfied uniformly in B, then also v' is analytic in B.

Thus, it seems quite natural to try to compute explicitly a few terms of the ε-expansion of a KAM-torus.

Consider (H) and let, for simplicity, $d=1$ and $f(x,t;\varepsilon) = \varepsilon g(x,t)$. Then inserting the series

$$u(\theta, t; \varepsilon) \equiv \sum_{l=1}^{\infty} u^{(l)}(\theta, t)\varepsilon^l$$

into (E) and comparing powers of ε, one gets

$$D^2 u^{(1)} = -g_x(\theta, t) \tag*{(EP)$_1$}$$

$$D^2 u^{(l+1)} = -\sum_{k \in K_l} (\partial_x^{k_1 + \ldots + k_l} g_x) \prod_{i=1}^{l} \frac{(u^{(i)})^{k_i}}{k_i!}, \qquad l \geqq 1, \tag*{(EP)$_l$}$$

where K_l is the set of all non-negative integer vectors $k = (k_1, \ldots, k_l) \in \mathbf{N}^l$ such that

$$\sum_{i=1}^{l} i k_i = l.$$

Notice that these are linear equations and that the right-hand side of (EP)$_l$ is a combination of $u^{(1)}, \ldots, u^{(l)}$. Thus, one can solve (EP)$_l$ iteratively. [The same formulae hold for the mapping case (M) if one drops the t.]

For example, in the (H1) case, one obtains immediately

$$u^{(1)} = -\left[\frac{1}{\omega^2}\sin\theta + \frac{1}{(\omega-1)^2}\sin(\theta-t)\right],$$

$$u^{(2)} = \frac{1}{2}\left[\frac{1}{4\omega^4}\sin 2\theta + \left(\frac{1}{\omega^2}+\frac{1}{(\omega-1)^2}\right)\frac{1}{(2\omega-1)^2}\sin(2\theta-t)\right.$$

$$\left. + \left(\frac{1}{\omega^2}-\frac{1}{(\omega-1)^2}\right)\sin t + \frac{\sin 2(\theta-t)}{4(\omega-1)^4}\right].$$

Our initial approximate solution will be

$$v^{(0)} \equiv \sum_{l=1}^{l_0} u^{(l)}\varepsilon^l \tag{IG}$$

with $l_0 = 24$ for (H1) and $l_0 = 38$ for the standard map.

(d) The Role of Computers

Even though the solution of (EP) is completely elementary, the concrete calculation of $u^{(l)}$ is not a trivial task: Even in the simple (H1)-case, computing (IG) with $l_0 = 24$ means to evaluate 2756 non-zero Fourier coefficients.

Here enters the aid of computers, which may be used to give rigorous lower and upper bounds on the result of (possibly) lengthy operations between real numbers. A possible way of using rigorously a computer is to perform the so-called interval-arithmetic in the fashion of [22] or [10]. This is the strategy that we followed in order to evaluate the Fourier coefficients of $v^{(0)}$, using a VAX 8600. Actually, we used mechanical computations also for the evaluation of the norms relative to the initial approximation and for the application of the KAM algorithm; however the latter computations are sensibly simpler and faster than the former.

The above choice of the "order" l_0 has been made so as to obtain a compromise between a (relatively) little amount of computations and "reasonable" quantitative results; compare, also, Remark 13 of Appendix D.

In proving our results, we will be careful in clearly separating the theoretical parts from the computational ones and we will provide and comment the main computer program that we used.

(e) Concluding Remark

The existence and construction of smooth but not analytic KAM surfaces for a given system is a relevant and difficult problem. Even in the case of the standard map it is not known whether a given KAM curve undergoes, as ε is increased, a gradual loss of smoothness or if the transition from analyticity to discontinuity is an "instantaneous" phenomenon. Such problems remain beyond the reach of the techniques presented here. However, let ϱ_m denote the "maximum radius of ε-analyticity" for a given KAM-torus, i.e.,

$$\varrho_m \equiv \inf_{(\theta,t)} \left\{\text{radius of convergence of } \sum u^{(l)}\varepsilon^l\right\}.$$

It is quite clear that implementing further the above ideas and techniques one could in principle prove existence of KAM surfaces for $|\varepsilon|<\varrho_m$. Our results and their numerical extrapolations (compare Sect. 9) seem to indicate that, at least for some special models, ϱ_m might actually coincide with the breakdown threshold. This would not only support the second hypothesis considered above, but would also show a deep connection between ε-singularities and dynamics.

2. Newton Method

In this section we describe the Newton iteration procedure and solve the associated linearized equation [$\equiv$Eq. (2.1) below]. To stress the algebraic character of this section we do not specify yet the functional spaces in which we will work.

Let v and e satisfy equation (AE) with $\mathcal{M}\equiv I+v_\theta$ invertible $\left(\text{as above } (v_\theta)_{ij}=\frac{\partial v_i}{\partial\theta_j}\right)$ and denote with a superscript T matrix transposition. Then one has the following

Lemma 1 (Moser-Salamon-Zehnder). *Let z be a solution of*

$$D(\mathcal{M}^T\mathcal{M}Dz)=-\mathcal{M}^Te. \tag{2.1}$$

Then, setting

$$w\equiv\mathcal{M}z,\qquad v'\equiv v+w,$$

the following equation holds:

$$D^2v'+f_x(\theta+v',t;\varepsilon)=e' \tag{2.2}$$

with

$$e'\equiv e_\theta z+q_1+q_2,$$

where, denoting by f_{xx} the matrix with entries $(f_{xx})_{ij}=\frac{\partial^2 f}{\partial x_i\partial x_j}$,

$$q_1\equiv f_x(\theta+v+w,t;\varepsilon)-f_x(\theta+v,t;\varepsilon)-f_{xx}(\theta+v,t;\varepsilon)w,$$

$$q_2\equiv(\mathcal{M}^T)^{-1}\mathcal{A}Dz,\qquad \mathcal{A}\equiv\mathcal{M}^TD\mathcal{M}-(D\mathcal{M}^T)\mathcal{M}.$$

Furthermore the matrix-valued function $\mathcal{A}$ satisfies

$$\begin{cases}\langle\mathcal{A}\rangle\equiv\int_{\mathbf{T}^{d+1}}\mathcal{A}\frac{d\theta dt}{(2\pi)^{d+1}}=0\\ D\mathcal{A}=\mathcal{M}^Te_\theta-e_\theta^T\mathcal{M}.\end{cases} \tag{2.3}$$

Remark 3. For Eq. (2.1) to make sense, $\mathcal{M}^Te$ must have vanishing mean value (over $\mathbf{T}^{d+1}$) and that this is indeed the case follows from (AE). In fact, denoting by δ_{ij} the Kronecker symbol,

$$\int_{\mathbf{T}^{d+1}}(\mathcal{M}^Te)_i=\sum_{l=1}^d\int\frac{\partial v_l}{\partial\theta_i}D^2v_l+\left(\delta_{li}+\frac{\partial v_l}{\partial\theta_i}\right)f_{x_l}(\theta+v,t;\varepsilon)$$

$$=\sum_{l=1}^d\int\left(D^2\frac{\partial v_l}{\partial\theta_i}\right)v_l+\int\frac{\partial}{\partial\theta_i}f(\theta+v,t;\varepsilon)=0.$$

Proof of Lemma 1. Using the definitions of w, v', and e', (2.2) can be rewritten as

$$D^2 v + D^2(\mathcal{M} z) + f_x + f_{xx}\mathcal{M} z = e_\theta z + q_2\,,$$

where f_x and f_{xx} are evaluated at $(\theta + v, t; \varepsilon)$. Then, equation (AE) implies

$$D^2(\mathcal{M} z) + f_{xx}\mathcal{M} z = -e + e_\theta z + q_2\,. \tag{2.4}$$

Take the gradient with respect to θ of (AE) to get

$$D^2\mathcal{M} + f_{xx}\mathcal{M} = e_\theta\,, \tag{2.5}$$

so that (2.4) becomes

$$D^2(\mathcal{M} z) - (D^2\mathcal{M})z = -e + q_2\,.$$

Now, use the definition of q_2 to get

$$\mathcal{M}^T D^2(\mathcal{M} z) - \mathcal{M}^T(D^2\mathcal{M})z = -\mathcal{M}^T e + (\mathcal{M}^T D\mathcal{M} - (D\mathcal{M}^T)\mathcal{M})Dz\,,$$

which will be easily recognized as Eq. (2.1).

It remains to prove (2.3). Integrating by parts one obtains, for any i, j,

$$\int_{\mathbf{T}^{d+1}} \mathcal{A}_{ij} = -2\int [(D\mathcal{M}^T)\mathcal{M}]_{ij} = 2\sum_{l=1}^{d} \int v_l D\,\frac{\partial^2}{\partial\theta_i\partial\theta_j}\,v_l = 0\,.$$

Finally, Eq. (2.5) and its transposed will yield easily the second equation in (2.3). □

We proceed now to solve (2.1), referring to the next sections for the precise assumptions and estimates. For a (vector or matrix-valued) function on $\mathbf{T}^{d+1}$, with vanishing mean value $\langle h\rangle$, we denote

$$(D^{-1}h)(\theta, t) \equiv \sum_{\substack{(n,m)\in\mathbf{Z}^{d+1}\\ (n,m)\neq 0}} \frac{\hat{h}_{(n,m)}}{i(\omega\cdot n + m)}\, e^{i(n\cdot\theta + mt)}\,,$$

where the hat denotes Fourier coefficients and the dot the standard inner product. Then, the unique solution of (2.1) for which

$$\langle w\rangle \equiv \int_{\mathbf{T}^{d+1}} w\,\frac{d\theta dt}{(2\pi)^{d+1}} \equiv \langle \mathcal{M} z\rangle = 0 \tag{2.6}$$

is given by

$$z = D^{-1}\{(\mathcal{M}^T\mathcal{M})^{-1}[c_0 - D^{-1}(\mathcal{M}^T e)]\} + c_1\,, \tag{2.7}$$

where the constant c_0 is chosen so as to be able to invert D the second time and c_1 so as to have (2.6):

$$\begin{aligned} c_0 &\equiv \langle (\mathcal{M}^T\mathcal{M})^{-1}\rangle^{-1}\langle (\mathcal{M}^T\mathcal{M})^{-1}D^{-1}(\mathcal{M}^T e)\rangle\,,\\ c_1 &\equiv -\langle \mathcal{M} D^{-1}\{(\mathcal{M}^T\mathcal{M})^{-1}[c_0 - D^{-1}(\mathcal{M}^T e)]\}\rangle\,. \end{aligned} \tag{2.8}$$

Remark 4 (The Mapping Case). To adapt this section to the mapping case (M), (EM), (DM) one needs simply to make the following modifications. Set $d=1$ and consider t-independent functions of $\theta\in\mathbf{T}$ (this corresponds to substitute $\mathbf{T}^{d+1}$ with

$\mathbf{T}$ in the above formulae); substitute the expression $\mathcal{M}^T\mathcal{M}$ [appearing in (2.1), (2.7), and (2.8)] with

$$\left(1+v_\theta\left(\theta+\frac{\omega}{2}\right)\right)\left(1+v_\theta\left(\theta-\frac{\omega}{2}\right)\right);$$

finally for h with mean-value (on $\mathbf{T}$) zero define $D^{-1}h$, in the obvious way, by setting

$$(D^{-1}h)(\theta)\equiv \sum_{\substack{n\in\mathbf{Z}\\ n\neq 0}} \frac{\hat{h}_n}{i2\sin\left(\frac{n\omega}{2}\right)} e^{in\theta}.$$

Notice that the hypothesis that $\omega/2\pi\in\mathcal{D}_1$ [see (DC)] implies

$$\left|\sin\left(\frac{n\omega}{2}\right)\right|^{-1} \leqq \frac{\gamma}{2}|n|^\tau, \qquad \forall n\neq 0. \tag{2.9}$$

After these modifications the whole section holds word-by-word for the present case, but notice that now, as well as in the Hamiltonian case with $d=1$, $\mathcal{A}$ and hence q_2 vanish identically.

3. Norms and Function Spaces

In order to provide the Newton procedure with estimates necessary to control the objects involved, we need to fix suitable function spaces with relative norms.

The choice of norms (for $d\geqq 2$) is a rather subtle point if one is interested in obtaining "optimal constants" and "optimal dependence" on the dimension d.

We will consider complex functions with values in a vector space $\mathcal{V}$, where $\mathcal{V}$ can be either $\mathbf{C}^d$, or the space of linear maps from $\mathbf{C}^d$ into itself, denoted by $\mathcal{L}(\mathbf{C}^d)$, or the space of linear maps from $\mathbf{C}^d$ into $\mathcal{L}(\mathbf{C}^d)$, denoted by $\mathcal{L}(\mathbf{C}^d, \mathcal{L}(\mathbf{C}^d))$.

The norms that we will use in the vector spaces $\mathcal{V}$ are the following.

If $c\equiv(c_1, \ldots, c_d)$ belongs to $\mathbf{C}^d$ (or to any subspace of $\mathbf{C}^d$) we set

$$|c|\equiv\left(\sum_{i=1}^d |c_i|^2\right)^{1/2}, \qquad |c|_1\equiv\sum_{i=1}^d |c_i|;$$

if $M\in\mathcal{L}(\mathbf{C}^d)$ and $T\in\mathcal{L}(\mathbf{C}^d, \mathcal{L}(\mathbf{C}^d))$ we set, respectively,

$$|M|\equiv\sup_{\substack{c\in C^d\\ |c|=1}}|Mc|, \qquad |T|\equiv\sup_{\substack{c\in C^d\\ |c|=1}}|Tc|.$$

Now, by $\mathcal{R}_p(\xi, \varrho; \mathcal{V})$ we denote the space of real-analytic functions

$$h: \Delta_{\xi,\varrho}\ni(\theta, t, \varepsilon)\to h(\theta, t; \varepsilon)\in\mathcal{V},$$

$$\Delta_{\xi,\varrho}\equiv\{(\theta, t, \varepsilon)\in\mathbf{C}^{d+2}: |\mathrm{Im}\,\theta_i|\leqq\xi\ (i=1, \ldots, d), |\mathrm{Im}\,t|\leqq\xi,\ |\varepsilon|\leqq\varrho\},$$

which are periodic (with period 2π) in each variable $\theta_1, \ldots, \theta_d, t$. We regard, then, $\mathcal{R}_p(\xi, \varrho; \mathcal{V})$ as a Banach space with respect to either the supremum norm

$$|h|_{\xi,\varrho}\equiv\left(\sum_{i=1}^d \sup_{\Delta_{\xi,\varrho}}|h_i|^2\right)^{1/2},$$

if h is $\mathbf{C}^d$-valued, or with respect to the supremum norm

$$|h|_{\xi,\varrho} \equiv \sup_{\substack{c\in C^d\\ |c|=1}} |hc|_{\xi,\varrho},$$

if h takes values in $\mathscr{L}(\mathbf{C}^d)$ or $\mathscr{L}(\mathbf{C}^d, \mathscr{L}(\mathbf{C}^d))$.

In treating the mapping case, without giving explicit notice, we will refer to the subspace of $\mathscr{R}_p(\xi, \varrho; \mathbf{C})$ of t-independent functions, which we will still denote by the same symbol $\mathscr{R}_p(\xi, \varrho; \mathbf{C})$.

Finally, we recall here, for convenience, the following standard notation. If h is a (smooth) $\mathbf{C}^d$-valued function defined on some domain of $\mathbf{C}^d$, h_x (or $\partial_x h$) denotes the matrix-valued function with entries

$$(h_x)_{ij} \equiv \frac{\partial h_i}{\partial x_j}$$

and h_{xx} (or $\partial_{xx} h$) the $\mathscr{L}(\mathbf{C}^d, \mathscr{L}(\mathbf{C}^d))$-valued function defined by setting, for any $c \in \mathbf{C}^d$,

$$(h_{xx}c)_{ij} \equiv \sum_{l=1}^{d} \frac{\partial^2 h_i}{\partial x_j \partial x_l} c_l .$$

4. Control of the Solution of the Linearized Equation

In order to estimate the solution z of the linearized equation (2.1) we need two technical lemmata, which will be proven in Appendix A. The first is a standard inequality for holomorphic functions and the second is a result à la Rüssmann [33, 34].

Before stating the lemmata, we define, for any $\delta > 0$ and $l = 0, 1$, the following "small-divisor series" for, respectively, the Hamiltonian and the mapping case:

$$\sigma_l(\delta) \equiv \left[2^{d+1} \sum_{\substack{(n,m)\in Z^{d+1}\\ (n,m)\neq 0}} \left(\frac{|n|^l}{\omega \cdot n + m} \right)^2 e^{-\delta(|n|_1 + |m|)} \right]^{1/2}, \tag{4.1}$$

$$\sigma_l(\delta) \equiv \left[\sum_{n=1}^{\infty} \left(\frac{n^l}{\sin\left(\frac{n\omega}{2}\right)} \right)^2 e^{-\delta n} \right]^{1/2}, \quad \text{(mapping case)}. \tag{4.2}$$

Remark 5. At this point the formalism for maps is completely unified with that for the Hamiltonian system (with $d = 1$) and we will not need to make any further distinctions between the two models provided one keeps in mind the adjustments listed in the preceding remark.

Lemma 2. *Let* $h \in \mathscr{R}_p(\xi, \varrho; \mathbf{C}^d)$. *Then for any* $0 < \delta \leqq \xi$,

$$|h_\theta|_{\xi-\delta,\varrho} \leqq |h|_{\xi,\varrho} \delta^{-1} .$$

Lemma 3. *Let* $h \in \mathscr{R}_p(\xi, \varrho; \mathbf{C}^d)$ *have mean value (on* $\mathbf{T}^{d+1}$*) zero. Then, for any* $0 < \delta \leqq \xi$, *and for* $l = 0, 1$,

$$|\partial_\theta^l D^{-1} h|_{\xi-\delta,\varrho} \leqq \sigma_l(2\delta)\, |h|_{\xi,\varrho} .$$

The same inequality, with $l = 0$, *holds if* $h \in \mathscr{R}_p(\xi, \varrho; \mathscr{L}(\mathbf{C}^d))$.

Now, let z be the solution of (2.1) given by (2.7), (2.8) and assume that v, $e \in \mathcal{R}_p(\xi, \varrho; \mathbf{C}^d)$ and that $\mathcal{M} \equiv (I + v_\theta)$ is invertible on $\Delta_{\xi,\varrho}$. Denote by M, $\tilde{M}$, and E upper bounds on, respectively, $|\mathcal{M}|_{\xi,\varrho}$, $|\mathcal{M}^{-1}|_{\xi,\varrho}$ and $|e|_{\xi,\varrho}$ and by $s_l(\delta)$ an upper bound on $\sigma_l(\delta)$. Then, one has, for any $0 < \delta < \xi$,

$$|z|_{\xi-\delta,\varrho} \leqq EM\tilde{M}^2 s_0(\delta)^2 b, \tag{4.3}$$

$$|z_\theta|_{\xi-\delta,\varrho} \leqq EM\tilde{M}^2 s_1(\delta) s_0(\delta) b_1, \tag{4.4}$$

where

$$b_1 \equiv 1 + (M\tilde{M})^2 \frac{s_0(2\xi)}{s_0(\delta)},$$

$$b \equiv b_1 + M\left(\frac{s_0(\xi)}{s_0(\delta)}\right)^2 \left[1 + (M\tilde{M})^2 \frac{s_0(2\xi)}{s_0(\xi)}\right]. \tag{4.5}$$

We remark in passing that usually (i.e., if δ is not too close to ξ) $\sigma_0(\delta) \gg \sigma_0(\xi)$.

Proof of (4.3) and (4.4). We start by estimating the constants c_0 and c_1 given in (2.8). The relations $|\mathcal{M}| = |\mathcal{M}^T|$, $|\mathcal{M}^{-1}| \geqq |\mathcal{M}|^{-1}$ and the positivity of the matrix $\mathcal{M}^T\mathcal{M}$ for $(\theta, t) \in \mathbf{T}^{d+1}$ imply the estimates

$$\tilde{M}^{-2} \leqq |\mathcal{M}^T\mathcal{M}| \leqq M^2, \quad (\theta, t, \varepsilon) \in \Delta_{\xi,\varrho},$$

$$M^{-2} \leqq |(\mathcal{M}^T\mathcal{M})^{-1}| \leqq \tilde{M}^2, \quad (\theta, t, \varepsilon) \in \Delta_{\xi,\varrho},$$

$$|\langle(\mathcal{M}^T\mathcal{M})^{-1}\rangle^{-1}| \leqq M^2,$$

where $\Delta_{\xi,\varrho}$ is defined in the preceding section and, as above, $\langle\cdot\rangle$ denotes average on the torus $\mathbf{T}^{d+1}$ (or on $\mathbf{T}$ for the mapping case). Now, applying Lemma 3 with $\delta = \xi$, one obtains

$$|c_0| \leqq M^3\tilde{M}^2 s_0(2\xi)E,$$

and, applying the same lemma twice with $\delta = \xi/2$, one obtains

$$\begin{aligned} |c_1| &\leqq M\tilde{M}^2 s_0(\xi)(|c_0| + |D^{-1}(\mathcal{M}^T e)|_{\xi/2,\varrho}) \\ &\leqq M\tilde{M}^2 s_0(\xi)\,[(M^3\tilde{M}^2 s_0(2\xi)E + Ms_0(\xi)E]. \end{aligned}$$

In the same fashion, applying Lemma 3 twice with δ replaced by $\delta/2$, one gets easily (4.3) and (4.4). □

Remark 6. If f and v are odd functions of (x, t) [as it will be the case in our applications to (H1) and to the standard map], c_1 vanishes being the average of an odd function. Thus, in such case, (4.3) holds with $b = b_1$.

5. KAM Algorithm

Maintaining the above notations and the assumptions

$$v, \mathcal{M}^{-1}, e \in \mathcal{R}_p(\xi, \varrho; \mathbf{C}^d),$$

we collect the main estimates relative to the Newton iteration procedure in the following

Lemma 4 (Inductive Lemma). *Let w, e' be as in Lemma 1 and set $\xi' \equiv \xi - \delta > 0$. Then one has*

$$|w|_{\xi',\varrho} \leqq W, \quad |w_\theta|_{\xi',\varrho} \leqq W_1, \quad |e'|_{\xi',\varrho} \leqq E'$$

with (b being as in (4.5)*)*

$$W \equiv Ea, \quad a \equiv b(M\tilde{M}s_0(\delta))^2,$$

$$W_1 \equiv Ea\left(\frac{V_1}{M}\delta^{-1} + \frac{s_1(\delta)}{s_0(\delta)}\right),$$

and, denoting by F_3 an upper bound on $|f_{xxx}|_{\xi'+V+W,\varrho}$,

$$E' \equiv E^2 a\left(\frac{aF_3}{2} + \frac{\delta^{-1}}{M} + \chi_d 2\delta^{-1}\tilde{M}\frac{s_0(2\delta)}{s_0(\delta)}\frac{b'}{b}\right), \tag{5.1}$$

where

$$\chi_d \equiv 1 \quad for \quad d \geqq 2, \quad \chi_1 \equiv 0, \quad b' \equiv 1 + (M\tilde{M})^2\frac{s_0(2\xi)}{s_0(2\delta)}.$$

Remark 7. If $\xi' + V + W$ exceeds the widths of the (θ, t)-analyticity domain of f then $F_3 \equiv +\infty$ and the lemma is trivially empty.

The proof of the lemma in $d=1$ is a straightforward application of the results of the preceding section, namely, Lemma 2, (4.3) and (4.4). But the same arguments (with the same constants!) work also for $d \geqq 2$ thanks to the definition of Sect. 3. To give an example, let us estimate the "tensor-valued" function $v_{\theta\theta}$ appearing in v'_θ:

$$v'_\theta \equiv v_\theta + w_\theta \equiv v_\theta + v_{\theta\theta}z + \mathscr{M}z_\theta.$$

$$\begin{aligned}
|v_{\theta\theta}|_{\xi',\varrho} &\equiv \sup_{\substack{c, c' \in C^d \\ |c| = |c'| = 1}} \left(\sum_{i=1}^d \left|\sum_{j,l=1}^d \frac{\partial^2 v_i}{\partial\theta_l \partial\theta_j} c_l c'_j\right|^2_{\xi-\delta,\varrho}\right)^{1/2} \\
&\leqq \sup_{|c| = |c'| = 1} \left(\sum_i \sum_j |c'_j|^2 \left|\frac{\partial}{\partial\theta_j}\sum_l \frac{\partial v_i}{\partial\theta_l} c_l\right|^2_{\xi-\delta,\varrho}\right)^{1/2} \\
&\leqq \delta^{-1} \sup_{|c| = |c'| = 1} \left(\sum_i \sum_j |c'_j|^2 \left|\sum_l \frac{\partial v_i}{\partial\theta_l} c_l\right|^2_{\xi,\varrho}\right)^{1/2} \\
&= \delta^{-1}|v_\theta|_{\xi,\varrho},
\end{aligned}$$

where the first estimate comes from Schwarz inequality and the second from Lemma 2 (with $d=1$). Finally, to estimate the q_2 term (appearing in the definition of e' for $d \geqq 2$) apply Lemma 3 to [see (2.3) and (2.7)]

$$\mathscr{A} = D^{-1}(\mathscr{M}^T e_\theta - e_\theta^T \mathscr{M})$$

and

$$Dz = (\mathscr{M}^T\mathscr{M})^{-1}[c_0 - D^{-1}(\mathscr{M}^T e)]. \quad \square$$

The KAM algorithm, referred to in the introduction, is obtained by iterative applications of Lemma 4, after having fixed a suitable Banach-space scale.

More precisely, assume to have some initial approximate solution of (AE), $v = v^{(0)}$, belonging to $\mathscr{R}_p(\xi_0, \varrho; \mathbf{C}^d)$ for some $\xi_0 > 0$. For any strictly monotone decreasing sequence of $\{\xi_j\}_{j \geqq 1}$, $\xi_1 < \xi_0$, $\xi_j > 0$, one can apply iteratively the

Newton procedure and the above inductive lemma to obtain a family of solutions of (AE), $v^{(j)}$ and $e^{(j)}$, belonging to $\mathscr{R}_p(\xi_j, \varrho; \mathbf{C}^d)$, provided $\mathscr{M}^{(j)}$ is invertible on $\Delta_{\xi_j,\varrho}$. To be completely explicit, one uses Lemma 4 substituting iteratively, for $j \geqq 0$, $(\xi, \xi', \delta, M, \tilde{M}, W, W_1, E, E')$ with $(\xi_j, \xi_{j+1}, \delta_j, M^{(j)}, \tilde{M}^{(j)}, W^{(j)}, W_1^{(j)}, E^{(j)}, E^{(j+1)})$, having defined

$$V^{(j+1)} \equiv V + \sum_{i=0}^{j} W^{(i)}, \qquad V_1^{(j+1)} \equiv V_1 + \sum_{i=0}^{j} W_1^{(i)},$$

$$M^{(j+1)} \equiv M + \sum_{i=0}^{j} W_1^{(i)},$$

$$\tilde{M}^{(j+1)} \equiv \begin{cases} \tilde{M}\left(1 - \tilde{M} \sum_{i=0}^{j} W_1^{(j)}\right)^{-1}, & \text{if } \sum_{i=0}^{j} W_1^{(i)} < 1 \\ \infty, & \text{if } \sum_{i=0}^{j} W_1^{(i)} \geqq 1. \end{cases}$$

That $\tilde{M}^{(j+1)}$ is a bound on $|(\mathscr{M}^{(j+1)})^{-1}|_{\xi_{j+1},\varrho}$ comes from

$$|(\mathscr{M}^{(j+1)})^{-1}| \equiv \left|\left(\mathscr{M} + \sum_{i=0}^{j} w_\theta^{(i)}\right)^{-1}\right| \equiv \left|\left(I + \mathscr{M}^{-1} \sum_{i=0}^{j} w_\theta^{(i)}\right)^{-1} \mathscr{M}^{-1}\right|$$
$$\leqq |\mathscr{M}^{-1}| \left(1 - |\mathscr{M}^{-1}| \sum_{i=0}^{j} |w_\theta^{(i)}|\right)^{-1}.$$

Now, the scale-sequence that we choose is simply given by

$$\xi_j \equiv \frac{\xi_0}{2^j}, \quad \text{i.e.,} \quad \delta_j \equiv \frac{\xi_0}{2^{j+1}}. \tag{5.2}$$

Remark 8. It would be rather lengthy to try to justify, on a general level, why (5.2) is a "good" choice and we content ourselves by just mentioning that such a choice is related to the "quadratic convergence" of the Newton procedure (compare [7, Appendix C]).

Remark 9. Notice that the estimates in the Inductive Lemma involve upper bounds $s_l(\delta)$ on the small-divisor series $\sigma_l(\delta)$ given in (4.1). Even though it is rather easy to give rough evaluations of $\sigma_l(\delta)$, it is very important, for the efficiency of the algorithm, to have *accurate* estimates on $\sigma_l(\delta)$. We will show below how one can obtain satisfactory results, employing computer-assisted estimations (compare Lemma 9 and the following comments).

We conclude this section by pointing out that, in applications, the above algorithm can be applied only a *finite number of times*. Thus, to establish the existence of solutions u one needs to combine the algorithm with a KAM theorem, which we proceed now to describe.

6. KAM Condition

Here, we prove a condition, which, if satisfied by $M^{(j_0)}$, $\tilde{M}^{(j_0)}$, and $E^{(j_0)}$ for some $j_0 \geqq 0$, yields the convergence of the KAM algorithm and hence the existence of KAM surfaces.

Remark 10. In order to get a general, simple and explicit condition (in the style of, e.g., [16]) we will need to make various estimates certainly not optimal (see, e.g., next Lemma 5). Thus, the use of a KAM condition, in connection with the problem of obtaining good stability bounds, makes sense only in a suitable combination with the KAM algorithm of Sect. 5.

Even though we will apply the KAM condition to $v^{(j_0)}$ and $e^{(j_0)}$, we can state it independently of the preceding section. In order to do this, we need to introduce three constants K_1, K_2, and k_0 related to the upper bounds $s_l(\delta)$ on the small-divisor series $\sigma_l(\delta)$:

Lemma 5. *Let* $\sigma_l(\delta)$ *be as in* (4.1) *with* $0<\delta\leqq 1/2$. *Then*

$$\sigma_l(\delta) < K_l \gamma \delta^{-k_l}, \tag{6.1}$$

where, denoting Euler's gamma function by Γ,

$$k_l \equiv \tau + l + \frac{d+1}{2},$$

$$K_0 \equiv \frac{7}{2}\cdot\left(\frac{5}{2}\right)^d \left(\frac{\Gamma(2\tau+d)}{\Gamma(d/2)}\right)^{1/2}, \qquad K_1 \equiv K_0 \sqrt{(2\tau+1+d)(2\tau+d)}\,.$$

The same inequality holds for $\sigma_l(\delta)$ *as in* (4.2) *(mapping case) setting* $d=0$ *in the definition of* k_l *and* $d=1$ *in the definition of* K_l.

A proof of this simple lemma is given in Appendix B.

Now, let v, $e \in \mathscr{R}_p(\xi_*, \varrho; \mathbf{C}^d)$ satisfy (AE) for some $\xi_*>0$. Assume that $\mathscr{M}^{-1} \equiv (I+v_\theta)^{-1} \in \mathscr{R}_p(\xi_*, \varrho; \mathbf{C}^d)$ and denote, as usual, by M, $\tilde{M}$, E upper bounds on $|\mathscr{M}|_{\xi_*,\varrho}$, $|\mathscr{M}^{-1}|_{\xi_*,\varrho}$, $|e|_{\xi_*,\varrho}$ and by F_3 an upper bound on $|f_{xxx}|_{\xi_*-V,\varrho}$. For simplicity assume also that ξ_*^{-1}, M, $\tilde{M}$ are greater or equal than one.

Lemma 6 (KAM Condition). *Let*

$$\mathscr{K}(\lambda,\mu,v,\chi) \equiv (10^3 \cdot 2^{15k_0+1} \sqrt{K_0 K_1} K_0^4 \gamma^5) \lambda^2 \sqrt{\mu^{21} v^{10k_0+1}}\,\chi\,.$$

If

$$\mathscr{K}(M, M\tilde{M}, \xi_*^{-1}, F_3)\,|e|_{\xi_*,\varrho} \equiv \mathscr{K} E \leqq 1\,, \tag{6.2}$$

then equation (E) *(respectively* (E1)*) has a unique solution* $u \in \mathscr{R}_p(\xi_*/2, \varrho; \mathbf{C}^d)$ *with* $\langle u\rangle = \langle v\rangle$. *Furthermore,* $I+u_\theta$ *is invertible on* $\Delta_{\xi_*/2,\varrho}$ *and one has*

$$|u-v|_{\xi_*/2,\varrho} < \mathscr{K} E\,\frac{\xi_*}{64}, \tag{6.3}$$

$$|u_\theta - v_\theta|_{\xi_*/2,\varrho} < \frac{\mathscr{K} E}{2\tilde{M}}\,. \tag{6.4}$$

Remark 11. In order to prove the convergence of the KAM algorithm of the preceding section, one has to check if, for $j_0=0,1,\ldots$, condition (6.2) is verified with ξ_*, $|\mathscr{M}|_{\xi_*,\varrho}$, $|\mathscr{M}^{-1}|_{\xi_*,\varrho}$, $|e|_{\xi_*,\varrho}$ replaced by, respectively, $\xi_0/2^{j_0}$, $M^{(j_0)}$, $\tilde{M}^{(j_0)}$, $E^{(j_0)}$. In case of convergence (6.3) and (6.4) hold with v replaced by $v^{(j_0)}$ and the final analyticity width in the periodic variables will be $\xi_*/2 \equiv \xi_0/2^{j_0+1}$.

Proof of Lemma 6. First of all observe that the invertibility of $(I+u_\theta)$ is a trivial consequence of (6.4) and (6.2). Now, let $v^{(j)}\equiv v^{(j-1)}+w^{(j-1)}$, $e^{(j)}$ be the functions obtained by iteratively applying Lemma 1 of Sect. 2 ($v^{(0)}\equiv v$, $e^{(0)}\equiv e$). Let

$$\xi_*^{(j)}\equiv\frac{\xi_*}{2}+\frac{\xi_*}{2^{j+1}},\qquad \delta_j\equiv\xi_*^{(j)}-\xi_*^{(j+1)}=\frac{\xi_*}{2^{j+2}}$$

and let $V^{(j)}$, $V_1^{(j)}$, $W^{(j)}$, $W_1^{(j)}$, $M^{(j)}$, $\tilde{M}^{(j)}$, and $E^{(j)}$ be the bounds on the corresponding norms yielded by the KAM algorithm described in the preceding section, with $s_k(\delta)=s_k(\delta_j)$ replaced by the right-hand side of (6.1). For simplicity, we replace $W_1^{(j)}$ by

$$W_1^{(j)}\equiv E^{(j)}a^{(j)}\left(\delta_j^{-1}+\frac{s_1(\delta_j)}{s_0(\delta_j)}\right),$$

which can be done recalling the original derivation of W_1 and using the bound

$$|v_{\theta\theta}|_{\xi',\varrho}=|\partial_\theta\mathscr{M}|_{\xi',\varrho}\leqq M\delta^{-1}.$$

We claim that condition (6.2) implies, for a suitable $\mathscr{K}_0<\mathscr{K}$ and for any j,

$$E^{(j)}<(\mathscr{K}_0E)^{2^j}, \tag{6.5$_j$}$$

$$\xi_*^{(j)}+V^{(j)}\leqq\xi_*+V, \tag{6.6$_j$}$$

$$\tilde{M}^{(j)}\leqq 2\tilde{M}. \tag{6.7$_j$}$$

Before proving the claim, observe that (6.5), (6.6) yield easily the first part of the lemma. In fact, since $\xi_*^{(j)}\downarrow\xi_*/2$, (6.5) and (6.6) imply the uniform convergence in $\mathscr{R}_p(\xi_*/2,\varrho;\mathbf{C}^d)$ of $v^{(j)}$ to a unique solution u with, by construction, $\langle u\rangle=\langle v\rangle$ (compare Lemma 1).

We proceed now by induction on j. For $j=0$ the claim is obvious. Assume the claim true for $0,1,\ldots,j$ and notice that $(6.7)_j$ is equivalent to

$$2\tilde{M}\sum_{i=0}^{j-1}W_1^{(i)}\leqq 1, \tag{6.8$_j$}$$

which, since M and $\tilde{M}$ are greater or equal than one, implies, for $0\leqq i\leqq j$,

$$M^{(j)}\leqq\tfrac{3}{2}M. \tag{6.9$_j$}$$

Now, by the estimates in Lemma 4, by $(6.6)_i$, $(6.7)_i$, and $(6.9)_i$ with $i\leqq j$, observing that $k_0\geqq 3/2$, $K_0>9$, $K_1\geqq\sqrt{12}K_0$ and $\gamma>2$, one obtains easily the following bounds for $i\leqq j$

$$E^{(i+1)}\leqq(E^{(i)})^2\beta\eta^i, \tag{6.10}$$

$$W^{(i)}\leqq E^{(i)}\beta_0\eta_0^i, \tag{6.11}$$

$$W_1^{(i)}\leqq E^{(i)}\beta_1\eta_1^i, \tag{6.12}$$

with

$$\beta\equiv 100\cdot K_0^4\gamma^4 2^{8k_0}M^2(M\tilde{M})^8F_3\xi_*^{-4k_0},\qquad \eta\equiv 2^{4k_0},$$
$$\beta_0\equiv 14\cdot K_0^2\gamma^2 2^{4k_0}M(M\tilde{M})^4\xi_*^{-2k_0},\qquad \eta_0\equiv 2^{2k_0},$$
$$\beta_1\equiv 73\cdot K_1K_0\gamma^2 2^{4k_0}M(M\tilde{M})^4\xi_*^{-(2k_0+1)},\qquad \eta_1\equiv 2^{2k_0+1}.$$

To give an example we derive (6.10) [the derivation of (6.11) and (6.12) are completely analogous]. By (5.1), observing that $b'/b \leqq 2^{k_0}$ and using the inductive assumptions together with $(6.9)_i$, one has for $i \leqq j$,

$$E^{(i+1)} \leqq (E^{(i)})^2 \frac{A^2 F_3}{2} \left[1 + \frac{2}{A} \frac{2^{i+2}}{\xi_*} + \frac{8}{A} \tilde{M} \frac{2^{i+2}}{\xi_*} \right],$$

where A [which is a bound on a of (5.1)] is given by

$$A \equiv \left[1 + \left(\frac{1}{8}\right)^{k_0} (3M\tilde{M})^2 + \frac{3}{2} M \left(\frac{1}{4}\right)^{2k_0} \left[1 + (3M\tilde{M})^2 \left(\frac{1}{2}\right)^{k_0} \right]\right] \times \left[3M\tilde{M}K_0\gamma \left(\frac{2^{i+2}}{\xi_*}\right)^{k_0} \right]^2 .$$

Since $\gamma > 2$, $k_0 \geqq 3/2$, $K_0 > 9$, and $\xi_* \leqq 1$,

$$432^2 \tilde{M}^4 \xi_*^{-2k_0} 2^{2k_0 i} < A < 14 K_0^2 \gamma^2 2^{4k_0} M (M\tilde{M})^4 \xi_*^{-2k_0} 2^{2k_0 i},$$

from which (6.10) follows.

Now, notice that $\mathscr{K}$ can be written as

$$\mathscr{K} = r\beta\eta\sqrt{2\tilde{M}\beta_1\eta_1} \tag{6.13}$$

with some $r > 11/10$.

To prove $(6.5)_{j+1}$, let $\mathscr{K}_0 \equiv \beta\eta \ (< \mathscr{K})$ and use (6.10) with $i = 1, \ldots, j$, to get

$$\begin{aligned} E^{(j+1)} &\leqq E^{2^{j+1}} \prod_{i=0}^{j} (\beta\eta^{j-i})^{2^i} \\ &= \left[E\beta^{\left(\sum_{i=1}^{j+1} 2^{-i}\right)} \eta^{\left(\sum_{i=1}^{j+1} (i-1)2^{-i}\right)} \right]^{2^{j+1}} \\ &< (\mathscr{K}_0 E)^{2^{j+1}} . \end{aligned}$$

In order to prove $(6.6)_{j+1}$, observe that conditions (6.2) and (6.13) imply the following bounds

$$\frac{\beta_0}{\mathscr{K}} < 10^{-11} \xi_*^5, \qquad \frac{\beta_0 \eta_0 \mathscr{K}_0^2}{\mathscr{K}^2} < \frac{\xi_*}{72}, \qquad \frac{\mathscr{K}}{\mathscr{K}_0 \sqrt{\eta_0}} > 4579 . \tag{6.14}$$

Now, $(6.6)_{j+1}$ is equivalent to

$$\sum_{i=0}^{j} W^{(i)} \leqq \xi_* \left(\frac{1}{2} - \frac{1}{2^{j+2}} \right),$$

and, by (6.11), (6.12), (6.14), and (6.2) one has

$$\begin{aligned} \sum_{i=0}^{j} W^{(i)} &< \beta_0 E + \beta_0 \sum_{i=1}^{\infty} (\mathscr{K}_0 E)^{2^i} \eta_0^i \\ &< \beta_0 E + \beta_0 (\mathscr{K}_0 E)^2 \eta_0 \left(1 + \frac{1}{\log(\mathscr{K}_0 E \sqrt{\eta_0})^{-1}} \right) \\ &< \mathscr{K} E \frac{\xi_*}{64} < \xi_* \left(\frac{1}{2} - \frac{1}{2^{j+2}} \right). \end{aligned} \tag{6.15}$$

Analogously, in order to prove $(6.8)_{j+1}$ use condition (6.2), (6.13) and the fact that

$$\left(\frac{K_1}{K_0}\right)^{1/2}\frac{1}{K_0^3}<\frac{1}{271}$$

to get the bounds

$$\frac{2\tilde{M}\beta_1}{\mathscr{K}}<10^{-9},\qquad \frac{\mathscr{K}_0\sqrt{\eta_1}}{\mathscr{K}}<\frac{1}{3238}.$$

Thus, recalling that $r>11/10$, one obtains

$$\begin{aligned}2\tilde{M}\sum_{i=0}^{j}W_1^{(i)}&<2\tilde{M}\beta_1E+2\tilde{M}\beta_1(\mathscr{K}_0E)^2\eta_1\left(1+\frac{1}{\log(\mathscr{K}_0E\sqrt{\eta_1})^{-1}}\right)\\&<\mathscr{K}E\left(10^{-9}+\frac{\mathscr{K}E}{r^2}\left(1+\frac{1}{\log 3238}\right)\right)<\mathscr{K}E.\end{aligned}\tag{6.16}$$

Finally, (6.15) and (6.16) imply immediately (6.3) and (6.4). □

7. Application to the Hamiltonian (H1)

In this section we apply the above KAM algorithm to the KAM-surfaces equation associated to the Hamiltonian (H1), namely

$$D^2u=\varepsilon[\sin(\theta+u)+\sin(\theta+u-t)],\qquad D\equiv\omega\frac{\partial}{\partial\theta}+\frac{\partial}{\partial t},\tag{7.1}$$

and prove the following

Theorem 1. *Let* $\omega=(\sqrt{5}-1)/2$ *and let* $\xi=(2^{11}\cdot 10)^{-1}$ $(\sim 4.8\cdot 10^{-5})$, $\varrho=0.015$. *Then Eq.* (7.1) *admits a unique solution* u *in* $\mathscr{R}_p(\xi,\varrho;\mathbf{C})$ *with vanishing mean-value*

$$\langle u\rangle\equiv\int_{\mathbf{T}^2}u(\theta,t;\varepsilon)\frac{d\theta dt}{(2\pi)^2}=0.$$

For such solution one has

$$\begin{gathered}|u|_{\xi,\varrho}<0.182,\\0.22<|u_\theta(\pi,0;0.015)|<|u_\theta|_{\xi,\varrho}<0.2419.\end{gathered}\tag{7.2}$$

Furthermore, if v *is the polynomial approximant,*

$$v(\theta,t;\varepsilon)\equiv\sum_{l=1}^{l_0}u^{(l)}(\theta,t)\varepsilon^l,\qquad \langle u^{(l)}\rangle=0,\tag{7.3}$$

where the $u^{(l)}$*'s are the unique trigonometric polynomials (of degree* l*) satisfying* (EP) *with* $g=\cos x+\cos(x-t)$, *then, for* $l_0=24$, *one has*

$$|u-v|_{\xi,\varrho}<6.84\cdot 10^{-5},\qquad |u_\theta-v_\theta|_{\xi,\varrho}<3.096\cdot 10^{-3}.$$

We split the proof of the theorem in four further lemmata, two of which can be proven with the aid of a computer. The first is a general estimate on the error function relative to approximants of type (7.3).

Lemma 7. *Let* $f(x,t;\varepsilon)=\varepsilon g(x,t)$, *let* v *be as in* (7.3) *and let* e *be the associated error function (cf.* (AE)*). Let* $\xi,\varrho>0$ *and denote by* P *the polynomial in* ϱ *given by*

$$P\equiv \sum_{l=1}^{l_0} |u^{(l)}|_\xi \varrho^l,$$

which is an upper bound on $|v|_{\xi,\varrho}$. *Then*

$$|e|_{\xi,\varrho}\leqq \varrho\Bigg[|\partial_x^{l_0+1}g|_{\xi+P}\frac{P^{l_0}}{l_0!} + \Big(\sup_{2\leqq l\leqq l_0}|\partial_x^l g|_\xi\Big)\sum_{l=1}^{l_0-1}\Big(\frac{P^l}{l!}-\varrho^l\sum_{k\in\mathscr{K}_l}\prod_{i=1}^{l}\frac{(|u^{(i)}|_\xi)^{k_i}}{k_i!}\Big)\Bigg], \tag{7.4}$$

where (as in (b) *of Sect. 1)* $\mathscr{K}_l\equiv\Big\{k\in\mathbf{N}^l:\ \sum_{i=1}^{l} ik_i=l\Big\}$.

Proof. By definition of $u^{(l)}$ one has

$$\begin{aligned} e&\equiv D^2v+\varepsilon g_x(\theta+v,t)\\ &=\varepsilon\left[g_x(\theta+v,t)-\sum_{l=0}^{l_0-1}\frac{\partial_x^l g_x}{l!}v^l\right]\\ &\quad+\varepsilon\sum_{l=1}^{l_0-1}\frac{\partial_x^l g_x}{l!}\sum_{h=l_0}^{l_0(l_0-1)}\varepsilon^h\sum_{\substack{k\in\mathscr{K}_h\\ |k|_1=l}} l!\prod_{i=1}^{h}\frac{(u^{(i)})^{k_i}}{k_i!}, \end{aligned}$$

which implies

$$\begin{aligned} |e|_{\xi,\varrho}\leqq\varrho&\Bigg[|\partial_x^{l_0+1}g|_{\xi+P,\varrho}\frac{P^{l_0}}{l_0!}+\Big(\sup_{2\leqq l\leqq l_0}|\partial_x^l g|_\xi\Big)\\ &\times\Bigg\{\sum_{l=1}^{l_0-1}\sum_{h=l_0}^{l_0(l_0-1)}\varrho^h\sum_{\substack{k\in\mathscr{K}_h\\ |k|_1=l}}\prod_{i=1}^{h}\frac{|u^{(i)}|_\xi^{k_i}}{k_i!}\Bigg\}\Bigg]. \end{aligned}$$

Now, the term in curly brackets in the above expression can be written as

$$\begin{aligned} &\sum_{l=1}^{l_0-1}\frac{P^l}{l!}-\sum_{l=1}^{l_0-1}\sum_{h=1}^{l_0-1}\varrho^h\sum_{\substack{k\in\mathscr{K}_h\\ |k|_1=l}}\prod_{i=1}^{h}\frac{|u^{(i)}|_\xi^{k_i}}{k_i!}\\ &=\sum_{l=1}^{l_0-1}\frac{P^l}{l!}-\sum_{h=1}^{l_0-1}\sum_{l=1}^{l_0-1}\varrho^l\sum_{\substack{k\in\mathscr{K}_l\\ |k|_1=h}}\prod_{i=1}^{l}\frac{|u^{(i)}|_\xi^{k_i}}{k_i!}\\ &=\sum_{l=1}^{l_0-1}\frac{P^l}{l!}-\sum_{l=1}^{l_0-1}\varrho^l\sum_{k\in\mathscr{K}_l}\prod_{i=1}^{l}\frac{|u^{(i)}|_\xi^{k_i}}{k_i!}.\qquad\square \end{aligned}$$

Lemma 8 (Computer-Assisted). *Let* $\xi_0=1/10$; *let* v *and* ϱ *be as in Theorem 1* ($l_0=24$) *and let* e *be the associated error function. Then*

$$|v|_{\xi_0,\varrho}<0.1819\equiv V,\qquad |v_\theta|_{\xi_0,\varrho}<0.2388\equiv V_1,\qquad |e|_{\xi_0,\varrho}<8.023\cdot10^{-10}\equiv E.$$

This lemma has been proven by the computer program "INITIAL" reported in Appendix D. The ideas on which the program is based are the following. The system (EP) is readily solved in Fourier representation. In fact

$$u^{(1)} = -\left[\frac{1}{\omega^2}\sin\theta + \frac{1}{(1-\omega)^2}\sin(\theta - t)\right]$$
$$= \frac{i}{2}\left[\frac{1}{\omega^2}(e^{i\theta} - e^{-i\theta}) + \frac{1}{(1-\omega)^2}(e^{i(\theta-t)} - e^{-i(\theta-t)})\right],$$

and the Fourier coefficients of $u^{(l+1)}$, for $l \geqq 1$, are given by

$$\hat{u}^{(l+1)}_{(n,m)} = \frac{-1}{(\omega n + m)^2}\sum_{k \in \mathscr{K}_l} \hat{\varphi}^{(k)}_{(n,m)}, \qquad \hat{u}^{(l+1)}_{(0,0)} \equiv 0,$$

where the $\varphi^{(k)}$ for $k \in \mathscr{K}_l$ are the trigonometric polynomials given by

$$\varphi^{(k)} \equiv \partial_\theta^{k_1 + \ldots + k_l}[\sin\theta + \sin(\theta - t)] \prod_{j=1}^{l} \frac{(u^{(j)})^{k_j}}{k_j!}$$
$$= \frac{1}{2}\left[((-1)^{|k|_1} e^{i\theta} - e^{-i\theta}) + ((-1)^{|k|_1} e^{i(\theta - t)} - e^{-i(\theta - t)})\right] \prod_{j=1}^{l} \frac{(u^{(j)}/i)^{k_j}}{k_j!}.$$

By induction on l, it is easy to see that

$$\frac{u^{(l)}}{i} = \sum_{\substack{|n| \leqq l \\ |m| \leqq l}} c^{(l)}_{(n,m)} e^{i(n\theta + mt)},$$

where the $c^{(l)}_{(n,m)}$ are real coefficients odd in (n,m), i.e.,

$$c^{(l)}_{(-n,-m)} = -c^{(l)}_{(n,m)}.$$

Now, the first and main part of the program "INITIAL" gives, using interval-arithmetic, an accurate evaluation of the numbers $c^{(l)}_{(n,m)}$, $l = 1, \ldots, l_0 = 24$. More precisely, it is proved in Appendix D that

$$c^{(l)}_{(n,m)} \in \Delta \equiv \Delta(c^{(l)}_{(n,m)}) \equiv (x_L, x_U),$$

where, for each $c^{(l)}_{(n,m)}$, x_L, and x_U are rational numbers (with a finite fractional part in binary representation), verifying

$$\frac{|x_U - x_L|}{\min(|x_L|, |x_U|)} < 4 \cdot 10^{-7}. \tag{7.5}$$

The second part of "INITIAL", which, from a computational point of view, is trivial with respect to the first one, evaluates the supremum norms according to the formulae

$$|u^{(l)}|_\xi \leqq \sum |c^{(l)}_{(n,m)}| e^{(|n| + |m|)\xi} \equiv U^{(l)}, \tag{7.6}$$

$$|u^{(l)}_\theta|_\xi \leqq \sum |n|\,|c^{(l)}_{(n,m)}| e^{(|n| + |m|)\xi} \equiv U^{(l)}_1, \tag{7.7}$$

$$|v|_{\xi,\varrho} \leqq \sum_{l=1}^{l_0} \varrho^l U^{(l)}, \qquad |v_\theta|_{\xi,\varrho} \leqq \sum_{l=1}^{l_0} \varrho^l U^{(l)}_1, \tag{7.8}$$

together with formula (7.4) with $|u^{(l)}|_\xi$ replaced by $U^{(l)}$, $\sup_{2 \leqq l \leqq l_0} (|\partial^l_x g|_\xi)$ and $|\partial_x^{l_0+1} g|_{\xi+P}$ replaced by $Ch(\xi) \equiv \cosh\xi + \cosh 2\xi$ and $Ch(\xi + P)$.

Remark 12. We point out that the program "INITIAL" yields actually much more than the bounds indicated in Lemma 8. In fact, it gives "explicitly" the Taylor and Fourier representation of the initial approximation v. [The quotation marks refer to the fact that the real numbers involved in such representation are given in terms of intervals (x_L, x_U) verifying (7.5).]

In order to apply the KAM algorithm we still have to provide accurate bounds $s_k(\delta)$ on the small divisor series $\sigma_k(\delta)$. To do this, we will use the following elementary lemma, which is proven in Appendix C.

Lemma 9. *Let $\omega \in (0,1)$ be a quadratic irrational number (i.e., $\omega \in \mathscr{D}_1$ with $\tau = 1$). Let $0 < \delta \leqq 1/2$ and let $\sigma_k(\delta)$ be as in (4.1). Then for any integer $N \geqq 1$ one has*

$$\sigma_k(\delta) < s_k^{(N)}(\delta) \equiv 2\left[\sum_{(n,m)\in A_N}\left(\frac{|n|^k}{\omega n - m}\right)^2 e^{-\delta(|n|+|m|)} + 2(\gamma^2 S_k^{(N)} + I_k)\right]^{1/2},$$

where

$$A_N \equiv \{(n,m) \in Z^2 \backslash (0,0) \text{ such that } -(N-1) \leqq n \leqq (N-1),\ \omega n - \tfrac{3}{2} < m < \omega n + \tfrac{3}{2}\},$$

and, setting $\alpha \equiv \delta(1+\omega)$,

$$S_0^{(N)} \equiv 3e^{\delta/2} e^{-\alpha(N-1)} \frac{1}{\alpha^3}(2 + (2N+1)\alpha + N^2\alpha^2),$$

$$S_1^{(N)} \equiv 3e^{\delta/2} e^{-\alpha(N-1)} \frac{1}{\alpha^5}(24 + (24N+36)\alpha + (12N^2+24N+14)\alpha^2 + (4N^3+6N^2+4N+1)\alpha^3 + N^4\alpha^4),$$

$$I_0 \equiv \frac{4}{\alpha}, \qquad I_1 \equiv \frac{8}{\alpha^3}\left(1 + \delta\frac{9}{8(1-\omega)^3}\right).$$

In order to apply this lemma in an effective way one should use it in conjunction with a straightforward computer-assisted evaluation of the finite sum over A_N appearing in the definition of $s_k^{(N)}(\delta)$. To give an example, one can prove bounds of the following type:

$$120 < \sigma_0(\tfrac{1}{20}) < s_0^{(170)}(\tfrac{1}{20}) < 122,$$

$$242 < \sigma_0(\tfrac{1}{40}) < s_0^{(350)}(\tfrac{1}{40}) < 244,$$

$$485 < \sigma_0(\tfrac{1}{80}) < s_0^{(1000)}(\tfrac{1}{80}) < 486,$$

where the left-hand-side values are obtained by replacing the series in σ_k by the finite sums over A_N.

Finally, we have

Lemma 10 (Computer-Assisted). *Let v, e, V, V_1 and E be as in Lemma 8 and set $M \equiv 1 + V_1$, $\tilde{M} \equiv (1 - V_1)^{-1}$. Let $v^{(j)}$ and $e^{(j)}$ ($v^{(0)} \equiv v$, $e^{(0)} \equiv e$) be the sequences of functions yielded by iteratively applying Lemma 1 of Sect. 2. Let $V^{(j)}$, $V_1^{(j)}$, and $E^{(j)}$ ($V^{(0)} \equiv V$, $V_1^{(0)} \equiv V_1$, $E^{(0)} \equiv E$) be the sequences of numbers obtained by applying the* KAM *algorithm of Sect. 5 with*

$$\xi_j \equiv \frac{\xi_0}{2^j}\left(\delta_j \equiv \xi_j - \xi_{j+1} = \frac{\xi_0}{2^{j+1}}\right), \qquad \xi_0 \equiv 1/10,$$

and with s_k *replaced by (see Lemma 9)* $s_k^{(N_k^{(j)})}(\delta_j)$ *with*

$$N_0^{(j)} \equiv 250 \cdot 2^j, \quad N_1^{(j)} \equiv 350 \cdot 2^j \ (j < 10); \quad N_k^{(j)} \equiv 50\,000 \ (j \geqq 10).$$

(Recall that in Lemma 4 one can take $b' = b$.*) Then the* KAM *algorithm converges. More precisely, if* $\mathscr{K}$ *is as in Lemma 6, setting* $\xi_* \equiv \xi_{10}$, *one has*

$$\mathscr{K} E^{(10)} < 2.884 \cdot 10^{-6}, \tag{7.9}$$

$$|u - v^{(10)}|_{\xi_*/2, \varrho} < 4.4 \cdot 10^{-12}, \tag{7.10}$$

$$|u_\theta - v_\theta^{(10)}|_{\xi_*/2, \varrho} < 1.093 \cdot 10^{-6}, \tag{7.11}$$

where $\mathscr{K}$ *is computed at* $(M^{(10)}, M^{(10)}\tilde{M}^{(10)}, \xi_*^{-1}, F_3)$ *with* $F_3 \equiv \max\{1, \varrho Ch(\xi_* + V^{(10)})\}$.

Theorem 1 is a corollary of this lemma with $\xi \equiv \xi_*/2$.

The proof of Lemma 10 is based on a straightforward translation in computer-language of the explicit formulae indicated above.

The upper bounds in (7.2) are obtained by observing that

$$|u|_{\xi,\varrho} < |u - v|_{\xi,\varrho} + |v|_{\xi,\varrho}, \qquad |u_\theta|_{\xi,\varrho} < |u_\theta - v_\theta|_{\xi,\varrho} + |v_\theta|_{\xi,\varrho},$$

and using Lemma 8. The lower bound in (7.2) is based on a computer-assisted evaluation of $v_\theta(\pi, 0; 0.015)$ and on the inequality

$$|u_\theta(\pi, 0; 0.015)| > |v_\theta(\pi, 0; 0.015)| - |u_\theta - v_\theta|_{\xi_*/2, \varrho}.$$

8. Application to the Standard Map

Here we consider, in a way completely analogous to the preceding section, the KAM-curve equation for the standard-map, i.e.,

$$D^2 u = \varepsilon \sin(\theta + u), \qquad Du \equiv u\left(\theta + \frac{\omega}{2}\right) - u\left(\theta - \frac{\omega}{2}\right). \tag{8.1}$$

Theorem 2. *Let* $\omega = (\sqrt{5} - 1)\pi$ *and let* $\xi = (2^9 \cdot 10)^{-1}$ $(\sim 1.96 \cdot 10^{-4})$, $\varrho = 0.65$. *Then Eq. (8.1) has a unique solution* $u \in \mathscr{R}_p(\xi, \varrho; \mathbf{C})$ *with vanishing mean-value on* $\mathbf{T}$. *For such a function one has*

$$|u|_{\xi,\varrho} < 0.2528,$$

$$0.35 < |u_\theta(\pi; 0.65)| < |u_\theta|_{\xi,\varrho} < 0.3684.$$

Furthermore, if v *is the analog of (7.3) in Theorem 1 with* $l_0 = 38$,

$$v(\theta; \varepsilon) \equiv \sum_{l=1}^{l_0} u^{(l)}(\theta)\varepsilon^l, \qquad \langle u^{(l)} \rangle = 0, \qquad (l_0 = 38),$$

then

$$|u - v|_{\xi,\varrho} < 2.512 \cdot 10^{-5}, \qquad |u_\theta - v_\theta|_{\xi,\varrho} < 1.588 \cdot 10^{-3}.$$

The proof of this theorem is obtained by following the strategy of the preceding section with few obvious changes. More precisely, Lemma 7 and its proof hold identically for e, v satisfying the analog of (AE), i.e.,

$$e = D^2 v + \varepsilon g_x(\theta + v). \tag{8.2}$$

As for the bounds $s_k^{(N)}(\delta)$ on the small-divisor series (4.2), one sees easily from the proof of Lemma 9 (see Appendix C) that one can take

$$s_k^{(N)}(\delta) \equiv \left(\sum_{n=1}^{N} \left(\frac{n^k}{\sin\left(\frac{n\omega}{2}\right)} \right)^2 e^{-\delta n} + \frac{\gamma^2}{4} S_k^{(N)} \right)^{1/2},$$

where the $S_k^{(N)}$ are as in Lemma 9.

Now the modifications indicated in Appendix D of the program "INITIAL" together with the KAM algorithm yields

Lemma 11 (Computer-Assisted). *Let v, ξ, and ϱ be as in Theorem 2 ($l_0 = 38$) and let e be given by* (8.2). *Then*

$$|v|_{\xi,\varrho} < 0.2527, \quad |v_\theta|_{\xi,\varrho} < 0.3668, \quad |e|_{\xi,\varrho} < 4.392 \cdot 10^{-9}.$$

Moreover, setting $\xi_* = \xi_8$ and $Ch(\xi) \equiv \cosh\xi$, Lemma 10 holds word-by-word if one substitutes (7.9), (7.10), (7.11) *with*

$$\mathscr{K} E^{(8)} < 2.247 \cdot 10^{-29},$$

$$|u - v^{(8)}|_{\xi_*/2,\varrho} < 1.372 \cdot 10^{-34},$$

$$|u_\theta - v_\theta^{(8)}|_{\xi_*/2,\varrho} < 7.1 \cdot 10^{-30}.$$

9. Two Numerical Hints

There are several numerical experiments that one can carry out in relation with the methods presented in this paper.

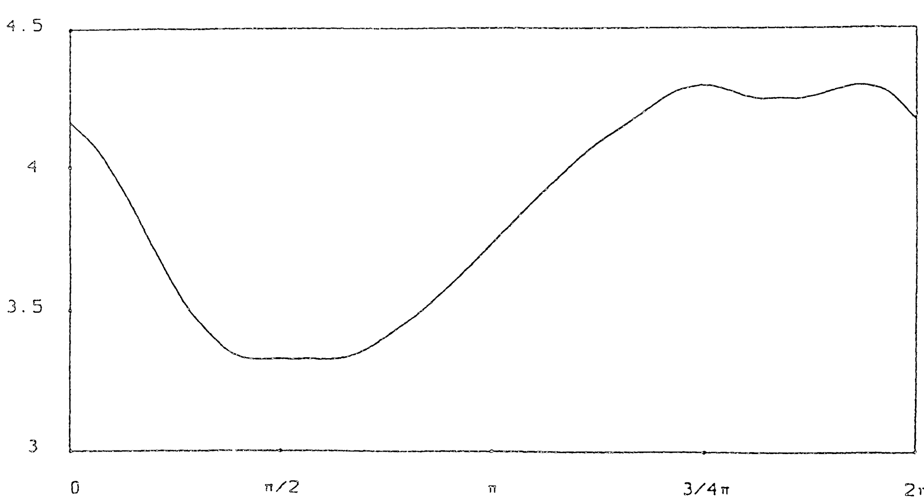

Fig. 1

To mention an example related to the (H1)-model, a numerical evaluation of

$$\sup_{(\theta,t)\in\mathbf{T}^2} \sup_{|\varepsilon|=\varrho} |e(\theta,t;\varepsilon)| \equiv M(\varrho),$$

where e is the error function associated to the polynomial approximation (7.3) with $l_0 = 38$, indicates that $M(\varrho) < 10^{-5}$ ($\ll \varrho$) for $\varrho < 0.026$, while for $\varrho \geqq 0.031$ $M(\varrho) \sim \varrho$, suggesting that a drastic phenomenon takes place for complex values of ε around the believed break-down threshold (~ 0.027).

Another type of experiment, related to the standard map, is synthesized by Figs. 1–4.

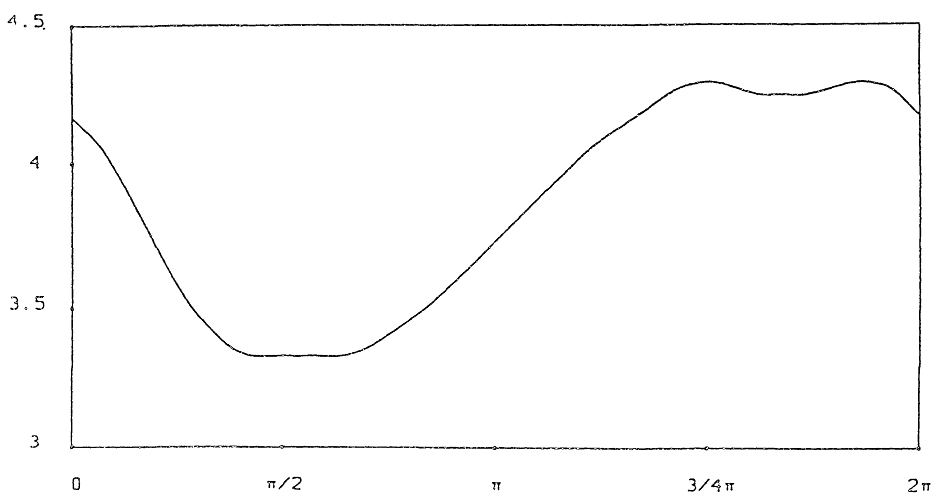

Fig. 2

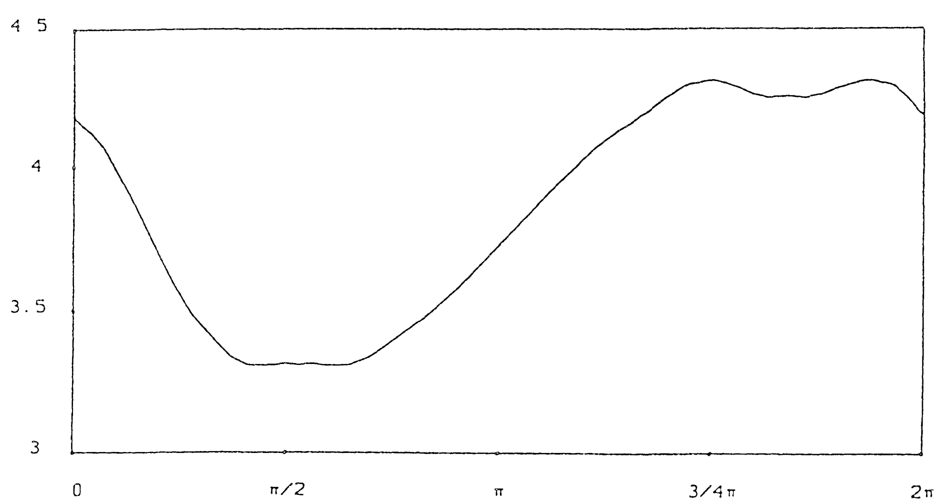

Fig. 3

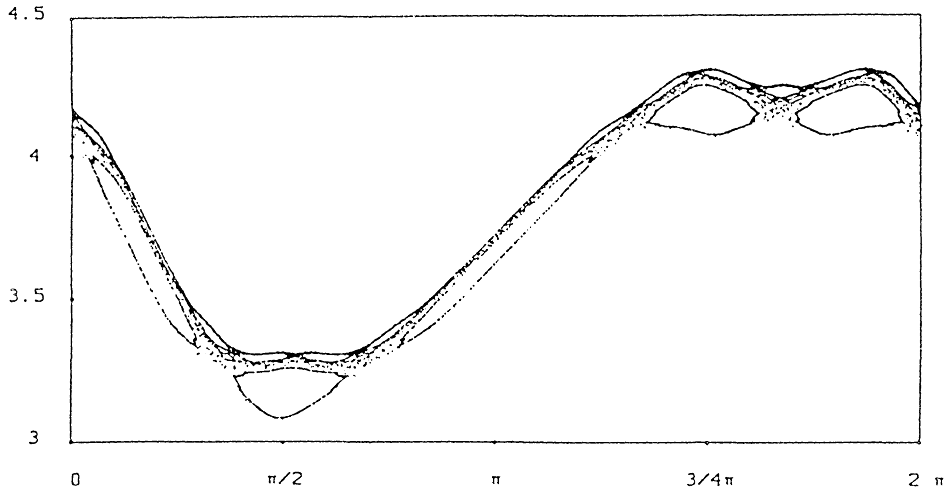

Fig. 4

Figures 1 and 3 reproduce the graphs of

$$(x, y)=(\theta+v(\theta;\varepsilon), \omega+v(\theta;\varepsilon)-v(\theta-\omega;\varepsilon))$$

for, respectively, $\varepsilon=0.97$ and $\varepsilon=1$, where v is the polynomial approximation of Theorem 2. In Figs. 2 and 4 we took some initial (x_0, y_0) lying on the graphs of, respectively, Figs. 1 and 3, and plot the evolution of such initial data according to the standard-map flow (20000 iterations). Figures 1 and 2 seem to be identical.

Appendix A

Proof of Lemma 2. Consider first a holomorphic function $h_0\colon \Delta_{\xi,\varrho}\subset\mathbf{C}^{d+2}\to\mathbf{C}$. Then, Cauchy's integral formula implies, for any j,

$$\left|\frac{\partial h_0(\theta,t;\varepsilon)}{\partial\theta_j}\right|_{\xi-\delta,\varrho}=\left|\frac{1}{2\pi i}\oint_{|\zeta_j-\theta_j|=\delta}\frac{h_0(\theta_1,\ldots,\zeta_j,\ldots,\theta_d,t;\varepsilon)}{(\zeta_j-\theta_j)^2}\,d\zeta_j\right|_{\xi-\delta,\varrho}\leqq\delta^{-1}|h|_{\xi,\varrho}. \tag{A1}$$

Now, if $h\in\mathscr{R}_p(\xi,\varrho;\mathbf{C}^d)$, (A1) implies

$$\begin{aligned}|h_\theta|^2_{\xi-\delta,\varrho}&\equiv\sup_{|c|=1}\sum_{i=1}^{d}\left|\sum_{j=1}^{d}\frac{\partial h_i}{\partial\theta_j}c_j\right|^2_{\xi-\delta,\varrho}\\&\leqq\sup_{|c|=1}\sum_i\left(\sum_j\left|\frac{\partial h_i}{\partial\theta_j}\right|_{\xi-\delta,\varrho}|c_j|\right)^2\\&\leqq\sup_{|c|=1}\sum_i\left(\sum_j|h_i|_{\xi,\varrho}\delta^{-1}|c_j|\right)^2\\&=\delta^{-2}|h|^2_{\xi,\varrho}.\quad\square\end{aligned}$$

Proof of Lemma 3. As above, let us first prove

$$|\partial_{\theta_j}^l D^{-1} h_0|_{\xi-\delta,\varrho} \leqq \sigma_l(2\delta)\,|h_0|_{\xi,\varrho} \tag{A2}$$

for a holomorphic function $h_0 : \Delta_{\xi,\varrho} \to \mathbf{C}$, with vanishing mean value.

Denote by $\|\cdot\|_{\xi,\varrho}$ the $\mathbf{L}^2$-norm

$$\|h_0\|_{\xi,\varrho}^2 \equiv \sup_{\substack{|a_1|,\ldots,|a_d| \leqq \xi \\ |b| \leqq \xi \\ |\varepsilon| \leqq \varrho}} \int_{\mathbf{T}^{d+1}} |h_0(\theta + ia, t + ib; \varepsilon)|^2 \frac{d\theta dt}{(2\pi)^{d+1}}\,.$$

Then, for any $v = (v_1, \ldots, v_d) \in \{-1, 1\}^d$, $\mu \in \{-1, 1\}$, one has

$$\sup_{|\varepsilon| \leqq \varrho} \sum_{(n,m)} e^{2(n \cdot v + m\mu)\xi} |\hat{h}_{0_{(n,m)}}(\varepsilon)|^2 \leqq \|h_0\|_{\xi,\varrho}^2\,. \tag{A3}$$

To prove (A3), let, first, $\xi' < \xi$ and consider the function

$$h'_0 \equiv h_0(\theta - iv\xi', t - i\mu\xi'; \varepsilon)\,.$$

Such function belongs to $\mathscr{R}_p(\xi - \xi', \varrho; \mathbf{C})$ and Cauchy's theorem implies

$$\hat{h}'_{0_{(n,m)}} = e^{\xi'(n \cdot v + m\mu)} \hat{h}_{0_{(n,m)}}\,.$$

Thus, Parseval's identity yields

$$\sum |\hat{h}_{0_{(n,m)}}|^2 e^{2\xi'(n \cdot v + m\mu)} = \int_{\mathbf{T}^{d+1}} |h'_0|^2 \frac{d\theta dt}{(2\pi)^{d+1}} \leqq \|h_0\|_{\xi,\varrho}^2\,.$$

Taking the supremum over $\xi' < \xi$ one obtains (A3). Now, consider first the differential case $D = \omega \cdot \partial_\theta + \partial_t$. From the maximum principle, Schwarz inequality, (DC) and (A3), it follows (dropping the index 0)

$$\begin{aligned}
|\partial_{\theta_j}^l D^{-1} h|_{\xi,\varrho} &= \left| \sum_{(n,m) \neq 0} \hat{h}_{(n,m)} \frac{n_j^l}{(\omega \cdot n + m)} e^{i(n \cdot \theta + mt)} \right|_{\xi-\delta,\varrho} \\
&= \sup_{|\varepsilon| \leqq \varrho} \sup_{(v,\mu) \in \{-1,1\}^{d+1}} \left| \sum_{(n,m) \neq 0} \hat{h}_{(n,m)} \frac{n_j^l}{(\omega \cdot n + m)} e^{(n \cdot v + m\mu)(\xi - \delta)} \right| \\
&\leqq \sup_{|\varepsilon| \leqq \varrho} \sum |\hat{h}_{(n,m)}| \left(\sum_{\mu,v} e^{2(n \cdot v + m\mu)\xi} \right)^{1/2} e^{-\delta(|n|_1 + |m|)} \frac{|n_j|^l}{|\omega \cdot n + m|} \\
&\leqq \sigma_l(2\delta) \sup_{|\varepsilon| \leqq \varrho} \left(\frac{1}{2^{d+1}} \sum_{(n,m)} |\hat{h}_{(n,m)}|^2 \sum_{\mu,v} e^{2(n \cdot v + m\mu)\xi} \right)^{1/2} \\
&\leqq \sigma_l(2\delta)\, \|h\|_{\xi,\varrho} \leqq \sigma_l(2\delta)\, |h|_{\xi,\varrho}\,.
\end{aligned}$$

The case of the finite difference operator D [see (DM)] is proved in exactly the same way substituting $\mathbf{T}^{d+1}$ with $\mathbf{T}$, $(\omega \cdot n + m)$ with $2\sin\left(\frac{n\omega}{2}\right)$ and using (2.9) in place of (DC).

Now, let $h\in\mathscr{R}_p(\xi,\varrho;\mathbf{C}^d)$ and $l=0$. Then, by (A2),

$$|D^{-1}h|^2_{\xi-\delta,\varrho}\equiv\sum_i|D^{-1}h_i|^2_{\xi-\delta,\varrho}\leqq\sigma_0(2\delta)^2\sum_i|h_i|^2_{\xi,\varrho}\equiv\sigma_0(2\delta)^2|h|^2_{\xi,\varrho}.$$

If $l=1$, then

$$\begin{aligned}|\partial_\theta D^{-1}h|^2_{\xi-\delta,\varrho}&\equiv\sup_{|c|=1}\sum_i\sum_j\left|D^{-1}\frac{\partial h_i}{\partial\theta_j}c_j\right|^2_{\xi-\delta,\varrho}\\&\leqq\sup_{|c|=1}\sum_i\left(\sum_j|\partial_{\theta_j}D^{-1}h_i|_{\xi-\delta,\varrho}|c_j|\right)^2\\&\leqq\sup_{|c|=1}\sum_i\left(\sum_j\sigma_1(2\delta)|h_i|_{\xi,\varrho}|c_j|\right)^2\\&=\sigma_1(2\delta)^2|h|_{\xi,\varrho}.\end{aligned}$$

Finally, if $h\in\mathscr{R}_p(\xi,\varrho;\mathscr{L}(\mathbf{C}^d))$, applying (A2) to the functions $\sum_i c_jh_{ij}$, one has

$$\begin{aligned}|D^{-1}h|^2_{\xi-\delta,\varrho}&=\sup_{|c|=1}\sum_i\left|\sum_j D^{-1}h_{ij}c_j\right|^2_{\xi-\delta,\varrho}\\&\leqq\sigma_0(2\delta)^2\sup_{|c|=1}\sum_i\left|\sum_j c_jh_{ij}\right|^2_{\xi,\varrho}\\&=\sigma_0(2\delta)^2|h|^2_{\xi,\varrho}.\quad\square\end{aligned}$$

Appendix B

The estimates of Lemma 5 (for both the Hamiltonian and the mapping case) are based on the following fact. Let $t\geqq 1$, $0<\delta\leqq\frac{1}{2}$, then

$$\sum_{n\in\mathbf{Z}^d}|n|^te^{-\delta|n|_1}<2(e\pi)^{d/2}\frac{\Gamma(t+d)}{\Gamma\left(\frac{d}{2}\right)}\delta^{-(t+d)}.\tag{B1}$$

To prove (B1), let $Q_n\equiv\{x\in\mathbf{R}^d:n_i\leqq x_i\leqq n_i+1,\ i=1,\dots,d\}$. Then

$$\begin{aligned}&\sum_{n\in\mathbf{Z}^d}|n|^te^{-\delta|n|_1}\\&\leqq 2^d\sum_{n\in\mathbf{N}^d}|n|^te^{-\delta|n|_1}\\&<2^d\sum_{n\in\mathbf{N}^d}e^{d\delta}\int_{Q_n}|x|^te^{-\delta|x|}dx\\&=e^{d\delta}\int_{\mathbf{R}^d}|x|^te^{-\delta|x|}dx\\&=e^{d\delta}\left(\int_{\mathbf{R}_+}r^{t+d-1}e^{-\delta r}dr\right)\cdot area\{x\in\mathbf{R}^d:|x|=1\}\\&=e^{d\delta}\frac{\Gamma(t+d)}{\delta^{t+d}}\frac{2\pi^{d/2}}{\Gamma\left(\frac{d}{2}\right)}.\end{aligned}$$

Now, in the Hamiltonian case, since $\gamma > 2$, $\tau \geqq d \geqq 1$

$$\sigma_l(\delta) < 2^{\frac{d+1}{2}} \left(\sum_{m \neq 0} \frac{1}{m^2} e^{-\delta|m|} + \gamma^2 \sum_{n \neq 0} |n|^{2(\tau+l)} e^{-\delta|n|_1} \sum_m e^{-\delta|m|} \right)^{1/2}$$

$$< 2^{\frac{d+1}{2}} \left(\frac{2}{\delta} + 2\gamma^2(\sqrt{e}+1)(e\pi)^{d/2} \frac{\Gamma(2(\tau+l)+d)}{\Gamma(d/2)} \delta^{-(2(\tau+l)+1+d)} \right)^{1/2}$$

$$< \gamma(4e\pi)^{d/4} 2(\sqrt{e}+1)^{1/2} \left(1 + \frac{1}{16(e+\sqrt{e})}\right)^{1/2} \left(\frac{\Gamma(2(\tau+l)+d)}{\Gamma(d/2)}\right)^{1/2} \delta^{-\left(\tau+l+\frac{d+1}{2}\right)}$$

$$< K_l \gamma \delta^{-k_l}.$$

For the mapping case, using (2.9) and again (B1) with $d=1$,

$$\sigma_l(\delta) < \frac{\gamma}{2} \left(\sum_{n=1}^{\infty} n^{2(\tau+l)} e^{-\delta n} \right)^{1/2}$$

$$< \frac{\gamma}{2} (\sqrt{e}\,\Gamma(2(\tau+l)+1))^{1/2} \delta^{-\left(\tau+l+\frac{1}{2}\right)},$$

which actually gives a better bound than the one indicated in the lemma.

Appendix C

Proof of Lemma 9. Let

$$A_N^* \equiv \{(n,m) \in \mathbf{Z}^2 : n \geqq N, \omega n - \tfrac{3}{2} < m < \omega n + \tfrac{3}{2}\},$$

$$B \equiv \{(n,m) \in \mathbf{Z}^2 : \omega n + \tfrac{3}{2} < m\},$$

$$a_{(n,m)}^{(k)} \equiv \left(\frac{|n|^k}{\omega n - m} \right)^2 e^{-\delta(|n|+|m|)}.$$

Then, one has

$$\sigma_k(\delta) = 2 \left(\sum_{\substack{(n,m) \in \mathbf{Z}^2 \\ (n,m) \neq (0,0)}} a_{(n,m)}^{(k)} \right)^{1/2}$$

$$= 2 \left(\sum_{A_N} a_{(n,m)}^{(k)} + 2 \sum_{A_n^*} a_{(n,m)}^{(k)} + 2 \sum_B a_{(n,m)}^{(k)} \right)^{1/2}.$$

Since, for any n,

$$\#\{m \in \mathbf{Z} : \omega n - \tfrac{3}{2} < m < \omega n + \tfrac{3}{2}\} = 3,$$

using (DC) with $\tau = 1$, one gets

$$\sum_{A_N^*} a_{(n,m)}^{(k)} < 3\gamma^2 e^{\delta/2} \sum_{n=N}^{\infty} e^{-\delta(1+\omega)n} n^{2(k+1)}.$$

This last sum can be computed explicitly, using the formula

$$\sum_{n=N}^{\infty} e^{-\beta n} n^l = (-1)^l \frac{d^l}{d\beta^l} \frac{e^{-\beta N}}{1-e^{-\beta}}, \qquad l \in \mathbf{N}.$$

Thus, using the estimate

$$\frac{e^{-\beta}}{1-e^{-\beta}}<\frac{1}{\beta}, \qquad \forall\beta>0,$$

one sees that

$$\sum_{A_N^*} a_{(n,m)}^{(k)}<S_k^{(N)}.$$

Now, let

$$D_{(n,m)}\equiv\{(x,y)\in\mathbf{R}^2: n-1\leqq x\leqq n,\\ \omega(x-n)+m-1\leqq y\leqq\omega(x-n)+m\}.$$

Then

$$\sum_B a_{(n,m)}^{(k)}<\sum_B e^{\delta}\int_{D_{(n,m)}}\frac{e^{-\delta(|x|+|y|)}}{|\omega x-y|}x^{2k}dxdy=e^{\delta}I_k',$$

where

$$I_k'\equiv\int_{\{y-\omega x\geqq\frac{1}{2}\}}\frac{e^{-\delta(|x|+|y|)}}{|\omega x-y|}x^{2k}dxdy.$$

Making the linear change of variables $(\xi,\eta)=(x,y-\omega x)$ one obtains

$$I_k'=\int_{1/2}^{\infty}\left(e^{\delta\eta}\int_{-\infty}^{-\eta/\omega}e^{\delta(1+\omega)\xi}\xi^{2k}d\xi+e^{-\delta\eta}\int_{-\eta/\omega}^{0}e^{\delta(1-\omega)\xi}\xi^{2k}d\xi\right.\\ \left.+e^{-\delta\eta}\int_0^{\infty}e^{-\delta(1+\omega)\xi}\xi^{2k}d\xi\right)\frac{1}{\eta^2}\,d\eta.$$

Thus, recalling that $\alpha\equiv\delta(1+\omega)$, one obtains

$$I_0'=\frac{2}{1-\omega^2}\int_{\delta/2}^{\delta/2\omega}\frac{e^{-\eta}}{\eta^2}\,d\eta<\frac{2}{1-\omega^2}\int_{\delta/2}^{\delta/2\omega}\frac{1}{\eta^2}\,d\eta=\frac{4}{\alpha}\equiv I_0.$$

Analogously, one obtains

$$I_1'=\frac{8}{\delta^3}\frac{1}{(1-\omega^2)^3}(e^{-\delta/2}(1+3\omega^2)-e^{-\delta/2\omega}(3\omega+\omega^3)-r)$$

with some $r>0$. Thus

$$I_1'<\frac{8}{\delta^3}\frac{1}{(1-\omega^2)^3}\left[(1-\omega)^3+\delta(1-\omega^2)+\frac{\delta^2}{8}(1+3\omega^2)\right]\\ <\frac{8}{\alpha^3}\left(1+\delta\frac{9}{8(1-\omega)^3}\right)\equiv I_1. \quad \square$$

Appendix D

Here, we report the program "INITIAL", which evaluates strict upper bounds on the norms relative to the initial approximant v.

The program is written in FORTRAN and must, because of the interval-arithmetic subroutines, be run on a VAX. The basic informations on the structure of a VAX machine can be found in [39].

We will try to maintain the notation as close as possible to that of Sect. 7, which illustrates the basic strategy of the program.

All the functions involved in the program will be trigonometric polynomials of the form (7.3) and, with abuse of language, we will refer to the real number c as to Fourier coefficients. A function "a" of the form (7.3) will be represented by (A, $N1$, $N2$, NA, $MAXA$), where A is a vector of length NA listing the Fourier coefficients of a, $N1$, and $N2$ correspond to the relative Fourier indices (n, m) and $MAXA$ is the maximum of $\{|n|, |m|\}$.

The functions $u(1), \ldots, u(24)$ will be represented by a unique quintuple (C, $N1$, $N2$, NC, $MAXC$); in this case NC represents the sum of the number of Fourier coefficients of all the u's and $MAXC$ is now also a vector [$MAXC(i)$ refers to $u(i)$].

All the real numbers R (or vectors) will be represented by a couple of numbers, RD and RU, which are left and right ends of an interval containing R. (In the comments R will refer, for short, to such a couple.)

Comments are preceded by "$C\ldots$". Rigorous bounds on results of a sequence of elementary operations will be obtained by successive calls of the relative interval-arithmetic subroutines and, to simplify the reading, these sequences of calls will be preceded by "$C-OP$." followed by the FORTRAN standard notation relative to the sequence of elementary operations in question.

```
      PROGRAM INITIAL
C
C...This program must be run in "G_floating" (compare FUNCTION UP).
C
      IMPLICIT DOUBLE PRECISION (A-H,O-Z)
      DIMENSION CD(5512),CU(5512),N1(5512),N2(5512),NC(0:24),MAXC(24)
      DIMENSION AD(5512),AU(5512),NA1(5512),NA2(5512)
      DIMENSION FACTD(0:24),FACTU(0:24),K(24)
      DIMENSION RMD(-24:24,-24:24),RMU(-24:24,-24:24)
      DIMENSION UD(24),UU(24),U1D(24),U1U(24)
      PARAMETER (P1=1.D+00,P2=2.D+00,P5=5.D+00,HALF=.5D+00)
C
C...The numbers RHO = 0.015 and CSI = 1/10 are given (Since such numbers
C   do not have a finite binary expansion we will substitute them
C   with upper bounds provided by the interval-arithmetic subroutines,
C   the relative lower bounds RHOD and CSID will never be used.)
C
      X=15.D+00
      Y=1000.D+00
      CALL DIV(X,X,Y,Y,RHOD,RHO)
      X=P1
      Y=10.D+00
      CALL DIV(X,X,Y,Y,CSID,CSI)
C
      J0=24
C
C...Table of the factorials used in the program: FACT(n)=n!
C
      FACTD(0)=1
      FACTU(0)=1
     DO 50 I=1,J0
C-OP.  FACT(I)=FACT(I-1)*I
      FACTD(I)=I
      FACTU(I)=I
      CALL MUL(FACTD(I-1),FACTU(I-1),FACTD(I),FACTU(I))
 50   CONTINUE
C
C...   Definition of omega : OM=(SQRT(5.D+00)-1)/2
      XD=P5
      XU=P5
      CALL SQR(XD,XU)
      CALL SUM(-P1,-P1,XD,XU)
      CALL DIV(XD,XU,P2,P2,OMD,OMU)
```

```
C
C...The function u(1) is given:
        NC(1)=4
        MAXC(1)=1
        N1(1)=1
        N2(1)=0
C-OP.   C(1)=.5/OM**2
        XD=OMD
        XU=OMU
        CALL SQ(XD,XU)
        CALL DIV(HALF,HALF,XD,XU,CD(1),CU(1))
        N1(2)=1
        N2(2)=-1
C-OP.   C(2)=.5/(1-OM)**2
        XD=P1
        XU=P1
        CALL SUM(-OMU,-OMD,XD,XU)
        CALL SQ(XD,XU)
        CALL DIV(HALF,HALF,XD,XU,CD(2),CU(2))
        N1(3)=-1
        N2(3)=0
        CD(3)=-CU(1)
        CU(3)=-CD(1)
        N1(4)=-1
        N2(4)=1
        CD(4)=-CU(2)
        CU(4)=-CD(2)
C
C...u(J) is constructed from u(J-1),u(J-2),...,u(1):
C
       DO 1 I=1,J0
        K(I)=0
 1      CONTINUE
       DO 2 J=2,J0
C
C...The vector K (defined in the next comment )is reset equal to
C   zero (for this it is enough to set K(1)=0,see FUNCTION NK):
C
        K(1)=0
C
C...The FUNCTION NK(K,N) provides iteratively the integer vectors
C   K(1),...,K(N), such that K(1)+2*K(2)+...+N*K(N)=N  (the first
C   call must be done with the vector K identically zero and the
C   first output is K=(0,...,0,1), the last one is K=(N,0,...,0);
C   NK is equal to zero after the last call, otherwise is one).
C
 1000   NNK=NK(K,J-1)
C
        MODK=0
       DO 100 I=1,J-1
        MODK=MODK+K(I)
 100    CONTINUE
C
C...the function A=(d/dtheta)**|k| [sin theta + sin(theta-t)] is given:
        NA=4
        MAXA=1
        NA1(1)=1
        NA2(1)=0
        AD(1)=HALF*(-1)**MODK
        AU(1)=AD(1)
        NA1(2)=-1
        NA2(2)=0
        AD(2)=-HALF
        AU(2)=AD(2)
        NA1(3)=1
        NA2(3)=-1
        AD(3)=AD(1)
        AU(3)=AD(3)
        NA1(4)=-1
        NA2(4)=1
        AD(4)=AD(2)
        AU(4)=AD(4)
```

```
C
        RD=P1
        RU=P1
       DO 3 I=1,J-1
C
C...The Fourier coefficients of the function phi(k) are computed:
C
       DO 4 II=1,K(I)
C
C...The subroutine FMULT computes the Fourier coefficients of the
C   product of a function A with u(I).The result is called again A.
C
        CALL FMUL(CD,CU,N1,N2,NC,MAXC,I,AD,AU,NA1,NA2,NA,MAXA)
 4      CONTINUE
C
C
C...Computation of  k1!*k2!*...*kJ!
C
       IF (K(I).GE.2) THEN
C-OP.   R=R*FACT(K(I))
        CALL MUL(FACTD(K(I)),FACTU(K(I)),RD,RU)
       ENDIF
C
 3      CONTINUE
C
C...The (n,m)-Fourier coefficients of phi(k) are added up in the
C   auxiliary matrix RM(n,m).
C
       DO 5 N=1,NA
C-OP.   RM(NA1(N),NA2(N))=RM(NA1(N),NA2(N))+A(N)/R
        CALL DIV(AD(N),AU(N),RD,RU,XD,XU)
        CALL SUM(XD,XU,RMD(NA1(N),NA2(N)),RMU(NA1(N),NA2(N)))
 5      CONTINUE
C
C...Update of MAXC(j):
C
       IF (MAXA.GT.MAXC(J)) MAXC(J)=MAXA
       IF (NNK.EQ.1) GOTO 1000
C
C...Definition of u(j) and clearing of the matrix RM:
C
        NN=NC(J-1)
C
C...Since we chose to work with functions with vanishing mean value, we
C   have  RM(0,0)=0 (the computer will actually give a value of about
C   + or - 10**-16 ,due to approximations involved in evaluating cancel-
C   lations):
C
        RMD(0,0)=0.
        RMU(0,0)=0.
C
C... (D**-2) phi(k) is computed:
C
       DO 6 I=-MAXC(J),MAXC(J)
       DO 6 II=-MAXC(J),MAXC(J)
C-OP.  IF (RM(I,II).NE.0.) THEN
       IF ((RMD(I,II).NE.0.).OR.(RMU(I,II).NE.0.)) THEN
        NN=NN+1
        N1(NN)=I
        N2(NN)=II
C-OP.   C(NN)=RM(I,II)/(OM*I+II)**2
        XD=I
        XU=I
        CALL MUL(OMD,OMU,XD,XU)
        R=II
        CALL SUM(R,R,XD,XU)
        CALL SQ(XD,XU)
        CALL DIV(RMD(I,II),RMU(I,II),XD,XU,CD(NN),CU(NN))
        RMD(I,II)=0.
        RMU(I,II)=0.
       ENDIF
 6      CONTINUE
        NC(J)=NN
 2      CONTINUE
```

```
C
C...The computation of the Fourier coefficients of u(j) (j=1,...,24) is
C   completed.
C...The second part of the program, where V, V1, E of Lemma 8 are compu-
C   ted according to formulae (7.6),(7.7),(7.8), and (7.1), follows.
C
C...     EXC=EXP(CSI)
         CALL EXPN(CSI,CSI,EXCD,EXCU)
        DO 7 J=1,J0
        DO 7 I=NC(J-1)+1,NC(J)
C
C...Since the u's are odd:
        IF ((N1(I).GT.0).OR.((N1(I).EQ.0).AND.(N2(I).GT.0))) THEN
C-OP.     EX=EXP((N1(I)+ABS(N2(I)))*CSI)
         M=N1(I)+ABS(N2(I))
         CALL POWER(M,EXCD,EXCU,EXD,EXU)
C-OP.     U(J)=U(J)+(EX+1/EX)*ABS(C(I))
         CALL DIV(P1,P1,EXD,EXU,XD,XU)
         CALL SUM(EXD,EXU,XD,XU)
         CALL MUL(ABS(CD(I)),ABS(CU(I)),XD,XU)
         CALL SUM(XD,XU,UD(J),UU(J))
C-OP.     U1(J)=U1(J)+(EX+1/EX)*ABS(C(I))*N1(I)
         X=N1(I)
         CALL MUL(X,X,XD,XU)
         CALL SUM(XD,XU,U1D(J),U1U(J))
        ENDIF
 7       CONTINUE
C
        DO 8 J=1,J0
C-OP.     V=V+RHO**J*U(J)
         CALL POWER(J,RHO,RHO,YD,YU)
         XD=YD
         XU=YU
         CALL MUL(UD(J),UU(J),XD,XU)
         CALL SUM(XD,XU,VD,VU)
C-OP.     V1=V1+RHO**J*U1(J)
         XD=YD
         XU=YU
         CALL MUL(U1D(J),U1U(J),XD,XU)
         CALL SUM(XD,XU,V1D,V1U)
 8       CONTINUE
        DO 9 J=1,J0-1
C-OP.     S=0.
         SD=0.
         SU=0.
         K(1)=0
 1010    NNK=NK(K,J)
C-OP.     P=1.
         PD=P1
         PU=P1
        DO 10 N=1,J
C-OP.     P=P*U(N)**K(N)/FACT(K(N))
         M=K(N)
         CALL POWER(M,UD(N),UU(N),XD,XU)
         BD=FACTD(K(N))
         BU=FACTU(K(N))
         CALL DIV(XD,XU,BD,BU,XD,XU)
         CALL MUL(XD,XU,PD,PU)
 10      CONTINUE
C-OP.     S=S+P
         CALL SUM(PD,PU,SD,SU)
         IF (NNK.EQ.1) GO TO 1010
C-OP.     E=E+V**J/FACT(J)-RHO**J*S
         CALL POWER(J,VD,VU,YD,YU)
         CALL POWER(J,RHO,RHO,ZD,ZU)
         XD=YD
         XU=YU
         BD=FACTD(J)
         BU=FACTU(J)
         CALL DIV(XD,XU,BD,BU,XD,XU)
         CALL MUL(SD,SU,ZD,ZU)
         CALL SUM(-ZU,-ZD,XD,XU)
         CALL SUM(XD,XU,ED,EU)
 9       CONTINUE
```

```
C-OP.   EX=EXP(CSI)
        EXD=EXCD
        EXU=EXCU
C-OP.   E=E*RHO*(EX+1/EX+EX**2+1/EX**2)/2
C-OP.   EX1=EXP(CSI+V)
C-OP.   E=E+(V**J0/FACT(J0))*RHO*(EX1+1/EX1+EX1**2+1/EX1**2)/2
        CALL MUL(VD,VU,YD,YU)
        BD=FACTD(J0)
        BU=FACTU(J0)
        CALL DIV(YD,YU,BD,BU,YD,YU)
        ZZD=YD
        ZZU=YU
        CALL MUL(RHO,RHO,ED,EU)
        X=2.D+00
        AAD=ED
        AAU=EU
        CALL DIV(AAD,AAU,X,X,ED,EU)
        CALL DIV(P1,P1,EXD,EXU,XD,XU)
        YD=XD
        YU=XU
        CALL SQ(YD,YU)
        ZD=EXD
        ZU=EXU
        CALL SQ(ZD,ZU)
        CALL SUM(EXD,EXU,XD,XU)
        CALL SUM(YD,YU,XD,XU)
        CALL SUM(ZD,ZU,XD,XU)
        CALL MUL(XD,XU,ED,EU)
        XXD=CSI+VD
        XXU=CSI+VU
        CALL EXPN(XXD,XXU,EX1D,EX1U)
        CALL DIV(P1,P1,EX1D,EX1U,XD,XU)
        YD=XD
        YU=XU
        CALL SQ(YD,YU)
        ZD=EX1D
        ZU=EX1U
        CALL SQ(ZD,ZU)
        CALL SUM(EX1D,EX1U,XD,XU)
        CALL SUM(YD,YU,XD,XU)
        CALL SUM(ZD,ZU,XD,XU)
        CALL MUL(XD,XU,ZZD,ZZU)
        CALL MUL(RHO,RHO,ZZD,ZZU)
        X=2.D+00
        AAD=ZZD
        AAU=ZZU
        CALL DIV(AAD,AAU,X,X,ZZD,ZZU)
        CALL SUM(ZZD,ZZU,ED,EU)
C
C...The computation of the intervals containing the numbers   V, V1, E
C   is completed. Now we let the computer convert the result in de-
C   cimal notation. That the numbers given in Lemma 8 are  (generous)
C   upper bounds on  VU, V1U,  EU  follows from, e.g., [39].
C
        WRITE(*,*)VD,V1D,ED
        WRITE(*,*)VU,V1U,EU
        END
C
C.FUN.
C
        FUNCTION NK(K,N)
        DIMENSION K(N)
C
        KS=K(1)
        K(1)=0
       DO 1 I=2,N
        IP=I
        IF (K(I).GT.0) GO TO 2
 1      CONTINUE
        K(N)=1
        NK=1
        IF (N.EQ.1) NK=0
        RETURN
```

```
 2       K(IP)=K(IP)-1
         KS=KS+IP
         IP=IP-1
        DO 10 I=IP,1,-1
         K(I)=KS/I
         KS=KS-I*K(I)
         IF(KS.EQ.0) GOTO 3
 10      CONTINUE
 3       NK=1
        IF (K(1).EQ.N) NK=0
         RETURN
         END
C.SUB.
         SUBROUTINE FMUL(CD,CU,N1,N2,NC,MAXC,I0,AD,AU,NA1,NA2,NA,MAXA)
         IMPLICIT DOUBLE PRECISION (A-H,O-Z)
         DIMENSION CD(5512),CU(5512),N1(5512),N2(5512),
      1NC(0:24),MAXC(24)
         DIMENSION AD(5512),AU(5512),NA1(5512),NA2(5512)
         DIMENSION RMD(-24:24,-24:24),RMU(-24:24,-24:24)
         MAXA=MAXC(I0)+MAXA
        DO 10 I=-MAXA,MAXA
        DO 10 II=-MAXA,MAXA
         RMD(I,II)=0.
         RMU(I,II)=0.
10       CONTINUE
        DO 1 N=NC(I0-1)+1,NC(I0)
        DO 1 M=1,NA
         I=N1(N)+NA1(M)
         II=N2(N)+NA2(M)
C-OP.    RM(I,II)=RM(I,II)+C(N)*A(M)
         XD=AD(M)
         XU=AU(M)
         CALL MUL(CD(N),CU(N),XD,XU)
         CALL SUM(XD,XU,RMD(I,II),RMU(I,II))
1        CONTINUE
         NA=0
        DO 2 I=-MAXA,MAXA
        DO 2 II=-MAXA,MAXA
C-OP.IF (RM(I,II).NE.0.) THEN
        IF ((RMD(I,II).NE.0.).OR.(RMU(I,II).NE.0.)) THEN
         NA=NA+1
         AD(NA)=RMD(I,II)
         AU(NA)=RMU(I,II)
         RMD(I,II)=0.
         RMU(I,II)=0.
         NA1(NA)=I
         NA2(NA)=II
        ENDIF
2        CONTINUE
         RETURN
         END
C
C
C...The subroutines for the interval-arithmetic follow.
C   We point out that the machine we used indicates automathically
C   several arithmetical errors like overflows or square roots of
C   negative numbers. Running this program on other machines might
C   require the addition of routines controlling such errors.
C
C.FUN.
         DOUBLE PRECISION FUNCTION UP(R)
C
C...This function gives the smallest strict upper bound on a real number
C   R  represented in G_floating (double precision) notation.
C
     INTEGER*2 KP(4)
     REAL*8 R,X
     EQUIVALENCE (X,KP(1))
     X=R
   IF (X.GT.0.) THEN
    IF (KP(4).EQ.32767) THEN
     KP(4)=-32768
     UP=X
     RETURN
```

```
     ENDIF
      KP(4)=KP(4)+1
     IF (KP(4).NE.0) THEN
      UP=X
      RETURN
     ENDIF
     IF (KP(3).EQ.32767) THEN
      KP(3)=-32768
      UP=X
      RETURN
     ENDIF
      KP(3)=KP(3)+1
     IF (KP(3).NE.0) THEN
      UP=X
      RETURN
     ENDIF
     IF (KP(2).EQ.32767) THEN
      KP(2)=-32768
      UP=X
      RETURN
     ENDIF
      KP(2)=KP(2)+1
     IF (KP(2).NE.0) THEN
      UP=X
      RETURN
     ENDIF
      KP(1)=KP(1)+1
      UP=X
      RETURN
    ELSE IF (X.LT.0.) THEN
     IF (KP(4).EQ.-32768) THEN
      KP(4)=32767
      UP=X
      RETURN
     ENDIF
      KP(4)=KP(4)-1
     IF (KP(4).NE.-1) THEN
      UP=X
      RETURN
     ENDIF
     IF (KP(3).EQ.-32768) THEN
      KP(3)=32767
      UP=X
      RETURN
     ENDIF
      KP(3)=KP(3)-1
     IF (KP(3).NE.-1) THEN
      UP=X
      RETURN
     ENDIF
     IF (KP(2).EQ.-32768) THEN
      KP(2)=32767
      UP=X
      RETURN
     ENDIF
      KP(2)=KP(2)-1
     IF (KP(2).NE.-1) THEN
      UP=X
      RETURN
     ENDIF
      KP(1)=KP(1)-1
      UP=X
      RETURN
    ELSE
      UP=X
      RETURN
    ENDIF
      END
C.FUN.
        DOUBLE PRECISION FUNCTION DOWN(R)
        IMPLICIT DOUBLE PRECISION (A-H,O-Z)
        DOWN=-UP(-R)
        RETURN
        END
```

```
C
C...The subroutines for the elementary operations follow. That taking,
C   basically, UP and DOWN of the results given by the computer is
C   enough to get rigorous results follows from [39], pag. 177.
C
C.SUB.
        SUBROUTINE SUM(AD,AU,BD,BU)
C
C...                     B = A + B
C
        IMPLICIT DOUBLE PRECISION (A-H,O-Z)
        BD=DOWN(AD+BD)
        BU=UP(AU+BU)
        RETURN
        END
C.SUB.
        SUBROUTINE MUL(AD,AU,BD,BU)
C
C...                     B = A * B
C
        IMPLICIT DOUBLE PRECISION (A-H,O-Z)
      IF (AD.GE.0.) THEN
       IF (BD.GE.0.) THEN
        BD=DOWN(AD*BD)
        BU=UP(AU*BU)
        RETURN
       ELSE IF (BU.LE.0.) THEN
        BD=DOWN(AU*BD)
        BU=UP(AD*BU)
        RETURN
       ELSE
        BD=DOWN(AU*BD)
        BU=UP(AU*BU)
        RETURN
       ENDIF
      ELSE IF (AU.LE.0.) THEN
       IF (BD.GE.0.) THEN
        B=DOWN(AD*BU)
        BU=UP(AU*BD)
        BD=B
        RETURN
       ELSE IF (BU.LE.0.) THEN
        B=DOWN(AU*BU)
        BU=UP(AD*BD)
        BD=B
        RETURN
       ELSE
        B=DOWN(AD*BU)
        BU=UP(AD*BD)
        BD=B
        RETURN
       ENDIF
      ELSE
       IF (BD.GE.0.) THEN
        BD=DOWN(AD*BU)
        BU=UP(AU*BU)
        RETURN
       ELSE IF (BU.LE.0.) THEN
        B=DOWN(AU*BD)
        BU=UP(AD*BD)
        BD=B
        RETURN
       ELSE
        B=DOWN(AD*BU)
        R=DOWN(AU*BD)
        IF (R.LT.B) B=R
        C=UP(AD*BD)
        R=UP(AU*BU)
        IF (R.GT.C) C=R
        BD=B
        BU=C
        RETURN
       ENDIF
      ENDIF
        END
```

```
C.SUB.
        SUBROUTINE DIV(AD,AU,BD,BU,CD,CU)
C
C...                     C = A / B
C
        IMPLICIT DOUBLE PRECISION (A-H,O-Z)
      IF (AD.GE.0.) THEN
       IF (BD.GT.0.) THEN
        CD=DOWN(AD/BU)
        CU=UP(AU/BD)
        RETURN
       ELSE IF (BU.LT.0.) THEN
        CD=DOWN(AU/BU)
        CU=UP(AD/BD)
        RETURN
       ENDIF
      ELSE IF (AU.LE.0.) THEN
       IF (BD.GT.0.) THEN
        CD=DOWN(AD/BD)
        CU=UP(AU/BU)
        RETURN
       ELSE IF (BU.LT.0.) THEN
        CD=DOWN(AU/BD)
        CU=UP(AD/BU)
        RETURN
       ENDIF
      ELSE
       IF (BD.GT.0.) THEN
        CD=DOWN(AD/BD)
        CU=UP(AU/BD)
        RETURN
       ELSE IF (BU.LT.0.) THEN
        CD=DOWN(AU/BU)
        CU=UP(AD/BU)
        RETURN
       ENDIF
      ENDIF
        R=0.
        CD=1/R
        RETURN
        END
C.SUB.
        SUBROUTINE SQ(AD,AU)
C
C...             A = A**2
C
        IMPLICIT DOUBLE PRECISION (A-H,O-Z)
       IF (AD.GE.0.) THEN
        AD=DOWN(AD*AD)
        AU=UP(AU*AU)
        RETURN
       ELSE IF (AU.LE.0.) THEN
        B=AD
        AD=DOWN(AU*AU)
        AU=UP(B*B)
        RETURN
       ELSE
        B=AD
        AD=DOWN(AD*AU)
        C=UP(B*B)
        AU=UP(AU*AU)
        IF (C.GT.AU) AU=C
        RETURN
       ENDIF
        END
C.SUB.
        SUBROUTINE SQR(AD,AU)
C
C...             A = SQRT(A)
C
        IMPLICIT DOUBLE PRECISION (A-H,O-Z)
        BD=AD
        BU=AU
        AD=SQRT(AD)
```

```
1       AD=DOWN(AD)
        IF (UP(AD*AD).GT.BD) GO TO 1
        AU=SQRT(AU)
2       AU=UP(AU)
        IF (DOWN(AU*AU).LT.BU) GO TO 2
        RETURN
        END

C.SUB.
        SUBROUTINE POWER(M,XD,XU,YD,YU)
C
C...             Y = Y**M , M positive integer:
C
        IMPLICIT DOUBLE PRECISION (A-H,O-Z)
        DIMENSION N(0:20)
       IF (M.EQ.0) THEN
        YD=1.D+00
        YU=1.D+00
        RETURN
       ENDIF
        CALL BIN(M,N,IMIN,IMAX)
        YD=XD
        YU=XU
       DO 1 I=1,IMIN
        CALL SQ(YD,YU)
1       CONTINUE
        ZD=YD
        ZU=YU
        I1=IMIN
       DO 2 I=IMIN+1,IMAX
       IF (N(I).NE.0) THEN
        I2=I
      DO 3 J=1,I2-I1
        CALL SQ(ZD,ZU)
3       CONTINUE
        CALL MUL(ZD,ZU,YD,YU)
        I1=I
       ENDIF
2       CONTINUE
        RETURN
        END
C.SUB.
        SUBROUTINE BIN(M,N,IMIN,IMAX)
C
C...This subroutine gives the binary decomposition of any strictly
C   positive integer M < 2**20.
C...The outcomes are the vector N , IMIN and IMAX:
C
C                         M = Sum(i=IMIN,..,IMAX) N(i)*2**i.
C
        DIMENSION N(0:20)
        M0=M
       DO 1 I=0,20
        N(I)=MOD(M0,2)
        M0=M0/2
1       CONTINUE
C
C...Computation of IMIN and IMAX:
C
        I=-1
2       I=I+1
       IF (N(I).EQ.0) GO TO 2
        IMIN=I
        I=21
3       I=I-1
       IF (N(I).EQ.0) GO TO 3
        IMAX=I
        RETURN
        END
C.SUB.
        SUBROUTINE EXPN(AD,AU,BD,BU)
C
C...             B = EXP(A) ,   0 < A < 1/2 :
```

```
C
      IMPLICIT DOUBLE PRECISION (A-H,O-Z)
      XD=1.D+00
      XU=1.D+00
      BD=1.D+00
      BU=1.D+00
C
C...  R = K!
C
      RD=1.D+00
      RU=1.D+00
     DO 1 K=1,15
      CALL MUL(AD,AU,XD,XU)
      C=K
      CALL MUL(C,C,RD,RU)
      CALL DIV(XD,XU,RD,RU,YD,YU)
      CALL SUM(YD,YU,BD,BU)
 1    CONTINUE
C
C...R = 10**-17 is an upper bound on Sum(k=16,17,...){(1/2)**k/k!} :
      R=1.D-17
      BU=UP(BU+R)
      RETURN
      END
```

We indicate, now, the modifications necessary to use the above program for the standard map case (for ease of notation we symbolically indicate intervals with numbers and sequences of calls of interval-arithmetic subroutines with the standard FORTRAN notation).

Suppress $N2$ and $NA2$; substitute (everywhere it appears) 5512 with 760, 24 with 38; RM becomes a $(-38, 38)$-vector.

Define OM as $(SQRT(D5)-1)*PI$, where "PI" denotes an interval containing 3.141592653589793....

Add a subroutine COSINE, which evaluates $\cos x$ with an accuracy of about $10^{**}-16$ in the fashion of the above SUBROUTINE EXPN.

The function $u(1)$ is given by the following sequence of instructions: $NC(1)=2$, $MAXC(1)=1$, $N1(1)=1$, $C(1)=-1/(4^*(\text{COS}(OM)-1))$, $N1(2)=-1$, $C(2)=-C(1)$.

The function $A=(d/d\,\text{theta})^{**}|k|$ (sin theta) (which substitutes the function $A=(d/d\,\text{theta})^{**}|k|$ [sin theta + sin(theta-t)] above) is given by the following sequence of instructions: $NA=2$, $MAXA=1$, $NA1(1)=1$, $A(1)=\text{HALF}^*(-1)^{**}(K+1)$, $NA1(2)=-1$, $A(2)=\text{HALF}$.

In the second part of the program one needs simply to recall that, now, Ch(csi) is defined as cosh(csi) (compare Lemma 11).

Remark 13. It is clear that the efficiency of "INITIAL" can be certainly improved, however in order to get "significantly" better results one should probably turn to more efficient computers (compare the data reported in Appendix E). Also, we tried to make the program as simple as possible so as to reduce the possibility of mechanical (and human, of course,) mistakes.

We conclude by mentioning that the running-time of "INITIAL" on the VAX 8600 of the E.T.H. in Zürich was about, respectively, 60 min of CPU time for the standard map and 140 min for (H1). The running-times of the "numerical" version of "INITIAL", i.e., a version without interval-arithmetic, were about, respectively, 8 and 14 minutes.

Appendix E

Here we report some data relative to the behaviour of the KAM algorithm (see Sect. 6) with respect to the initial approximation v in the Hamiltonian case (H1).

Let $v_{l_0}=v$ be the polynomial approximant (7.3) of "order l_0".

In the following table we report "with four significant digits" (see below) and for $l_0=1, 2, \ldots, 24$, the maximum ϱ for which the KAM algorithm with initial approximant v_{l_0} converges yielding a KAM torus analytic in $\mathscr{R}_p(\xi, \varrho; \mathbf{C})$ for some $\xi>0$. "With four significant digits" means that the KAM algorithm diverges if one increases the values of ϱ by 1/10000. (The value of j at which the algorithm diverges is, in the present situation, between 5 and 18.)

l_0	ϱ	ξ	l_0	ϱ	ξ
1	0.0008	$0.8/2^{10}$	13	0.0108	$0.16/2^{13}$
2	0.0023	$0.6/2^{11}$	14	0.0116	$0.15/2^{11}$
3	0.0033	$0.5/2^{10}$	15	0.0122	$0.15/2^{16}$
4	0.005	$0.5/2^{13}$	16	0.0126	$0.14/2^{16}$
5	0.0055	$0.4/2^{12}$	17	0.0132	$0.14/2^{11}$
6	0.0068	$0.3/2^{11}$	18	0.0135	$0.13/2^{11}$
7	0.008	$0.3/2^{11}$	19	0.0140	$0.13/2^{11}$
8	0.0078	$0.25/2^{11}$	20	0.0145	$0.12/2^{13}$
9	0.009	$0.2/2^{13}$	21	0.0138	$0.106/2^{14}$
10	0.0097	$0.19/2^{11}$	22	0.0144	$0.105/2^{13}$
11	0.0104	$0.18/2^{11}$	23	0.0146	$0.103/2^{11}$
12	0.0113	$0.17/2^{12}$	24	0.015	$0.1/2^{11}$

Acknowledgements. We are deeply indebted to G. Gallavotti for having introduced us to these topics and for his continuous help and encouragement. We are grateful to J. Moser for many helpful suggestions and for the kind hospitality at the Forschungsinstitut für Mathematik of the ETH – Zürich. Finally, we thank A. Giorgilli and P. Wittwer for useful comments on the numerical part.

References

1. Arnold, V.I.: Proof of a Theorem by A.N. Kolmogorov on the invariance of quasiperiodic motions under small perturbations of the Hamiltonian. Russ. Math. Surv. **18**, 9 (1963)
2. Arnold, V.I., Avez, A.: Ergodic problems of classical mechanics. New York: Benjamin 1968
3. Aubry, S., Le Daeron, P.Y.: The discrete Frenkel-Kontorova model and its extensions. I. Physica **8**D, 381 (1983)
4. 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. Commun. Math. Phys. **113**, 87 (1987)

5. Benettin, G., Gallavotti, G.: Stability of motions near resonances in quasi-integrable Hamiltonian systems. J. Stat. Phys. **44**, 293 (1986)
6. Braess, D., Zehnder, E.: On the numerical treatment of a small divisor problem. Numer. Math. **39**, 269 (1982)
7. Celletti, A., Chierchia, L.: Rigorous estimates for a computer-assisted KAM theory. J. Math. Phys. **28**, 2078 (1987)
8. Celletti, A., Falcolini, C., Porzio, A.: Rigorous numerical stability estimates for the existence of KAM tori in a forced pendulum. Ann. Inst. Henri Poincaré **47**, 85 (1987)
9. Chirikov, B.V.: A universal instability of many dimensional oscillator systems. Phys. Rep. **52**, 263 (1979)
10. Eckmann, J.-P., Wittwer, P.: Computer methods and Borel summability applied to Feigenbaum's equation. Lecture Notes in Physics, Vol. 227. Berlin, Heidelberg, New York: Springer 1985
11. Eliasson, H.: Absolutely convergent series expansions for quasiperiodic motions, preprint. Univ. of Stockholm (1987)
12. Escande, D.F.: Stochasticity in classical Hamiltonian systems: Universal aspects. Phys. Rep. **121**, 165 (1985)
13. Escande, D.F.: Private communication
14. Escande, D.F., Doveil, F.: Renormalization method for computing the threshold of the large-scale stochastic instability in two degrees of freedom Hamiltonian systems. J. Stat. Phys. **26**, 257 (1981)
15. Falcolini, C.: Private communication
16. Gallavotti, G.: Perturbation theory for classical Hamiltonian systems. In: Scaling and self-similarity in physics. Fröhlich, J. (ed.) PPh. 7, Boston: Birkhäuser 1984
17. Greene, J.M.: A method for determining a stochastic transition. J. Math. Phys. **20**, 1183 (1979)
18. Hénon, M., Heiles, C.: The applicability of the third integral of motion: Some numerical experiments. Astron. J. **69**, 73 (1964)
19. Herman, M.: Sur le courbes invariantes par le difféomorphismes de l'anneau. Astérisque **2**, 144 (1986)
20. Iooss, G., Helleman, R.H.G., Stora, R. (eds.): Chaotic behaviour of deterministic systems. Amsterdam: North-Holland 1983
21. Kolmogorov, A.N.: On the conservation of conditionally periodic motions under small perturbation of the Hamiltonian. Dokl. Akad. Nauk. SSR **98**, 469 (1954)
22. Lanford III, O.E.: Computer assisted proofs in analysis. Phys. A **124**, 465 (1984)
23. MacKay, R.S.: Transition to chaos for area-preserving maps. Lecture Notes in Physics, Vol. 247, p. 390. Berlin, Heidelberg, New York: Springer 1985
24. MacKay, R.S., Percival, I.C.: Converse KAM: Theory and practice. Commun. Math. Phys. **98**, 469 (1985)
25. Mather, J.N.: Existence of quasiperiodic orbits for twist homeomorphisms of the annulus. Topology **21**, 457 (1982)
26. Mather, J.N.: Non-existence of invariant circles. Ergodic Theory Dyn. Syst. **4**, 301 (1984)
27. Moser, J.: On invariant curves of area-preserving mappings of an annulus. Nach. Akad. Wiss. Göttingen, Math. Phys. Kl. II **1**, 1 (1962)
28. Moser, J.: A rapidly convergent iteration method and non-linear partial differential equations. Ann. Scuola Norm. Sup. Pisa **20**, 265 (1966)
29. Moser, J.: Minimal solutions of variational problems on a torus. Ann. Inst. Henri Poincaré **3**, 229 (1986)
30. Moser, J.: Minimal foliations on a torus. Forschungsinstitut für Mathematik, ETH Zürich, September 1987
31. Nekhoroshev, N.N.: An exponential estimate of the time of stability of nearly integrable Hamiltonian systems. Russ. Math. Surv. **32**, 1 (1977)
32. Percival, I.C.: Variational principles for invariant tori and cantori, in nonlinear dynamics and the beam-beam interaction. AIP Conference Proceedings 57, p. 302. M. Month, J.C. Herrera (eds.). (1980)

33. Rüssmann, H.: On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus. Lecture Notes in Physics, Vol. 38, p. 598. Berlin, Heidelberg, New York: Springer 1975
34. Rüssmann, H.: On the existence of invariant curves of twist mappings of an annulus. Lecture Notes in Mathematics, Vol. 1007, p. 677. Berlin, Heidelberg, New York: Springer 1983
35. Salamon, D., Zehnder, E.: KAM theory in configuration space. Preprint 1987
36. Wayne, C.E.: The KAM theory of systems with short range interactions. I. Commun. Math. Phys. **96**, 311 (1984)
37. Wayne, C.E.: Bounds on the trajectories of a system of weakly coupled rotators. Commun. Math. Phys. **104**, 21 (1986)
38. Wisdom, J., Peale, S.J.: The chaotic rotation of Hyperion. Icarus **58**, 137 (1984)
39. (no author listed) Vax architecture handbook. Digital Equipment Corporation (1981)

Communicated by J.-P. Eckmann

Received January 19, 1988; in revised form March 15, 1988