Rigorous estimates for a computer-assisted KAM theory

Alessandra Celletti and Luigi Chierchia[a)]
Forschungsinstitut für Mathematik, ETH-Zentrum, CH-8092 Zürich, Switzerland

(Received 11 August 1986; accepted for publication 6 May 1987)

Nonautonomous Hamiltonian systems of one degree of freedom close to integrable ones are considered. Let ϵ be a positive parameter measuring the strength of the perturbation and denote by ϵ_c the critical value at which a given KAM (Kolmogorov–Arnold–Moser) torus breaks down. A computer-assisted method that allows one to give rigorous lower bounds for ϵ_c is presented. This method has been applied in Celletti–Falcolini–Porzio (to be published in Ann. Inst. H. Poincaré) to the Escande and Doveil pendulum yielding a bound which is within a factor 40.2 of the value indicated by numerical experiments.

I. INTRODUCTION

A problem that has been extensively investigated, both in physics and mathematics, is the stability of invariant surfaces for perturbed integrable systems.[1,2]

Roughly speaking, most of the invariant surfaces for an integrable system are preserved under perturbation if the strength ϵ of the perturbation is sufficiently small. But when ϵ exceeds a certain critical value ϵ_c, these smooth surfaces disappear.

We are interested in analytical tools that allow one to give rigorous and nevertheless realistic lower bounds for ϵ_c in the case of Hamiltonian systems.

For relatively simple dynamical systems, such as holomorphic mappings of the plane ("Siegel's center problem") or some special examples of area preserving diffeomorphisms of an annulus, rather complete results are now available.[3–9] To the best of our knowledge, the methods used in obtaining these results have not been extended to Hamiltonian flows for which the only general tools rely on classical perturbation theory and on KAM theory.[10–16]

For concreteness, and in view of an application we will mention below, we will consider only nonautonomous Hamiltonian systems with one degree of freedom. We remark, however, that our techniques extend easily to the general higher-dimensional situation.

To be more precise, let us consider a Hamiltonian

$$H_0 \equiv h_0(A) + \epsilon f_0(A,\phi,t), \quad \epsilon \geqslant 0,$$

which is a real analytic function defined on a complex domain of the form $D_{R_0}(A_0)\times S^2_{\xi_0}$, where $A_0\in\mathbb{R}$, $\xi_0>0$, $R_0>0$, $D_{R_0}(A_0)$ is the complex disk

$$\{A\in\mathbb{C}:\ |A-A_0|\leqslant R_0\},$$

and $S^2_{\xi_0}$ is the two-dimensional complex strip

$$\{(\phi,t)\in\mathbb{C}^2:\ |\mathrm{Im}\,\phi|\leqslant\xi_0,\ |\mathrm{Im}\,t|\leqslant\xi_0\}.$$

We assume that the perturbation f_0 has a period 2π both in the "angular" variable ϕ and in the time variable t. In other words, the phase space of our system is the product of a real interval with the standard two-dimensional torus $\mathbb{T}^2\equiv\mathbb{R}^2/2\pi\mathbb{Z}^2$.

The integrable part h_0 is assumed to be nondegenerate, i.e., for any $A\in D_{R_0}(A_0)$,

$$h_0''\equiv\frac{d^2h_0}{dA^2}(A)\neq 0.$$

Finally the center A_0 is assumed to be such that the frequency $\omega\equiv h_0'(A_0)$ verifies the Diophantine condition

$$|\omega\nu_1+\nu_2|^{-1}\leqslant C\,|\nu_1|^\tau \tag{1}$$

for some $C>0$, $\tau\geqslant 1$, and every $(\nu_1,\nu_2)\equiv\nu\in\mathbb{Z}^2$ with $\nu_1\neq 0$. For $\epsilon=0$ the torus $\mathfrak{T}^{(0)}(\omega)\equiv\{A_0\}\times\mathbb{T}^2$ is invariant for h_0 and the flow is simply given by

$$(A_0,\phi_0,t_0)\overset{S_t}{\rightarrow}(A_0,\phi_0+\omega t,t_0+t). \tag{2}$$

From KAM theory one knows that, for ϵ sufficiently small, there exists, in an ϵ-neighborhood of $\mathfrak{T}^{(0)}(\omega)$, a (unique) analytic torus $\mathfrak{T}^{(\epsilon)}(\omega)$ invariant for H_0 and on which the flow is still given, in suitable coordinates, by (1). Numerical experiments (see e.g., Refs. 17–20 as well as Refs. 1 and 2) have shown that these KAM tori break down when, as mentioned above, ϵ reaches a critical value ϵ_c. We remark that the lower bounds obtained from standard KAM theory have always turned out to be several order of magnitude away from the numerical evidence.

In this paper we are concerned with the problem of obtaining "reasonable" (i.e., "in reasonable agreement with numerical evidence") rigorous lower bounds ϵ_* of ϵ_c so as to insure the existence of KAM tori for $\epsilon<\epsilon_*$.

The method that we are going to present is based on the scheme used by Arnold in his proof of the theorem on conservation, under perturbations, of quasiperiodic motions.[11] We recall briefly this scheme.

One constructs a sequence of Hamiltonians H_j of the form

$$h_j(A';\epsilon)+\epsilon^{2^j}f_j(A',\phi',t';\epsilon), \quad t'=t,$$

defined in shrinking domains $D_{R_j}(A_j)\times S^2_{\xi_j}$. The centers $A_j\in\mathbb{R}$ are chosen so as to keep the frequencies fixed, i.e., $h_j'(A_j)=\omega$. The Hamiltonian H_{j+1}, for $j=0,1,\ldots$, is obtained from H_j with the aid of a real analytic canonical transformation

$$C_j:\ D_{R_{j+1}}(A_{j+1})\times S^2_{\xi_{j+1}}\rightarrow D_{R_j}(A_j)\times S^2_{\xi_j}$$

close to the identity transformation. In order to construct C_j,

[a)] Permanent address: Dipartimento di Matematica, IIa Universita' di Roma, 00173 Rome, Italy.

there are a certain number of smallness conditions (in the literature usually referred to as "inductive hypotheses") that ϵ has to satisfy. If ϵ is small enough one can show that all the inductive hypotheses are verified and that, in a suitable sense, H_j becomes, as j goes to infinity, closer to an integrable Hamiltonian. From this one can conclude that the invariant torus $\mathfrak{T}^{(\epsilon)}(\omega)$ is obtained as a limit of the composed transformations

$$C_0 \cdot C_1 \cdot \cdots \cdot C_{j-1}(\{A_j\} \times \mathbb{T}^2).$$

The inductive hypotheses consist of a set of estimates needed to control all the quantities entering in the scheme sketched above. These estimates involve also arbitrary choices of various auxiliary parameters. It is natural to try to obtain better stability estimates by varying these parameters. It will turn out that, in our situation, the dependence of the estimates on the auxiliary parameters is very simple so that, in concrete applications, it will be easy to make good choices. There is, however, a delicate choice that concerns the amount of shrinking of the analyticity domain in the periodic variables ϕ and t. We will discuss this point in detail in Appendix C below.

For the purpose of this introduction, let us denote by $\mathscr{I}_j$ the set of specific conditions that will form our inductive hypotheses at the jth perturbative step leading from H_j to H_{j+1}. In this context, the weakest condition that ϵ has to satisfy in order to insure the existence of the KAM tori $\mathfrak{T}^{(\epsilon)}(\omega)$ is

$$\epsilon < \epsilon_\infty \equiv \sup\{\epsilon > 0\colon \mathscr{I}_j \text{ are satisfied for every } j = 0,1,2,\ldots\}.$$

Of course, such a condition has little practical interest since it involves checking an infinite number of estimates. So, we will introduce, for any preassigned integer j_0, a new set of estimates $\mathscr{I}^*_{j_0}$, which will imply all the $\mathscr{I}_j$ for $j > j_0$. Then, for any j_0,

$$\epsilon_{j_0} \equiv \sup\{\epsilon > 0\colon \mathscr{I}_0, \mathscr{I}_1, \ldots, \mathscr{I}_{j_0}, \mathscr{I}^*_{j_0} \text{ are satisfied}\}$$

will provide a concrete lower bound for ϵ_c. In fact, with our choices, ϵ_{j_0} will form a strictly increasing sequence in j_0, so that one obtains better lower bounds by taking larger values of j_0. Now, for $j_0 \sim 20$ the number of elementary operations (i.e., additions, multiplications,..., taking powers, exponentials,...) needed in checking that $\epsilon < \epsilon_{j_0}$ is of the order of 10^5–10^6. Thus, in carrying out these estimates, one is naturally led to the use of computers.

Our method has been applied in Ref. 21, in conjunction with other rigorous numerical computations, in order to give rigorous stability estimates in the following case. Let

$$H_0 \equiv A^2/2 + \epsilon(\cos\phi + \cos(\phi - t)),$$

and consider the stability of the golden-mean torus, i.e., the torus which for $\epsilon = 0$ is given by $\{A_0\} \times \mathbb{T}^2$ with $A_0 \equiv \omega \equiv (1+\sqrt{5})/2$. In Ref. 20, Escande and Doveil gave numerical, as well as (nonrigorous) theoretical evidence in order to show that this torus disappears for $\epsilon = \epsilon_c \sim 1/29.41$. In Ref. 21 it is proved that the golden-mean torus exists and is analytic for $\epsilon < \epsilon_* \equiv \frac{1}{1180}$, and comparing with the experimental results one has

$$\epsilon_c/\epsilon_* \leqslant 40.2.$$

In Ref. 21, it is also pointed out that the best result it was possible to obtain by replacing the method of this paper with the more standard KAM techniques gives a lower bound $\epsilon_{**} = \frac{1}{7930}$ for which $\epsilon_c/\epsilon_{**} \leqslant 2458$.

We conclude this introduction by remarking that the role of computers in obtaining the bounds discussed above is merely to perform lengthy operations with real numbers. By now it is well known how to employ computers in the evaluation of rigorous estimates using, for example, "interval analysis." For more information on this point we refer the reader to Refs. 22–24 and to the literature cited there.

The content of the rest of the paper is as follows. Section II contains the inductive scheme, Sec. III the KAM theorem, Sec. IV the inductive hypotheses $\mathscr{I}^*_{j_0}$, and Sec. V rigorous numerical estimates; and Appendix A contains the self-contained description of the KAM algorithm constructed in this paper, Appendix B the implicit function theorems and a transcedental inequality, and Appendix C the choice of the analyticity-loss sequence $\{\delta_j\}$.

II. INDUCTIVE SCHEME

In this section, maintaining the notations and assumptions of Sec. I, we show how to construct the canonical transformation C_j.

Let us denote by F_j, G_j, and L_j upper bounds on, respectively, $\sup|f_j|$, $\sup|h_j''|$, and $\sup|h_j''|^{-1}$, where the supremum is taken over the domains of definitions $D_{R_j}(A_j) \times S^2_{\xi_j} \equiv D_j \times S^2_{\xi_j}$ and the bars denote the standard norm of complex numbers.

The analyticity assumptions imply the following estimates on the Fourier coeffiecients of f_j:

$$|\hat{f}_{j,\nu}(A)| \leqslant F_j e^{-\xi_j\|\nu\|}, \quad \nu = (\nu_1,\nu_2) \in \mathbb{Z}^2, \tag{3}$$

where for integer vectors ν, $\|\nu\| \equiv |\nu_1| + |\nu_2|$, and

$$\hat{f}_{j,\nu}(A) \equiv \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} f_j(A,\phi,t) e^{-i(\nu_1\phi + \nu_2 t)}\, d\phi\, dt.$$

Another fundamental property of holomorphic functions, which we will often use, is the following. If g is holomorphic on a (smooth) domain D, then for any subdomain $D' \subset D$ one has

$$\sup_{D'} |g'| \leqslant \sup_{D} |g| \; [\mathrm{dist}(\partial D', \partial D)]^{-1}. \tag{4}$$

This estimate follows easily from Cauchy's integral formula for g' taking as contour of integration a circle of radius $\mathrm{dist}(\partial D', \partial D)$ and center $z_0 \in D'$.

Now, following Arnold, we fill in the technical details necessary to carry over the scheme sketched in Sec. I.

Cutoff: Let us split the Fourier expansion of f_j in the following way:

$$f_j = \hat{f}_{j,0} + f_j^{(1)} + f_j^{(2)},$$

where

$$f_j^{(1)} \equiv \sum_{0<\|\nu\|<N_j} \hat f_{j,\nu}(A) e^{i(\nu_1\phi + \nu_2 t)},$$

$$f_j^{(2)} \equiv \sum_{\|\nu\|>N_j} \hat f_{j,\nu}(A) e^{i(\nu_1\phi + \nu_2 t)},$$

with N_j, to be exactly determined later, such that $f_j^{(2)} \sim O(\epsilon^{2^j})$.

Hamilton–Jacobi perturbative step: Following classical perturbation theory[14] we remove (formally) the perturbation to order $O(\epsilon^{2^{j+1}})$ via the generating function

$$\tilde\Phi_j \equiv A'\phi + \epsilon^{2^j}\Phi_j(A',\phi,t;\epsilon),$$

$$\Phi_j \equiv \sum_{0<\|\nu\|<N_j} \frac{\hat f_{j,\nu}(A')}{-i[\nu_1 h_j'(A') + \nu_2]} e^{i(\nu_1\phi + \nu_2 t)}.$$

In this case the new integrable part will be

$$h_{j+1}(A') \equiv h_j(A') + \epsilon^{2^j}\hat f_{j,0}(A').$$

Analyticity loss in the action variables and the $(j+1)$th approximation to the invariant torus: To make rigorous the formal step described above we have to take care of the small divisors appearing in Φ_j and to do this we have to restrict the analyticity domain in the new action variables. Let $\gamma>1$ and N_j be such that

$$R'_{j+1} \equiv (1-1/\gamma)(CG_jN_j^{1+\tau})^{-1} \leqslant R_j. \tag{5}$$

From (5) it follows easily that for each $A'\in D_{R'_{j+1}}(A_j)$ and $\|\nu\| \leqslant N_j$, $\nu_1 \neq 0$,

$$|\nu_1 h_j'(A') + \nu_2|^{-1} \leqslant \gamma C |\nu_1|^\tau. \tag{6}$$

Using an elementary implicit function theorem (Lemma 1 of Appendix B), we can determine the $(j+1)$th approximation to the ω-torus: If, for some $\gamma_1>1$,

$$[\gamma_1^2/(\gamma_1-1)]\epsilon^{2^j}F_jL_jR_j^{-2} < 1, \tag{7}$$

then there exists a unique $A_{j+1}\in B_{[(\gamma_1-1)/\gamma_1]R_j}(A_j)$ such that

$$h'_{j+1}(A_{j+1}) \equiv h_j'(A_{j+1}) + \epsilon^{2^j}\hat f'_{j,0}(A_{j+1}) = h_j'(A_j) = \omega;$$

moreover one has

$$|A_{j+1} - A_j| \leqslant \gamma_1\epsilon^{2^j}F_jL_jR_j^{-1}. \tag{8}$$

The numbers γ and γ_1 are the first "auxiliary parameters" (Sec. I) which we introduce in order to have complete control of the quantities entering in the estimates.

Now, defining

$$R_{j+1} \equiv R'_{j+1} - \gamma_1\epsilon^{2^j}F_jL_jR_j^{-1},$$

it follows from (8) and (5) that

$$D_{j+1} \equiv D_{R_{j+1}}(A_{j+1}) \subset D_{R'_{j+1}}(A_j) \subset D_j.$$

In order to have complete control on R_{j+1} we require

$$\gamma_1\epsilon^{2^j}F_jL_jR_j^{-1} < (1-1/\gamma_1)(1/\gamma_2)(CG_jN_j^{1+\tau})^{-1}, \quad 1/\gamma_2 \equiv 1 - 1/\gamma, \tag{9}$$

and obtain the bounds

$$(\gamma_3CG_jN_j^{1+\tau})^{-1} \leqslant R_{j+1} < (\gamma_2CG_jN_j^{1+\tau})^{-1}, \tag{10}$$

with $\gamma_3 \equiv \gamma_1\gamma_2$. Notice that (5) and (9) imply (7).

Control of grad Φ_j and analyticity loss in the periodic variables: In order to control the derivatives of Φ_j we have to restrict the analyticity domain in the (ϕ,t)-variables. Let $\delta_j < \xi_j$ and $A'\in D_{R'_{j+1}}(A_j)$, $(\phi,t)\in S^2_{\xi_j-\delta_j}$. Using (6) and the estimates (3) and (4) we obtain

$$\left|\frac{\partial\Phi_j}{\partial\phi}(A',\phi,t)\right| \leqslant \gamma CF_j \sum_{\nu\neq 0} |\nu_1|^{1+\tau} e^{-\delta_j\|\nu\|} \equiv k_j^{(1)}CF_j, \tag{11}$$

$$\left|\frac{\partial\Phi_j}{\partial A'}(A',\phi,t)\right| < \lambda_j'\gamma CF_jR_j^{-1}\left(\sum_{\nu\neq 0}|\nu_1|^\tau e^{-\delta_j\|\nu\|} + \sum_{\nu_2\neq 0}\frac{e^{-\delta_j|\nu_2|}}{|\nu_2|}\right) + \gamma^2C^2F_jG_j\sum_{\nu\neq 0}|\nu_1|^{1+2\tau}e^{-\delta_j\|\nu\|}$$
$$\equiv k_j^{(2)}CF_jR_j^{-1} + k_j^{(3)}C^2F_jG_j, \tag{12}$$

where λ_j' denotes a (strict) upper bound on $[(1-R'_{j+1}/R_j)]^{-1}$. Analogously one gets

$$\left|\frac{\partial^2\Phi_j}{\partial\phi\,\partial A'}\right| \leqslant k_j^{(4)}CF_jR_j^{-1} + k_j^{(5)}C^2F_jG_j \tag{13}$$

with

$$k_j^{(4)} \equiv \lambda_j'\gamma\sum_{\nu\neq 0}|\nu_1|^{1+\tau}e^{-\delta_j\|\nu\|},$$

$$k_j^{(5)} \equiv \gamma^2\sum_{\nu\neq 0}|\nu_1|^{2+2\tau}e^{-\delta_j\|\nu\|}.$$

The jth canonical transformation: At a fixed time t, the canonical transformation generated by $\tilde\Phi_j$ is obtained by inverting the functions of mixed variables

$$A = A' + \epsilon^{2^j}\frac{\partial\Phi_j}{\partial\phi}(A',\phi,t), \quad \phi' = \phi + \epsilon^{2^j}\frac{\partial\Phi_j}{\partial A'}(A',\phi,t). \tag{14}$$

First of all we have to make sure that $A\in D_j$ if $A'\in D_{j+1}$. Using (8), (10), and (11) it is readily checked that this will be achieved by requiring

$$(\gamma_2CG_jN_j^{1+\tau})^{-1} + \gamma_1\epsilon^{2^j}F_jL_jR_j^{-1} + k_j^{(1)}\epsilon^{2^j}CF_j \leqslant R_j. \tag{15}$$

Notice that (15) implies (5), thus (15) and (9) imply (5) and (7). Now, using (13) and (10), it is readily checked that

$$\epsilon^{2^j}(k_j^{(4)}CF_jR_j^{-1} + k_j^{(5)}C^2F_jG_j) \leqslant 1 \tag{16}$$

implies the injectivity on D_{j+1} of the first map in (14), which can therefore be inverted in the form

$$A\in\tilde D_j \to A + \epsilon^{2^j}\tilde\Gamma_j(A,\phi,t) = A'\in D_{j+1},$$

where $\tilde D_j$ denotes the image of D_{j+1} under the direct map. Moreover

$$\sup|\tilde\Gamma_j| \leqslant \sup\left|\frac{\partial\Phi_j}{\partial\phi}\right|.$$

In order to invert the second map in (14) we have to allow another analyticity loss in the angle variables (however, in practical applications, this second analyticity loss will turn out to be irrelevant with respect to the first one). Let $\dot\delta_j > 0$ be such that $\xi_{j+1} \equiv \xi_j - \delta_j - \dot\delta_j > 0$. Then, using another elementary implicit function theorem (Lemma 2 of Appendix B), we have that if

$$\epsilon^{2^j}(k_j^{(2)}CF_jR_j^{-1}+k_j^{(3)}C^2F_jG_j)\dot{\delta}_j^{-1}\leqslant 1 \tag{17}$$

the second map in (14) is inverted by

$$\phi'\in S^1_{\xi_{j+1}}\to\phi'+\epsilon^{2^j}\Delta_j(A',\phi',t)=\phi\in S^1_{\xi_j-\delta_j}$$

with

$$\sup|\Delta_j|\leqslant\sup\left|\frac{\partial\Phi_j}{\partial A'}\right|.$$

We can finally define the canonical transformation C_j and its inverse $\widetilde{C}_j$,

$$\widetilde{C}_j:\ (A,\phi,t)\to(A',\phi',t)=(A+\epsilon^{2^j}\widetilde{\Gamma}_j(A,\phi,t),\ \phi+\epsilon^{2^j}\widetilde{\Delta}_j(A,\phi,t),\ t),$$

$$C_j:\ (A',\phi',t)\to(A,\phi,t)=(A'+\epsilon^{2^j}\Gamma_j(A',\phi',t),\ \phi'+\epsilon^{2^j}\Delta_j(A',\phi',t),\ t),$$

with

$$\widetilde{\Delta}_j(A,\phi,t)\equiv\frac{\partial\Phi_j}{\partial A'}(A+\epsilon^{2^j}\widetilde{\Gamma}_j(A,\phi,t),\ \phi,\ t),$$

$$\Gamma_j(A',\phi',t)\equiv\frac{\partial\Phi_j}{\partial\phi}(A',\ \phi'+\epsilon^{2^j}\Delta_j(A',\phi',t),\ t);$$

the domain of holomorphy in the new variables being D_{j+1} and $S^2_{\xi_{j+1}}$.

Estimates on H_{j+1}: The new Hamiltonian $H_{j+1}(A',\phi',t;\epsilon)$ is given by

$$h_j\left(A'+\epsilon^{2^j}\frac{\partial\Phi_j}{\partial\phi}\right)+\epsilon^{2^j}f_j\left(A'+\epsilon^{2^j}\frac{\partial\Phi_j}{\partial\phi},\phi,t\right)+\epsilon^{2^j}\frac{\partial\Phi_j}{\partial t},$$

where Φ_j is evaluated at (A',ϕ,t) with $\phi=\phi'+\epsilon^{2^j}\Delta_j(A',\phi',t)$ so that $f_{j+1}(A',\phi',t)$ will be defined by

$$\epsilon^{-2^{j+1}}\left[h_j\left(A'+\epsilon^{2^j}\frac{\partial\Phi_j}{\partial\phi}\right)-h_j(A')\right]+\epsilon^{-2^j}\left[\hat{f}_{j,0}\left(A'+\epsilon^{2^j}\frac{\partial\Phi_j}{\partial\phi}\right)-\hat{f}_{j,0}(A')\right]$$

$$+\epsilon^{-2^j}\sum_{0<\|\nu\|<N_j}\hat{f}_{j,\nu}\left(A'+\epsilon^{2^j}\frac{\partial\Phi_j}{\partial\phi}\right)e^{i(\nu_1\phi+\nu_2t)}+\epsilon^{-2^j}\sum_{\|\nu\|>N_j}\hat{f}_{j,\nu}\left(A'+\epsilon^{2^j}\frac{\partial\Phi_j}{\partial\phi}\right)\cdot e^{i(\nu_1\phi+\nu_2t)}+\epsilon^{-2^j}\frac{\partial\Phi_j}{\partial t}.$$

Denoting by λ_j an upper bound on $[1-R_j^{-1}(R_{j+1}+\epsilon^{2^j}k_j^{(1)}CF_j)]^{-1}$, a straightforward computation yields

$$\sup_{D_{j+1}\times S^2_{\xi_{j+1}}}|f_{j+1}(A',\phi',t)|\leqslant G_j(CF_jk_j^{(1)})^2+k_j^{(6)}CF_j^2R_j^{-1}+F_j\epsilon^{-2^j}\sum_{\|\nu\|>N_j}e^{-\delta_j\|\nu\|}, \tag{18}$$

where

$$k_j^{(6)}\equiv\lambda_j[k_j^{(1)}/(1-e^{-\delta_j})^2].$$

At this point we choose N_j. Let $\alpha>0$ be a new auxiliary parameter and set

$$N_j\equiv\delta_j^{-1}[\log Q_j+2\log(k_j^{(7)}+\log Q_j)],$$

where

$$Q_j\equiv(\alpha k_j^{(8)}\epsilon^{2^j}C^2F_jG_j)^{-1},\quad k_j^{(8)}\equiv(k_j^{(1)})^2/4\beta_j\delta_j^{-1},$$

$$B_j\equiv e^{-\delta_j}/(1-e^{-\delta_j}),\quad k_j^{(7)}\equiv(\beta_j+1)\delta_j.$$

With these definitions it is easy to see that (Lemma 3, Appendix B)

$$F_j\epsilon^{-2^j}\sum_{\|\nu\|>N_j}e^{-\delta_j\|\nu\|}\leqslant\alpha G_j(CF_jk_j^{(1)})^2 \tag{19}$$

provided

$$16e^{-k_j^{(7)}}(\alpha k_j^{(8)}\epsilon^{2^j}C^2F_jG_j)\leqslant 1. \tag{20}$$

Notice that since $\delta_j<1$, $k_j^{(7)}<(1-e^{-1})^{-1}$ and $16e^{-k_j^{(7)}}>e$; hence (20) implies $N_j\geqslant\delta_j^{-1}$.

Denoting by P_j an upper bound on $\epsilon^{2^j}C^2F_jG_j$ and using (18), (19), and (10) we get the basic recurrence relation

$$P_{j+1}\equiv\begin{cases}P_0^2[\sigma_0+\tau_0/(CGR)], & j=0,\\ P_j^2[\sigma_j+\tau_jN_{j-1}^{1+\tau}], & j\geqslant 1,\end{cases}$$

with

$$\sigma_j\equiv\frac{G_{j+1}}{G_j}(1+\alpha)(k_j^{(1)})^2,\quad \tau_j\equiv\begin{cases}G_1k_0^{(6)}/G_0, & j=0,\\ G_{j+1}\gamma_3k_j^{(6)}/G_j, & j\geqslant 1.\end{cases}$$

Next we indicate the necessary bounds on h_j'',

$$\sup|h_j''|\leqslant G_j+\epsilon^{2^j}\lambda_j^2F_jR_j^{-2}\equiv G_{j+1},$$

$$\sup|h_j''|^{-1}\leqslant L_j(1-\epsilon^{2^j}\lambda_j^2L_jF_jR_j^{-2})^{-1}\equiv L_{j+1},$$

provided

$$\epsilon^{2^j}\lambda_j^2L_jF_jR_j^{-2}<1. \tag{21}$$

As for the λ's one can set

$$\lambda_j'=\begin{cases}[1-(\gamma_2CGRN_0^{1+\tau})^{-1}]^{-1}, & j=0,\\ [1-\gamma_1\cdot(N_{j-1}/N_j)^{1+\tau}]^{-1}, & j\geqslant 1,\end{cases}$$

$$\lambda_j=\begin{cases}[1-(\gamma_2CGRN_0^{1+\tau})^{-1}-k_0^{(1)}\epsilon CFR^{-1}]^{-1}, & j=0,\\ [1-\gamma_1\cdot(N_{j-1}/N_j)^{1+\tau}-P_jk_j^{(1)}\gamma_3N_{j-1}^{1+\tau}]^{-1}, & j\geqslant 1.\end{cases}$$

III. KAM THEOREM

We need now to fix the analyticity-loss sequences $\{\delta_j\}$ and $\{\dot{\delta}_j\}$. First notice that we must have

$$\sum_0^\infty(\delta_j+\dot{\delta}_j)<\xi. \tag{22}$$

Let δ be an auxiliary parameter such that $0<\delta<\xi$. Set

$$\delta_j\equiv\delta/2^{j+1},\quad j\geqslant 0,$$

$$\dot{\delta}_j\equiv\begin{cases}\epsilon k_0^{(2)}CFR^{-1}+k_0^{(3)}P_0, & j=0,\\ P_j(\gamma_3k_j^{(2)}N_{j-1}^{1+\tau}+k_j^{(3)}), & j\geqslant 1,\end{cases}$$

and require the condition

$$\sum_0^j\dot{\delta}_n<\xi-\delta. \tag{23}$$

Then it is clear that (22) and (17) are automatically verified.

Remark: In principle any sequence $\{\delta_j\}$ such that

$$\delta_j > 0, \quad \sum \delta_j < \xi, \quad \sum 2^{-j} \log \delta_j^{-1} < \infty$$

is admissible. Our choice is related to the "quadratic" character of the inductive scheme that we are following. For a fuller discussion, see Appendix C. As in Sec. I, let us denote by $\mathscr{I}_j$ the inductive hypotheses (9), (15), (16), (20), (21), and (23) and by ϵ_∞ the number $\sup\{\epsilon > 0: \mathscr{I}_j$ are verified for every integer $j = 0,1,2,...\}$.

KAM Theorem: If $\epsilon < \epsilon_\infty$ then the map

$$(\phi',t') \in \mathbb{T}^2 \to J(\phi',t') \equiv \lim_{j\to\infty} C_0 \cdot \cdots \cdot C_{j-1}(A_j,\phi',t') \in \mathbb{R}\times\mathbb{T}^2$$

yields an (analytic) embedding of $\mathbb{T}^2$ into the (generalized) phase space of H_0 so that $J(\mathbb{T}^2)$ is invariant for the flow S_t generated by H_0 and

$$S_t(J(\phi',t')) = J(\phi' + \omega t, t + t'). \tag{24}$$

Proof of (24): Since $C_0 \cdot \cdots \cdot C_{j-1}$ is a canonical transformation, denoting by $S_t^{(j)}$ the flow generated by H_j, we have

$$\begin{aligned} S_t(J(\phi',t')) &\equiv \lim_{j\to\infty} S_t(C_0 \cdot \cdots \cdot C_{j-1}(A_j,\phi',t')) \\ &= \lim C_0 \cdot \cdots \cdot C_{j-1}(S_t^{(j)}(A_j,\phi',t')) \\ &= \lim C_0 \cdot \cdots \cdot C_{j-1}(A_j + O(\epsilon^{2^j} F_j \delta_j^{-1})t, \\ &\qquad \phi' + \omega t + O(\epsilon^{2^j} F_j G_j \delta_j^{-1})t^2 \\ &\qquad + O(\epsilon^{2^j} F_j R_j^{-1})t,\ t' + t) \\ &= \lim C_0 \cdot \cdots \cdot C_{j-1}(A_j, \phi' + \omega t, t' + t) \\ &= J(\phi' + \omega t, t' + t). \end{aligned}$$

The step before the last identity follows from the inductive hypotheses.

IV. THE INDUCTIVE HYPOTHESES $\mathscr{I}_{j_0}^*$

Assume that the inductive hypotheses are satisfied for $0 \leqslant j \leqslant j_0$, $j_0 > 1$; we present now a method to control them for $j > j_0$. To do this we have to simplify the conditions at the expense of stronger requirements.

Remark: This step, in standard KAM theory, is taken at $j_0 = 0$ and is one of the reasons for the inaccuracy of standard estimates.

Let us start by imposing

$$\left(\frac{k_{j-1}^{(7)}}{k_j^{(7)}}\right)^2 \frac{Q_{j-1}}{Q_j} \equiv \left(\frac{k_{j-1}^{(7)}}{k_j^{(7)}}\right)^2 \frac{k_j^{(8)}}{k_{j-1}^{(8)}} [P_{j-1}(\sigma_{j-1} + \tau_{j-1} N_{j-2}^{1+\tau})] \leqslant 1, \quad j > j_0. \tag{25}$$

Since $k_j^{(7)} \downarrow 1$ one sees that Q_j is increasing in j so that (25) implies easily

$$N_{j-1}/N_j \leqslant \tfrac{1}{2}, \quad j > j_0. \tag{26}$$

Next we split condition (15) in two pieces: Let γ_4 be a new auxiliary parameter such that $\gamma_4 < \gamma_1^{-1} 2^{1+\tau}$ and let $\gamma_5 \equiv (1 - \gamma_4^{-1})$. If we require

$$\gamma_4(\gamma_1(1/2^{1+\tau}) + P_j k_j^{(1)} \gamma_3 N_{j-1}^{1+\tau}) \leqslant 1, \tag{27}$$

$$\gamma_5 P_j \gamma_1 \gamma_3^2 L_j G_j N_{j-1}^{2+2\tau} \leqslant 1, \tag{28}$$

it is clear that, by (26) and the choice of γ_4 and γ_5, (15) is recovered. Moreover

$$\lambda_j \leqslant \lambda \equiv \gamma_4/(\gamma_4 - 1), \quad \lambda_j' \leqslant \lambda' \equiv \gamma_5/(\gamma_5 - 1).$$

Finally we strengthen (21) and to do this we introduce the last auxiliary parameter. Let $l > 1$ and require

$$[l/(l-1)] P_j (\lambda\gamma_3)^2 G_j L_j N_{j-1}^{2+2\tau} \leqslant 1. \tag{29}$$

This is done so that, since $G_j L_j \geqslant 1$, one gets

$$L_{j+1}/L_j \leqslant 1, \quad G_{j+1}/G_j \leqslant g \equiv 2 - 1/l. \tag{30}$$

At this point we need a simple upper bound on N_j $(j > j_0)$. To do this we disregard "logarithmic corrections": Use $\sigma_j + \tau_j N_{j-1}^{1+\tau} > (k_j^{(1)})^2$ to get, for $j \geqslant 1$,

$$P_{j+j_0}^{1/2^j} > P_{j_0} \Psi_j^*, \quad \Psi_j^* \equiv \Psi_j^*(j_0) \equiv \prod_1^j{}_k\, (k_{j_0+k-1}^{(1)})^{2^{1-k}}. \tag{31}$$

Now use (31) to check

$$N_{j+j_0} \leqslant 4^{j+1} \chi_{j+1}, \quad j \geqslant 1, \tag{32}$$

with

$$\begin{aligned} 0 < \chi_{j+1} &\equiv \chi_{j+1}(P_{j_0}) \equiv 2^{j_0-1} \delta^{-1} \log\left[(\alpha k_{j_0+j}^{(8)})^{1/2^j} P_{j_0} \Psi_j^*\right]^{-1} \\ &\qquad + (1/2^{j-1}) \log\{k_{j_0+j}^{(7)} \\ &\qquad + 2^j \log\left[\alpha k_{j_0+j}^{(8)})1/2^j P_{j_0} \Psi_j^*\right]^{-1}\}. \end{aligned}$$

Finally using (32) we obtain the estimate

$$P_{j_0+j}^{1/2^j} \leqslant P_{j_0} \Psi_j \tag{33}$$

with

$$\Psi_1 \equiv (\sigma_{j_0} + \tau_{j_0} N_{j_0-1}^{1+\tau})^{1/2},$$

$$\Psi_2 \equiv \Psi_1 (\sigma'_{j_0+1} + \tau'_{j_0+1} N_{j_0}^{1+\tau})^{1/4},$$

$$\Psi_j \equiv \Psi_2 \prod_3^{j-1} \left[\sigma'_{j_0+k} + \tau'_{j_0+k} (4^k \chi_k)^{1+\tau}\right]^{1/2^{k+1}}, \quad j \geqslant 3,$$

and

$$\sigma_j' \equiv g(1+\alpha)(k_j^{(1)})^2,$$

$$\tau_j' \equiv g\gamma_3 \lambda k_j^{(1)} (1 - e^{-\delta_j})^{-2}, \quad j > j_0.$$

Notice that Ψ_j^*, χ_j, and Ψ_j converge monotonically and very rapidly as $j \uparrow \infty$; we will denote the corresponding limits by Ψ^*, χ, and Ψ,

$$\Psi_j^* \uparrow \Psi^*, \quad \chi_j \downarrow \chi, \quad \Psi_j \uparrow \Psi.$$

We are now in a position to control easily all the inductive hypotheses [(9), (27), (28), (16), (20), (29), (23), and (25)] for $j \geqslant j_0 + 1$. Consider, for example, (9), which can be rewritten as

$$[\gamma_1/(\gamma_1 - 1)] \gamma_3^2 P_j G_j L_j (N_{j-1} N_j)^{(1+\tau)} < 1. \tag{34}$$

Using (33), (30), and (32) one sees that, for $j = j_0 + n$ and $n \geqslant 2$, (34) is implied by

$$P_{j_0} \theta_n^{(1)} \psi_n \leqslant 1, \tag{35}$$

where

$\theta_n^{(1)} \equiv [(\gamma_1/(\gamma_1-1))\gamma_3^2 G_{j_0} L_{j_0} (gl)^n (4^{2n+1}\chi_n\chi_{n+1})^{1+\tau}]^{1/2^n}$.

Now, it is not hard to see that $\theta_n^{(1)} \downarrow 1$ and that the function $n \to \theta_n^{(1)}\psi_n$ $(n \geqslant 2)$ has a unique maximum that will be achieved for some value $n = n_1^*$. Therefore (9) will be implied, for any $j \geqslant j_0 + 2$, by

$$P_{j_0}\theta_{n_1^*}^{(1)}\psi_{n_1^*} \leqslant 1.$$

Completely analogous considerations apply to the rest of the inductive hypotheses; for a complete and explicit list of all the conditions entering in $\mathscr{I}_{j_0}^*$, see Appendix A.

V. RIGOROUS NUMERICAL ESTIMATES

The condition $\epsilon < \epsilon_\infty$ in the KAM theorem of Sec. III can now be replaced by the more practical condition $\epsilon < \epsilon_{j_0}$, where j_0 is any integer greater than 2 and, as in Sec. I,

$$\epsilon_{j_0} \equiv \sup\{\epsilon > 0\colon \mathscr{I}_0, \mathscr{I}_1, ..., \mathscr{I}_{j_0}, \mathscr{I}_{j_0}^* \text{ are verified}\}.$$

From the preceding sections it follows that ϵ_{j_0} is a strictly increasing function of j_0, so that, in concrete applications, one is interested in taking j_0 as large as possible. Already for j_0 greater than, say, 5, it will be readily realized that, in order to check that $\epsilon < \epsilon_{j_0}$, the use of a computer becomes necessary (in applications a reasonable choice might be $j_0 \sim 30$; compare Ref. 21. In this case one can proceed as follows.

Let us denote by $\boldsymbol{\alpha}$ the set of auxiliary parameters $\{\delta, \alpha, \gamma, \gamma_1, \gamma_4, l\}$ and, to stress that the estimates depend on the choice of $\boldsymbol{\alpha}$, let us replace ϵ_{j_0} by and $\epsilon_{j_0}(\boldsymbol{\alpha})$. One can then write a program that, for any choice of $\boldsymbol{\alpha}$, checks if a given number ϵ verifies or not the conditions $\mathscr{I}_0, \mathscr{I}_1, ... \mathscr{I}_{j_0}, \mathscr{I}_{j_0}^*$. By "trial and error," it will be easy to find a (close) lower estimate, $\epsilon_*(\boldsymbol{\alpha})$ of $\epsilon_{j_0}(\boldsymbol{\alpha})$. At this point, varying $\boldsymbol{\alpha}$, one can "maximize" $\epsilon_*(\boldsymbol{\alpha})$ so as to obtain the final result. Because of the simple dependence of ϵ_{j_0} on $\boldsymbol{\alpha}$, this latter operation will turn out to be rather straightforward.

Important remark: Our method, as well as all KAM theorems, deals with very general situations and, *a fortiori*, does not exploit the peculiarities of the system at hand; such peculiarities might include the geometry of the phase space, singularities in the action variables, special properties in Fourier space, symmetries, etc. Thus, before applying our method, one might use the more flexible finite-order perturbation theory to conjugate the given Hamiltonian to a new one with a smaller perturbation and which, in general, having lost all its special properties, will be closer to a "generic" Hamiltonian. For a detailed discussion and illustration of these ideas we refer the reader to Ref. 21.

ACKNOWLEDGMENTS

The authors are deeply indebted to Professor G. Gallavotti for having started them in this topic and for having been a continuous source of inspiration and help. They wish also to thank Professor J. Moser for helpful comments.

Finally, they thank respectively, the Department of Planetary Sciences and the Department of Mathematics of the University of Arizona for the kind hospitality during the period in which part of this work has been completed.

APPENDIX A: SELF-CONTAINED DESCRIPTION OF THE KAM ALGORITHM CONSTRUCTED IN THIS PAPER

Let

$$H_0(A,\phi,t;\epsilon) \equiv h_0(A) + \epsilon f_0(A,\phi,t),\ (A,\phi,t) \in B_{R_0}(A_0) \times \mathbb{T}^2,$$

where

$$B_{R_0}(A_0) \equiv \{A \in \mathbb{R}\colon |A - A_0| \leqslant R_0\},\quad \mathbb{T}^2 \equiv \mathbb{R}^2/2\pi\mathbb{Z}^2$$

and $A_0 \in \mathbb{R}$ is such that $\omega \equiv h_0'(A_0)$ satisfies

$$|\omega\nu_1 + \nu_2|^{-1} \leqslant C|\nu_1|^\tau$$

for any $(\nu_1,\nu_2) \in \mathbb{Z}^2$, $\nu_1 \neq 0$ and for some $C, \tau > 0$.

Assume that H_0 can be extended to a holomorphic function on

$$D_0 \times S_0 \equiv D_{R_0}(A_0) \times S_{\xi_0}^2$$
$$\equiv \{A \in \mathbb{C}\colon |A - A_0| \leqslant R_0\}$$
$$\times \{(z_1,z_2) \in \mathbb{C}^2\colon |\operatorname{Im} z_i| \leqslant \xi_0,\ i = 1,2\}$$

and denote by F_0, G_0, L_0 upper bounds on, respectively,

$$\sup_{D_0 \times S_0} |f_0|,\quad \sup_{D_0} |h_0''|,\quad \sup_{D_0} |h_0''|^{-1}.$$

Finally, let $j_0 \geqslant 2$ be a fixed integer.

1. KAM theorem (compare Secs. III and IV)

If $\mathscr{I}_j$ $(j = 0,1,...,j_0)$ and $\mathscr{I}_{j_0}^*$ are the inductive hypotheses described below and if $\epsilon < \epsilon_{j_0} \equiv \sup\{\epsilon > 0\colon \mathscr{I}_0, \mathscr{I}_1, ..., \mathscr{I}_{j_0}$ and $\mathscr{I}_{j_0}^*$ are verified$\}$ then there exists an analytic torus ϵ close to $\{A_0\} \times \mathbb{T}^2$ invariant for the flow generated by H_0. On such a torus the flow is given (in suitable coordinates) by

$$(\phi', t') \in \mathbb{T}^2 \to (\phi' + \omega t, t' + t).$$

The rest of this appendix is devoted to the description of the conditions $\mathscr{I}_j$ and $\mathscr{I}_{j_0}^*$. These conditions are expressed in terms of recursive objects. To introduce such objects we start with the following.

2. Definition of the auxiliary parameters

Let δ, α, γ, γ_1, γ_4, l, be such that $\delta < \xi_0$, $\alpha > 0$, $\gamma > 1$, $\gamma_1 > 1$, $\gamma_4 < \gamma_1^{-1} 2^{1+\tau}$, $l > 1$.

Now, define

$$\gamma_2 \equiv (1 - 1/\gamma)^{-1},\quad \gamma_3 \equiv \gamma_1\gamma_2,$$
$$\gamma_5 \equiv (1 - 1/\gamma_4)^{-1},\quad g \equiv 2 - 1/l,$$
$$\lambda \equiv \gamma_4(\gamma_4 - 1)^{-1},\quad \lambda' \equiv \gamma_5(\gamma_5 - 1)^{-1}.$$

Then, for $j \geqslant 0$, we set

$$\delta_j \equiv \delta/2^{j+1},\quad s_j(\rho) \equiv \sum_{\nu \in \mathbb{Z}^2} |\nu_1|^\rho e^{-\delta_j\|\nu\|}\quad (\rho > 0),$$
$$k_j^{(1)} \equiv \gamma s_j(1+\tau),\quad k_j^{(3)} \equiv \gamma^2 s_j(1+2\tau),$$
$$k_j^{(5)} \equiv \gamma^2 s_j(2+2\tau),\quad \beta_j \equiv e^{-\delta_j}(1 - e^{-\delta_j})^{-1},$$
$$k_j^{(7)} \equiv (\beta_j + 1)\delta_j,$$
$$k_j^{(8)} \equiv (k_j^{(1)})^2/(4\beta_j\delta_j^{-1}).$$

Next we will introduce the recursive quantities P_j, Q_j, N_j, λ_j, λ_j', $k_j^{(2)}$, $k_j^{(4)}$, $k_j^{(6)}$, G_j, L_j, σ_j, τ_j. These quantities will be computable according to the following "computational sequence" ("$\cdots \to X \to Y$" means "from the set of quantities

X and the quantities known before the computation of X one can compute the set of quantities Y"):

$\epsilon, C, F_0, R_0, G_0, L_0 \to$
$P_0 \to Q_0 \to N_0 \to \lambda_0 \to \lambda'_0 \to k_0^{(2)}, k_0^{(4)}, k_0^{(6)}, G_1, L_1, \to \sigma_0, \tau_0 \to$
$P_1 \to Q_1 \to N_1 \to \lambda_1, \lambda'_1 \to k_1^{(2)}, k_1^{(4)}, k_1^{(6)}, G_2, L_2 \to \sigma_1, \tau_1 \to$
$\vdots$
$P_j \to Q_j \to N_j \to \lambda_j, \lambda'_j \to k_j^{(2)}, k_j^{(4)}, k_j^{(6)}, G_{j+1}, L_{j+1} \to \sigma_j, \tau_j \to .$

3. Definition of the recursive quantities

We have

$$P_0 \equiv \epsilon C^2 F_0 G_0, \quad Q_0 \equiv (\alpha k_0^{(8)} P_0)^{-1},$$
$$N_0 \equiv \delta_0^{-1} [\log Q_0 + 2\log(k_0^{(7)} + \log Q_0)],$$
$$\lambda_0 \equiv [1 - (\gamma_2 C G_0 R_0 N_0^{1+\tau})^{-1} - k_0^{(1)} \epsilon C F_0 R_0^{-1}]^{-1},$$
$$\lambda'_0 \equiv [1 - (\gamma_2 C G_0 R_0 N_0^{1+\tau})^{-1}]^{-1},$$
$$k_0^{(2)} \equiv \lambda'_0 \gamma \left(s_0(\tau) + 2 \sum_{n=1}^{\infty} \frac{e^{-\delta_0 n}}{n} \right),$$
$$k_0^{(4)} \equiv \lambda'_0 \gamma s_0 (1+\tau), \; k_0^{(6)} \equiv \lambda_0 k_0^{(1)} (1 - e^{-\delta_0})^{-2},$$
$$G_1 \equiv G_0 + \epsilon \lambda_0^2 F_0 R_0^{-2}), \quad L_1 \equiv L_0 (1 - \epsilon \lambda_0^2 L_0 F_0 R_0^{-2})^{-1},$$
$$\sigma_0 \equiv (G_1/G_0)(1+\alpha)(k_0^{(1)})^2, \quad \tau_0 \equiv G_1 k_0^{(6)}/G_0.$$

For $j \geqslant 1$ we set

$$P_j \equiv \begin{cases} P_0^2 [\sigma_0 + \tau_0/(C G_0 R_0)^{-1}], & j=1, \\ P_{j-1}^2 [\sigma_{j-1} + \tau_{j-1} N_{j-2}^{1+\tau}], & j \geqslant 2, \end{cases}$$
$$Q_j \equiv (\alpha k_j^{(8)} P_j)^{-1},$$
$$N_j \equiv \delta_j^{-1} [\log Q_j + 2\log(k_j^{(7)} + \log Q_j)],$$
$$\lambda_j = [1 - \gamma_1 \cdot (N_{j-1}/N_j)^{1+\tau} - P_j k_j^{(1)} \gamma_3 N_{j-1}^{1+\tau}]^{-1},$$
$$\lambda'_j = [1 - \gamma_1 \cdot (N_{j-1}/N_j)^{1+\tau}]^{-1},$$
$$k_j^{(2)} \equiv \lambda'_j \gamma \left(s_j(\tau) + 2 \sum_{n=1}^{\infty} \frac{e^{-\delta_j n}}{n} \right),$$
$$k_j^{(4)} \equiv \lambda'_j \gamma s_j (1+\tau), \quad k_j^{(6)} \equiv \lambda_j k_j^{(1)} (1 - e^{-\delta_j})^{-2},$$
$$G_{j+1} \equiv G_j [1 + (\lambda_j \gamma_3)^2 P_j N_j^{2+2\tau}],$$
$$L_{j+1} \equiv L_j [1 - (\lambda_j \gamma_3)^2 P_j N_j^{2+2\tau} G_j L_j]^{-1},$$
$$\sigma_j \equiv (G_{j+1}/G_j)(1+\alpha)(k_j^{(1)})^2,$$
$$\tau_j \equiv (G_{j+1}/G_j) \gamma_3 k_j^{(6)}.$$

Remark: At the moment, some of the above quantities may be ill defined but this will not be the case as soon as the correspondent conditions $\mathscr{I}_j$ are verified.

4. The inductive hypotheses $\mathscr{I}_j$ $(0 \leqslant j \leqslant j_0)$

The following set of inequalities, (A1)–(A5), constitute the set of inductive hypotheses $\mathscr{I}_j$ with $j = 0, 1, \ldots, j_0$:

$$[\gamma_1/(\gamma_1 - 1)] \gamma_3 \epsilon F_0 L_0 R_0^{-1} (C G_j N_j^{1+\tau}) < 1 \quad (j=0),$$
$$[\gamma_1/(\gamma_1 - 1)] \gamma_3^2 P_j G_j L_j (N_{j-1} N_j)^{1+\tau} < 1 \quad (1 \leqslant j \leqslant j_0), \tag{A1}$$
$$(\gamma_2 C G_0 N_0^{1+\tau})^{-1} + \gamma_1 \epsilon F_0 L_0 R_0^{-1} + k_0^{(1)} \epsilon C F_0 \leqslant R_0 \quad (j=0),$$
$$\gamma_1 [(N_{j-1}/N_j^{1+\tau} + \gamma_3^2 P_j G_j L_j N_{j-1}^{2+2\tau} + \gamma_2 k_j^{(1)} P_j N_{j-1}^{1+\tau}] \leqslant 1 \quad (1 \leqslant j \leqslant j_0), \tag{A2}$$
$$\epsilon k_0^{(4)} C F_0 R_0^{-1} + k_0^{(5)} P_0 \leqslant 1 \quad (j=0),$$
$$[\gamma_3 k_j^{(4)} N_{j-1}^{1+\tau} + k_j^{(5)}] P_j \leqslant 1 \quad (1 \leqslant j \leqslant j_0), \tag{A3}$$
$$16 e^{-k_j^{(7)}} \alpha k_j^{(8)} P_j \leqslant 1 \quad (1 \leqslant j \leqslant j_0), \tag{A4}$$
$$\epsilon \lambda_0^2 L_0 F_0 R_0^{-2} < 1 \quad (j=0),$$
$$(\gamma_3 \lambda_j)^2 P_j G_j L_j N_{j-1}^{2+2\tau} \leqslant 1 \quad (1 \leqslant j \leqslant j_0), \tag{A5}$$
$$k_0^{(2)} \epsilon C F_0 R_0^{-1} + k_0^{(3)} P_0 + \sum_{1}^{j}{}_n (k_n^{(2)} N_{n-1}^{1+\tau} + k_n^{(3)}) P_n \leqslant \xi - \delta. \tag{A6}$$

[(A1)–(A6) correspond to, respectively, (9), (15), (16), (20), (21), and (23) of Secs. II and III.]

5. The inductive hypotheses $\mathscr{I}_{j_0}^*$

In order to describe the set of conditions in $\mathscr{I}_{j_0}^*$ we need the following definitions:

$$\sigma'_j \equiv g(1+\alpha)(k_j^{(1)})^2, \quad \tau'_j \equiv g \gamma_3 \lambda k_j^{(1)} (1 - e^{-\delta_j})^{-2}, \quad j > j_0, \quad \Psi_n^* \equiv \prod_{1}^{n}{}_m (k_{j_0+m-1}^{(1)})^{1/2^{m-1}},$$
$$\chi_n \equiv 2^{j_0 - 1} \delta^{-1} \log[(\alpha k_{j_0+n-1}^{(8)})^{1/2^{n-1}} P_{j_0} \Psi_n^*]^{-1} + (1/2^{n-2}) \log\{k_{j_0+n-1}^{(7)} + 2^{n-1} \log[\alpha k_{j_0+n-1}^{(8)})^{1/2^{n-1}} P_{j_0} \Psi_n^*]^{-1}\},$$
$$\Psi_1 \equiv (\sigma_{j_0} + \tau_{j_0} N_{j_0-1}^{1+\tau})^{1/2}, \quad \Psi_2 \equiv \Psi_1 (\sigma'_{j_0+1} + \tau'_{j_0+1} N_{j_0}^{1+\tau})^{1/4}, \quad \Psi_n \equiv \Psi_2 \prod_{2}^{n-1}{}_k [\sigma'_{j_0+k} + \tau'_{j_0+k} (4^k \chi_k)^{1+\tau}]^{1/2^{k+1}}, \quad n \geqslant 3,$$
$$\tilde{k}_{j_0+n}^{(2)} \equiv \lambda' \gamma s_{j_0+n}(\tau), \quad \tilde{k}_{j_0+n}^{(4)} \equiv \lambda' \gamma s_{j_0+n}(1+\tau),$$
$$\theta_n^{(1)} \equiv [(\gamma_1/(\gamma_1 - 1)) \gamma_3^2 G_{j_0} L_{j_0} (gl)^n (4^{2n+1} \chi_n \chi_{n+1})^{1+\tau}]^{1/2^n}, \quad \theta_n^{(2)} \equiv [\gamma_3 k_{j_0+n}^{(1)} (4^n \chi_n)^{1+\tau} (1 - \gamma_4 \gamma_1 / 2^{1+\tau})^{-1}]^{1/2^n},$$
$$\theta_n^{(3)} \equiv [\gamma_1 \gamma_3^2 \gamma_5 G_{j_0} L_{j_0} (lg)^n (4^n \chi_n)^{2+2\tau}]^{1/2^n}, \quad \theta_n^{(4)} \equiv [\gamma_3 \tilde{k}_{j_0+n}^{(4)} (4^n \chi_n)^{1+\tau} + k_{j_0+n}^{(5)}]^{1/2^n}, \quad \theta_n^{(5)} \equiv [16 \alpha k_{j_0+n}^{(8)} e^{-k_{j_0+n}^{(7)}}]^{1/2^n},$$

$$\theta_n^{(6)} \equiv \left[\left(k_{j_0+n}^{(7)} / k_{j_0+n+1}^{(7)} \right)^2 \left(k_{j_0+n+1}^{(8)} / k_{j_0+n}^{(8)} \right) \left[\sigma'_{j_0+n} + \tau'_{j_0+n} (4^n \chi_n)^{1+\tau} \right] \right]^{1/2^n},$$

$$\theta_n^{(7)} \equiv \left[(l/(l-1))(\lambda\gamma_3)^2 G_{j_0} L_{j_0} (gl)^n (4^n \chi_n)^{2+2\tau} \right]^{1/2^n}, \quad \theta_n^{(8)} \equiv \left[\tilde{k}_{j_0+n}^{(2)} 4^n \chi_n + k_{j_0+n}^{(3)} \right]^{1/2^n}.$$

Now, denoting by Ψ the limit of the Ψ_n's, one has that

$$\Psi_n \uparrow \Psi \text{ and } \theta_n^{(i)} \downarrow 1 \quad (i = 1,2,\ldots,8);$$

moreover the functions $n \to \Psi_n \theta_n^{(i)}$ $(n \geqslant 2)$ have a unique maximum achieved at some value $n = n_i^*$.

The following set of inequalities, (A7)–(A14), constitutes the set of inductive hypotheses $\mathscr{I}_{j_0}^*$:

(A1) with $j = j_0 + 1$,

$$P_{j_0} \Psi_{n_1^*} \theta_{n_1^*}^{(1)} \leqslant 1, \tag{A7}$$

$$\gamma_4 \gamma_1 2^{-1(1+\tau)} + \gamma_3 P_{j_0+1} k_{j_0+1}^{(1)} N_{j_0}^{1+\tau} \leqslant 1,$$

$$P_{j_0} \Psi_{n_2^*} \theta_{n_2^*}^{(2)} \leqslant 1, \tag{A8}$$

$$\gamma_1 \gamma_3^2 \gamma_5 P_{j_0+1} G_{j_0+1} L_{j_0+1} N_{j_0}^{2+2\tau} \leqslant 1,$$

$$P_{j_0} \Psi_{n_3^*} \theta_{n_3^*}^{(3)} \leqslant 1, \tag{A9}$$

(A3) with $j = j_0 + 1$,

$$P_{j_0} \Psi_{n_4^*} \theta_{n_4^*}^{(4)} \leqslant 1, \quad n \geqslant 2, \tag{A10}$$

(A4) with $j = j_0 + 1$,

$$P_{j_0} \Psi_{n_5^*} \theta_{n_5^*}^{(5)} \leqslant 1, \quad n \geqslant 2, \tag{A11}$$

$$(k_j^{(7)} / k_{j+1}^{(7)})^2 (k_{j+1}^{(8)} / k_j^{(8)}) [P_j(\sigma_j + \tau_j N_{j-1}^{1+\tau})] \leqslant 1 \quad \text{with } j = j_0, j_0 + 1,$$

$$P_{j_0} \Psi_{n_6^*} \theta_{n_6^*}^{(6)} \leqslant 1, \quad n \geqslant 2, \tag{A12}$$

$$[l/(l-1)] P_{j_0+1} (\gamma\lambda_3)^2 G_{j_0+1} L_{j_0+1} N_{j_0}^{2+2\tau} \leqslant 1,$$

$$P_{j_0} \Psi_{n_7^*} \theta_{n_7^*}^{(7)}, \quad n \geqslant 2\ , \tag{A13}$$

$$k_0^{(2)} \epsilon C F_0 R_0^{-1} + k_0^{(3)} P_0 + \sum_{m=1}^{j_0+1} (k_m^{(2)} N_{m-1}^{1+\tau} + k_m^{(3)}) P_m + \sum_{m=2}^{n_8^*-1} (P_{j_0} \Psi_m \theta_m^{(8)})^{2^m} + \sum_{m=n_8^*}^{\infty} (P_{j_0} \Psi_{n_8^*} \theta_{n_8^*}^{(8)})^{2^m} \leqslant \xi_0 - \delta. \tag{A14}$$

Remark: Because the convergence of Ψ_n and $\theta_n^{(i)}$ to their limits takes place at a very fast rate, it is clear that to find explicitly the values n_i^* in concrete applications is not a difficult task.

APPENDIX B: IMPLICIT FUNCTION THEOREMS AND A TRANSCENDENTAL INEQUALITY

Lemma 1: Let I be the interval $(x_0 - r, x_0 + r)$, let g be a continuous function on I, and let h be a differentiable function on I.

If $(\sup_I |g|) \cdot (\sup_I |h'|^{-1}) < r$ then there exists a unique point $x_1 \in I$ s.t.

$$h(x_1) + g(x_1) = h(x_0).$$

Moreover $|x_1 - x_0| \leqslant (\sup |g|) \cdot (\sup |h'|^{-1})$.

Proof: The map $x \in I \to h^{-1} \circ (h(x_0) - g(x))$ is a contraction from I into I.

Lemma 2: Let g be a holomorphic map on S_ξ and denote by $|\cdot|_\xi$ the sup norm on S_ξ. If

$$\max\{|g'|_\xi, |g|_\xi \delta^{-1}\} < 1$$

then the map $z \in S_\xi \to z + g(z)$ is one-to-one from S_ξ onto $S_{\xi-\delta}$ and the inverse map $z' \in S_{\xi-\delta} \to z' + h(z') \in S_\xi$ satisfies $|h|_{\xi-\delta} \leqslant |g|_\xi$.

Proof: Injectivity is plain from

$$|z + g(z) - [z' + g(z')]| \geqslant |z - z'| (1 - |g'|_\xi), \quad z, z' \in S_\xi.$$

To prove surjectivity let $w \in S_{\xi-\delta}$. Then the map

$$z \in B \equiv \{z \in C : |z - w| < \delta\} \to w - g(z)$$

is a contraction from B into itself.

Lemma 3: If $e^{x_0} a \geqslant 16$ then $e^x (x + x_0)^{-1} \geqslant a$ for any $x \geqslant \log a + 2 \log(x_0 + \log a)$.

The proof is elementary and is omitted.

APPENDIX C: ON THE CHOICE OF THE ANALYTICITY-LOSS SEQUENCE $\{\delta_l\}$

The size of the perturbation f_{j+1} at the $(j+1)$th stage is given inductively by $P_{j+1} = P_j^2(\sigma_j + \tau_j N_{j-1}^{1+\tau})$, N_{j-1} being a logarithmic correction in P_{j-1}. If we disregard such logarithmic correction we get $P_{j+1} \cong P_j^2 \sigma_j$.

Let us assume, for the moment, that $\xi_0 < 1$. Then $\delta_j < 1$ for each j and

$$\sigma_j \cong s\delta_j^{-n}, \quad \text{some } s > 0 \text{ and } n \in \mathbb{Z},$$

so that

$$P_{j+1} \cong P_j^2 \sigma_j \cong P_j^2 s \delta_j^{-n} \cong P_0^{2^{j+1}} \prod_0^j {}_k (s\delta_k^{-n})^{2^{j-k}}.$$

From this one deduces that the best (up to the above logarithmic corrections) choice of $\{\delta_j\}$ is the one that minimizes the functional

$$\prod_0^\infty \delta_k^{-1/2^k}$$

over sequences satisfying $\Sigma \delta_k = \xi_0$. This is an easy minimum problem that can be immediately solved using Lagrange multipliers obtaining

$$\delta_k = \xi_0/2^{k+1}.$$

Now, if $\xi_0 \geqslant 1$, one can replace the auxiliary parameters α of Sec. V by $\alpha' \equiv \{\alpha, j', \delta_0, \delta_1, \ldots, \delta_{j'}\}$, where j' and $\delta_0, \ldots, \delta_{j'}$ are new auxiliary parameters such that

$$\xi' \equiv \xi_0 - \sum_{j=0}^{j'} \delta_j < 1.$$

Then, for $j > j'$ one can repeat the above argument and set $\delta_{j'+1} \equiv \xi'/2^k$ for $k \geqslant 1$.

[1] *Chaotic Behaviour of Deterministic Systems*, edited by G. Iooss, R. H. G. Helleman, and R. Stora (North-Holland, Amsterdam, 1983).

[2] I. Percival, Lect. Notes Phys. **247**, 12 (1985); R. S. MacKay, *ibid.* **247**, 390 (1985); J. Moser, *ibid.* **247**, 492 (1985).

[3] M. R. Herman, Lectures given at Les Houches, July 1981.

[4] M. R. Herman, Asterisque **1**, 103 (1985).

[5] R. de la Llave, J. Math. Phys. **24**, 2118 (1983).

[6] R. de la Llave and D. Rana, "Proof of accurate bounds in small denominators problems," preprint 1986.

[7] C. Liverani and G. Turchetti, Lett. Nuovo Cimento **39**, 17 (1984).

[8] J. N. Mather, Ergodic Theory Dynamical Systems **2**, 397 (1982).

[9] R. S. MacKay and I. C. Percival, Commun. Math. Phys. **98**, 469 (1985).

[10] A. N. Kolmogorov, Dokl. Akad. Nauk. SSR **98**, 527 (1954).

[11] V. I. Arnold, Russ. Math. Surveys **18**, 9 (1963).

[12] J. Moser, Nach. Akad. Wiss. Göttingen, Math. Phys. Kl. II, **1**, 1 (1962).

[13] H. Rüssmann, Lect. Notes Phys. **38**, 598 (1976).

[14] G. Gallavotti, *The Elements of Mechanics* (Springer, New York 1983).

[15] G. Gallavotti, *Perturbation Theory for Classical Hamiltonian Systems, Progress in Physics*, edited by J. Fröhlich (Birkhäuser, Boston, 1982).

[16] A. Giorgilli and L. Galgani, "Rigorous estimates for the series expansions of Hamiltonian perturbation theory," preprint 1986.

[17] B. Chirikov, Phys. Rep. **52**, 263 (1979).

[18] J. M. Greene, J. Math. Phys. **20**, 1183 (1979).

[19] G. Benettin, M. Casartelli, L. Galgani, A. Giorgilli, and J. M. Strelcyn, Nuovo Cimento B **50**, 211 (1979).

[20] D. F. Escande and F. Doveil, J. Stat. Phys. **26**, 257 (1981).

[21] A. Celletti, C. Falcolini, and A. Porzio, "Rigorous numerical stability estimates for the existence of KAM tori in a forced pendulum," to appear in Ann. Inst. H. Poincaré.

[22] O. E. Lanford III, "Computer assisted proofs in analysis," Physica A **124**, 465 (1984).

[23] O. E. Lanford III, "A computer assisted proof of Feigenbaum's conjectures," Bull. Am. Math. Soc. **6**, 427 (1982).

[24] J.-P. Eckmann, H. Koch, and P. Wittwer, "A computer assisted proof of universality in area preserving maps," Mem. Am. Math. Soc. **47**, 289 (1984).