C. K. R. T. Jones U. Kirchgraber H. O. Walther

*(Managing Editors)*

# Dynamics Reported

## Expositions in Dynamical Systems

*New Series: Volume 4*

With Contributions of
A.M. Blokh, A. Celletti, L. Chierchia,
C. Liverani, M.P. Wojtkowski, T. Wanner

Springer-Verlag
Berlin Heidelberg New York
London Paris Tokyo
Hong Kong Barcelona
Budapest

ISBN 3-540-58391-2 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-58391-2 Springer Verlag New York Berlin Heidelberg

Library of Congress Cataloging-in-Publication Data
Dynamics reported: expositions in dynamical systems/C.K.R.T. Jones, U. Kirchgraber, H.O. Walther, managing editors: with contributions of R. Fournier ... [et al.]. p. cm.
ISBN 3-540-58391-2 (Springer-Verlag Berlin Heidelberg New York: alk. paper).
ISBN 0-387-58391-2 (Springer Verlag New York Berlin Heidelberg: alk. paper)
I. Differentiable dynamical systems. I. Kirchgraber, Urs, 1945– . II. Walther, Hans-Otto.
III. Bielawsi, R. QA614.8D96 1991 003'.85–dc20 91-23213 CIP



Typesetting with TEX: Lewis & Leins, Berlin.
Produktion: PRODUserv Springer Produktions-Gesellschaft, Berlin
SPIN 10100470 41/3020-5 4 3 2 1 0 – Printed on acid-free paper

# Preface

*DYNAMICS REPORTED* reports on recent developments in dynamical systems.

Dynamical systems of course originated from ordinary differential equations. Today, dynamical systems cover a much larger area, including dynamical processes described by functional and integral equations, by partial and stochastic differential equations, etc. Dynamical systems have involved remarkably in recent years. A wealth of new phenomena, new ideas and new techniques are proving to be of considerable interest to scientists in rather different fields. It is not surprising that thousands of publications on the theory itself and on its various applications are appearing

*DYNAMICS REPORTED* presents carefully written articles on major subjects in dynamical systems and their applications, addressed not only to specialists but also to a broader range of readers including graduate students. Topics are advanced, while detailed exposition of ideas, restriction to *typical* results – rather than the *most general* ones – and, last but not least, lucid proofs help to gain the utmost degree of clarity.

It is hoped, that *DYNAMICS REPORTED* will be useful for those entering the field and will stimulate an exchange of ideas among those working in dynamical systems

Summer 1991

Christopher K.R.T Jones
Urs Kirchgraber
Hans-Otto Walther

*Managing Editors*

# Table of Contents

**The "Spectral" Decomposition for One-Dimensional Maps**

*Alexander M. Blokh*

# A Constructive Theory of Lagrangian Tori and Computer-assisted Applications

*Alessandra Celletti*[1] *and Luigi Chierchia*[2]

[1] *Dipartimento di Matematica, Università dell'Aquila, 67100–Coppito, L'Aquila (Italy)*
[2] *Dipartimento di Matematica, Terza Università di Roma, via C. Segre 2, 00146–Roma (Italy)*

## 1. Introduction

Perturbative techniques are among the most powerful tools in the theory of conservative dynamical systems. Besides giving finite time predictions (something well known to the astronomers of the eighteenth century), perturbation methods may be used to establish the existence of regular motions. H. Poincaré used thoroughly such methods in his investigation in Celestial Mechanics [Po], obtaining, e.g., his celebrated results on periodic orbits for Hamiltonian systems. A more recent success of perturbation ideas is the so called "KAM (Kolmogorov [Ko]-Arnold [A1]-Moser [Mo1]) theory", which ensures, under suitable smoothness assumptions, the survival under a small perturbation of "most" of the invariant maximal tori which foliate the phase-space of "integrable" conservative systems (see [B] for review and exhaustive references and [ChG] for recent developments).

One of the main themes we shall discuss here is *how small* the size of the perturbation has to be for the tori to persist.

The interest in such a problem is not only motivated by practical purposes (trying to apply KAM theory to concrete situations), but also by purely abstract questions.

Let $\mathcal{T}_\mu$ be one of the tori surviving the effect of a perturbation "of size" $\mu$. On $\mathcal{T}_\mu$ the motion is regular and past and future are known; in low dimension ($d \leq 2$) such tori constitute obstructions for the dynamics and confine the motions in regions where exact predictions may be impossible; even in high dimensions motions starting nearby $\mathcal{T}_\mu$ will remain close to it for extremely long time.

Here are three basic questions which we shall try to answer:

(a) Is it possible to give explicit approximations of $\mathcal{T}_\mu$ keeping track of approximation errors?
(b) Can one hope to deal with sizes of parameter values coming from actual observations?
(c) Experiments (see, e.g., the discussion in [Mo5]) and some theoretical results ([Ma], [AL], [MKP]) suggest that $\mathcal{T}_\mu$ breaks down as $\mu$ is increased. Do the perturbation techniques contain the elements for explaining the break-down of the invariant tori?

In this paper we shall discuss a general theory (in the *real analytic setting*), which follows recent developments in KAM theory ([Mo4], [SZ], [CC2]), concerning constructive existence results for Lagrangian tori. This theory, based on a Lagrangian formalism, has the advantage of freeing the older formalism from infinitely many changes of variables and instead deals directly with the tori equations in a spirit close to the "hard implicit function theorems" à la Nash-Moser (see [B] for review and references therein and in particular [Mo4], [Z] and [Ha]). The method is highly quantitative and we shall work out



all the estimates in full detail, keeping explicit track of the different quantities involved. We shall then apply the methods to various situations including the spin-orbit resonance problem of Celestial Mechanics.

In these applications we shall make use of computer-assisted calculations (see [L] for general informations). The need for machines comes in, e.g., for the accurate evaluation of the norm of *approximate solutions*, which are the starting point for the Newton iteration leading to the construction of true solutions.

In the final section we briefly discuss the theory of invariant curves for area-preserving (symplectic) diffeomorphisms of the cylinder, using a "direct" approach developed in [CC2], [CC3], [CC4] (see also [LR] for different techniques).

The range of applicability of the method covers parameter values of concrete interest and, in three-dimensional models, is shown to be within 70% "from optimal" (the first results in this direction, inspired by [G1], were obtained in [CC1] and [CFP]) while in symplectic map models we have been able to reach 86% of optimal [CC4].

Question (c) above is by far the most difficult to attack and we shall content ourselves by pointing out a direction ([BC], [BCCF]) which connects the disappearance of the tori with *complex singularities* in the space of the $\mu$-parameter.

The theoretical part of this paper is partly new, while most of the applications are selected from various papers of the authors.

The purpouse of our exposition is ambitiously twofold: we provide (with the help of numerous appendices) complete details so that non-specialists or graduate students could acquire a working knowledge of the main ideas and techniques in KAM theory; but, in so doing, we tried to mantain the exposition concise so that researchers active in the field can find elements of novelties before getting bored.

**Acknowledgment.** It is a pleasure to thank U. Kirchgraber for giving us the opportunity of attempting to achieve the just mentioned project.

## 2. Quasi-Periodic Solutions and Invariant Tori for Lagrangian Systems: Algebraic Structure

In this and in the following paragraph we consider the equation for *maximal* invariant tori and show how to solve it by means of a Newton-KAM method provided a "good enough" *non-degenerate approximate solution* is given.

### 2.1. Setup and Definitions

Let $\mathcal{L}(y, x, t)$ be a real-analytic function of $(y, x, t) \in Y \times \mathbb{T}^{d+1}$, where $Y$ is an open set in $\mathbb{R}^d$ and $\mathbb{T}^{d+1}$ is the standard $(d+1)$-dimensional torus with periods $2\pi$: $\mathbb{T}^{d+1} \equiv \mathbb{R}^{d+1}/(2\pi\mathbb{Z})^{d+1}$; in other words $\mathcal{L}$ is a real-analytic function of $2d+1$ variables, $2\pi$-periodic in $x_1, ..., x_d, t$. The Euler-Lagrange equations for the motions $t \to x(t) = (x_1(t), ..., x_d(t))$ associated to $\mathcal{L}$ (see [A2] for generalities) are given by

$$\frac{d}{dt} \mathcal{L}_y(\dot{x}, x, t) = \mathcal{L}_x(\dot{x}, x, t) , \tag{2.1}$$

where $\mathcal{L}_y \equiv \partial_y \mathcal{L}$ and $\mathcal{L}_x \equiv \partial_x \mathcal{L}$ denote the gradients of $\mathcal{L}$ with respect to $y$ and $x$. An important class of solutions of (2.1) is given by *quasi-periodic* solutions:

**Definition 2.1.** *A solution $x(t) \equiv (x_1(t), ..., x_d(t))$ of (2.1) is said to be quasi-periodic with frequencies $\omega \in Y \subset \mathbb{R}^d$ if $\omega$ is rationally independent with 1 (i.e. $\omega \cdot n + m = 0$ for some $n \in \mathbb{Z}^d$, $m \in \mathbb{Z}$ implies $n = 0$, $m = 0$) and if there exists a periodic function twice differentiable $u$: $(\theta, t) \in \mathbb{T}^{d+1} \to u(\theta, t) \in \mathbb{R}^d$, such that*

$$x_i(t) \equiv \omega_i t + u_i(\omega t, t) , \qquad (\text{mod } 2\pi) . \tag{2.2}$$

**Remark 2.2.** If $\mathcal{L}$ does not depend explicitly on the time $t$ one would replace, in the above definition, "$\omega$ rationally independent with 1" with "$\omega$ rationally independent" (i.e. $\omega \cdot n = 0 \Rightarrow n = 0$) and $u(\theta, t)$ with $u(\theta)$.

**Remark 2.3.** If $\mathcal{L}$ is independent of $(x, t)$, i.e. $\mathcal{L} = \mathcal{L}(y)$, $\mathcal{L}$ is said to be *integrable*: the Euler-Lagrange equations are trivial and all solutions are of the form $x(t) = x_0 + \omega t$, $\omega \equiv \dot{x}_0$. Thus, up to a set of Lebesgue measure zero of initial data ($\equiv \{(x_0, \dot{x}_0) : \dot{x}_0$ is rationally dependent$\}$) all solutions of integrable Lagrangians are quasi-periodic.

To any quasi-periodic solution there is naturally associated a family of solutions parametrized by $(d + 1)$ phase. $(\theta, \tau) \in \mathbb{T}^{d+1}$. In fact, since $(\omega, 1)$ is rationally independent the flow $(\omega t, t)$ is dense on $\mathbb{T}^{d+1}$; therefore it is easy to check that (2.2) is solution of (2.1) if and only if

$$\theta + \omega(t - \tau) + u(\theta + \omega(t - \tau), \ t - \tau) \tag{2.3}$$

is a solution for any $(\theta, \tau) \in \mathbb{T}^{d+1}$. This, in turn, is equivalent to require that $u(\theta, t)$ is solution of the following second order degenerate nonlinear system of partial differential equations on $\mathbb{T}^{d+1}$:

$$D\mathcal{L}_{y_j}(\omega + Du, \ \theta + u, \ t) = \mathcal{L}_{x_j}(\omega + Du, \ \theta + u, \ t) , \quad j = 1, ..., d, \tag{2.4}$$

where $D \equiv D_\omega$ denotes derivative along $(\omega, 1)$:

$$D \equiv D_\omega \equiv \omega \cdot \partial_\theta + \partial_t , \qquad Du \equiv (Du_1, ...Du_d) ,$$
$$Du_i \equiv \sum_{j=1}^{d} \omega_j \frac{\partial u_i}{\partial \theta_j} + \frac{\partial u_i}{\partial t} . \tag{2.5}$$

As an example, consider a planar mechanical system, made up by two interacting particles of masses $m_i$, constrained on concentric circles of radii $r_i$, whose center moves on a (coplanar) circle of radius $\rho$ with angular velocity $\dot{\lambda}(t) = \dot{\lambda}(t + 2\pi)$, the interaction being ruled by a potential energy depending on the squared distance of the two particles. Up to an additive time-dependent function (which does not contribute to the Euler-Lagrange equation) the Lagrangian of this system is given by

$$\mathcal{L}(y_1, y_2, x_1, x_2, t) = \frac{1}{2} \sum_{i=1}^{2} m_i \Big[ r_i^2 y_i^2 + 2 r_i y_i \rho \dot{\lambda} \ \cos(x_i - \lambda) \Big] + \\ - \ V(\cos(x_1 - x_2)) , \tag{2.6}$$

where $V$ is related to the true potential energy $U$ by $V(\xi) \equiv U(r_1^2 + r_2^2 - 2 r_1 r_2 \xi)$. For such a system, a quasi-periodic solution with frequencies $(\omega_1, \omega_2)$, $x(t) = \omega t + u(\omega t)$,



satisfies the system ($i = 1, 2$):

$$m_i r_i^2 \left[ D^2 u_i + \frac{\rho}{r_i} \ddot{\lambda} \cos(\theta_i + u_i - \lambda) \right] = \\ = - \ m_i r_i \rho \dot{\lambda}^2 \ \sin(\theta_i + u_i - \lambda) \\ + (-1)^{i-1} V' \Big( \cos(\theta_1 - \theta_2 + u_1 - u_2) \Big) \sin(\theta_1 - \theta_2 + u_1 - u_2) , \tag{2.7}$$

where $u_i = u_i(\theta_1, \theta_2, t)$ and $Du_i \equiv \omega_1 \partial_{\theta_1} u_i + \omega_2 \partial_{\theta_2} u_i + \partial_t u_i$.

**Remark 2.4.** Equation (2.4), and its variational formulation, has been introduced by Percival ([Pe]).

Quasi-periodic solutions span invariant tori; to be more precise we need a definition: denote by $\mathbb{1}$ the $(d \times d)$ identity matrix and let $(u_\theta)_{ij} \equiv \frac{\partial u_i}{\partial \theta_j}$, then

**Definition 2.5.** *We shall say that a quasi-periodic solution is non-degenerate if* $\forall (\theta, t) \in \mathbb{T}^{d+1}$

$$\det(\mathbb{1} + u_\theta) \neq 0 . \tag{2.8}$$

If $x(t)$ is a non-degenerate quasi-periodic solution, the map $(\theta, t) \to (\theta + u(\theta, t), \ t)$ yields a *non-contractible embedding* of $\mathbb{T}^{d+1}$ into itself; in other words, non-degenerate quasi-periodic solutions correspond to homotopically non trivial invariant tori of maximal dimension $d + 1$ run by a linear flow.

## 2.2. Approximate Solutions and Newton Scheme

Let us begin by setting up the notations. We shall think of vectors as of *column vectors* identifying $m$-vectors with $m \times 1$ matrices. If $f$ is a vector function, $f : \mathbb{R}^m \to \mathbb{R}^n$, the derivative of $f$ is the $n \times m$ matrix $\partial_x f \equiv \frac{\partial f}{\partial x} \equiv f_x$ with entries $(f_x)_{ij} = \frac{\partial f_i}{\partial x_j}$ ( so that $\frac{d}{d\varepsilon}|_{\varepsilon=0} f(x + \varepsilon\sigma) = f_x \sigma$). With these conventions the gradient of a scalar function has to be interpreted as a *row vector*, introducing a funny transpose in our basic equation (2.4) which we rewrite as

$$\mathscr{E}(u) \equiv D\mathcal{L}_y^T(\omega + Du, \theta + u, t) - \mathcal{L}_x^T(\omega + Du, \theta + u, t) = 0 . \tag{2.9}$$

**Definition 2.6.** *A real-analytic function $v(\theta, t)$ on $\mathbb{T}^{d+1}$ is called a non-degenerate approximate solution of (2.9) (in short approximate solution) if*

(*i*) there exists a $(2d + 1)$-neighbourhood ($v$-dependent) $\mathcal{N} \subset Y \times \mathbb{T}^{d+1}$ of the set

$$\left\{ (y, x, t) = (\omega + Dv, \theta + v, t) | \ (\theta, t) \in \mathbb{T}^{d+1} \right\} ,$$

such that, on it, the matrix $\mathcal{L}_{yy}$ is positive definite:

$$\mathcal{L}_{yy} > 0 , \quad \forall (y, x, t) \in \mathcal{N} ; \tag{2.10}$$

(*ii*) for each fixed $t$ the map $\theta \to \theta + v(\theta, t)$ is non-singular i.e.:

$$\det(\mathbb{1} + v_\theta) \neq 0 , \quad \forall (\theta, t) \in \mathbb{T}^{d+1} \qquad ([v_\theta]_{ij} = \frac{\partial v_i}{\partial \theta_j}) . \tag{2.11}$$

To an approximate solution we will always associate an *error-function* $\varepsilon = \varepsilon(\theta, t)$ by setting

$$\varepsilon(\theta, t) \equiv \mathscr{E}(v) \equiv D\mathscr{L}_y^T(\omega + Dv, \theta + v, t) - \mathscr{L}_x^T(\omega + Dv, \theta + v, t) \,. \tag{2.12}$$

Now, the basic idea is, roughly speaking, to try to solve (2.9) by *linearizing* the operator $\mathscr{E}$ at $v$ and finding a *better approximation* $v'$ such that as in *Newton schemes*

$$w \equiv v' - v \sim O(|\varepsilon|) \,, \qquad \varepsilon' \equiv \mathscr{E}(v') \sim O(|\varepsilon|^2) \,, \tag{2.13}$$

in suitable norms to be defined below. Iterating this procedure one may try to get a solution of the form $v + \sum w_j$.
Under suitable number-theoretical assumptions on the frequencies $\omega$, we shall see that this strategy is successfull *provided* one starts with an approximate solution for which the *error term is "small"* enough.

An *important example* is the following. Let $\mathscr{L}$ be a *nearly integrable* Lagrangian, i.e.

$$\mathscr{L}(y, x, t) = \mathscr{L}_0(y) + \mu\mathscr{L}_1(y, x, t), \qquad 0 < \mu \ll 1 \tag{2.14}$$

and assume that the hessian matrix of $\mathscr{L}_0$ is positive definite:

$$\partial_y^2 \mathscr{L}_0 \equiv \left[\frac{\partial^2 \mathscr{L}_0}{\partial y_i \partial y_j}\right] > 0 \,; \tag{2.15}$$

then $v \equiv 0$ is an approximate solution ($\mathscr{N}$ will be a neighbourhood of the torus $\{\omega\} \times \mathbb{T}^{d+1}$) with error function proportional to $\mu$:

$$\varepsilon(\theta, t) \equiv \mu\Big[D\partial_y\mathscr{L}_1(\omega, \theta, t) - \partial_x\mathscr{L}_1(\omega, \theta, t)\Big] \,. \tag{2.16}$$

**Remark 2.7.** The construction we p esent here works for a somewhat more general class of approximate solutions satisfying

$$\det\Big\{(\mathbb{1} + v_\theta)^T \, \mathscr{L}_{yy} \, (\mathbb{1} + v_\theta)\Big\} \neq 0 \,, \qquad \det \int_{\mathbb{T}^d} \Big\{[\mathbb{1} + v_\theta]^T \, \mathscr{L}_{yy} \, [\mathbb{1} + v_\theta]\Big\} \neq 0 \,, \tag{2.17}$$

in place of (2.10), the argument of $\mathscr{L}_{yy} = \partial_y^2\mathscr{L}$ being $(\omega + Dv, \theta + v, t)$ (see [SZ]).

The matrix appearing in (2.17) is an important quantity and it deserves a name:

**Definition 2.8.** *For a given non-degenerate approximate solution $v$ we shall call the matrix*

$$\mathscr{T} \equiv \mathscr{T}_v \equiv [\mathbb{1} + v_\theta]^T \, \mathscr{L}_{yy}(\omega + Dv, \theta + v, t) \, [\mathbb{1} + v_\theta] \tag{2.18}$$

*the* twist matrix *of $v$.*

By our non-degeneracy assumption

$$\mathscr{T} = \mathscr{T}(\theta, t) > 0 \,, \qquad (\theta, t) \in \mathbb{T}^{d+1} \,. \tag{2.19}$$

Because of the frequent occurence of the map $(\theta, t) \in \mathbb{T}^{d+1} \to (\omega + Dv, \theta + v, t) \in \mathbb{R}^d \times \mathbb{T}^{d+1}$, we shall give to it a name too:



**Definition 2.9.** *Given a non-degenerate approximate solution $v$ the map*

$$(\theta, t) \in \mathbb{T}^{d+1} \to \phi(\theta, t) \equiv \phi_v(\theta, t) \equiv (\omega + Dv(\theta, t), \; \theta + v(\theta, t), \; t)$$

*will be called the $v$-embedding map.*

## 2.3. The Linearized Equation

Let $v$ be an approximate solution of (2.9) [see Definition 2.6] and let $\varepsilon(\theta, t)$ be the associated error function defined in (2.12). We want to find a (vector-) function $w(\theta, t)$ such that

$$\mathscr{E}(v + w) = \varepsilon' \,, \tag{2.20}$$

with $\varepsilon'$ *quadratic* in $\varepsilon$: the exact meaning of "quadratic" will be clear in the next paragraph where the quantitative analysis is carried out. However, intuitively speaking, it means that if $\varepsilon$ is replaced by $\mu\varepsilon$ ($\mu \in \mathbb{R}$), then $\varepsilon'$ should have the form $\mu^2\varepsilon'$.

Linearizing (2.12) at $v$ one finds

$$\mathscr{E}(v + w) = \mathscr{E}(v) + \mathscr{E}'(v)w + q_1 \;\equiv\; \varepsilon + \mathscr{E}'(v)w + q_1 \,, \tag{2.21}$$

where $q_1$ [defined by the first equality in (2.21)] is quadratic in $w$ and $\mathscr{E}'(v)$ is the second order linear-differential operator:

$$\begin{aligned} \mathscr{E}'(v) &\equiv D\Big[\mathscr{L}_{yy}D + \mathscr{L}_{yx}\Big] - \mathscr{L}_{xy}D - \mathscr{L}_{xx} \,, && \text{i.e.} \,, \\ \mathscr{E}'(v)g &\equiv D\Big[\mathscr{L}_{yy}Dg + \mathscr{L}_{yx}g\Big] - \mathscr{L}_{xy}Dg - \mathscr{L}_{xx}g \,, && \forall g : \mathbb{T}^{d+1} \to \mathbb{R}^d \,, \end{aligned} \tag{2.22}$$

where

$$[\mathscr{L}_{yx}]_{ij} = \frac{\partial^2 \mathscr{L}}{\partial y_i \partial x_j} \qquad \text{and} \qquad \mathscr{L}_{xy} = \mathscr{L}_{yx}^T \,, \tag{2.23}$$

all derivatives (with respect to $x$, $y$) being evaluated at $(y, x, t) = \phi_v(\theta, t) \equiv (\omega + Dv, \theta + v, t)$. The explicit expression for $q_1$ is:

$$\begin{aligned} q_1 \equiv\, & D\Big[\mathscr{L}_y^T \circ \phi_{v'} - \mathscr{L}_y^T \circ \phi_v - \mathscr{L}_{yy} \circ \phi_v Dw - \mathscr{L}_{yx} \circ \phi_v w\Big] \\ & - \Big[\mathscr{L}_x^T \circ \phi_{v'} - \mathscr{L}_x^T \circ \phi_v - \mathscr{L}_{xy} \circ \phi_v Dw - \mathscr{L}_{xx} \circ \phi_v w\Big] \,, \end{aligned} \tag{2.24}$$

where $\phi_{v'} \equiv \phi_{v+w}$.
Thus, the linear equation to be solved is

$$\mathscr{E}'(v)w + \varepsilon \;=\; q \tag{2.25}$$

for some (real-analytic) function $q$ quadratic in $\varepsilon$ (or, what is the same, in $w$) and in this way we would have

$$\mathscr{E}(v + w) \;=\; q_1 + q \;\equiv\; \varepsilon' \,. \tag{2.26}$$

**Remarks 2.10.**

(*i*) At first sight, equation (2.25) does not look very promising, the operator $\mathcal{E}'$ being a non-constant coefficient, degenerate second order operator on $\mathbb{T}^{d+1}$.

(*ii*) It is important to introduce the "extra error" $q$; in fact the equation $\mathcal{E}'(v)w+\varepsilon=0$, in general, does *not* admit any solution.

The most delicate part of the whole method is the reduction of (2.25) to a *constant coefficient* equation "explicitly" solvable.

## 2.4. Solution of the Linearized Equation

Taking the $\theta$-gradient of (2.12) brings in naturally the operator $\mathcal{E}'(v)$: denoting by $\mathcal{M}$ the (invertible) matrix $\mathbb{1}+v_\theta$ (see (2.11)) one finds

$$\varepsilon_\theta=\mathcal{E}'(v)\mathcal{M}\,,\qquad \mathcal{M}\equiv\mathbb{1}+v_\theta\,, \tag{2.27}$$

where as usual $[\varepsilon_\theta]_{ij}=\frac{\partial\varepsilon_i}{\partial\theta_j}$. This suggests to look for $w$ in the form

$$w\equiv\mathcal{M}z \tag{2.28}$$

for some vector-function $z=z(\theta,t)$ to be determined. Thus

$$\begin{aligned}\mathcal{E}'(v)w+\varepsilon\;&\equiv\;\mathcal{E}'(v)(\mathcal{M}z)+\varepsilon\\&=(\mathcal{E}'(v)\mathcal{M})z+\mathcal{L}_{yy}D\mathcal{M}Dz+D(\mathcal{L}_{yy}\mathcal{M}Dz)\\&\quad+(\mathcal{L}_{yx}-\mathcal{L}_{xy})\mathcal{M}Dz+\varepsilon\\&=\varepsilon_\theta z+\mathcal{L}_{yy}\,\mathcal{M}Dz+D(\mathcal{L}_{yy}\mathcal{M}Dz)\\&\quad+(\mathcal{L}_{yx}-\mathcal{L}_{xy})\mathcal{M}Dz+\varepsilon\\&\equiv\;q_2+\mathcal{L}_{yy}D\mathcal{M}Dz+D(\mathcal{L}_{yy}\mathcal{M}Dz)+\mathcal{L}^A_{yx}\mathcal{M}Dz+\varepsilon\,,\end{aligned} \tag{2.29}$$

where the superscript $A$ denotes the antisymmetric part of a matrix $[B^A\equiv B-B^T]$ and

$$q_2\equiv\varepsilon_\theta z \tag{2.30}$$

is quadratic in $\varepsilon$, $w$.

Some more algebra is needed: denoting by $\mathcal{A}$ the antisymmetric part of $\mathcal{M}^T\partial_\theta\mathcal{L}_y^T\equiv\mathcal{M}^T\partial_\theta(\mathcal{L}_y^T(\omega+Dv,\theta+v,t))$:

$$\begin{aligned}\mathcal{A}&\equiv(\mathcal{M}^T\partial_\theta\mathcal{L}_y^T)^A\equiv\mathcal{M}^T\partial_\theta\mathcal{L}_y^T-(\partial_\theta\mathcal{L}_y^T)^T\mathcal{M}\\&=\mathcal{M}^T\mathcal{L}_{yy}D\mathcal{M}+\mathcal{M}^T\mathcal{L}_{yx}\mathcal{M}-D\mathcal{M}^T\mathcal{L}_{yy}\mathcal{M}-\mathcal{M}^T\mathcal{L}_{xy}\mathcal{M}\end{aligned} \tag{2.31}$$

and recalling the definition of twist matrix $\mathcal{T}$ (cfr. (2.18)) we see that (2.29) can be rewritten as:

$$\mathcal{E}'(v)w+\varepsilon\;=\;q_2+\mathcal{M}^{-T}\Big[D(\mathcal{T}Dz)+\mathcal{A}Dz\Big]+\varepsilon\,, \tag{2.32}$$

where $\mathcal{M}^{-T}\equiv(\mathcal{M}^T)^{-1}$.

To proceed, we have to bring in a key element: the operator $D\equiv\omega\cdot\partial_\theta+\partial_t$ acting on the space of real-analytic functions of $(\theta,t)\in\mathbb{T}^{d+1}$ is *invertible* on its range *provided* the $(d+1)$-vector $(\omega,1)$ satisfies the "Diophantine condition":



$$|\omega\cdot n+m|\ge\frac{1}{\gamma|n|^\tau}\,,\qquad\forall n\in\mathbb{Z}^d\backslash\{0\},\ m\in\mathbb{Z}\,, \tag{2.33}$$

for some $\gamma,\tau>0$.

**Assumption 2.11.** *We assume that the vector* $\omega$ *entering in (2.9) through* $D\equiv\omega\cdot\partial_\theta+\partial_t$ *is a Diophantine vector, i.e. satisfies (2.33).*

**Remark 2.12.**

(*i*) Because of the rational independence of $(\omega,1)$ the range of $D$ consists of functions with vanishing average on $\mathbb{T}^{d+1}$.

(*ii*) We shall assume that in (2.33) it is:

$$\tau\ge d\quad\text{and}\quad\gamma\ge\frac{\sqrt{5}+3}{2}\,; \tag{2.34}$$

$\tau\ge d$ is implied by a classical theorem by Liouville; the second inequality is assumed for simplicity ($\gamma=(\sqrt{5}+3)/2$ is the diophantine constant for the golden mean $\omega=\omega_g\equiv(\sqrt{5}-1)/2$). In the case $\tau>d$ almost all (in the sense of Lebesgue measure) $\omega$'s in $\mathbb{R}^d$ are Diophantine.

(*iii*) In the time independent case one would just suppress $m$ in (2.33) and assume that $\tau\ge d-1$; for $\tau>d-1$ one has a set of full measure.

Using Fourier expansions we see immediately that the *unique solution with zero average* of

$$Dg\;=\;h(\theta,t)\,,\qquad h\equiv\sum_{(n,m)\neq0}\hat{h}_{n,m}\,e^{i(n\cdot\theta+mt)} \tag{2.35}$$

for a given analytic function $h$ with zero average is given by

$$g=\sum_{(n,m)\neq0}\frac{\hat{h}_{n,m}}{i(\omega\cdot n+m)}\,e^{i(n\cdot\theta+mt)}\;\equiv\;D^{-1}h \tag{2.36}$$

and in general all solutions of (2.35) are given by $c+D^{-1}h$ for a constant $c$. Analyticity of $h$ implies that the Fourier coefficients decay exponentially fast in $|n|+|m|$ and, therefore, by (2.36) $D^{-1}h$ is also analytic (and real-analytic if so is $g$). On the other hand, there exist Liouville vectors $\omega$ which can be approximated by rational vectors arbitrarily fast [e.g. $\exists\omega$: $0<|\omega\cdot n+m|\le\exp\,[\exp-(|n|+|m|)]$ $\forall(n,m)\neq0$]; for such vectors the expansion (in (2.36)) may not make sense and we see that the assumption (2.33) is essential.

There is one more step in order to describe explicitly the solution of the linearized equation and it consists in recognizing that $\mathcal{A}$ is given by the formula

$$\mathcal{A}\;=\;D^{-1}(\mathcal{M}^T\varepsilon_\theta)^A\,, \tag{2.37}$$

showing that the term $\mathcal{A}Dz$ is quadratic in $\varepsilon$, $w$ so that

$$\begin{aligned}\mathcal{E}'(v)w+\varepsilon\;&=\;\mathcal{M}^{-T}\Big[D(\mathcal{T}Dz)+\mathcal{M}^T\varepsilon\Big]+q_2+q_3\,,\\q_3&\equiv\mathcal{M}^{-T}D^{-1}(\mathcal{M}^T\varepsilon_\theta)^A\,Dz\,.\end{aligned} \tag{2.38}$$

The proof of (2.37) is given in Appendix 1; notice however that (2.37) implies in particular that the entries of $\mathcal{A}$ and of $(\mathcal{M}^T\varepsilon_\theta)^A$ are functions with zero average:

$$\langle \mathcal{A}\rangle = 0\,, \quad \langle \mathcal{M}^T\varepsilon_\theta - \varepsilon_\theta^T\mathcal{M}\rangle = 0\,, \qquad \langle\cdot\rangle \equiv \int_{\mathbb{T}^{d+1}} \cdot \frac{d\theta dt}{(2\pi)^{d+1}}\,. \tag{2.39}$$

Now, since

$$\begin{aligned}\int \mathcal{M}^T\varepsilon\, d\theta dt &= \int \Big[(D\mathcal{L}_y - \mathcal{L}_x)\mathcal{M}\Big]^T \\ &= -\int \Big[\mathcal{L}_y D\mathcal{M} + \mathcal{L}_x\mathcal{M}\Big]^T = -\int \Big[\partial_\theta \mathcal{L}\Big]^T = 0\,,\end{aligned} \tag{2.40}$$

we see that by our assumption on $v$ [see (*i*), (*ii*), Definition 2.6] and on $\omega$ [Assumption 2.11], the equation

$$D(\mathcal{T}Dz) = -\mathcal{M}^T\varepsilon\,, \qquad \mathcal{T} \equiv \mathcal{M}^T\mathcal{L}_{yy}\mathcal{M}\,, \tag{2.41}$$

can be solved and admits the *general solution*

$$z = D^{-1}\Big\{\mathcal{T}^{-1}\Big[c_0 - D^{-1}(\mathcal{M}^T\varepsilon)\Big]\Big\} + c_1\,, \tag{2.42}$$

with

$$c_0 \equiv \langle\mathcal{T}^{-1}\rangle^{-1}\,\langle\mathcal{T}^{-1}D^{-1}(\mathcal{M}^T\varepsilon)\rangle\,, \tag{2.43}$$

so that (cfr. (2.32))

$$\mathcal{E}'(v)w + \varepsilon = q_2 + q_3 \equiv q\,. \tag{2.44}$$

Notice that the choice of the "integration constant" $c_0$ is enforced by the fact that $\langle Dz\rangle = 0$, while $c_1$ is arbitrary. We normalize $w = \mathcal{M}z$ by requiring that

$$\langle w\rangle \equiv \langle\mathcal{M}z\rangle = 0 \;\Leftrightarrow\; c_1 \equiv -\langle\mathcal{M}\, D^{-1}\Big\{\mathcal{T}^{-1}\Big[c_0 - D^{-1}(\mathcal{M}^T\varepsilon)\Big]\Big\}\rangle\,. \tag{2.45}$$

We collect the results of this section in the following

**Proposition 2.13.** *Let $\omega$ satisfy Assumption 2.11, let $v$ be a (non-degenerate) approximate solution of the equation (2.9) and let $\varepsilon(\theta, t)$ be the associated error function: $\varepsilon = \mathcal{E}(v)$ (see (2.12)). If we set $w \equiv \mathcal{M}z \equiv (\mathbb{1} + v_\theta)z$ with $z$ defined in (2.42), (2.43), (2.45) (see (2.41) for the definition of $\mathcal{T}$), then $w(\theta, t)$ is a real-analytic function with zero average and setting $v' \equiv v + w$, one has*

$$\mathcal{E}(v') = q_1 + q_2 + q_3 \equiv \varepsilon'\,, \tag{2.46}$$

*with*

$$q_1 \equiv \mathcal{E}(v+w) - \mathcal{E}(v) - \mathcal{E}'(v)w\,, \quad q_2 \equiv \varepsilon_\theta z\,, \quad q_3 \equiv (\mathcal{M}^T)^{-1}\, D^{-1}\Big[(\mathcal{M}^T\varepsilon_\theta)^A\Big]\, Dz\,, \tag{2.47}$$

*$\mathcal{E}'(v)$ being defined in (2.22).*



# 3. Quasi-Periodic Solutions and Invariant Tori for Lagrangian Systems: Quantitative Analysis

Here we introduce the scale of function spaces necessary to carry out the quantitative analysis and prove the main estimates.

## 3.1. Spaces of Analytic Functions and Norms

The linearized operator $\mathcal{E}'$ (see (2.22)) involves the degenerate vector field $D \equiv \omega\cdot\partial_\theta + \partial_t$ and, as we already noticed, in order for the inverse $D^{-1}$ to make sense in general the rationally independent vector $(\omega, 1)$ has to satisfy suitable number-theoretical requirements (see Assumption 2.11). However, even in such a case, the fact that it may happen that $|\omega\cdot n_i + m_i| = O(\frac{1}{|n_i|^\tau})$ for suitable sequences $\{(n_i, m_i)\}$, shows that $D^{-1}h$ is *less differentiable* than $h$ [not of course in the direction $(\omega, 1)$ where $D^{-1}h$ *gains* differentiability].

This problem is as old as the modern foundation of mechanics ([A2] and references therein) and is known as the *small divisors problem.* It was only with Carl Siegel [S] in 1942 (in the simpler context of linearization of complex maps around a fixed point) and later with Kolmogorov, Arnold and Moser, that it was possible to overcome technically this problem for the first time (see also Eliasson [El] for a remarkable proof avoiding the Newton method and [Her1], [Her2] for "non local" methods).

The basic technical idea of KAM theory is the following (see [Mo1], [Mo2], [Z], [B], [G1], [G2] for other introductive discussions). One picks a monotone family of Banach spaces of periodic functions on $\mathbb{T}^{d+1}$

$$\mathcal{B}_\xi \subset \mathcal{B}_{\xi'} \quad \text{if} \quad \xi' \le \xi\,, \tag{3.1}$$

where the real parameter $\xi$ measures the regularity of the functions, so that if $\langle h\rangle = 0$ and $h \in \mathcal{B}_{\xi'}$, then

$$|D^{-1}h|_{\xi'} \le K|h|_\xi \tag{3.2}$$

for a suitable constant $K$ depending on $\xi'$. The unboundness of $D^{-1}$ reflects in $K \uparrow \infty$ as $\xi' \uparrow \xi$. In a Newton scheme like the one described in the preceding section, the constant $K$, which will necessarily appear in estimating the new error $\varepsilon'$, will be compared with $|\varepsilon|^2$ and one hopes that this square will under iteration eventually beat the divergence due to $K$.

Let us begin the concrete work.

Denote by $\mathcal{B}_\xi$ the Banach space of real-analytic (periodic) functions on $\mathbb{T}^{d+1}$ which admits an analytic continuation to a domain containing the complex neighbourhood

$$\Delta_\xi \equiv \Big\{(\theta, t) \in \mathbb{C}^{d+1}:\ |\mathrm{Im}\,\theta_i| \le \xi,\ |\mathrm{Im}\,t| \le \xi\Big\}$$

and let $|\cdot|_\xi$ denote the sup norm on $\Delta_\xi$:

$$|h|_\xi \equiv \sup_{|\mathrm{Im}\theta_i|\le\xi,\,|\mathrm{Im}t|\le\xi} |h(\theta, t)|\,. \tag{3.3}$$

Let $\mathfrak{B}^d_\xi$ denote the space of real-analytic vector valued functions $u : \Delta_\xi \to \mathbb{C}^d$ with the norm

$$|u|_\xi \equiv \sum_{i=1}^{d} |u_i|_\xi \equiv \sum_{i=1}^{d} \sup_{\Delta_\xi} |u_i| . \tag{3.4}$$

**Remark 3.1.** It will be important for us to consider functional equations containing one or more parameters $\mu$ belonging to some compact subset $\mathcal{P}$ of $\mathbb{C}^m$, e.g. $\mathcal{P} = \{\mu \in \mathbb{C},\ |\mu| \le \mu_0\}$ for some $\mu_0 > 0$. In such a case, the above "$\sup_{\Delta_\xi}$" will be replaced by "$\sup_{\Delta_\xi \times \mathcal{P}}$" and the uniformity in the estimates will yield, as byproduct, regular (e.g. analytic) dependence upon the parameter(s) $\mu \in \mathcal{P}$. However, since the set $\mathcal{P}$ will not change in our iterations we shall often denote indifferently

$$\sup_{\Delta_\xi \times \mathcal{P}} |\cdot| \equiv |\cdot|_\xi \equiv |\cdot|_{\xi,\mathcal{P}} . \tag{3.5}$$

The norm of matrix/tensor-valued functions will then be defined by the standard operator norm: if

$$\mathcal{M} : \Delta_\xi \supset \mathbb{T}^{d+1} \to L(\mathbb{C}^d) \equiv L(\mathbb{C}^d, \mathbb{C}^d) \equiv \{d \times d - \text{matrices}\} , \tag{3.6}$$

then we set

$$|\mathcal{M}|_\xi = \sup_{|c|=1} |\mathcal{M}c|_\xi \qquad (c \in \mathbb{C}^d ,\quad |c| \equiv \sum_{i=1}^{d} |c_i|) \tag{3.7}$$

($\mathcal{M}c$ is a $\mathbb{C}^d$ valued function in $\Delta_\xi$ and therefore the $|\cdot|_\xi$ is defined in (3.4) above) and in general by induction if

$$\mathcal{T} : \Delta_\xi \to \underbrace{L(\mathbb{C}^d, L(\mathbb{C}^d, \ldots, L(\mathbb{C}^d, \mathbb{C}^d), \mathbb{C}^d), \ldots)}_{p \text{ times}} \equiv L^p(\mathbb{C}^d) \tag{3.8}$$

then, for $c \in \mathbb{C}^d$, $\mathcal{T}c \in L^{p-1}(\mathbb{C}^d)$ and we set

$$|\mathcal{T}|_\xi \equiv \sup_{|c|=1} |\mathcal{T}c| \qquad \left(c \in \mathbb{C}^d, |c| = \sum_{i=1}^{d} |c_i|\right) . \tag{3.9}$$

For example if $u : \Delta_\xi \to \mathbb{C}^d$, $u_\theta : \Delta_\xi \to L(\mathbb{C}^d)$ and

$$\begin{aligned} |u_\theta|_\xi &\equiv \sup_{|c|=1} |u_\theta c|_\xi \equiv \sup_{|c|=1} \sum_{i=1}^{d} |\sum_{j=1}^{d} \frac{\partial u_i}{\partial \theta_j} c_j|_\xi \\ &\equiv \sup_{|c|=1} \sum_{i=1}^{d} \sup_{\Delta_\xi} |\sum_{j=1}^{d} \frac{\partial u_i}{\partial \theta_j} c_j| \end{aligned} \tag{3.10}$$

or if $\phi$ is a map of $\Delta_\xi$ into the domain of $\mathcal{L}(y, x, t)$ then

$$|\mathcal{L}_{yxy} \circ \phi|_\xi = \sup_{|c|=|b|=1} \sum_{i=1}^{d} \sup_{\Delta_\xi} |\sum_{j,k=1}^{d} \mathcal{L}_{y_i x_j y_k} \circ \phi\, b_j c_k| . \tag{3.11}$$



## 3.2. Analytic Tools

The basic technical tools go back to Cauchy and give the possibility of estimating the derivative of a holomorphic function in a domain $\Omega$ by the supremum of the function in a bigger domain $\Omega'$ divided by the distance between the boundaries. In formulae:

**Lemma 3.2.** *Let $h$ be an analytic map from $\Omega \times \mathcal{P} \to \mathbb{C}$, where $\Omega$ is a (smooth) domain in $\mathbb{C}^d$ and $\mathcal{P} \subset \mathbb{C}^k$ a space of parameters. Then for any subdomain $\Omega' \subset \Omega$ with $\mathrm{dist}(\Omega', \partial\Omega) \equiv \delta > 0$ and for any multi index $m = (m_1, \ldots, m_d) \in \mathbb{N}^d$ one has*

$$\sup_{\Omega' \times \mathcal{P}} |\partial_z^m h| \equiv \sup_{\Omega' \times \mathcal{P}} |\frac{\partial^{|m|} h}{\partial z_1^{m_1} \ldots \partial z_d^{m_d}}| \le m!\, \delta^{-|m|} \sup_{\Omega \times \mathcal{P}} |h| \tag{3.12}$$

($|m| = m_1 + \ldots + m_d$).

Let $h$ be an analytic map $h : \Omega \times \mathcal{P} \to L^p(\mathbb{C}^d)$ for some $p \in \mathbb{N}$ ($L^0(\mathbb{C}^d) \equiv \mathbb{C}^d$); then, $\forall l \in \mathbb{Z}_+$, $\partial_z^l h \in L^{p+l}(\mathbb{C}^d)$ and

$$\sup_{\Omega' \times \mathcal{P}} |\partial_z^l h| \le l!\, \delta^{-l} \sup_{\Omega \times \mathcal{P}} |h| . \tag{3.13}$$

The proof of this simple lemma, which is based on Cauchy's integral formula, is given in Appendix 2. A consequence of this lemma is that if $h : \Delta_\xi \to \mathbb{C}^d$, then

$$|\partial_\theta^l h|_{\xi-\delta} \le l! \delta^{-l} |h|_\xi ; \tag{3.14}$$

notice that in the last inequality it would not be necessary to reduce the domain of $t$, which is simply playing the role of a parameter.
Similar statements hold also for the operator $D^{-1}$.

**Lemma 3.3.** *Let $h = h(\theta, t; \mu)$ be a real-analytic map of $\Delta_\xi \times \mathcal{P}$ into $\mathcal{H}$, where $\mathcal{H}$ is either $\mathbb{C}$, or $\mathbb{C}^d$ or $L^p(\mathbb{C}^d)$ and let $l \ge 1$. Then ($|\cdot|_{\xi,\mathcal{P}} \equiv \sup_{\Delta_\xi \times \mathcal{P}} |\cdot|$)*

$$|D^{-1} \partial_\theta^l h|_{\xi-\delta,\mathcal{P}} \le \sigma_l(2\delta)\, |h|_{\xi,\mathcal{P}} , \tag{3.15}$$

*where*

$$\sigma_l(\rho) = \Big[2^{d+1} \sum_{(n,m) \in \mathbb{Z}^{d+1} \setminus (0,0)} \big(\frac{\|n\|^l}{\omega \cdot n + m}\big)^2 e^{-\rho(|n|+|m|)}\Big]^{\frac{1}{2}} , \tag{3.16}$$

($\|n\| \equiv (\sum_{i=1}^{d} |n_i|^2)^{1/2}$, $|n| \equiv \sum_{i=1}^{d} |n_i|$). *The same estimate holds for $l = 0$ provided $h$ has vanishing mean value over $\mathbb{T}^{d+1}$. If $(\omega, 1)$ verifies Assumption 2.11 then*

$$\sigma_l(\rho) < K_l \gamma \delta^{-(\tau+l)} , \qquad K_l \equiv 2^{d+2-(\tau+l)} \sqrt{\Gamma(2(\tau+l)+1)} , \tag{3.17}$$

*$\Gamma$ being Euler's gamma function.*

The proof is given in Appendix 3. Notice that if $(\omega, 1)$ satisfies (2.33) it is very easy to check (3.15) with $\sigma$ replaced by $K\gamma\delta^{-(\tau+l+d)}$. For example, in the case $l = 0$ and $\mathcal{H} = \mathbb{C}$, recalling that the Fourier coefficients, $h_{(n,m)}$, of an analytic function $h$ decay exponentially:

$$|h_{(n,m)}| \le e^{-(|n|+|m|)\xi} |h|_\xi , \tag{3.18}$$

one finds immediately ($h_{(0,0)} \equiv \langle h \rangle = 0$):

$$
\begin{aligned}
|D^{-1}h|_{\xi-\delta} &= |\sum_{(n,m)\neq 0} \frac{h_{(n,m)}}{i(\omega\cdot n+m)} e^{i(n\cdot\theta+mt)}|_{\xi-\delta} \\
&\leq |h|_\xi\,\gamma \sum_{(n,m)\neq 0} |n|^\tau e^{-\xi} e^{(|n|+|m|)(\xi-\delta)} \\
&= \gamma |h|_\xi \sum_{(n,m)\neq 0} |n|^\tau e^{-\delta(|n|+|m|)} \\
&\leq K\gamma |h|_\xi \delta^{-(\tau+d)} ,
\end{aligned}
\tag{3.19}
$$

for some positive constant $K$ depending on $\tau$, $d$.

The fact that $\tau+d$ in (3.19) can be actually replaced by $\tau$ is a quite remarkable fact due to Rüssmann ([R1], [R2], [R3]) (see also Appendix 3).

There is also another reason for leaving the explicit expression of $\sigma_l$ in (3.16) and is related to our computer-assisted technique: in trying to establish "sharp" numerical bounds there will be delicated points where we shall estimate $\sigma$ "accurately" with the aid of the computer rather than using the simple (and necessarily "non sharp") bound in (3.17).

## 3.3. Norm-Parameters

Let us go back to (2.9) and assume that a non degenerate solution $v$ is given (see Definition 2.6). In this section we introduce several positive numbers controlling the norms of the relevant objects. We need to define the following complex domains of $\mathbb{C}^{d+1}$ (recall that $\Delta_\xi \equiv \{(\theta, t) \in \mathbb{C}^{d+1} : |\mathrm{Im}\,(\theta_i)| \leq \xi,\ |\mathrm{Im}\,(t)| \leq \xi\}$)

$$
\mathcal{D}^\xi \equiv \Phi_v \Delta_\xi \equiv \{(y, x, t) \in \mathbb{C}^{2d+1} \mid y = \omega + Dv(\theta, t),\ x = \theta + v(\theta, t),\ (\theta, t) \in \Delta_\xi\}
\tag{3.20}
$$

and for any $\rho = (\rho_1, \rho_2) \in \mathbb{R}^2_+$ $(\Omega = |(\omega, 1)| = \sum_{i=1}^d |\omega_i| + 1)$:

$$
\mathcal{D}^\xi_\rho \equiv \{(y, x, t) = (y_0 + y_1, x_0 + x_1, t_0) \mid (y_0, x_0, t_0) \in \mathcal{D}^\xi,\ |y_1| \leq \Omega\rho_1,\ |x_1| \leq \rho_2\} .
\tag{3.21}
$$

Now, because of Definition 2.6 and the analyticity of $\mathcal{L}$ we can assume that there exist $\xi$, $\alpha > 0$ such that

$$
\begin{aligned}
&(i)\ \mathcal{D}^\xi_{(\alpha,\alpha)} \subset \text{analyticity domain of } \mathcal{L} \\
&(ii)\ \sup_{\mathcal{D}^\xi_{(\alpha,\alpha)}} |\mathcal{L}_{yy}^{-1}| < \infty \\
&(iii)\ \sup_{\Delta_\xi} |(I + v_\theta)^{-1}| \equiv \sup_{\Delta_\xi} |\mathcal{M}^{-1}| < \infty .
\end{aligned}
\tag{3.22}
$$



We then set ($|\cdot|_\xi \equiv \sup_{\Delta_\xi} |\cdot|$ and notice that $\mathcal{D}^\xi \subset \mathcal{D}^\xi_\alpha$)

$$
\begin{aligned}
&|\mathcal{L}_{yy}|_{\mathcal{D}^\xi} \leq L , \qquad |\mathcal{L}_{yy}^{-1}|_{\mathcal{D}^\xi} \leq \overline{L} , \\
&|\mathbb{1} + v_\theta|_\xi \equiv |\mathcal{M}|_\xi \leq M , \qquad |\mathcal{M}^{-1}|_\xi \leq \overline{M} , \\
&|v_t|_\xi \leq S , \\
&|\varepsilon|_\xi \equiv |\mathcal{E}(v)|_\xi \leq E .
\end{aligned}
\tag{3.23}
$$

We need more norms relative to the derivatives of $\mathcal{L}$ (in the following formula $|\cdot|$ is short for $|\cdot|_{\mathcal{D}^\xi}$)

$$
\begin{aligned}
&\max\{\Omega^2|\mathcal{L}_{xyy}|,\ \Omega|\mathcal{L}_{xxy}|,\ |\mathcal{L}_{xxx}|\} \leq L_3 , \\
&\max\{\Omega^2|\mathcal{L}_{yyy}|,\ \Omega|\mathcal{L}_{yyx}|,\ |\mathcal{L}_{yxx}|\}\,\Omega \leq L'_3 , \\
&\max\{\Omega^3|\mathcal{L}_{yyyy}|,\ \Omega^2|\mathcal{L}_{yyyx}|,\ \Omega^2|\mathcal{L}_{yyxy}|,\ \Omega|\mathcal{L}_{yyxx}|,\ \Omega|\mathcal{L}_{yxxy}|,\ |\mathcal{L}_{yxxx}|\}\,\Omega \leq L_4 , \\
&\max\{\Omega^2|\mathcal{L}_{yyyt}|,\ \Omega|\mathcal{L}_{yyxt}|,\ |\mathcal{L}_{yxxt}|\}\,\Omega \leq L'_4 , \qquad (\text{where } |\cdot| \equiv |\cdot|_{\mathcal{D}^\xi}) .
\end{aligned}
\tag{3.24}
$$

The various powers of $\Omega \equiv |(\omega, 1)|$ have been introduced for later convenience. It might appear somewhat strange that for a quadratic scheme for an equation involving the first derivative of $\mathcal{L}$, we introduce the fourth derivatives, but, as we shall see, these "extra" derivatives allow us to avoid one more loss in the analyticity domain, which is a very costly operation from the point of view of accurate bounds.
Finally, we will also denote by $s_i(\rho)$, $\rho > 0$, an upper bound on $\sigma_i(\rho)$ (cfr. (3.16)) and for simplicity we assume that $\rho \to s_i(\rho)$ is decreasing:

$$
\sigma_i(\rho) \leq s_i(\rho) \qquad (\rho \to s_i(\rho)\ \text{ decreasing}) .
\tag{3.25}
$$

We end up this section noticing the following simple relations among the above parameters:

$$
L\overline{L} \geq 1 , \qquad M \geq 1 , \qquad \overline{M} \geq 1 .
\tag{3.26}
$$

The first inequality is obvious; in fact:

$$
1 = |\mathbb{1}| = |\mathcal{L}_{yy}\,\mathcal{L}_{yy}^{-1}|_{\mathcal{D}^\xi} \leq L\overline{L} .
\tag{3.27}
$$

Next:

$$
M \equiv |\mathcal{M}|_\xi = |\mathcal{M}^T|_\xi > |\mathcal{M}|_0 = |\mathcal{M}^T|_0 ;
\tag{3.28}
$$

then, if $e_1$ denotes the $d$-vector $(1, 0, ..., 0)^T$ $(|e_1| = 1)$,

$$
\mathcal{M}^T e_1 = e_1 + (\partial_\theta v_1)^T .
\tag{3.29}
$$

Therefore if $(\theta_0, t_0)$ is a critical point of the periodic function $v_1(\theta, t)$ then

$$
\mathcal{M}^T(\theta_0, t_0)\, e_1 = e_1 , \qquad \mathcal{M}^{-T}(\theta, t)\, e_1 = e_1 .
\tag{3.30}
$$

Thus, from (3.28) and (3.30) it follows immediately the second and third inequalities of (3.26).

## 3.4. Bounds on the Solution of the Linearized Equation

Here we shall provide bounds on $|w|_{\xi-\delta}$ and $|w_\theta|_{\xi-\delta}$, where $w$ is the solution of the linearized equation (2.44) and $\delta$ is (at the moment) an arbitrary number such that

$$0 < \delta \le \xi \,. \tag{3.31}$$

Since

$$|w|_{\xi-\delta} \equiv |\mathcal{M}z|_{\xi-\delta} \le |\mathcal{M}|_\xi |z|_\xi \le M|z|_\xi \tag{3.32}$$

and, by (3.13) applied to $\mathcal{M} : \Delta_\xi \to L(\mathbb{C}^d)$,

$$|w_\theta|_{\xi-\delta} \equiv |\mathcal{M}_\theta z + \mathcal{M} z_\theta|_{\xi-\delta} \le M(\delta^{-1}|z|_{\xi-\delta} + |z_\theta|_{\xi-\delta}) \,. \tag{3.33}$$

we see that we have to estimate $|z|_{\xi-\delta}$ and $|z_\theta|_{\xi-\delta}$.

**Remark 3.4.** Obviously, once a bound on $|z|_{\xi-\delta}$ is established one could immediately estimate $|z_\theta|_{\xi-\delta}$ in, say, $\xi - 2\delta$ by using Cauchy estimates. However, with some more work, it is possible to estimate $z_\theta$ *directly* in $\Delta_{\xi-\delta}$. Restricting the domain of analyticity (or, better, the domain where it is possible to estimate the sup-norms) is a *very costly operation* from the point of view of "optimal bounds", and it is, therefore, important to avoid unnecessary analyticity losses.

Let us begin by estimating the constants $c_0$ and $c_1$ appearing in the definition of $z$ (cfr. (2.42), (2.43), (2.45)). We need some properties of the twist matrix $\mathcal{T}(\theta, t)$ for $(\theta, t)$ real [recall that $|\cdot|_0 \equiv \sup_{\mathbf{T}^{d+1}} |\cdot|$, $\phi_v \equiv (\omega + Dv(\theta, t),\ \theta + v(\theta, t),\ t)$]:

**Lemma 3.5.** *Let $\mathcal{T} \equiv \mathcal{M}^T \mathcal{L}_{yy} \mathcal{M} \equiv \mathcal{M}^T \mathcal{L}_{yy} \circ \Phi_v \mathcal{M}$ be the twist matrix of a non degenerate approximate solution $v$ and let $M$, $\overline{M}$, $L$, $\overline{L}$ denote upper bounds on (respectively) $|\mathcal{M}|_\xi$, $|\mathcal{M}^{-1}|_\xi$, $|\mathcal{L}_{yy}|_{\mathcal{D}\xi}$, $|\mathcal{L}_{yy}^{-1}|_{\mathcal{D}\xi}$: then*

$$\text{(i) } \overline{M}^{-2}\overline{L}^{-1} \le |\mathcal{T}|_0 \le |\mathcal{T}|_\xi \le M^2 L$$

$$\text{(ii) } M^{-2}L^{-1} \le |\mathcal{T}^{-1}|_0 \le |\mathcal{T}^{-1}|_\xi \le \overline{M}^2\overline{L}$$

$$\text{(iii) } |\langle \mathcal{T}^{-1}\rangle^{-1}|_0 \le |\mathcal{T}|_0 \qquad \left(\langle\cdot\rangle \equiv \int_{\mathbf{T}^{d+1}} \cdot \frac{d\theta dt}{(2\pi)^{d+1}}\right).$$

The proof of this simple lemma is given in Appendix 4. From this lemma it follows immediately that

$$|c_0| \le (M\overline{M})^2 (L\overline{L})\ |D^{-1}(\mathcal{M}^T \varepsilon)|_0 \tag{3.34}$$

and from (3.15) (used here with $l = 0$ and $\delta = \xi$) it follows

$$|c_0| \le M(M\overline{M})^2(L\overline{L})\ s_0(2\xi)E \,. \tag{3.35}$$



The estimate of $c_1$ is analogous. Using twice (3.15) (with $\delta = \xi/2$ each time) one obtains:

$$\begin{aligned} |c_1| &\le M\ |D^{-1}\ [(\mathcal{M}^T\mathcal{L}_{yy}\mathcal{M})^{-1}[c_0 - D^{-1}(\mathcal{M}^T\varepsilon)]]|_0 \\ &\le M s_0(\xi) \quad |(\mathcal{M}^T\mathcal{L}_{yy}\mathcal{M})^{-1}[c_0 - D^{-1}(\mathcal{M}^T\varepsilon)]|_{\xi/2} \\ &\le M s_0(\xi)\ |(\mathcal{M}^T\mathcal{L}_{yy}\mathcal{M})^{-1}|_{\xi/2}\ (|c_0| + |D^{-1}(\mathcal{M}^T\varepsilon)|_{\xi/2}) \\ &\le (M\overline{M})^2 s_0(\xi)^2\ E\overline{L}\left[1 + (M\overline{M})^2\ L\overline{L}\ \frac{s_0(2\xi)}{s_0(\xi)}\right]. \end{aligned} \tag{3.36}$$

The estimate of $|z|_{\xi-\delta}$ proceeds along the same lines, using (3.15) twice (with $\delta$ replaced here by $\delta/2$ twice):

$$\begin{aligned} |z|_{\xi-\delta} &\le s_0(\delta)\ |(\mathcal{M}^T\mathcal{L}_{yy}\mathcal{M})^{-1}(c_0 - D^{-1}(\mathcal{M}^T\varepsilon))|_{\xi-\delta/2} + |c_1| \\ &\le s_0(\delta)|(\mathcal{M}^T\mathcal{L}_{yy}\mathcal{M})^{-1}|_\xi\ (|c_0| + |D^{-1}(\mathcal{M}^T\varepsilon)|_{\xi-\delta/2}) + |c_1| \\ &\le s_0(\delta)|(\mathcal{M}^T\mathcal{L}_{yy}\mathcal{M})^{-1}|_\xi\ (|c_0| + s_0(\delta)|\mathcal{M}^T\varepsilon|_\xi) + |c_1| \\ &\le E\overline{L}s_0(\delta)^2(M\overline{M})\overline{M}\ \left\{1 + (M\overline{M})^2\ L\overline{L}\ \frac{s_0(2\xi)}{s_0(\delta)}\right. \\ &\qquad \left. + M\left(\frac{s_0(\xi)}{s_0(\delta)}\right)^2\left[1 + (M\overline{M})^2\ L\overline{L}\ \frac{s_0(2\xi)}{s_0(\xi)}\right]\right\} \\ &\equiv E\overline{L}\ \overline{M}\ (M\overline{M})\ s_0(\delta)^2\ b \\ &\equiv E\overline{L}a\ M^{-1}\,, \end{aligned} \tag{3.37}$$

where last identities define the parameters $a$ and $b$. Also the estimate of $z_\theta$ is similar as long as one uses (3.15) twice but the first term with $l = 1$ (and $\delta$ replaced by $\delta/2$):

$$\begin{aligned} |z_\theta|_{\xi-\delta} &= |\frac{\partial}{\partial\theta}\ D^{-1}\left\{(\mathcal{M}^T\mathcal{L}_{yy}\mathcal{M})^{-1}\ \left[c_0 - D^{-1}(\mathcal{M}^T\varepsilon)\right]\right\}|_{\xi-\delta} \\ &\le s_1(\delta)|(\mathcal{M}^T\mathcal{L}_{yy}\mathcal{M})^{-1}|_\xi\ \ (|c_0| + |D^{-1}(\mathcal{M}^T\varepsilon)|_{\xi-\delta/2}) \\ &\le E\overline{L}M\overline{M}^2 s_0(\delta)s_1(\delta)\ \left\{1 + (M\overline{M})^2\ L\overline{L}\ \frac{s_0(2\xi)}{s_0(\delta)}\right\} \\ &= E\overline{L}\ M\ (M\overline{M})\ s_0(\delta)s_1(\delta)\ b_1\,, \end{aligned} \tag{3.38}$$

where the last identity defines the parameter $b_1$.
Finally, by (3.32) and (3.33) we find:

$$\begin{aligned} &|w|_{\xi-\delta} \le E\overline{L}a\,, \qquad a \equiv s_0(\delta)^2\ (M\overline{M})^2 b\,, \\ &b \equiv 1 + (M\overline{M})^2 L\overline{L}\frac{s_0(2\xi)}{s_0(\delta)} + M\left(\frac{s_0(\xi)}{s_0(\delta)}\right)^2\left[1 + (M\overline{M})^2 L\overline{L}\frac{s_0(2\xi)}{s_0(\xi)}\right] \end{aligned} \tag{3.39}$$

and

$$\begin{aligned} &|w_\theta|_{\xi-\delta} \le E\overline{L}a\left(\delta^{-1} + \frac{s_1(\delta)}{s_0(\delta)}\frac{b_1}{b}\right), \\ &b_1 \equiv 1 + (M\overline{M})^2 L\overline{L}\ \frac{s_0(2\xi)}{s_0(\delta)}\,. \end{aligned} \tag{3.40}$$

## 3.5. Bounds on the New Error Term

At this point all the machinery is set up and we are ready to estimate the new error term $\varepsilon' \equiv \mathscr{E}(v')$ (cfr. Proposition 2.13) and to check that

$$|\varepsilon'|_{\xi'} \equiv |\varepsilon'|_{\xi-\delta} \le K\overline{L}E^2 , \tag{3.41}$$

for a suitable constant $K$ depending on $\xi, \delta$ and on the norm-parameters introduced in §3.3. The purpouse of this section (and relative appendices where all details are carried out) is to provide an explicit and accurate expression for the constant $K$.

It is fairly clear and straightforward how to proceed; however the care (apparently quite excessive) we shall put in determining $K$ is justified by *(i)* the need for complete explicitness (expecially in view of concrete applications) and *(ii)* the need of keeping track of the quantitative roles all the different parameters play in the scheme so as, e.g., to avoid dangerous (from the point of view of accuracy) approximations.
Of course in order to get a manageable theorem we shall need to do several semplifications at the expense of accuracy; but one of the points of the present work (and of [CC1], [CC2], [CC3], [CC4], [CFP], [CG]) is that the stringent smallness requirement of such a theorem can be mitigated by a previous iterative application of the set of accurate estimates we are working out here. We will come back on this crucial point in the next section.
After these premises we formulate (without shame) the main result of this section.

**Proposition 3.6.** *Let $\varepsilon' = \mathscr{E}(v') \equiv q_1 + q_2 + q_3$ as in (2.46), (2.47) of Proposition 2.13; let $0 < \delta < \xi$ and let $M, \overline{M}, L, \overline{L}, E, s_i(\rho)$ $(i = 0, 1)$, $L_3, L_4, L_3', L_4'$ be as in (3.23), (3.24), (3.25) of §3.3. Then (3.41), i.e.*

$$|\varepsilon'|_{\xi'} \le K\overline{L}E^2 , \qquad \xi' = \xi - \delta , \tag{3.42}$$

*holds with*

$$\begin{aligned} K \equiv a \Big\{ & \frac{\delta^{-1}}{M} + \frac{a}{2}(c+1)^2 \, \overline{L}\Big[L_3 + L_4(S+M)(1+\delta^{-1}) + \frac{4}{3}E\overline{L}a\delta^{-2}L_4 \\ & \cdot \Big(\frac{\delta}{4} + g + \frac{\delta^2}{4}\frac{s_1(\delta)}{s_0(\delta)}\frac{b_1}{b}\Big) + L_4'\Big] \\ & + 4L_3' \, a\overline{L}(c+1)\delta^{-2} \Big(\frac{\delta}{4} + g + \frac{\delta^2}{4}\frac{s_1(\delta)}{s_0(\delta)}\frac{b_1}{b}\Big) \\ & + \chi_d \cdot 4\delta^{-1}\overline{M}s_0(\delta) \, \frac{a'}{a} \Big\} , \end{aligned}$$

*where*

$$\begin{cases} \chi_d = 0 & d = 1 \\ \chi_d = 1 & d \ge 2, \end{cases}$$

$$a \equiv (M\overline{M})^2 s_0(\delta)^2 b , \qquad b_1 \equiv 1 + (M\overline{M})^2 L\overline{L} \, \frac{s_0(2\xi)}{s_0(\delta)} ,$$

$$b \equiv b_1 + M \left(\frac{s_0(\xi)}{s_0(\delta)}\right)^2 \left[1 + (M\overline{M})^2 L\overline{L} \, \frac{s_0(2\xi)}{s_0(\xi)}\right] , \qquad c \equiv \delta^{-1} + \frac{a'}{\Omega a} .$$



$$a' \equiv (M\overline{M})^2 s_0(2\delta) b_1' , \qquad b_1' \equiv 1 + (M\overline{M})^2 L\overline{L} \, \frac{s_0(2\xi)}{s_0(2\delta)} ,$$

$$a'' \equiv (M\overline{M})^2 s_0(\delta) b_1 , \qquad g \equiv 1 + \frac{\delta}{4}\frac{a'}{a} + \frac{\delta}{4}s_1(\delta)\frac{a''}{a} + \frac{\delta}{2}\frac{a''}{a} . \tag{3.43}$$

**Proof:** The easiest term to treat is $q_2$: by (3.37) one sees that

$$|q_2|_{\xi'} \equiv |\varepsilon_\theta z|_{\xi'} \le |\varepsilon_\theta|_{\xi'} \, |z|_{\xi'} \le |\varepsilon|_{\xi}|z|_{\xi'} \, \delta^{-1} \le E^2\overline{L} \, \frac{a}{M\delta} , \tag{3.44}$$

with

$$a \equiv (M\overline{M})^2 s_0(\delta)^2 b , \qquad b_1 \equiv 1 + (M\overline{M})^2 L\overline{L} \, \frac{s_0(2\xi)}{s_0(\delta)}$$

and

$$b \equiv b_1 + M \left(\frac{s_0(\xi)}{s_0(\delta)}\right)^2 \left[1 + (M\overline{M})^2 L\overline{L} \, \frac{s_0(2\xi)}{s_0(\xi)}\right] .$$

Next, recalling the definiton of $v'$ (Proposition 2.13) and of $\phi_v, \phi_v'$, we see that

$$\begin{aligned} |q_1|_{\xi'} \le & |D\Big[\mathscr{L}_y^T \circ \phi_{v'} - \mathscr{L}_y^T \circ \phi_v - \mathscr{L}_{yy} \circ \phi_v Dw - \mathscr{L}_{yx} \circ \phi_v w\Big]|_{\xi'} \\ & + |\mathscr{L}_x^T \circ \phi_{v'} - \mathscr{L}_x^T \circ \phi_v - \mathscr{L}_{xy} \circ \phi_v Dw - \mathscr{L}_{xx} \circ \phi_v w|_{\xi'} \\ \equiv & \ |Dq_1^{(y)}|_{\xi'} + |q_1^{(x)}|_{\xi'} , \end{aligned} \tag{3.45}$$

where $q_1^{(y)}$ and $q_1^{(x)}$ have been here defined (in the obvious way). Then using the integral formula for the remainder of Taylor's formula (at the second order) and recalling the definition of $L_3$ one finds

$$\begin{aligned} |q_1^{(x)}|_{\xi'} & \le \frac{1}{2}\Big[|\mathscr{L}_{xyy}|_{\xi'} \, |Dw|_{\xi'}^2 + 2|\mathscr{L}_{xxy}|_{\xi'} \, |w|_{\xi'}|Dw|_{\xi'} + |\mathscr{L}_{xxx}|_{\xi'} \, |w|_{\xi'}^2\Big] \\ & \le \frac{L_3}{2} \left[\frac{|Dw|_{\xi'}^2}{\Omega^2} + 2|w|_{\xi'}\frac{|Dw|_{\xi'}}{\Omega} + |w|_{\xi'}^2\right] , \end{aligned} \tag{3.46}$$

where $|\mathscr{L}_\cdot|_{\xi'}$ is short for $|\mathscr{L}_\cdot \circ \phi_v|_{\xi'}$. We need therefore a bound on $Dw$:

$$Dw = D(\mathcal{M}z) = (D\mathcal{M})z + \mathcal{M}Dz . \tag{3.47}$$

First of all observe, in general, that if $f : \theta \in \mathbb{R}^m \to \mathbb{R}^d$ and $\alpha \in \mathbb{R}^m$, then

$$(\alpha \cdot \partial_\theta) f = \partial_\theta f \, \alpha \quad \Rightarrow \quad |(\alpha \cdot \partial_\theta) f| \le |\alpha| \, |\partial_\theta f| , \tag{3.48}$$

so that by Lemma 3.2 one finds

$$|D\mathcal{M}|_{\xi'} \le \Omega\delta^{-1}M . \tag{3.49}$$

Recalling (3.37) one obtains

$$\begin{aligned}
|Dz|_{\xi'} &= |(\mathcal{M}^T \mathcal{L}_{yy}\mathcal{M})^{-1}\left[c_0 - D^{-1}(\mathcal{M}^T e)\right]|_{\xi'} \\
&\le \overline{M}^2\overline{L}\left[|c_0| + s_0(2\delta)ME\right] \\
&\le \overline{M}^2\overline{L}\left[M^3\overline{M}^2 L\overline{L} s_0(2\xi)E + s_0(2\delta)ME\right] \qquad (3.50)\\
&\equiv E\overline{L}M\overline{M}^2\, s_0(2\delta)b_1' \\
&\equiv E\overline{L}\,\frac{a'}{M}\,,
\end{aligned}$$

where we have introduced the constants

$$a' \equiv (M\overline{M})^2 s_0(2\delta)b_1'\,, \qquad b_1' \equiv 1 + (M\overline{M})^2 L\overline{L}\,\frac{s_0(2\xi)}{s_0(2\delta)}\,. \tag{3.51}$$

Thus, form (3.47), (3.49), (3.50):

$$\begin{aligned}
|Dw|_{\xi'} &\le \Omega\delta^{-1}E\overline{L}a + E\overline{L}a' = E\overline{L}a\left(\Omega\delta^{-1} + \frac{a'}{a}\right) \qquad (3.52)\\
&\equiv E\,\overline{L}\,a\,\Omega\,c\,,
\end{aligned}$$

where

$$c = \delta^{-1} + \frac{a'}{a\Omega}\,.$$

And finally,

$$|q_1^{(x)}|_{\xi'} \le \frac{L_3}{2}\left[E^2\overline{L}^2a^2c^2 + 2E^2\overline{L}^2a^2c + E^2\overline{L}^2a^2\right] = \frac{L_3}{2}E^2\overline{L}^2a^2\,(c+1)^2\,. \tag{3.53}$$

To estimate $Dq_1^{(y)}$ we could treat $q_1^{(y)}$ similarly to $q_1^{(x)}$ and then use Lemma 3.2 to estimate $Dq_1^{(y)}$; but this would lead to an extra loss in the domain of analyticity in $\theta$, which we want to avoid. Instead we compute explicitly $Dq_1^{(y)}$ and calling here

$$f = f(\omega+Dv, \theta+v, t) = \mathcal{L}_y^T(\omega+Dv, \theta+v, t)\,, \qquad f^+ = f(\omega+Dv+Dw, \theta+v+w, t) \tag{3.54}$$

we find

$$\begin{aligned}
|Dq_1^{(y)}|_{\xi'} &\equiv |D\left[f^+ - f - f_yDw - f_xw\right]|_{\xi'} \\
&\equiv |\omega\cdot\partial_\theta(f^+ - f - f_yDw - f_xw) + \partial_t(f^+ - f - f_yDw - f_xw)|_{\xi'} \\
&\le |\partial_\theta(f^+ - f - f_yDw - f_xw)\,\omega|_{\xi'} + |\partial_t(f^+ - f - f_yDw - f_xw)|_{\xi'} \\
&\equiv A_1 + A_2\,.
\end{aligned} \tag{3.55}$$

And by the integral formula for the remainder of Taylor's expansion one has:

$$\begin{aligned}
A_1 &= |\partial_\theta(f^+ - f - f_yDw - f_xw)\,\omega|_{\xi'} \\
&= |\partial_\theta\left\{\int_0^1(1-\beta)\left[f_{yy}DwDw + f_{yx}Dw\,w + f_{xy}w\,Dw + f_{xx}ww\right]d\beta\right\}\,\omega|_{\xi'}
\end{aligned} \tag{3.56}$$

and

$$A_2 = |\partial_t\left\{\int_0^1(1-\beta)\left[f_{yy}DwDw + f_{yx}Dw\,w + f_{xy}wDw + f_{xx}ww\right]d\beta\right\}|_{\xi'}\,.$$



where the derivatives of $f$ are evaluated at $(\omega + Dv + \beta Dw, \theta + v + \beta w, t)$. Performing the $\theta$ and $t$ derivatives one sees that $A_i$ involves, besides derivatives of $\mathcal{L}$ and quantities already controlled, also $Dw_\theta$, $w_t$, $Dw_t$. To estimate $Dw_\theta$ observe that

$$Dw_\theta = D(\mathcal{M}z_\theta + \mathcal{M}_\theta z) = (D\mathcal{M})z_\theta + \mathcal{M}\,Dz_\theta + (D\mathcal{M}_\theta)z + \mathcal{M}_\theta Dz \tag{3.57}$$

and

$$|Dz_\theta|_{\xi'} \le 2\delta^{-1}|Dz|_{\xi-\frac{\delta}{2}} = 2\delta^{-1}\frac{E\overline{L}a''}{M}\,, \qquad a'' = (M\overline{M})^2 s_0(\delta)b_1\,.$$

Therefore, from the inequality

$$|D\mathcal{M}|_{\xi'} \le \Omega\delta^{-1}M\,, \qquad \Omega \equiv \sum_{i=1}^d \omega_i + 1\,,$$

one has

$$|Dw_\theta|_{\xi'} \le \Omega E\overline{L}\delta^{-1}\left[a''s_1(\delta) + 2a'' + 4\delta^{-1}a + a'\right] = 4E\overline{L}\delta^{-2}ag\Omega\,,$$

with

$$g \equiv 1 + \frac{\delta}{4}\frac{a'}{a} + \frac{\delta}{4}s_1(\delta)\frac{a''}{a} + \frac{\delta}{2}\frac{a''}{a}\,.$$

Analogously one obtains

$$\begin{aligned}
|w_t|_{\xi'} &\le E\overline{L}a\left(\delta^{-1} + \frac{s_1(\delta)}{s_0(\delta)}\frac{b_1}{b}\right)\,,\\
|Dw_t|_{\xi'} &\le 4E\overline{L}\delta^{-2}ag\Omega\,.
\end{aligned}$$

With a bit of patience one will now obtain proposition Proposition 3.6 above: see Appendix 5 for complete details. □

## 4. KAM Algorithm

In the preceeding two sections we saw how, starting from a given approximate solution, one can construct a new approximation leading to a new error term which is quadratically smaller than the original one. If the new approximation is non-degenerate (in the sense of definition Definition 2.6) one can iterate. This procedure will lead to an algorithm that, given a *set of positive numbers* ($\equiv$ the norm-parameters relative to a given approximate solution), produces a *new set of positive numbers* ($\equiv$ the norm-parameters relative to the new approximate solution).

Eventhough the estimates we established look complicated (because we avoided arbitrary approximations), any computer will have little trouble in performing the iteration for us (of course a control of the errors introduced by the machine is needed: see §8).

If the norm of the error term converges to zero (in a suitable way), the solution we are after will be constructed. Of course, in order to establish the convergence of the scheme in a finite number of steps we still need a theorem: such a theorem will be discussed in §5.

## 4.1. A Self-Contained Description of the KAM Algorithm

Given an initial non-degenerate approximate solution $v \equiv v^{(0)}$, let $\xi_0 \equiv \xi$ and $\alpha$ be as in (3.22). Fix now a sequence $\delta_j$ of positive numbers and let

$$\sum_{j=0}^{\infty} \delta_j < \xi_0 , \qquad \xi_{j+1} \equiv \xi_j - \delta_j . \tag{4.1}$$

At this level the choice of the "analyticity-loss-sequence" $\{\delta_j\}$ is rather arbitrary; however it might already be clear that asymptotically it will have to satisfy certain requirements (we shall come back on this point, see §6).

Let $j \geq 0$ and let

$$\mathcal{N}_j \equiv \{M_j,\ \overline{M}_j,\ S_j,\ E_j,\ \rho_j \equiv (\rho_{1j}, \rho_{2j}),\ L_j,\ \overline{L}_j,\ L_{3j},\ L_{4j},\ L'_{3j},\ L'_{4j}\} \tag{4.2}$$

be the set of positive numbers controlling the norms of $v^{(j)}$ on $\Delta_{\xi_j}$ and of $\mathcal{L}$ on $\mathfrak{D}_{\rho_j}^{\xi_j}$ [i.e., $M_j \geq |\mathbb{1} + v_\theta^{(j)}|_{\xi_j} \equiv |\mathbb{1} + v_\theta^{(j)}|_j$, $\overline{M}_j \geq |(\mathbb{1} + v_\theta^{(j)})^{-1}|_{\xi_j}$, $S_j \geq |v_t^{(j)}|_{\xi_j}$, $E_j \geq |\mathcal{E}(v^{(j)})|_{\xi_j}$, $L_j \geq |\mathcal{L}_{yy}|_{\mathfrak{D}_{\rho_j}^{\xi_j}} \equiv |\mathcal{L}_{yy}|_j$, $\overline{L}_j \geq |\mathcal{L}_{yy}^{-1}|_{\mathfrak{D}_{\rho_j}^{\xi_j}}$, the remaining parameters being as in (3.24) with $|\cdot|$ replaced by $|\cdot|_{\mathfrak{D}_{\rho_j}^{\xi_j}}$ ]. The estimates of §3 can be encoded in the following rule defining $\mathcal{N}_{j+1}$ in terms of $\mathcal{N}_j$:

$$M_{j+1} \equiv M_j + E_j \overline{L} a_j \left(\delta_j^{-1} \quad \frac{s_1(\delta_j)}{s_0(\delta_j)} \frac{b_{1j}}{b_j}\right)$$

$$\overline{M}_{j+1} \equiv \begin{cases} \overline{M}_j \left(1 - \overline{M}_j\, E_j \overline{L} a_j \left(\delta_j^{-1} + \frac{s_1(\delta_j)}{s_0(\delta_j)} \frac{b_{1j}}{b_j}\right)\right)^{-1} & \text{if } |w_\theta^{(j)}|_j < 1 \\ \infty & \text{if } |w_\theta^{(j)}|_j \geq 1 \end{cases}$$

$$S_{j+1} \equiv S_j + E_j \overline{L} a_j \left(\delta_j^{-1} + \frac{s_1(\delta_j)}{s_0(\delta_j)} \frac{b_{1j}}{b_j}\right)$$

$$E_{j+1} \equiv K_j E_j^2 \overline{L} , \qquad \text{with}$$

$$\begin{aligned} K_j \equiv a_j \Big\{ & \frac{\delta_j^{-1}}{M_j} + \frac{a_j}{2}(c_j+1)^2\, \overline{L}\Big[L_3 + L_4(S_j + M_j)(1+\delta_j^{-1}) + \frac{4}{3} E_j \overline{L} a_j \delta_j^{-2} L_4 \\ & \cdot \Big(g_j + \frac{\delta_j}{4} + \frac{\delta_j^2}{4} \frac{s_1(\delta_j)}{s_0(\delta_j)} \frac{b_{1j}}{b_j}\Big) + L'_4\Big] \\ & + 4L'_3\, \overline{L} a_j (c_j+1)\delta_j^{-2} \left(g_j + \frac{\delta_j}{4} + \frac{\delta_j^2}{4} \frac{s_1(\delta_j)}{s_0(\delta_j)} \frac{b_{1j}}{b_j}\right) \\ & + \chi_d \cdot 4\delta_j^{-1} \overline{M}_j s_0(\delta_j)\, \frac{a'_j}{a_j} \Big\} , \end{aligned}$$

$$\rho_{1,j+1} \equiv \rho_{1j} + \Omega E_j \overline{L} a_j \left(\delta_j^{-1} + \frac{s_1(\delta_j)}{s_0(\delta_j)} \frac{b_{1j}}{b_j}\right) ,$$

$$\rho_{2,j+1} \equiv \rho_{2j} + E_j \overline{L} a_j , \tag{4.3}$$

where $s_\lambda(\delta)$ are upper bounds on $\sigma_l(\delta)$ (cfr. (3.16), (3.25)) and $\chi_d$, $a_j$, $b_j$, $c_j$, $g_j$ are as in (3.43) with $\delta, M, \overline{M}, \ldots$ replaced by $\delta_j, M_j, \overline{M}_j, \ldots$ .



The definitions of $M_{j+1}$, $S_{j+1}$ and $E_{j+1}$ come immediately from from $v^{(j+1)} = v^{(j)} + w^{(j)}$ and from (3.42) (obviously here $v^{(j)}$ plays the role of $v$, $\varepsilon^{(j)}$ of $\varepsilon$, etc.). One easily obtains $\overline{M}_{j+1}$ observing that

$$\begin{aligned} (\mathcal{M}^{(j+1)})^{-1} = (\mathcal{M}^{(j)} + w_\theta^{(j)})^{-1} &= (\mathcal{M}^{(j)})^{-1}\, (\mathbb{1} + (\mathcal{M}^{(j)})^{-1} w_\theta^{(j)}) \\ &\leq |(\mathcal{M}^{(j)})^{-1}|\, (1 - |(\mathcal{M}^{(j)})^{-1}|\, |w_\theta^{(j)}|)^{-1} . \end{aligned}$$

Finally, one has $\rho_{1,j+1} = \rho_{1j} + \Omega |w_\theta^{(j)}|_{\xi_j}$ and $\rho_{2,j+1} = \rho_{2j} + |w^{(j)}|_{\xi_j}$. Now, if $\overline{M}_{j+1} < \infty$, the function $v^{(j+1)} = v^{(j)} + w^{(j)}$ (again $w^{(j)}$ here plays the role of $w$ of §2–§3) is a non-degenerate approximate solution and iteration is possible.
Clearly the problem is now to give criteria for the convergence of

$$v^{(j)} \equiv v^{(0)} + \sum_{i=0}^{j-1} w^{(i)} \tag{4.4}$$

to the solution $u(\theta, t)$ we are after. Here, we establish a criterium which however is not practical (involving the check of infinitely many conditions) and we postpone to the next sections a complete discussion.

**Proposition 4.1.** *Let $p \geq 2$, $\xi_\infty \equiv \xi_0 - \sum_{i=0}^{\infty} \delta_i > 0$ and recall the definition of $\alpha$ (see (3.22)). Then, if*

$$\overline{M}_j < \infty \qquad (\forall j \geq 0) ,$$

$$\sum_{j=1}^{\infty} \rho_{1j} \leq \Omega\alpha , \qquad \sum_{j=1}^{\infty} \rho_{2j} \delta_j^{-p} < \infty ,$$

*then the KAM algorithm converges, i.e., $v^{(j)} = v^{(0)} + \sum_{i=0}^{j-1} w^{(i)}$ converges in the $C^p$-norm on $\Delta_{\xi_\infty}$ to a solution $u$ of (2.4).*

After realizing that $\rho_{1j}$ is a bound on $|w_\theta^{(j)}|_j$, that $\rho_{2j} \leq \rho_{1j}$ and that (by Lemma 3.2)

$$\sum_{i=0}^{j-1} |\partial_\theta^p w^{(i)}|_{\xi_\infty} \leq \sum_{i=0}^{j-1} \frac{1}{(\xi_i - \xi_\infty)^p}\, |w^{(i)}|_{\xi_i} \leq \sum_{i=0}^{j-1} \delta_i^{-p}\, |w^{(i)}|_{\xi_i}$$

the above statement becomes obvious.

## 5. A KAM Theorem

Here we prove a theorem which provides a "simple" quantitative criterion for the existence (and local uniqueness) of a solution of (2.4) close to an approximate solution $v$. The theorem is formulated in a way that makes it possible to apply it to approximate solutions obtained via the KAM algorithm of §4 (in which case $v$ below would correspond to $v_N$ if the KAM algorithm has been applied $N$ times, $M$ to $M_N$, etc.).

**Theorem 5.1.** *Let $\omega$ satisfy Assumption 2.11 and let $v$ be a real-analytic approximate solution of (2.4). Let $r \equiv 1/67^4$ and let $M$, $\overline{M}$, $L$, $\overline{L}$, $S$, $L_3$, $L_3'$, $L_4$, $L_4'$, $E$, be as in §3.3, (3.23), (3.24): the norms relative to $\mathcal{L}$ in (3.24) are $|\cdot| \equiv |\cdot|_{\mathfrak{D}^{\xi}_{(r,r)}}$; let $K_l$ be as in (3.17), assume (for simplicity) $0 < \xi < 1$; finally define*

$$K \equiv 1664^2 \cdot M^2 (M\overline{M})^8 (L\overline{L})^2 (S+M) K_1 K_0^3 \lambda \gamma^4 \, \xi^{-4\tau-3} \, 2^{12\tau} \,, \tag{5.1}$$

*where $\lambda \equiv \max\,(\overline{L}L_3, \overline{L}L_3', \overline{L}L_4, \overline{L}L_4', 1)$ and $\gamma$, $\tau$ are the diophantine constants of $\omega$ (see (2.33)). If*

$$KE\overline{L} \leq 1 \,, \tag{5.2}$$

*then (2.4) has a unique real-analytic solution $u$ with $\langle u \rangle = \langle v \rangle$ admitting an analytic extension to $\Delta_{\xi/2}$, such that $\mathbb{1} + u_\theta$ is invertible on $\Delta_{\frac{\xi}{2}}$ and*

$$|u - v|_{\frac{\xi}{2}} < KE\overline{L} \, \frac{\xi}{321^2} \,, \tag{5.3}$$

$$|u_\theta - v_\theta|_{\frac{\xi}{2}} < \frac{KE\overline{L}}{125^2} \,. \tag{5.4}$$

*Local uniqueness holds in the following sense. If $u$ and $u'$ are two non-degenerate solutions in $\Delta_\xi$ with vanishing mean-value ($\langle u \rangle = \langle u' \rangle = 0$) and if the parameters $M$, $\overline{M}$,... in the definition of the above constant $K$ are, here, defined replacing $v$ with $u$, then*

$$\sqrt{K}\, |u - u'|_\xi \leq 1 \quad \Longrightarrow \quad u \equiv u' \quad (\text{in } \Delta_\xi) \,. \tag{5.5}$$

*If $\mathcal{L}$ depends analytically on parameters $\mu \in \mathcal{P}$ ($\mathcal{P}$ being a compact subset of $\mathbb{C}^m$) and the above norms ($M$, $\overline{M}$, etc.) are defined replacing $\Delta_\xi$ by $\Delta_\xi \times \mathcal{P}$, then the solution $u$ is analytic also in $\mu \in \mathcal{P}$ and (5.3), (5.4) hold on $\Delta_\xi \times \mathcal{P}$.*

**Proof:** Let $v^{(0)} \equiv v$, $\varepsilon^{(0)} \equiv \varepsilon$ and, for $j \geq 0$, let

$$\xi_j \equiv \frac{\xi}{2} + \frac{\xi}{2^{j+1}} \,, \qquad \delta_j \equiv \xi_j - \xi_{j+1} = \frac{\xi}{2^{j+2}} \,. \tag{5.6}$$

and (cfr. Lemma 3.3 ):

$$\begin{aligned} s_0(\delta) &= K_0 \gamma \delta^{-\tau} \,, & K_0 &= 2^{d+2-\tau} \sqrt{\Gamma(2\tau+1)} \,, \\ s_1(\delta) &= K_1 \gamma \delta^{-(\tau+1)} \,, & K_1 &= 2^{d+1-\tau} \sqrt{\Gamma(2\tau+3)} \,. \end{aligned} \tag{5.7}$$

We claim that if (5.2) holds, then we can construct via Proposition 2.13 a sequence of non-degenerate approximate solutions $v^{(j)} \equiv v^{(j-1)} + w^{(j-1)}$ for all $j \geq 1$ (this means that we can apply iteratively for $j \geq 1$ Proposition 2.13 with $v = v^{(j-1)}$, $w = w^{(j-1)}$, $\varepsilon = \varepsilon^{(j-1)}$ and $v' = v^{(j)}$, $\varepsilon' = \varepsilon^{(j)}$, $\mathcal{M} = \mathcal{M}^{(j)} \equiv (\mathbb{1} + v_\theta^{(j)})$, etc.). Moreover if $V_j$, $V_{1j}$, $W_j$, $W_{1j}$, $W_{tj}$, $M_j$, $\overline{M}_j$, $E_j$ denote bounds on the corresponding norms [i.e. $V_j \geq |v^{(j)}|_{\xi_j}$, $V_{1,j} \geq |v_\theta^{(j)}|_{\xi_j}$, $W_j \geq |w^{(j)}|_{\xi_j}$, $W_{t,j} \geq |w_t^{(j)}|_{\xi_j}$, $M_j \geq |\mathcal{M}^{(j)}|_{\xi_j}$, $\overline{M}_j \geq |(\mathcal{M}^{(j)})^{-1}|_{\xi_j}$], then $\overline{M}_j < \infty$ for all $j$ and the following estimates are true for every $j$:

$$E_j \overline{L} \leq (KE\overline{L})^{2^j} \tag{5.8$_j$}$$

$$|\sum_{k=0}^{j} Dw^{(k)}|_{\xi_j} \leq r\Omega \,, \qquad r \equiv \frac{1}{67^4} \tag{5.9$_j$}$$



$$\sum_{k=0}^{j} W_k \leq r \tag{5.10$_j$}$$

$$M_j \leq 2M \tag{5.11$_j$}$$

$$\overline{M}_j \leq 2\overline{M} \tag{5.12$_j$}$$

$$S_j + M_j \leq 2S + 2M \,. \tag{5.13$_j$}$$

Observe that (5.9) and (5.10) show the consistency of the choice of the domain $\mathfrak{D}^{\xi}_{(r,r)}$, which can be kept fixed during the iteration.
To check the claim, observe first, that from the recursive definition of $V_j$, $V_{1j}$, $M_j$, $\overline{M}_j$, $S_j$, it is:

$$\begin{aligned} V_{j+1} &\equiv V + \sum_{i=0}^{j} W_i \\ V_{1,j+1} &\equiv V_1 + \sum_{i=0}^{j} W_{1i} \\ M_{j+1} &\equiv M + \sum_{i=0}^{j} W_{1i} \\ \overline{M}_{j+1} &\equiv \begin{cases} \overline{M} \cdot (1 - \overline{M} \sum_{i=0}^{j} W_{1i})^{-1} & \text{if } \sum_{i=0}^{j} W_{1i} < 1 \\ \infty & \text{if } \sum_{i=0}^{j} W_{1i} \geq 1 \end{cases} \\ S_{j+1} &\equiv S + \sum_{i=0}^{j} W_{ti} \,. \end{aligned}$$

We now want to prove (5.8),...,(5.13) by induction on $j$: $j = 0$ is obvious. Assume the claim true for $0,\ldots,j$; we want to prove it for $j+1$. Let $\lambda \equiv \max(\overline{L}L_3, \overline{L}L_3', \overline{L}L_4, \overline{L}L_4', 1)$; it is not difficult (just a bit tedious) to check that for $i \leq j$:

$$\begin{aligned} E_{i+1}\overline{L} &\leq (E_i \overline{L})^2 \beta_0 {\gamma_0}^i \\ W_i &\leq E_i \overline{L} \beta_1 {\gamma_1}^i \\ W_{1i} &\leq E_i \overline{L} \beta_2 {\gamma_2}^i \\ |Dw^{(i)}| &\leq E_i \overline{L} \beta_3 {\gamma_3}^i \, \Omega \end{aligned} \tag{5.14}$$

with

$$\begin{aligned} \beta_0 &\equiv 8 \cdot 208^2 M^2 (M\overline{M})^8 (L\overline{L})^2 (S+M) K_1 K_0^3 \lambda \gamma^4 2^{8\tau} \xi^{-4\tau-3} \,, & \gamma_0 &\equiv 2^{4\tau+3} \\ \beta_1 &\equiv 13 \cdot M (M\overline{M})^4 L\overline{L} \, K_0^2 \gamma^2 2^{4\tau+1} \xi^{-2\tau} \,, & \gamma_1 &\equiv 2^{2\tau} \\ \beta_2 &\equiv 165 \cdot M (M\overline{M})^4 L\overline{L} \, K_1 K_0 \gamma^2 2^{4\tau} \xi^{-2\tau-1} \,, & \gamma_2 &\equiv 2^{2\tau+1} \\ \beta_3 &\equiv 130 \cdot M (M\overline{M})^4 L\overline{L} \, K_0^2 \gamma^2 2^{4\tau} \xi^{-2\tau-1} \,, & \gamma_3 &\equiv 2^{2\tau+1} \end{aligned}$$

(see Appendix 6 for details).

Next: $(5.12)_j$ is equivalent to

$$2\overline{M}\sum_{i=0}^{j-1} W_{1i} \le 1\,, \qquad (5.15)_j$$

and this condition implies $(5.11)_j$. In view of these comments, to prove the claim we have to show that

$$E_j\overline{L} \le (K E\overline{L})^{2^j}\,, \qquad K \equiv \beta_0\gamma_0\,, \qquad (5.16)$$

$$|\sum_{i=0}^{j} Dw^{(i)}|_{\xi_j} \le r\Omega\,, \qquad r \equiv \frac{1}{67^4}\,, \qquad (5.17)$$

$$\sum_{k=0}^{j} W_k \le r\,, \qquad (5.18)$$

$$2\overline{M}\sum_{k=0}^{j-1} W_{1k} \le 1\,, \qquad (5.19)$$

$$|\sum_{k=0}^{j} w_t^{(k)}| \le S\,. \qquad (5.20)$$

Proof of (5.16):

$$E_{j+1}\overline{L} \le (E_j\overline{L})^2\beta_0\gamma_0{}^j \;\le\; (E\overline{L})^{2^{j+1}}\prod_{i=0}^{j}(\beta_0\gamma_0{}^{j-i})^{2^i}$$

$$= \Big[E\overline{L}\;\beta_0^{\sum_{i=1}^{j+1}\frac{1}{2^i}}\;\gamma_0^{\sum_{i=1}^{j+1}\frac{i-1}{2^i}}\Big]^{2^{j+1}} \;<\; (E\overline{L}\beta_0\gamma_0)^{2^{j+1}}\,,$$

and, since

$$\sum_{i=1}^{j+1}\frac{1}{2^i} = 1-\frac{1}{2^{j+1}}\,, \qquad \sum_{i=1}^{j+1}\frac{i-1}{2^i} = 1-\frac{j+2}{2^{j+1}}\,,$$

one obtains:

$$E_{j+1}\overline{L} \;\le\; \Big[E\overline{L}\beta_0{}^{1-\frac{1}{2^{j+1}}}\;\gamma_0{}^{1-\frac{j+2}{2^{j+1}}}\Big]^{2^{j+1}} = \frac{(E\overline{L}\beta_0\gamma_0)^{2^{j+1}}}{\beta_0\gamma_0{}^{j+2}}\,.$$



Proof of (5.17):

$$|\sum_{i=0}^{j} Dw^{(i)}|_{\xi_j} \le \sum_{i=0}^{j}|Dw^{(i)}|_{\xi_j} \le \beta_3\Omega\sum_{i=0}^{j}E_i\overline{L}\,\gamma_3{}^i$$

$$< \frac{\beta_3}{\beta_0\gamma_0}\Omega\sum_{i=0}^{j}(E\overline{L}\beta_0\gamma_0)^{2^i}\Big(\frac{\gamma_3}{\gamma_0}\Big)^i < \frac{\beta_3}{\beta_0\gamma_0}\Omega\sum_{i=0}^{j}\Big(\frac{1}{2^4}\Big)^i$$

$$< \frac{16\beta_3}{15\beta_0\gamma_0}\Omega \le r\Omega\,, \qquad r \equiv \frac{1}{67^4}\,,$$

where (5.2) and $\frac{16\beta_3}{15\beta_0\gamma_0} \le r$ have been used.

Proof of (5.18):

$$\sum_{k=0}^{j} W_k \le \beta_1\sum_{i=0}^{j}E_i\overline{L}\gamma_1{}^i$$

$$< \frac{\beta_1}{\beta_0\gamma_0}\sum_{i=0}^{j}(E\overline{L}\beta_0\gamma_0)^{2^i}\Big(\frac{\gamma_1}{\gamma_0}\Big)^i < \frac{\beta_1}{\beta_0\gamma_0}\sum_{i=0}^{j}\Big(\frac{1}{2^5}\Big)^i$$

$$< \frac{32\beta_1}{31\beta_0\gamma_0} \le r\,.$$

Proof of (5.19):

$$2\overline{M}\sum_{i=0}^{j-1}W_{1i} \le 2\overline{M}\,\beta_2\sum_{i=0}^{j-1}E_i\overline{L}\gamma_2{}^i$$

$$< \frac{2\overline{M}\beta_2}{\beta_0\gamma_0}\sum_{i=0}^{j-1}(E\overline{L}\beta_0\gamma_0)^{2^i}\Big(\frac{\gamma_2}{\gamma_0}\Big)^i < \frac{2\overline{M}\beta_2}{\beta_0\gamma_0}\sum_{i=0}^{j-1}\Big(\frac{1}{2^4}\Big)^i$$

$$< \frac{32\overline{M}\beta_2}{15\beta_0\gamma_0} \le 1\,.$$

Proof of (5.20): since

$$S_j + M_j \le S + \sum_{i=0}^{j-1}|w_t{}^{(i)}| + M + \sum_{i=0}^{j-1}|w_\theta{}^{(i)}|\,;$$

we have to show that

$$\sum_{i=0}^{j-1}|w_t{}^{(i)}| + \sum_{i=0}^{j-1}|w_\theta{}^{(i)}| \;\le\; S+M\,.$$

From $w^{(i)} = \mathcal{M}^{(i)}z^{(i)}$, one has $w_t^{(i)} = \mathcal{M}_t^{(i)}z + z_t^{(i)}\mathcal{M}$, and

$$|w_t^{(i)}| \le \delta_i^{-1}S_i\frac{E_i\overline{L}a_i}{M_i} \;+\; E_i\overline{L}a_i\,\frac{s_1(\delta_i)}{s_0(\delta_i)}\frac{b_{1i}}{b_i}\,;$$

hence

$$\begin{aligned}
&\sum_{i=0}^{j-1}|w_t{}^{(i)}|+\sum_{i=0}^{j-1}|w_\theta{}^{(i)}| \le \\
&\le \sum_{i=0}^{j-1}\frac{E_i\overline{L}a_i\delta_i^{-1}S_i}{M_i}+\sum_{i=0}^{j-1}E_i\overline{L}a_i\delta_i^{-1}\left(1+2\delta_i\,\frac{s_1(\delta_i)}{s_0(\delta_i)}\frac{b_{1i}}{b_i}\right)\\
&\le 2\sum_{i=0}^{j-1}E_i\overline{L}a_i\delta_i^{-1}\left[\frac{S+M}{M_i}\;+\;\delta_i\,\frac{s_1(\delta_i)}{s_0(\delta_i)}\frac{b_{1i}}{b_i}\right]\\
&\le S+M\,.
\end{aligned}$$

The inductive argument is complete and (5.8) ÷ (5.13) hold for all $j$.
From the above estimates it now follows that

$$|v^{(j)}-v|_{\xi_j}\le|\sum_{i=0}^{j-1}w^{(i)}|_{\xi_j}\;<\;\frac{32\beta_1}{31\beta_0\gamma_0}\,KE\overline{L}<KE\overline{L}\,\frac{\xi}{321^2} \tag{5.21}$$

and

$$|v_\theta^{(j)}-v_\theta|\le|\sum_{i=0}^{j-1}w_\theta{}^{(i)}|_{\xi_j}\;<\;\frac{16\beta_2}{15\beta_0\gamma_0}\,KE\overline{L}\;<\;\frac{KE\overline{L}}{125^2}\,. \tag{5.22}$$

This shows that the series $v+\sum_{i\ge0}w^{(i)}$ converges uniformly in the complex compact domain $\Delta_{\xi_\infty}\equiv\Delta_{\xi/2}$ in the $C^1$-norm to a (real-)analytic function $u\equiv v+\sum_{i\ge0}w^{(i)}$, which by (5.21) and (5.22) satisfies (5.3) and (5.4). Obviously, by construction, $u$ will satisfy (2.4) and by (5.12)

$$|(\mathbb{1}+u_\theta)^{-1}|_{\xi/2}=\lim_{j\to\infty}|(\mathbb{1}+v_\theta^{(j)})^{-1}|_{\xi/2}\le\overline{M}_j\le2\overline{M}\,, \tag{5.23}$$

showing that $u$ is a non-degenerate solution of (2.4).
The condition

$$E\overline{L}\,\beta_0\,\gamma_0\;<\;1$$

is equivalent to $KE\overline{L}\le1$ if one sets:

$$K\;\equiv\;\beta_0\gamma_0\;\equiv\;1664^2\cdot M^2(M\overline{M})^8(L\overline{L})^2(S+M)K_1K_0^3\lambda\gamma^4\;\xi^{-4\tau-3}2^{12\tau}\,.$$

We sketch now the proof of *local uniqueness*. Let $u, u'$ be two non-degenerate solutions of (2.4) with $\langle u\rangle=\langle u'\rangle=0$, real-analytic in some $\Delta_\xi$ ($\xi<1$) and let $w\equiv u'-u$ and set $\eta\equiv|w|_\xi$.
Let $M,\overline{M}$,...be bounds on $|(\mathbb{1}+u_\theta)|_\xi$, $|(\mathbb{1}+u_\theta)^{-1}|_\xi$...and let $\xi_j$, $\delta_j$ be as in (5.6). Since $u$ and $u'$ are both solutions we see that

$$0\equiv\mathscr{E}(u')=\mathscr{E}(u)+\mathscr{E}'(u)w+q_1\,,\qquad\text{i.e.}\quad\mathscr{E}'(u)w+q_1=0\,, \tag{5.24}$$

where $q_1$ is as in (2.24) with $v, v'$ replaced by, respectively, $u$, $u+w\equiv u'$. Now, $q_1$ can be estimated on $\Delta_{\xi_0-\delta_0/2}$ by (cfr. (3.46), (3.48) ) $[30\cdot L_3\Omega\eta^2\xi^{-3}]$; here we are disregarding that many terms in (2.24) actually cancel. On the other hand one can check that setting $z\equiv\mathcal{M}^{-1}w$, one has (cfr. (2.32) and note that in the present case $q_2=0$, $\mathcal{A}=0$):



$$\mathscr{E}'(u)w=\mathcal{M}^{-T}[D(\mathcal{T}Dz)]\,, \tag{5.25}$$

where, of course, the twist-matrix $\mathcal{T}$ is defined as in (2.18) with $u$ replacing $v$. Thus:

$$w=-\mathcal{M}\,D^{-1}\Big\{\mathcal{T}^{-1}[D^{-1}(\mathcal{M}^Tq_1)+c_1]\Big\}+\mathcal{M}c_2\,, \tag{5.26}$$

with $c_1$ defined so that the term in curly brackets has vanishing mean-value and $c_2$ so that $\langle w\rangle=0$ (recall that $u, u'$ have vanishing mean-value and that $\langle\mathcal{M}\rangle=1$):

$$c_1=\langle\mathcal{T}^{-1}\rangle^{-1}\langle\mathcal{T}^{-1}D^{-1}[\mathcal{M}^Tq_1]\rangle\,,\qquad c_2=\langle\mathcal{M}D^{-1}\Big\{\mathcal{T}^{-1}[D^{-1}(\mathcal{M}^Tq_1)+c_1]\Big\}\rangle\,. \tag{5.27}$$

But then one can check that $w$ can be estimated on $\Delta_{\xi_1}$ by (compare with (3.39)):

$$|w|_{\xi_1}\le|q_1|_{\xi_0}\overline{L}a\,, \tag{5.28}$$

where $a$ is defined in (3.39) with $\xi=\xi_0$ and $\delta=\delta_0$. Thus we see that the size of $w$ in $\Delta_{\xi_1}$ is *quadratically* smaller than its size in $\Delta_{\xi_0}$. Now, iterating such argument, mimicking the estimates in this section, one can easily check that

$$|w|_{\xi_j}\le\left(K\eta^2\overline{L}\right)^{2^j}\,, \tag{5.29}$$

which shows that, if (5.2) holds, then $|w|_{\xi_\infty}=0$, so that, by analyticity we can conclude that $w\equiv(u'-u)\equiv0$ in $\Delta_\xi$.
Finally, the last statement in Theorem 5.1 on the dependence on parameters $\mu$ is obvious by the uniformity of the bounds. □

## 6. Application of the KAM Algorithm to Problems with Parameters

In this section we shall describe how one may apply the machinery of §2–§5 in an efficient way to problems with parameters.

### 6.1. Convergent Power-Series (Lindstedt-Poincaré-Moser Series)

In many physical problems one deals with a one-parameter family of Lagrangians $\mathcal{L}(y,x,t;\mu)$ such that for $\mu=0$ an explicit solution $u_0(\theta,t)$ of (2.4) is known. For example if

$$\mathcal{L}\;\equiv\;\mathcal{L}_0(y)\;+\;\mu\mathcal{L}_1(y,x,t)$$

is a nearly-integrable Lagrangian, $u_0\equiv0$ is the trivial solution of (2.4) with $\mu=0$ and any $\omega$. If we assume that $\mathcal{L}$ depends analytically on $\mu\in\mathcal{P}$, where $\mathcal{P}$ is some compact domain in $\mathbf{C}$ containing $\mu=0$ and if all the sup-norms $|\cdot|_.$ of §2–§5 are replaced by $\sup_{\mathcal{P}}|\cdot|_.$, if $E=\sup_{\mathcal{P}}\sup_{\Delta_\xi}|\mathscr{E}(\theta;\mu)|$ satisfies (5.2) then by the uniformity of all the limits, we can conclude that $u=u(\theta,t;\mu)$ *will be real-analytic* in $\Delta_\xi\times\mathcal{P}$; therefore the solution $u$ *admits* a convergent power series expansion in the *complex parameter* $\mu$.
For example in the case of nearly-integrable Lagrangians, as already noted, we can take $v=0$ (which clearly is non-degenerate) so that $\varepsilon=\mu\mathcal{L}_1\equiv\mathcal{L}_1(\omega,\theta,t)$; now if we take

$\mathcal{P} = \{\mu \in \mathbb{C},\ |\mu| \le \mu_0\}$ one sees that

$$E = \mu_0 \, |\mathcal{L}_1|_\xi .$$

Therefore if

$$\mu_0 < \left(K \, |\mathcal{L}_1|_\xi \, \overline{L}\right)^{-1} . \tag{6.1}$$

we can conclude by Theorem 5.1 the existence of $u(\theta, t; \mu)$ real-analytic on $\Delta_{\xi/2} \times \{\mu \in \mathbb{C} \ : \ |\mu| \le \mu_0\}$ and $\mu_0$ close to 0.

The problem of the convergence of such series in $\mu$ is as old as Celestial Mechanics and was considered by Lindstedt (see also [A2] for more informations). Poincaré longly studied this problem, which he considered as one of the central problems in Celestial Mechanics, but did not arrive to any conclusion (actually he thought quite unlikely the convergence of such series but did not exclude it). It was only with the use of KAM techniques that it was possible to answer positively the question. In particular J. Moser devoted a beautiful paper setting completely such a problem ([Mo3]; for different "direct" proofs see [EkI] and [ChF2]).

As one can see from (6.1) the smallness requirement dictated by a *tout-court* application of the KAM theorem gives a radius of convergence absurdely small and certainly of little practical interest. Instead one can proceed as follows.

## 6.2. Improving the Lower Bound on the Radius of Convergence

*First step.* Compute as many as you can (and if you are not Gauss or Delaunay you may want to use a computer) Taylor-Fourier coefficients of the expansion

$$u(\theta, t; \mu) \equiv \sum_{k \ge 1} u_k(\theta, t) \, \mu^k \equiv \sum_{k \ge 1} \ \sum_{(n,m) \in \mathbb{Z}^{d+1}} u_{kn} \, e^{i(n \cdot \theta + mt)} \mu^k .$$

Notice that expanding in $\mu$ the tori equation and equating the $k^{th}$-coefficient in $\mu$ will yield a linear equation for $u_k$ of the form:

$$D^2 u_k(\theta, t) = \Phi_k(\theta, t) ,$$

where $\Phi_k$ is a function (with vanishing mean value on $\mathbb{T}^{d+1}$) depending upon $u_1, \dots, u_{k-1}$ and of the derivatives of $\partial_{y_i} \partial_{x_j} J_{|\mu=0}$ (explicit formulae and a self-contained proof of $\langle \Phi_k \rangle = 0$ is given in §7.1 below).

*Second step.* Let

$$v \equiv v^{(0)} \equiv \sum_{k \le k_0} \ \sum_{|(n,m)| \le N_k} u_{kn} \, e^{i(n \cdot \theta + mt)} \mu^k \tag{6.2}$$

($k_0$ and $N_k$ depending on your computational ability). Make a guess $\mu_0$ for the true radius of convergence (eventually using numerical methods to get a hint). Fix a $\xi$ so that $\Phi_v(\Delta_\xi)$ is contained in the holomorphic domain of $\mathcal{L}$ and evaluate the starting parameters for the KAM algorithm, i.e. evaluate

$$M = \sup_{|\mu| \le \mu_0} \sup_{\Delta_\xi} |\mathbb{1} + v_\theta| \qquad \overline{M} = \sup_{|\mu| \le \mu_0} \sup_{\Delta_\xi} |(\mathbb{1} + v_\theta)^{-1}| , \qquad \text{etc.} ,$$



making sure that $\overline{M} < \infty$ (otherwise reduce the value of $\mu$ and/or $\xi$). This step may also involve computer-aided calculations.

*Third step.* Fix $N$ (a good starting choice may be $10 \le N \le 20$), fix $\delta_0, \dots, \delta_N$ so that $\xi - \sum_{i=0}^{N} \delta_i > 0$ and apply recursively the KAM algorithm, i.e. compute $\mathcal{N}_j$, $0 \le j \le N$ provided, of course, $\overline{M}_j < \infty \ \forall j$. In spite of the complications of the formulae involved, the application of the KAM algorithm is trivial from the computational point of view, *apart* from the evaluation of $s_i(\delta_j)$ ($\equiv$ upper bound on $\sigma_i(\delta_j)$, see (3.17)). Clearly it is possible by using truncations of $\sigma_i(\delta)$ to give an upper bound estimate $s_i(\delta)$ as close as one wishes to the actual value; however such an operation is in general quite time-consuming.

*Fourth step.* If $\overline{M}_j < \infty \ \forall 0 \le j \le N$, plug the values $M_N$, $\overline{M}_N$, $E_N$, $\xi_N = \xi_*$, $L$, $\overline{L}$, etc., in the formula for $K$ (see (5.1)) of the KAM theorem. If $K E_N \, \overline{L} \le 1$ we can apply the theorem and conclude the existence of a true solution $u$, $(K E_N \, \overline{L})$-close to the approximant $v^{(N)}$. Otherwise go back to the second step and vary the parameters $\mu_0$, $\xi$, $N$, $\delta_0, \dots, \delta_N$ (typically one would reduce $\mu_0$, etc.).

Let us be a bit more formal. Fixed the starting approximation $v$ (e.g. as in (6.2)) the above procedure can be viewed as a finite algorithm

$$\Lambda_N(\mu_0; \xi, \delta_0, \dots, \delta_N) ,$$

where $\Lambda_N = 1$ if $K E_N \, \overline{L} \le 1$, $\Lambda_N = 0$ otherwise. What we are after is

$$\mu_N \equiv \sup \{\mu_0 : \ \Lambda_N(\mu_0; \xi, \delta_0, \dots, \delta_N) = 1 \quad \text{for some } \xi, \delta_0, \dots, \delta_N\} \tag{6.3}$$

and since it is fairly clear that $\mu_N$ is increasing with $N$, one wants to get a good approximation of

$$\mu_\infty \equiv \mu_{KAM} \equiv \sup_N \mu_N .$$

Experience teaches us that $N = 20$ is usually a good approximation of $\infty$; however approximating $\mu_N$ is a difficult nonlinear programing problem which we believe is interesting by itself ([ChF]). Notice that in principle (i.e. if we can compute $v$ for any $k_0$ and $N_k$, if we can estimate efficiently the norms relative to $v$, etc.), $\sup_{k_0, N_k} \mu_\infty = \rho_a$, where

$$\rho_a \equiv \inf_\theta \{\text{radius of convergence of } \sum_{k \ge 1} u_k \mu^k\} . \tag{6.4}$$

It is therefore a pure (and highly non-trivial) computational problem to give accurate lower bounds on $\rho_a$.

There is however a much deeper theoretical problem beyond this approach. Namely, let

$$\rho_c \equiv \sup\{\rho : \exists \text{ a solution } u(\theta, t, \mu) \text{ of (2.4) which is } C(\mathbb{T}^{d+1} \times [0, \rho)) \cap C^2(\mathbb{T}^{d+1})\} . \tag{6.5}$$

*What is the relation between $\rho_a$ and $\rho_c$?* There is experimental evidence that in simple models (e.g. the so-called standard map) it is $\rho_a = \rho_c$ (see [BC]); however in general this will not be the case ([BCCF]). An obvious weaker and more realistic approach would be to replace $\rho_a$ above with

$$\rho_r \equiv \sup\{\rho > 0 : \exists \text{ real} - \text{analytic extension of } u(\theta, t, \mu) \text{ to } \mathbb{T}^{d+1} \times [0, \rho)\} . \quad (6.6)$$

Of course, in this case, one would have to replace the above computation of $v$ with other methods taking into account the possibility of $\mu$-analyticity domains different from circles.
Note that, obviously $\rho_a \le \rho_r \le \rho_c$.

One may also think of solving the equation at *given fixed* $\mu$, simplifying, therefore, significantly the first step above, which is by far the most time-consuming; this approach (which has been pursued in [LR]) however yields no informations on the $\mu$-dependence of the solution and in particular cannot be used to give lower estimates on the above critical parameters $\rho$..

## 7. Power Series Expansions and Estimate of the Error Term

As already noted if one considers a one-parameter family of analytic Lagrangians, such that for $\mu = 0$ a solution $u_0$ is known, it follows that there exists for $\mu$ sufficiently small an analytic solution $u(\theta, t; \mu)$ of the Euler-Lagrange equation. This solution can be expanded in power series of the perturbing parameter $\mu$. An approximate solution of the Euler-Lagrange equation can be obtained truncating the power series. The approximate solution will satisfy the Euler-Lagrange equation up to an error term. An indication of estimating this error is provided in the last part of this paragraph.

### 7.1. Power Series Expansions

Let us consider, for simplicity, a special class of nearly integrable Lagrangians (cfr. [CC2]) given by

$$\mathcal{L}(y, x, t) \equiv \frac{1}{2} \sum_{i=1}^{d} \frac{y_i^2}{2} + \mu V(x, t) , \quad (7.1)$$

where $y \in \mathbb{R}^d$ and $(x, t) \in \mathbb{T}^{d+1}$. For this Lagrangian, equation (2.4) takes the form

$$D^2 u = \mu\, V_x(\theta + u, t) , \quad (7.2)$$

where $V_x \equiv (\frac{\partial V}{\partial x_1}, \dots, \frac{\partial V}{\partial x_d})$. As mentioned in §6, if $\omega$ is Diophantine a non-degenerate solution $u$ exists and is analytic in the parameter $\mu$ in a small neighbourhood of the origin; therefore one can expand $u$ in power series as

$$u(\theta, t) \equiv \sum_{k=1}^{\infty} u_k(\theta, t)\mu^k . \quad (7.3)$$

We proceed now to describe a method for finding a recursive relation among the coefficients $u_k$. Let $f(\theta, t) \equiv V_x(\theta, t)$ and expand $f(\theta, t)$ in Fourier series as

$$f(\theta, t) \equiv \sum_{(n,m)\in \mathbb{Z}^{d+1}\setminus\{0\}} \hat{f}_{(n,m)}\, e^{i(n\cdot\theta + mt)} . \quad (7.4)$$



Following an idea that we learned in [Her3] (cfr. also [Go]), we define for any $(n, m) \in \mathbb{Z}^{d+1}$ with $(n, m) \neq 0$ the complex-analytic functions $b_k^{(n,m)}(\theta, t)$ as the coefficients of the series expansion in powers of $\mu$ of $e^{i(n\cdot(\theta+u)+mt)}$:

$$e^{i(n\cdot(\theta+u)+mt)} \equiv \sum_{k=0}^{\infty} b_k^{(n,m)}(\theta, t)\mu^k . \quad (7.5)$$

Differentiating (7.5) with respect to $\mu$ one has:

$$in\cdot u'\, e^{i(n\cdot(\theta+u)+mt)} = \sum_{k=1}^{\infty} k b_k^{(n,m)} \mu^{k-1} ,$$

namely

$$i\, n \cdot \sum_{h=1}^{\infty}\sum_{k=0}^{\infty} h u_h b_k^{(n,m)} \mu^{h-1}\mu^k = \sum_{k=1}^{\infty} k b_k^{(n,m)}\, \mu^{k-1}$$

or

$$i \sum_{k=1}^{\infty}\Big(n \cdot \sum_{h=1}^{\infty} h u_h b_{k-h}^{(n,m)}\Big)\mu^{k-1} = \sum_{k=1}^{\infty} k b_k^{(n,m)}\, \mu^{k-1} . \quad (7.6)$$

A comparison between terms of the same order in $\mu$ in the equality (7.6) shows that

$$b_0^{(n,m)} = e^{i(n\cdot\theta + mt)}$$
$$b_k^{(n,m)} = \frac{i}{k}\, n \cdot \sum_{h=1}^{k} h u_h b_{k-h}^{(n,m)} , \quad k \ge 1 . \quad (7.7)$$

Therefore, by (7.2)

$$D^2 u = \mu \sum_{(n,m)\in\mathbb{Z}^{d+1}\setminus\{0\}} \hat{f}_{(n,m)}\, e^{i(n\cdot(\theta+u)+mt)} ,$$

and by (7.5)

$$D^2 u_k = \sum_{(n,m)\in\mathbb{Z}^{d+1}\setminus\{0\}} \hat{f}_{(n,m)}\, b_{k-1}^{(n,m)} . \quad (7.8)$$

This equation makes sense provided that the right hand side has mean average zero over $\mathbb{T}^{d+1}$. Actually, we already know that this fact is true, since we proved in §6 that if $\mu_0$ is small enough then $\sum_{k\ge 1} u_k \mu^k$ is an absolutely convergent series for $|\mu| \le \mu_0$. However, for completeness we shall give now a purely algebraic check of the vanishing of the r.h.s. of (7.8). Let us denote by $[\cdot]_k$ the $k$-th coefficient of the $\mu$-power series expansion:

$$\text{if } \; g \equiv \sum_{k=1}^{\infty} g_k\, \mu^k \quad \text{then} \quad [g]_k = g_k .$$

From equation (7.2) we can rewrite (7.8) as

$$D^2 u_k = [\mu V_x(\theta + u, t)]_k .$$

**Proposition 7.1.** *Let $u_0 \equiv constant$ and let $k \geq 1$. Assume that there exist $u_0, \ldots, u_{k-1}$ such that for every $0 \leq l \leq k-1$ one has*

$$D^2 u_l = [\mu\, V_x(\theta + \sum_{i=1}^{l-1} u_i \mu^i, t)]_l \,. \tag{7.9}$$

*Then*

$$\int_{\mathbf{T}^{d+1}} [\mu\, V_x(\theta + \sum_{i=1}^{k-1} u_i \mu^i, t)]_k \, d\theta dt = 0 \,. \tag{7.10}$$

**Proof:** Notice that for $k = 1$, $u_0 \equiv$ const. does satisfy (7.9).
Now, for any function $G = G(x, t)$,

$$\int_{\mathbf{T}^d} \partial_\theta [G(\theta + u, t)] d\theta = 0 = \int_{\mathbf{T}^d} (\mathbb{1} + u_\theta)\, G_x(\theta + u, t) d\theta \,.$$

Let $G(x, t) \equiv \mu V_x(x, t)$; then

$$\int_{\mathbf{T}^d} (\mathbb{1} + u_\theta)\, \mu V_x(\theta + u, t) d\theta = 0$$

and since $[\cdot]$ is a linear operator, for any $l \geq 0$

$$\begin{aligned} [\int_{\mathbf{T}^d} (\mathbb{1} + u_\theta)\, \mu V_x(\theta + u, t) d\theta]_l &= 0 \\ = \int_{\mathbf{T}^d} [\mu V_x]_l d\theta &+ \int_{\mathbf{T}^d} [\mu u_\theta V_x]_l d\theta \,. \end{aligned} \tag{7.11}$$

Now by (7.9) and recalling that $u_0$ is independent of $\theta$, one sees that

$$\begin{aligned} \int_{\mathbf{T}^{d+1}} [\mu u_\theta V_x]_k \, d\theta dt &= \sum_{\substack{j+i=k \\ 1 \leq j \leq k,\ 0 \leq i \leq k-1}} \int_{\mathbf{T}^{d+1}} (u_j)_\theta \, [\mu V_x]_i \, d\theta dt \\ &= \sum_{j=1}^{k} \int_{\mathbf{T}^{d+1}} (u_j)_\theta \, D^2 u_{k-j} = \sum_{j=1}^{k-1} \int_{\mathbf{T}^{d+1}} (u_j)_\theta \, D^2 u_{k-j} \\ &= \frac{1}{2} \sum_{j=1}^{k-1} \int_{\mathbf{T}^{d+1}} \Big\{ (u_j)_\theta \, D^2 u_{k-j} + (u_{k-j})_\theta \, D^2 u_j \Big\} d\theta dt \,. \end{aligned}$$

Finally, integrating by parts three times one finds

$$\int_{\mathbf{T}^{d+1}} (u_j)_\theta \, D^2 u_{k-j} \, d\theta dt = - \int_{\mathbf{T}^{d+1}} D^2 u_j \, (u_{k-j})_\theta \, d\theta dt$$

(notice that in this identity we have to integrate over $t$ too). Therefore

$$\int_{\mathbf{T}^{d+1}} [\mu u_\theta V_x]_k \, d\theta dt = 0$$



and integrating (7.11) over $t$, one obtains (7.10). □

Thus one can invert the operator $D^2$ in (7.8) to get

$$u_k = D^{-2} \sum_{(n,m) \in \mathbb{Z}^{d+1} \setminus \{0\}} \hat{f}_{(n,m)} \; b_{k-1}^{(n,m)} \,, \tag{7.12}$$

which defines $u_k$ ($\langle u_k \rangle = 0$), in terms of the preceding functions $u_0, \ldots, u_{k-1}$. Notice that it is not legitimate to interchange the order of the summation and of $D^{-2}$ as the functions $b_k^{(n,m)}$ *may not* have vanishing mean-value.

## 7.2. Truncated Series as Initial Approximations and the Majorant Method

We choose now, as initial approximate solution of (7.2) a truncation of the $\mu$-expansion of $u$ and discuss the estimates on the associated error function.
Thus, if $\{u_l\}$, $l \geq 1$ ($u_0 = 0$ as $\langle u_i \rangle = 0$), are the functions defined in the previous paragraph, we set

$$v^{(0)}(\theta, t) \equiv \sum_{l=1}^{l_0} u_l(\theta, t) \mu^l \,, \tag{7.13}$$

for a suitable $l_0 \in \mathbb{Z}_+$. Notice that if $V$ is a trigonometric polynomial, then so are the $u_l$ and the computation of $v^{(0)}$ reduces to a finite number of steps. In general, one can introduce truncations in Fourier space according to the desidered accuracy.

Recalling from §3.3 the definition of the norm-parameters, we see that in the present case [(7.1)], $L = \bar{L} = 1$, $L_3' = L_4 = L_4' = 0$; the vector $\rho$ can be replaced by $\rho_2$ (as no geometry in the $y$-variables comes really in). Thus, the only parameters we have to evaluate are $S_0$, $M_0$, $\overline{M}_0$, $E_0$, which are upper bounds on the norms of $v_t^{(0)}$, $\mathbb{1} + v_\theta^{(0)}$, $(\mathbb{1} - v_\theta^{(0)})^{-1}$, $\varepsilon^{(0)}$ (where $\varepsilon^{(0)} \equiv \mathscr{E}(v^{(0)})$) in the domain $\Delta_\xi \times \mathscr{P} \equiv \Delta_\xi \times \{\mu \in \mathbb{C} : |\mu| \leq \mu_0\}$; we also need to evaluate the parameter $L_3 \geq \mu_0 \, |V_{xxx}(\theta + v^{(0)}, t)|_{\Delta_{\rho_2}^\xi}$ so that $\lambda \equiv \max\{L_3, 1\}$. The estimate of $v_t^{(0)}$ can be obtained using

$$|\partial_t v^{(0)}|_{\xi, \mu_0} \leq \sum_{l=1}^{l_0} |\partial_t u_l|_{\xi, \mu_0} \, \mu_0^l \,, \qquad |\cdot|_{\xi, \mu_0} \equiv \sup_{\Delta_\xi \times \mathscr{P}} |\cdot| \,,$$

and analogously for $v_\theta^{(0)}$. Then $M_0$ and $\overline{M}_0$ can be estimated respectively by $1 + |v_\theta^{(0)}|_{\xi, \mu_0}$ and $(1 - |v_\theta^{(0)}|_{\xi, \mu_0})^{-1}$, *provided*

$$|v_\theta^{(0)}|_{\xi, \mu_0} < 1 \,.$$

We also set

$$V_0 \geq \sum_{l=1}^{l_0} |u_l|_\xi \, \mu_0^l \geq |v^{(0)}|_{\xi, \mu_0}$$

and

$$\rho \equiv \rho_2 \equiv \xi + V + r \qquad (r \equiv \frac{1}{67^4})$$

and we take $L_3 \geq \mu_0 |V_{xxx}|_{\Delta_\rho}$.
It remains to estimate the error-term $\varepsilon^{(0)}$. Let $f(\theta, t) \equiv V_x(\theta, t)$; by (7.9)

$$\begin{aligned}\varepsilon^{(0)} &= D^2 v^{(0)} - \mu f(\theta + v^{(0)}, t)\\ &= \sum_{l=1}^{l_0} D^2 u_l\, \mu^l - \mu f(\theta + v^{(0)}, t)\\ &= \sum_{l=1}^{l_0} D^2 u_l\, \mu^l - \mu \sum_{(n,m)\neq 0} \hat{f}_{(n,m)}\, e^{i(n\cdot(\theta+v^{(0)})+mt)}\\ &= \sum_{l=1}^{l_0} D^2 u_l\, \mu^l - \sum_{(n,m)\neq 0} \hat{f}_{(n,m)} \sum_{h=1}^{\infty} d_{h-1}^{(n,m)}(\theta, t)\mu^h\,,\end{aligned} \tag{7.14}$$

where the functions $d_h^{(n,m)}(\theta, t)$ are defined as the coefficients of the power series expansion

$$e^{i(n\cdot(\theta+v^{(0)})+mt)} \equiv \sum_{h\geq 0} d_h^{(n,m)}(\theta, t)\mu^h\,. \tag{7.15}$$

Therefore, one has:

$$\begin{aligned}\varepsilon^{(0)}(\theta, t) &= \Big[\sum_{l=1}^{l_0} \mu^l \Big(D^2 u_l - \sum_{(n,m)\neq 0} \hat{f}_{(n,m)} d_{l-1}^{(n,m)}\Big)\Big]\\ &\quad - \Big[\mu \sum_{(n,m)\neq 0} \hat{f}_{(n,m)} \sum_{l=l_0}^{\infty} \mu^l d_l^{(n,m)}\Big]\\ &\equiv F_{l_0} + R_{l_0} = R_{l_0}\,,\end{aligned} \tag{7.16}$$

since $F_{l_0} = 0$ because of the definition (7.13) of $v^{(0)}$. (If $V$ is not a trigonometric polynomial one can replace $|F_{l_0}|$ with an arbitrarily small positive number.)
To estimate $|R_{l_0}|$ we shall make use of an old technique, which we shall refer to as the *majorant method*. Such a technique, used e.g. by C. L. Siegel in [S], consists, roughly speaking, in comparing the supremum of an analytic function with the *value* of another analytic function with positive coefficients. More precisely:

**Lemma 7.2.** *Let $v^{(0)}(\theta, t)$ and $d_h^{(n,m)}(\theta, t)$ be as in (7.13) and (7.15) respectively. For any $\xi > 0$ define the sequence $\{a_l^{(n,m)}(\xi)\}$, $0 \leq l \leq l_0$, by*

$$\begin{aligned}a_0^{(n,m)} &= (|n| + |m|)\xi\\ a_l^{(n,m)} &\geq |n \cdot u_l(\theta, t)|_\xi\,, \quad 1 \leq l \leq l_0\end{aligned} \tag{7.17}$$

*and for $l \geq 0$ let $\delta_l^{(n,m)}$ be defined by the identity*

$$\exp\Big(\sum_{l=0}^{l_0} a_l^{(n,m)} \mu^l\Big) = \sum_{l=0}^{\infty} \delta_l^{(n,m)} \mu^l\,. \tag{7.18}$$

*Then one has*

$$|d_l^{(n,m)}(\theta, t)|_\xi \leq \delta_l^{(n,m)} \tag{7.19}$$



*and for any $\mu_0 > 0$,*

$$|\sum_{l=l_0}^{\infty} d_l^{(n,m)}(\theta, t)\mu^l|_{\xi,\mu_0} \leq \exp\Big(\sum_{l=0}^{l_0} a_l^{(n,m)} \mu_0^l\Big) - \sum_{l=0}^{l_0-1} \delta_l^{(n,m)} \mu_0^l\,, \tag{7.20}$$

*where, as above, $|\cdot|_{\xi,\mu_0} \equiv \sup_{\Delta_\xi \times \{|\mu| \leq \mu_0\}}$.*

**Proof:** As in the discussion on the $b$'s (cfr. (7.5) ÷ (7.7)), it is easy to check that $d_k^{(n,m)}$ and $\delta_k^{(n,m)}$ verify the recursive relations

$$\begin{aligned}d_0^{(n,m)}(\theta, t) &= e^{i(n\cdot\theta+mt)}\\ d_l^{(n,m)}(\theta, t) &= \frac{i}{l} n \cdot \sum_{h=1}^{\min(l,l_0)} h u_h(\theta, t) d_{l-h}^{(n,m)}\,, \quad l \geq 1\,,\\ \delta_0^{(n,m)} &= e^{(|n|+|m|)\xi}\\ \delta_l^{(n,m)} &= \frac{1}{l} \sum_{h=1}^{\min(l,l_0)} h a_h^{(n,m)} \delta_{l-h}^{(n,m)}\,, \quad l \geq 1\,.\end{aligned} \tag{7.21}$$

We prove (7.19) by induction on $l$. For $l = 0$,

$$|d_0^{(n,m)}(\theta, t)|_\xi \leq e^{(|n|+|m|)\xi} \equiv \delta_0^{(n,m)}\,.$$

Let now $l \geq 1$ and *assume* (7.19) for $0 \leq h \leq l-1$; by (7.17) and the inductive hypotheses one has:

$$\begin{aligned}|d_l^{(n,m)}(\theta, t)|_{\xi,\mu_0} &\leq \frac{1}{l} \sum_{h=1}^{\min(l,l_0)} h\, |n \cdot u_h(\theta, t)|_\xi |d_{l-h}^{(n,m)}(\theta, t)|_\xi\\ &\leq \frac{1}{l} \sum_{h=1}^{\min(l,l_0)} h a_h^{(n,m)} \delta_{l-h}^{(n,m)} \equiv \delta_l^{(n,m)}\,.\end{aligned}$$

Inequality (7.20) now follows from (7.19) and the definition of the $a_l$'s in (7.18). □

Therefore the estimate of the error term (7.16) on the domain $\Delta_\xi \times \{\mu \in \mathbb{C} : |\mu| \leq \mu_0\}$ is given by

$$|\varepsilon^{(0)}(\theta, t)|_{\xi,\mu_0} \leq \mu_0 \sum_{(n,m)\neq 0} |\hat{f}_{(n,m)}| \Big\{\exp\Big(\sum_{l=0}^{l_0} a_l^{(n,m)} \mu_0^l\Big) - \sum_{l=0}^{l_0-1} \delta_l^{(n,m)} \mu_0^l\Big\}\,, \tag{7.22}$$

where $a_l^{(n,m)}$ and $\delta_l^{(n,m)}$ are given in (7.17), (7.21).

The strategy outlined in this section was carried out in [CC2] and led to existence estimates (for the standard map and a forced pendulum) that are away from the experimentally observed "break-down values" by (respectively) a factor ~ 0.55 and 0.67.
However a serious computational hindrance is hidden in this approach. In fact, in order to carry out the above steps one needs to *rigorously* control the computational errors introduced by mechanical calculations. To do this one may use the so-called *interval arithmetic* (see §8 below for a more detailed discussion), which consists, basically, in

trapping the result of a computation performed by a computer in an interval whose end-points are representable by the machine and which is sure to contain the actual result of the given computation. Now, the computation of the $u_l$'s contains a lot of division by "small divisors" ($\omega \cdot n + m$) which have the effect of spreading very quickly the size of the intervals controlling the $u_l$'s.

To avoid such a problem one would have to turn to "arbitrary" accuracy computations, which are, obviously, very time-consuming (see [CC3] for more informations on this phenomenon).

A different approach is the following.

## 7.3. Numerical Initial Approximations

A quite different approach is to *compute numerically* (i.e. without caring about errors) the functions $u_l$ by means of (7.12) and (7.7) and *define* the initial approximation as

$$v^{(0)}(\theta, t) \equiv \sum_{l=1}^{l_0} \tilde{u}_l(\theta, t)\mu^l , \quad l_0 \in \mathbb{Z}_+ ,$$

where the $\tilde{u}_l$'s are the result, given by a computer, of the implementation of (7.12) and (7.7). With this choice of $v^{(0)}$, one has

$$F_{l_0} \equiv \sum_{l=1}^{l_0} \mu^l \left( D^2\tilde{u}_l - \sum_{(n,m)\neq 0} \hat{f}_{(n,m)}\, d_{l-1}^{(n,m)} \right) ,$$

which, *eventhough will not vanish anymore*, can be estimated with a finite number of operations. The estimate of $R_{l_0}$ is obtained, instead, applying directly Lemma 7.2, with $u_l$ replaced by $\tilde{u}_l$.

The advantage of this approach is that interval arithmetic is not used *directly* to control small divisors.

This strategy has been implemented in [CC4] on various models and yields indeed sensibly better results; for example, the existence of the "golden-mean" invariant curve for the standard map is established for values of the parameter away by a factor 1.16 from optimal (see also §10).

# 8. Computer Assisted Methods

In order to apply the method outlined in the preceding sections, one may have to perform lengthy but straightforward calculations (e.g. to calculate a "good" initial approximation together with the associated norms), in which case the use of computers may be helpful. In this section we briefly discuss the so-called *interval arithmetic*, the implementation of which allows to take care of rounding-off and propagation errors introduced by computers.

## 8.1. Representable Numbers and Intervals

A computer can represent exactly a finite set of numbers, which we shall call here the set of *representable numbers* $\mathfrak{R}$ (of course $\mathfrak{R}$ depends on the particular computer we



are considering). Such numbers are encoded by strings of "bits", i.e., 0's or 1's. For example, if $x = \sum_{j=1}^{N} \epsilon_j 2^{-j}$, ($\epsilon_j = 0$ or 1) is the binary expansion of the rational number $x \in [0, 1)$, one can identify $x$ with $\{\epsilon_1, ..., \epsilon_N\}$. To represent other rational numbers in $\mathbb{R}$, the computer uses extra bits in a symbolic way (see next section for a discussion of how VAXes handle that). Now, to deal with *real* numbers without making approximations one can try to trap them within the "smallest" possible interval whose end-points are in $\mathfrak{R}$. In this way, operations among numbers are replaced (in a quite straightforward way) by operations among intervals.

Before going into more details, let us discuss how computers perform "elementary" operations (i.e., additions, subtractions, multiplications and divisions).

In general, the result of an elementary operation between representable numbers is *not* a representable number and therefore the computer will, in general, approximate such a result. For some computers (like VAXes) the approximation rule is the following. Let us call *rounding bit* the first bit lost in the truncation of the theoretical result. Then:

*(i)* if the rounding bit is 0, the rounded result is equal to the chopped number;

*(ii)* if the rounding bit is 1, the chopped result is increased by one bit.

It is therefore clear that by modifying suitably certain bits one can find intervals of representable numbers which contain the theoretical result.

## 8.2. Intervals on VAXes

As an example, let us consider a VAX. The procedure to create *upper* and *lower* bounds on the result of an elementary operation depends on the structure of data one is working with. Therefore we start by illustrating the different kind of precisions available on a VAX (see [Vax]).

Real numbers are represented in "floating point" notation by a sign, an exponent and a fraction. The size of the floating point data may be of 32, 64, 128 bits. Correspondingly one distinguishes between F-floating (i.e. *simple* precision), D or G-floating (i.e. *double* precision) and H-floating (i.e. *quadruple* precision). The difference between D and G-floating is that G-floating reserves more space to the exponent (allowing numbers in the range of $\sim 0.56 \cdot 10^{-308}$ to $\sim 0.9 \cdot 10^{308}$) with consequent loss of precision of the fraction. In our computations we use G-floating data which we are going to describe in full detail. A G-floating datum is composed by 64 bits (i.e. a set of 0's and 1's), labelled from 0 to 63. The first bit denotes the sign of the number; bits 1 to 11 correspond to the exponent (one bit is reserved for the sign of the exponent), while the remaining 52 bits individuate the fraction. The precision of a G-floating number is approximately of 15 decimal digits. Moreover, there are two extra hidden *guard* bits which guarantee the result of an elementary operation "up to 1/2 of the last significant bit" (see [Vax], appendix H; the quoted sentence is related to the approximation rule discussed above and roughly speaking it means that the bit before the last one is always correct, while the last one is used to get "the closest guess" to the true result and therefore may not coincide with the corresponding bit of the theoretical outcome).

With this representation of real numbers, it can be shown that the upper and lower bounds on the result of an elementary operation are obtained, respectively, *by increasing or decreasing by one bit the last bit of the mantissa, taking eventually care of the* *propagation of the carry.* (The Fortran procedures to create upper and lower bounds of a real number are described in Appendix 7 ).

### 8.3. Interval Operations

We end this section by discussing in more detail interval operations.
Let $\odot$ denote one of the four elementary operations and, for $a, b \in \mathfrak{R}$, let $\langle a \odot b \rangle$ be the *result produced* by a VAX; let Up($a$), Down($a$) be the representable number obtained by, respectively, increasing or decreasing the last significant bit (taking care of the possible carry). Then it is clear that:

$$a \odot b, \ \langle a \odot b \rangle \in (\mathrm{Down}(\langle a \odot b \rangle), \mathrm{Up}(\langle a \odot b \rangle)), \qquad a, b \in \mathfrak{R} .$$

Now, consider first additions. If $a \in (a_-, a_+)$ and $b \in (b_-, b_+)$ where $a, b \in \mathbb{R}$ and $a_\pm, b_\pm \in \mathfrak{R}$, then, as above,

$$a + b \in (\mathrm{Down}(\langle a_- + b_- \rangle), \mathrm{Up}(\langle a_+ + b_+ \rangle)) \equiv (a_-, a_+) + (b_-, b_+) ,$$

which serves as definition of addition (and subtraction) between intervals.
The multiplication is slightly more complicated and several subcases must be considered in order to properly define $(a_-, a_+) * (b_-, b_+) \equiv (c_-, c_+)$, $(a_\pm, b_\pm, c_\pm \in \mathfrak{R})$. In computer-like language:

*(i):* Let $a_- \geq 0$.
If $b_- \geq 0$ then $(c_-, c_+) \equiv (\mathrm{Down}\langle a_- * b_- \rangle, \mathrm{Up}\langle a_+ * b_+ \rangle)$;
if $b_+ \leq 0$ then $(c_-, c_+) = (\mathrm{Down}\langle a_+ * b_- \rangle, \mathrm{Up}\langle a_- * b_+ \rangle)$.
Finally, if $b_- < 0$ and $b_+ > 0$, then $(c_-, c_+) = (\mathrm{Down}\langle a_+ * b_- \rangle, \mathrm{Up}\langle a_+ * b_+ \rangle)$.

*(ii):* Let $a_+ \leq 0$.
If $b_- \geq 0$ then $(c_-, c_+) = (\mathrm{Down}\langle a_- * b_+ \rangle, \mathrm{Up}\langle a_+ * b_- \rangle)$;
if $b_+ \leq 0$ then $(c_-, c_+) = (\mathrm{Down}\langle a_+ * b_+ \rangle, \mathrm{Up}\langle a_- * b_- \rangle)$.
Finally, if $b_- < 0$ and $b_+ > 0$, then $(c_-, c_+) = (\mathrm{Down}\langle a_- * b_+ \rangle, \mathrm{Up}\langle a_- * b_- \rangle)$.

*(iii):* Let $a_- < 0$ and $a_+ > 0$.
If $b_- \geq 0$ then $(c_-, c_+) = (\mathrm{Down}\langle a_- * b_+ \rangle, \mathrm{Up}\langle a_+ * b_+ \rangle)$;
if $b_+ \leq 0$ then $(c_-, c_+) = (\mathrm{Down}\langle a_+ * b_- \rangle, \mathrm{Up}\langle a_- * b_- \rangle)$.
Finally, if $b_- < 0$ and $b_+ > 0$, then
$(c_-, c_+) = (\min\{\mathrm{Down}\langle a_- * b_+ \rangle, \mathrm{Down}\langle a_+ * b_- \rangle, \}, \ \max\{\mathrm{Up}\langle a_- * b_- \rangle, \mathrm{Up}\langle a_+ * b_+ \rangle\})$ .

The division is treated in an analogous way.
Elementary transcendental functions (exponentials, logarithms, trigonometric functions, etc.) may be approximated by a finite sequence of elementary operations using Taylor expansions and simple inequalities to truncate the expansion at a certain (arbitrarily defined) order; for example:

$$\sum_{n=0}^{N-1} \frac{x^n}{n!} < e^x < \sum_{n=0}^{N-1} \frac{x^n}{n!} + \frac{x^N}{N!} \frac{N}{N-1}, \qquad \forall\, N \geq 2, \quad \forall\, 0 < x \leq 1 .$$



## 9. Applications: Three-Dimensional Phase Space Systems

Here and in the following section 10, we briefly illustrate the computer-assisted application of the above theory to a few concrete models. Some of the results presented here are new (compare end of §9.1), while most of them were obtained in preceding works of the authors: see [CC1], [CC2], [CC3], [CC4], [C1], [C2], to which we refer also for complete details.

### 9.1. A Forced Pendulum

One of the simplest non-integrable conservative system is a periodically forced classical pendulum described by the one-dimensional, time-dependent Lagrangian

$$\mathcal{L}(y, x, t) \equiv \frac{y^2}{2} - \mu\Big[\cos x + \cos(x - t)\Big] , \tag{9.1}$$

where $y \in \mathbb{R}$ and $(x, t) \in \mathbb{T}^2$.
This model, which is the central object of the renormalization theory of Escande and Doveil ([ED], [Es]) can also be viewed as describing the motion of a particle with charge $\mu$, in the field of a potential of two longitudinal (electrostatic) waves.
Here, we apply the method presented in the previous sections to (9.1) in order to construct the "*golden-mean* torus" $\mathcal{T}(\omega_g) \equiv \mathcal{T}(\frac{\sqrt{5}-1}{2})$, for values of the non-linearity parameter $\mu$ of the same order of magnitude of the expected "break-down" threshold (see below for definitions and for an experimental recipe to compute such a threshold).
The interest in the stability properties of this particular torus comes from various considerations: the main being that the golden-mean $\omega_g$, which satisfies (2.33) with $\gamma = (\sqrt{5} + 3)/2$ and $\tau = 1$ (cfr. Appendix 8), is, in a suitable sense, the "most irrational" number in (0, 1) and one expects this fact to show up in the $\mu$-power series, which involves the small divisors $(\omega_g n + m)$. Of course, this is a rather naive observation that might lead to the belief that the $\omega_g$-torus is always the most stable, while (cfr. §10) there are examples pointing in different directions.
Numerical methods for determining the critical value $\rho_c$ (see §6.2) have been developed in [Gr], [C], [ED]. The remarkable method developed by J. Greene in [Gr] is based on the following idea. Let $\{p_j/q_j\}$ be the sequence of rational approximants to the irrational number $\omega$ (see Appendix 8 for more informations) and let $\mathcal{P}(p_j/q_j)$ denote a periodic orbit with period $q_j$ and rotation number $p_j/q_j$. Greene conjectures that the disappearance of the torus $\mathcal{T}(\omega)$ is related to a sudden change, from stability to instability, of the periodic orbits $\mathcal{P}(p_j/q_j)$, which, as $j \to \infty$, approximates the torus $\mathcal{T}(\omega)$. This criterion applied to the forced pendulum (9.1) indicates that $\mu_c(\omega_g)$ is $\simeq 0.027$.

Now a brief history. Our first attempt ([CC1],[CFP]) to obtain stability estimates for $\mathcal{T}_\mu(\omega_g)$ in the above model was based on refining Arnold's version of the KAM theorem ([A1], see also [G1]). This strategy allowed us to establish existence for $0 \leq \mu \leq 6.75 \cdot 10^{-4}$, a value which was later increased up to $1.42 \cdot 10^{-3}$ in [CG]. However, this approach, which, as is well known, is based on a sequence of canonical (symplectic) transformations, presents an intrinsic difficulty related to the geometry of the domain where the the canonical transformations are defined. In fact, in order to control the resonances (i.e., phase points where the frequencies are rationally dependent),

such domains have to be taken smaller and smaller as the iteration is carried out and to obtain sharp quantitative results one is led to the difficult analysis of the domain of holomorphy of each canonical transformation.
This problem is bypassed by considering directly the parametric equation for the tori as discussed in the present work. The first implementation of this new strategy was carried out in [CC2] where the Euler-Lagrange equation

$$D^2 u = \mu \left[\sin(\theta + u) + \sin(\theta + u - t)\right] \tag{9.2}$$

is solved using as initial approximation the finite power series

$$v^{(0)}(\theta, t) \equiv \sum_{l=1}^{l_0} u_l(\theta, t)\mu^l , \tag{9.3}$$

the $u_l$'s being the Taylor coefficients of the (convergent) expansion around $\mu = 0$ of the solution (see §7.2).
Since, for $|\mu| < \rho_a$ (cfr. §6.2), $v^{(0)} \to u$ as $l_0 \to \infty$, we shall get better initial approximation by taking $l_0$ large. However the number of Fourier coefficients of $u_l$ grows rapidly with the order $l$: $u_1$ has 4 Fourier coefficients, $u_{10}$ has 120 coefficients and $u_{60}$ has 3720 coefficients. Therefore, computer-time limitations (if nothing else) forces to stop at relatively small orders. In [CC2] we computed the functions $u_l$ using the general formula

$$D^2 u_{l+1} = \sum_{h \in \mathcal{H}_l} \partial_\theta^{|h|} f \prod_{i=1}^{l} \frac{u_i^{h_i}}{h_i!} , \tag{9.4}$$

where $\mathcal{H}_l \equiv \{h \in \mathbb{N}^l : h_1 + 2h_2 + ... + lh_l = l\}$, with $f \equiv \sin\theta + \sin(\theta - t)$. Such a general formula presents, however, serious combinatorial problems as $l$ gets large. Using (9.4) (and about two hours of CPU on a VAX 8600) we computed, *using interval arithmetic*, (9.3) for $l_0 = 24$ and proved the existence of $\mathcal{T}(\frac{\sqrt{5}-1}{2})$ for $|\mu| < 0.015$.
In [CC3], using formulae (7.7) and (7.12), which reduce considerably the combinatorics problems, we could compute (with about the same computer time) (9.3) with $l_0 = 40$. At this order the existence of the golden-mean torus can be established (as above, via a KAM algorithm very similar to the one presented in the present work) for $|\mu| \le 0.018$. Finally, using the strategy described in §7.3, we computed numerically (i.e. without interval arithmetic) $v_0$ up to order $l_0 = 60$ and establish the existence of the golden mean torus for $|\mu| \le 0.019$. This result, which is new and to our present knowledge is the best rigorous result in a hamiltonian setting, is in agreement of the 70% with the numerical guess provided by Greene's method.

To obtain this existence result we solve the Euler-Lagrange equation (9.2) using the initial approximate solution $v^{(0)} = \sum_1^{l_0} \bar{u}_l \mu^l$ where the functions $\bar{u}_l(\theta, t)$ are *numerically computed* using the recursive formulae of §7, namely

$$b_0^{(n,m)} = e^{i(n\theta + mt)}$$

$$b_l^{(n,m)} = \frac{i}{l}\, n \cdot \sum_{h=1}^{l} h \bar{u}_h b_{l-h}^{(n,m)} , \quad l \ge 1$$



and

$$\bar{u}_l = \frac{1}{2i} D^{-2}\left[b_{l-1}^{(1,0)} - b_{l-1}^{(-1,0)} + b_{l-1}^{(1,-1)} - b_{l-1}^{(-1,1)}\right]$$

(which are easily recognized as (7.12) with $f(x, t) = \sin(x) + \sin(x - t)$).
The computation of the functions $\bar{u}_l$ has been performed on a VAX 6000; the computer time necessary to evaluate the function $u_{60}$ was about 24 minutes.
Then, following the (quite straightforward) steps of §6.2, we obtained the following

**Theorem 9.1.** *Let* $\omega = \frac{\sqrt{5}-1}{2}$ *and let* $\xi = 0.08$, $\rho = 0.019$. *Then equation (9.2) admits a locally unique real-analytic solution* $u(\theta, t; \mu)$ *with* $\langle u \rangle = 0$ *on* $\mathbb{T}^2$, *analytic in* $\Delta_\xi \times \{\mu \in \mathbb{C} : |\mu| \le \rho\}$. *Moreover, one can construct a polynomial approximation* $v(\theta, t; \mu) \equiv \sum_{l=1}^{60} \bar{u}_l(\theta, t)\mu^l$, *where* $\bar{u}_l$ *are trigonometric polynomials, satisfying*

$$|u - v|_{\xi,\rho} < 0.2526 , \qquad |u_\theta - v_\theta|_{\xi,\rho} < 0.3824$$

*where* $|\cdot|_{\xi,\rho} \equiv \sup_{\Delta_\xi, |\mu| \le \rho} |\cdot|$.

## 9.2. Spin-Orbit Coupling in Celestial Mechanics

We discuss now an example drawn from Celestial Mechanics, which has been investigated in [C1], [C2], [CF] using the methods developped in [CC2].

One of the most astonishing phenomena in the mechanics of our solar system is that all the *evolved* satellites of the solar system always point the same face toward the host planet, as in the Moon-Earth case. The only exception to this rule is the Mercury-Sun system, as radar observations have shown that the period of revolution of Mercury around the Sun is $\frac{3}{2}$ of the period of rotation about its spin-axis.
Exact commensurabilities between the period of rotation and the period of revolution go under the name of *spin-orbit resonances* ([GP], [He], [W]). More precisely, for an oblate satellite $S$ orbiting around a central body $P$ one has a $p : q$ resonance (for any $p, q \in \mathbb{Z}_+$) when the motion of $S$ is periodic and the ratio between the periods of revolution and rotation of the satellite is $\frac{p}{q}$; we shall denote by $\mathcal{P}(p/q)$ the set of all such orbits.
A natural question is whether the motion of satellites observed in spin-orbit resonance is stable or not.
Following [C1], [C2], [CF], we shall use the theory of invariant surfaces to give a positive answer to such a question under suitable simplifying assumptions.

Let us start by introducing the model we want to study.
Let $S$ be a triaxial homogeneous ellipsoidal satellite with principal moments of inertia $A < B < C$ subject to the gravitational attraction of a (fixed) central planet $P$ and assume that:
*i)* the orbit of the center of mass of $S$ around the central body $P$ is a fixed Keplerian ellipse;
*ii)* the spin-axis of $S$ coincides with its shortest physical axis and is perpendicular to the orbit plane;
*iii)* all the dissipative forces as well as perturbations due to other bodies are negligible (and therefore ignored).

Then, the equations of motion can be derived from the standard Euler's equations for a rigid body ([D]). Normalizing the period of revolution to $2\pi$, one obtains

$$\ddot{x} + \frac{3}{2}\,\frac{B-A}{C}\left(\frac{a}{r}\right)^3 \sin(2x-2f) \;=\; 0\,, \tag{9.5}$$

where $a$ is the semimajor axis of the Keplerian ellipse, $r$ the orbital "radius" (i.e. the distance between the centers of mass of $S$ and $P$), $f$ the true anomaly (i.e. the angle between the planet-satellite direction and the periapsis line) and $x$ is the angle between the longest axis of the ellipsoid and the periapsis line.

From assumption $i)$ and the theory of the two-body problem, it follows that the quantities $r$ and $f$ are periodic functions of the time, $r(t+2\pi)=r(t)$, $f(t+2\pi)=f(t)$, and that they are analytic functions of the orbital eccentricity $e$. Therefore expanding the second term of (9.5) in series, one obtains:

$$\ddot{x} + \mu \sum_{m\neq 0, m=-\infty}^{\infty} V(\frac{m}{2}, e)\, \sin(2x-mt) \;=\; 0\,, \tag{9.6}$$

where $\mu \equiv \frac{3}{2}\frac{B-A}{C}$ and the coefficients $V(\frac{m}{2}, e)$ are analytic functions of the eccentricity $e$:

$$V(\frac{m}{2}, e) \;\equiv\; e^{|m-2|} \sum_{k\geq 0} a_k e^{2k}\,,$$

for suitable $a_k \in \mathbb{R}$.

In writing (9.5) (or equivalently (9.6)) we have ignored the dissipative forces acting on the system. The major dissipative contribution is originated by the internal non-rigidity of the satellite and goes under the name of "tidal torque".

Having ignored such a force allow us to ignore *all* the quantities which are of comparable size. This leads us to consider equations of the form:

$$\ddot{x} + \mu \sum_{m\neq 0, m=N_1}^{N_2} W(\frac{m}{2}, e)\, \sin(2x-mt) \;=\; 0\,, \quad N_1, N_2 \in \mathbb{Z}\,, \tag{9.7}$$

where $W(\frac{m}{2}, e)$ are truncations, to a suitable order in the eccentricity, of the coefficients $V(\frac{m}{2}, e)$: of course, such truncations will depend upon the specific model at hand, on the *physical* values of $\mu$, $e$ and on the size of the observed tidal torque.

Under the above simplifications, we can investigate the stability of the system using the following argument.

The phase space $\mathcal{S}$ associated to (9.7)

$$\mathcal{S} \;=\; \{(y, x, t):\; y=\dot{x} \in \mathbb{R},\; (x,t)\in \mathbb{T}^2\}$$

has dimension three. Therefore bidimensional invariant surfaces $\mathcal{T}(\omega)$ divide the phase space into invariant compartments, with the property that any orbit starting in one of these regions would remain forever in it. Now, one can show that, in the parameter regions we shall consider below, the Poincarè map $(y, x) \to \phi^{2\pi}(y, x)$ [ $\phi^t(y, x) \equiv$ solution at time $t$ starting at $x(0)=x$, $\dot{x}(0)=y$] is a smooth "monotone twist map": $(\partial x'/\partial y) > 0$ (see [Mo5], [MK] for general informations). Invariant tori for (9.7) correspond to *invariant circles* for $\phi^{2\pi}$ and the Poincarè rotation number for such a circle coincide with the frequency $\omega$ associated to the invariant torus. It is not difficult to show (see, e.g., [Her1])



that if $\{z_i\}_{i=1,\ldots,q}$, $z_i \equiv (y_i, x_i)$, is a periodic orbit with rotation number $p/q$ and if $\Gamma_\omega$ is an invariant circle with rotation number $\omega >$ (resp. $<$) $p/q$, then $\Gamma_\omega$ lies above (resp. below) $\{z_i\}$, i.e., $\Gamma_\omega \cap \{x=x_i\} >$ (resp. $<$) $y_i$.

We make use of this property to trap a periodic orbit $\mathcal{P}(\frac{p}{q})$, associated to the $p:q$ resonance, between invariant surfaces $\mathcal{T}(\omega_1)$ and $\mathcal{T}(\omega_2)$, with $\omega_1 < \frac{p}{q} < \omega_2$. Obviously one is interested in taking $\omega_i$ as close as possible to $p/q$.

Therefore we select the two sequences of irrational frequencies

$$\Gamma_k^{(p/q)} \equiv \frac{p}{q} - \frac{1}{k+\omega_g}, \qquad \Delta_k^{(p/q)} \equiv \frac{p}{q} + \frac{1}{k+\omega_g}, \qquad k\in\mathbb{Z},\; k\geq 2$$

($\omega_g \equiv \frac{\sqrt{5}-1}{2}$), which approach $\frac{p}{q}$ from below and, respectively, above and satisfy the diophantine condition (Assumption 2.11) with the constant $\gamma \equiv \gamma_k = q^2(k+\omega_g)$.

Let us consider now the Moon-Earth system (Moon $\equiv S$, Earth $\equiv P$) and let us look at the synchronous resonance $\mathcal{P}(1/1)$, which human kind has been observing for quite a while.

According to our simplifications, we are led to study the Lagrangian

$$\begin{aligned}\mathcal{L}(y,x,t) \;\equiv\; & \frac{y^2}{2} + \mu\,[(-\frac{e}{4}+\frac{e^3}{32})\,\cos(2x-t) \\ & +(\frac{1}{2}-\frac{5}{4}e^2+\frac{13}{32}e^4)\,\cos(2x-2t) \;+\; (\frac{7}{4}e-\frac{123}{32}e^3)\,\cos(2x-3t) \\ & +(\frac{17}{4}e^2-\frac{115}{12}e^4)\,\cos(2x-4t) \;+\; (\frac{845}{96}e^3-\frac{32525}{1536}e^5)\,\cos(2x-5t) \\ & +\frac{533}{32}e^4\cos(2x-6t) \;+\; \frac{228347}{7680}e^5\,\cos(2x-7t)\;]\,,\end{aligned} \tag{9.8}$$

where the physical value of the perturbing parameter $\mu \equiv \frac{3}{2}\frac{B-A}{C}$ is $3.45\cdot 10^{-4}$ and that of the eccentricity $e$ is 0.0549. Using the techiques of §1 ÷ §9 (with the choice in §7.2) one can construct the surfaces $\mathcal{T}(\Gamma_k^{(1)})$ and $\mathcal{T}(\Delta_k^{(1)})$ for $k=2, 3, \ldots, 35$ and from the above discussion it then follows that the motion of the Moon, *as ruled by the approximate Lagrangian (9.8)*, will be forever trapped in the region enclosed by $\mathcal{T}(\Gamma_{35}^{(1)})$ and $\mathcal{T}(\Delta_{35}^{(1)})$, which, in turns, is shown to be a subset of $\{(y,x,t)\;:\;(x,t)\in\mathbb{T}^2\,,\,0.97\leq y\leq 1.03\}$.

Let us now consider the system Mercury-Sun, which is complicated by the fact that its eccentricity is relatively large: $e=0.2056$. Therefore we have to retain a larger number of terms in (9.7); in particular we consider the lagrangian function

$$\mathcal{L}(y,x,t) \;\equiv\; \frac{y^2}{2} + \frac{\mu}{2}\sum_{m\neq 0, m=-11}^{3} W(\frac{m}{2}, e)\cos(2x-mt)\,. \tag{9.9}$$

We are still able to conclude the stability of the 3:2 resonance, in which Mercury is actually observed for the astronomical value of the perturbing parameter, i.e. $\mu = 1.5\cdot 10^{-4}$. The tori closest to the periodic orbits in $\mathcal{P}(\frac{3}{2})$ are those with rotation numbers $\Gamma_{70}^{(3/2)}$ and $\Delta_{70}^{(3/2)}$; the outcoming trapping region is contained in $\{(y,x,t)\;:\;(x,t)\in\mathbb{T}^2\,,\,1.48\leq y\leq 1.52\}$.

## 10. Applications: Symplectic Maps

In this section we briefly discuss how the KAM techniques of §1 ÷ §8 can be adapted to deal with symplectic diffeomorphisms of plane regions.

Eventhough there is a tight connection (via Poincarè maps) between symplectic diffeomorphisms and Hamiltonian flows (cfr. [Do], [Mo6], [SZ]), it is interesting and often useful, to have *direct* formalisms and methods. The direct method discussed below was introduced in [CC2]; see also [CC4].

### 10.1. Formalism

Here we shall consider a special class of symplectic (area-preserving) twist diffeomorphisms of the cylinder $\mathscr{C} \equiv \mathbb{R} \times S^1$, $(S^1 \equiv \mathbb{T}^1 \equiv \mathbb{R}/2\pi\mathbb{Z})$, namely:

$$F : (y, x) \in \mathscr{C} \longmapsto (y', x') \equiv (y + f(x), x + y + f(x)) \in \mathscr{C}, \qquad \langle f \rangle = 0, \tag{10.1}$$

where $f$ is a real-analytic function on $S^1$ (i.e., a real-analytic function on $\mathbb{R}$ with period $2\pi$) with vanishing mean-value. As above, we shall also consider one-parameter families obtained by replacing $f$ with $\mu f$.

The word "twist" in the definition of the present model refers to the following property. If we look at the universal covering, $\mathbb{R}^2$, of the cylinder and consider a lift $\tilde{F}$ of $F$ (for example, replace, in (10.1), $\mathscr{C}$ with $\mathbb{R}^2$), then $\tilde{F}$ maps vertical lines $\{x = x_0\}$ into graphs of increasing functions of the $x$-variable; analytically: $(\partial x'/\partial y) > 0$. In our case, $(\partial x'/\partial y) = 1$.

The problem is to study the behaviour of the orbits $(y_n, x_n) \equiv F^n(y_0, x_0)$, where $F^n$ denotes $F$ composed with itself $n$ times.

The observation that, for (10.1), $y' = x' - x$, allows to eliminate the $y$-variable: $(y_n, x_n)$ is an $F$-orbit if and only if the sequence $\{x_n\}$ satisfies

$$x_{n+1} - 2x_n + x_{n-1} = f(x_n) ; \tag{10.2}$$

obviously, given a solution $\{x_n\}$ of (10.2) the associated $F$-orbit is simply $(x_n - x_{n-1}, x_n)$.

Analogously to Definition 2.1 we give the following

**Definition 10.1.** *A solution $\{x_n\}$ of (10.2) is called quasi-periodic with frequency $\omega \in \mathbb{R}$, if $\omega/2\pi$ is irrational and if there exists a continuous periodic function $u$: $\theta \in S^1 \to u(\theta) \in \mathbb{R}$, such that*

$$x_n \equiv \omega\, n + u(\omega n), \qquad (\text{mod } 2\pi). \tag{10.3}$$

*and, analogously to Definition 2.5:*

**Definition 10.2.** *We shall say that a quasi-periodic solution is non-degenerate if $\forall \theta \in S^1$*

$$(1 + u_\theta) \neq 0. \tag{10.4}$$

As for flows, non-degenerate quasi-periodic solutions correspond to invariant surfaces, which, in the present case, are *invariant circles*: the map

$$\theta \in S^1 \longmapsto \Big(\omega + u(\theta) - u(\theta - \omega), \theta + u(\theta)\Big) \tag{10.5}$$



yields a non-contractible embedding of $S^1$ into $\mathscr{C}$.

To require that $u(\theta)$ is a non-degenerate quasi-periodic solution of (10.2) with frequency $\omega$ is equivalent to require that $u$ satisfies the following non-linear finite-difference equation:

$$D^2 u = f(\theta + u(\theta)), \qquad (1 + u_\theta \neq 0), \tag{10.6}$$

where $D$, here, denotes the symmetrized finite-difference operator with step $\omega$

$$Du \equiv u(\theta + \frac{\omega}{2}) - u(\theta - \frac{\omega}{2}) \equiv u^+(\theta) - u^-(\theta). \tag{10.7}$$

The abuse of language in denoting with the same symbol different objects will be forgiven in view of the complete analogy of the present situation with the Lagrangian case.

### 10.2. The Newton Scheme, the Linearized Equation, etc.

The strategy of §2 can be carried out in this context and it is actually simpler because of the dimension and of the pecularity of the maps we are considering.

First notice that, since for any periodic function $g$ the $S^1$-average of $Dg$ vanishes, in dealing with the equation

$$Dg = h \tag{10.8}$$

one has to require that $\langle h \rangle = 0$, and in such a case, the unique solution with vanishing mean-value of (10.8) is given by:

$$g = D^{-1} h \equiv \sum_{n \in \mathbb{Z}, n \neq 0} \frac{\hat{h}_n}{2i \sin\left(\frac{n\omega}{2}\right)} e^{in\theta}, \qquad \left(h \equiv \sum_{n \in \mathbb{Z}, n \neq 0} \hat{h}_n e^{in\theta}\right), \tag{10.9}$$

and we see the reason for the irrationality of $\omega/2\pi$ in Definition 10.1.

However, $n\omega/2$ will come arbitrarily close to $0, \pi$ (mod $2\pi$) and we need a Diophantine asumption; from now on we shall assume that $\omega$ satisfies, for some $\gamma, \tau \geq 1$,

$$\left| \frac{\omega}{2\pi} n + m \right| \geq \frac{1}{\gamma |n|^\tau}, \qquad \forall n \in \mathbb{Z}\backslash\{0\}, \forall m \in \mathbb{Z}. \tag{10.10}$$

Let us now introduce the function spaces. Let $\Delta^1_\xi \equiv \{\theta \in \mathbb{C} : |\text{Im}\,\theta| \leq \xi\}$ and denote by $|\cdot|_\xi$ the sup norm over $\Delta^1_\xi$. Besides Lemma 3.2, which will be used with $d = 1$, we need the analogous of Lemma 3.3. For $l \geq 0$ we let $s_l(\delta)$ be an upper bound on the small-divisor series

$$\Big[ \sum_{n=1}^{\infty} \Big(\frac{n^l}{\sin \frac{n\omega}{2}}\Big)^2 e^{-\delta n} \Big]^{1/2} \leq s_l(\delta). \tag{10.11}$$

Then:

**Lemma 10.3.** *Let $h$ be a real-analytic function on $\Delta^1_\xi \times \mathscr{P}$, ($\mathscr{P}$ being a compact set of $\mathbb{C}$) and let $l \geq 1$. Then ($|\cdot|_{\xi,\mathscr{P}} \equiv \sup_{\Delta^1_\xi \times \mathscr{P}} |\cdot|$)*

$$|D^{-1} \partial^l_\theta h|_{\xi-\delta,\mathscr{P}} \leq s_l(2\delta)\, |h|_{\xi,\mathscr{P}}. \tag{10.12}$$

*The same estimate holds for $l = 0$ provided $h$ has vanishing mean value over $S^1$.*

An explicit estimate of $s_l(\rho)$ for $\omega$'s satisfying (10.10) is given in Appendix 9.

Let us now discuss the *Newton scheme.* As above, we shall call $v$ a *non-degenerate approximate solution* of (10.6) any real-analytic periodic function such that $1+v_\theta \neq 0$, and we shall associate to it its *error function* $\mathcal{E}(v) \equiv \varepsilon(\theta)$, given by

$$\mathcal{E}(v) \equiv \varepsilon(\theta) \equiv D^2 v - f(\theta+v) \,. \tag{10.13}$$

With the proficiency acquired in the more complicate case of §2, §3, the reader will have no trouble in checking the following Proposition (cfr. Proposition 2.13):

**Proposition 10.4.** *Let $\omega$ satisfy (10.10), let $v$ be a (non-degenerate) approximate solution of equation (10.6) and let $\varepsilon(\theta)$ be the associated error-function. Let*

$$\mathcal{M} \equiv 1 + v_\theta \,. \tag{10.14}$$

*Then $\langle \mathcal{M}\varepsilon\rangle = 0$ and if we set:*

$$w \equiv \mathcal{M}\Big\{D^{-1}[(\mathcal{M}^+\mathcal{M}^-)^{-1}(-D^{-1}(\mathcal{M}\varepsilon)+c_1)]+c_2\Big\} \,. \tag{10.15}$$

*with*

$$c_1 \equiv \frac{\langle(\mathcal{M}^+\mathcal{M}^-)^{-1}[-D^{-1}(\mathcal{M}\varepsilon)]\rangle}{\langle(\mathcal{M}^+\mathcal{M}^-)^{-1}\rangle} \,, \qquad c_2 \equiv \langle \mathcal{M}D^{-1}[(\mathcal{M}^+\mathcal{M}^-)^{-1}(-D^{-1}(\mathcal{M}\varepsilon)+c_1)]\rangle \,. \tag{10.16}$$

*and $v' \equiv v + w$, then $\langle w\rangle = 0$ and*

$$\mathcal{E}(v') \equiv \varepsilon' = \varepsilon_\theta \mathcal{M}^{-1} w - \{f(\theta+v') - f(\theta+v) - f_x(\theta+v)w\} \tag{10.17}$$

The estimates leading to the KAM algorithm and to the existence KAM theorem for the present situation, can be obtained, at this point, in a completely straightforward way; however, for completeness and convenience of the reader we shall collect the main estimates in Appendix 10.

## 10.3. Results

The above methods, together with the strategy outlined in §7.3, have been used in [CC4], to study the stability of various invariant circles for the following one-parameter families of twist maps:

$$F : (y,x) \in \mathcal{C} \longmapsto (y',x') \equiv (y+\mu f(x), x+y+\mu f(x)) \in \mathcal{C} \,, \tag{10.18}$$

with

$$\begin{aligned} &(M) \qquad f = \sin x \,, \qquad \text{(standard map)} \,, \\ &(M') \qquad f = \sin x + \frac{1}{50}\sin(5x) \,. \end{aligned} \tag{10.19}$$

The rotation numbers considered are:

$$\frac{\omega_1}{2\pi} \equiv \frac{\sqrt{5}-1}{2} \,, \qquad \frac{\omega_2}{2\pi} \equiv \frac{\sqrt{5}+5}{10} \,, \qquad \frac{\omega_3}{2\pi} \equiv \frac{\sqrt{2}}{2} \,. \tag{10.20}$$

Now, let $\Gamma_\mu(\omega_k)$, respectively $\Gamma'_\mu(\omega_k)$, be the invariant circles for $(M)$, resp. $(M')$. The stability results are summarized in the following table:



| *Curve* | $l_0$ | $\rho$ | $N$ | $\xi$ |
|---|---|---|---|---|
| $\Gamma_\mu(\omega_1)$ | 190 | 0.838 | 6 | $5.07\cdot 10^{-3}$ |
| $\Gamma_\mu(\omega_2)$ | 190 | 0.77 | 5 | $5.15\cdot 10^{-3}$ |
| $\Gamma_\mu(\omega_3)$ | 160 | 0.76 | 5 | $5.15\cdot 10^{-3}$ |
| $\Gamma'_\mu(\omega_1)$ | 60 | 0.4 | 7 | $5.03\cdot 10^{-3}$ |
| $\Gamma'_\mu(\omega_2)$ | 60 | 0.39 | 7 | $5.03\cdot 10^{-3}$ |

where $l_0$ is the order of the polynomial initial guess (cfr. §7.3) $v = \sum_{l=1}^{l_0} \bar{u}_l(\theta)\mu^l$, $\rho$ and $\xi$ measure the size of the analyticity domain of the solution $u$: $u(\theta;\mu)$ is real-analytic on $\Delta^1_\xi \times \{\mu \in \mathbb{C} : |\mu| \le \rho\}$, finally $N$ denotes the number of times the KAM algorithm has been used before applying the KAM theorem.

These results should be compared with the experimental prediction given by Greene's method discussed above;

| *Curve* | *Greene's threshold* |
|---|---|
| $\Gamma_\mu(\omega_1)$ | 0.9716 |
| $\Gamma_\mu(\omega_2)$ | 0.9044 − 0.9045 |
| $\Gamma_\mu(\omega_3)$ | 0.908 − 0.909 |
| $\Gamma'_\mu(\omega_1)$ | 0.6013 − 0.6014 |
| $\Gamma'_\mu(\omega_2)$ | 0.7213 − 0.7214 |

Hence, our theoretical results are in an agreement ranging within 86% ÷ 54 % . The reason for the more sensible discrepancy for the map $(M')$ seems to be related to the distribution of the $\mu$ singularities ($\theta$ almost everywhere in $S^1$) of the solution $u$. In particular there are numerical evidences that $\mu$-domain of analyticity of the solutions $u$ for the maps $(M)$ and $(M')$ (for the given rotation numbers) has (for almost every $\theta$) a natural boundary.

For the map $(M')$, it seems that $d_r > d_i$, if $d_{r/i}$ denotes the distance from the origin with the first real/purely immaginary singularity: see [BC], [BCCF].

## Appendix 1: Proof of (2.37)

We prove in this appendix the formula (2.37) of §2, i.e.

$$\mathcal{A} = D^{-1}(\mathcal{M}^T\varepsilon_\theta)^A \,, \tag{10.21}$$

where $\mathcal{A}$ is defined in (2.31) as

$$\mathcal{A} \equiv (\mathcal{M}^T \partial_\theta \mathcal{L}_y^T)^A$$

and the superscript $A$ denotes the antisymmetric part of a matrix, namely

$$\begin{aligned} D^{-1}(\mathcal{M}\varepsilon_\theta)^A &\equiv D^{-1}(\mathcal{M}^T\varepsilon_\theta) - D^{-1}(\varepsilon_\theta^T\mathcal{M}) \\ &= D^{-1}\Big[\mathcal{M}^T\varepsilon_\theta - \varepsilon_\theta^T\mathcal{M}\Big] \,. \end{aligned} \tag{10.22}$$

Taking the gradient with respect to $\theta$ of the definition (2.12) of the error function

$$\varepsilon(\theta,t) = D\mathcal{L}_y^T(\omega+Dv,\theta+v,t) - \mathcal{L}_x^T(\omega+Dv,\theta+v,t) \,, \tag{10.23}$$

one has

$$\varepsilon_\theta = D[\mathcal{L}_{yy}D\mathcal{M} + \mathcal{L}_{yx}\mathcal{M}] - \mathcal{L}_{xy}D\mathcal{M} - \mathcal{L}_{xx}\mathcal{M} . \tag{10.24}$$

Multiplying (10.24) by $\mathcal{M}^T$ and taking the antisymmetric part, one has:

$$\begin{aligned}\mathcal{M}^T\varepsilon_\theta - \varepsilon_\theta^T\mathcal{M} &= \mathcal{M}^T\, D[\mathcal{L}_{yy}D\mathcal{M} + \mathcal{L}_{yx}\mathcal{M}] - \mathcal{M}^T\mathcal{L}_{xy}D\mathcal{M}\\ &\quad - D[D\mathcal{M}^T\mathcal{L}_{yy} + \mathcal{M}^T\mathcal{L}_{xy}]\mathcal{M} + D\mathcal{M}^T\mathcal{L}_{yx}\mathcal{M}\\ &= D[\mathcal{M}^T\mathcal{L}_{yy}D\mathcal{M} + \mathcal{M}^T\mathcal{L}_{yx}\mathcal{M}] - D(D\mathcal{M}^T\mathcal{L}_{yy}\mathcal{M} + \mathcal{M}^T\mathcal{L}_{xy}\mathcal{M}) .\end{aligned}$$

Finally, recalling the definition of $\mathcal{A}$

$$\begin{aligned}\mathcal{A} &\equiv (\mathcal{M}^T\partial_\theta\mathcal{L}_y^T)^A \equiv \mathcal{M}^T\partial_\theta\mathcal{L}_y^T - (\partial_\theta\mathcal{L}_y^T)^T\mathcal{M}\\ &= \mathcal{M}^T\mathcal{L}_{yy}D\mathcal{M} + \mathcal{M}^T\mathcal{L}_{yx}\mathcal{M} - D\mathcal{M}^T\mathcal{L}_{yy}\mathcal{M} - \mathcal{M}^T\mathcal{L}_{xy}\mathcal{M} ,\end{aligned} \tag{10.25}$$

one obtains

$$\mathcal{M}^T\varepsilon_\theta - \varepsilon_\theta^T\mathcal{M} = D\mathcal{A} . \ \square$$

## Appendix 2: Proof of Lemma 3.2

In this appendix we prove the Lemma 3.2 of §3.2.

**Lemma:** *Let $h$ be an analytic map from $\Omega \times \mathcal{P} \to \mathbb{C}$, where $\Omega$ is a (smooth) domain in $\mathbb{C}^d$ and $\mathcal{P} \subset \mathbb{C}^k$ a space of parameters; then for any subdomain $\Omega' \subset \Omega$ with* $\mathrm{dist}(\Omega', \partial\Omega) \equiv \delta > 0$ *and for any multi index $m = (m_1, ..., m_d) \in \mathbb{N}^d$ one has*

$$\sup_{\Omega'\times\mathcal{P}} |\partial_z^m h| \equiv \sup_{\Omega'\times\mathcal{P}} \Big|\frac{\partial^{|m|}h}{\partial z_1^{m_1}...\partial z_d^{m_d}}\Big| \le m!\, \delta^{-|m|} \sup_{\Omega\times\mathcal{P}} |h| \tag{10.26}$$

($|m| = m_1 + ... + m_d$). *Moreover, if $h$ is an analytic map, $h : \Omega \times \mathcal{P} \to L^p(\mathbb{C}^d)$ for some $p \in \mathbb{N}$ ($L^0(\mathbb{C}^d) \equiv \mathbb{C}^d$), then $\forall l \in \mathbb{Z}_+$, $\partial_z^l h \in L^{p+l}(\mathbb{C}^d)$ and*

$$\sup_{\Omega'\times\mathcal{P}} |\partial_z^l h| \le l!\, \delta^{-l} \sup_{\Omega\times\mathcal{P}} |h| . \tag{10.27}$$

**Proof:** Consider first a holomorphic function $h_0 : \Omega \times \mathcal{P} \to \mathbb{C}$. Then, Cauchy's integral formula implies

$$\begin{aligned}&\sup_{\Omega'\times\mathcal{P}} \Big|\frac{\partial^{|m|}h_0}{\partial z_1^{m_1}...\partial z_d^{m_d}}\Big| =\\ &= \sup_{\Omega'\times\mathcal{P}} \Big| \frac{m!}{(2\pi i)^d} \oint_{|\zeta_1 - z_1| = \delta, ..., |\zeta_d - z_d| = \delta} \frac{h_0(\zeta_1, ..., \zeta_d)}{\prod_{k=1}^d (\zeta_k - z_k)^{m_k+1}}\, d\zeta_1 ... d\zeta_d \Big| \\ &\le m!\, \delta^{-|m|} \sup_{\Omega\times\mathcal{P}} |h_0| .\end{aligned} \tag{10.28}$$



Now, if $h : \Omega \times \mathcal{P} \to L^p(\mathbb{C}^d)$ for some $p \in \mathbb{N}$, then (10.28) implies

$$\begin{aligned}\sup_{\Omega'\times\mathcal{P}} |\partial_z^l h| &\equiv \sup_{|c_1| = ... = |c_l| = 1}\ \sup_{\Omega'\times\mathcal{P}} |\partial_z^l h\ c_1 ... c_l|\\ &\le \sup_{|c_1| = ... = |c_l| = 1}\ \sup_{\Omega'\times\mathcal{P}} \Big( |\partial_z^l h|\ |c_1| ... |c_l| \Big)\\ &\le \sup_{|c_1| = ... = |c_l| = 1} \Big( l!\, \delta^{-l} \sup_{\Omega\times\mathcal{P}} |h|\ |c_1| ... |c_l| \Big)\\ &= l!\, \delta^{-l} \sup_{\Omega\times\mathcal{P}} |h| . \ \square\end{aligned}$$

## Appendix 3: Proof of Lemma 3.3

This appendix is devoted to the proof of the Lemma 3.3 of §3.2.

**Lemma:** *Let $h = h(\theta, t; \mu)$ be a real-analytic map of $\Delta_\xi \times \mathcal{P}$ into $\mathcal{H}$, where $\mathcal{H}$ is either $\mathbb{C}$, or $\mathbb{C}^d$ or $L^p(\mathbb{C}^d)$ and let $l \ge 1$. Then*

$$|D^{-1}\partial_\theta^l h|_{\xi-\delta,\mathcal{P}} \le \sigma_l(2\delta)\ |h|_{\xi,\mathcal{P}} , \tag{10.29}$$

*where*

$$\sigma_l(\rho) = \Big[2^{d+1} \sum_{(n,m)\in\mathbb{Z}^{d+1}\setminus(0,0)} \Big(\frac{\|n\|^l}{\omega\cdot n + m}\Big)^2 e^{-\rho(|n|+|m|)}\Big]^{\frac{1}{2}} , \tag{10.30}$$

($\|n\| \equiv (\sum_{i=1}^d |n_i|^2)^{1/2}$, $|n| \equiv \sum_{i=1}^d |n_i|$). *Moreover the same estimate holds for $l = 0$ provided $h$ has vanishing mean value over $\mathbb{T}^{d+1}$. If $(\omega, 1)$ verifies Assumption 2.11 then*

$$\sigma_l(\rho) < K_l\gamma\delta^{-(\tau+l)} , \qquad K_l \equiv 2^{d+2-(\tau+l)}\sqrt{\Gamma(2(\tau+l)+1)} ,$$

*$\Gamma$ being Euler's gamma function.*

**Proof:** We prove first (10.29) for a holomorphic function $h_0 : \Delta_\xi \times \mathcal{P} \to \mathbb{C}$ with vanishing mean value. Denote by $\|\cdot\|_{\xi,\mathcal{P}}$ the $L^2$-norm

$$\|h_0\|^2_{\xi,\mathcal{P}} \equiv \sup_{\mathcal{P}}\ \sup_{\substack{|a_1|,...,|a_d| \le \xi\\ |b| \le \xi}} \int_{\mathbb{T}^{d+1}} |h_0(\theta + ia, t + ib)|^2 \frac{d\theta dt}{(2\pi)^{d+1}} .$$

Then, for any $\nu = (\nu_1, ..., \nu_d) \in \{-1, 1\}^d$, $\lambda \in \{-1, 1\}$, one has

$$\sup_{\mathcal{P}} \sum_{(n,m)} e^{2(n\cdot\nu + m\lambda)\xi}\ |\hat h_{0_{(n,m)}}|^2 \le \|h_0\|^2_{\xi,\mathcal{P}} . \tag{10.31}$$

To prove (10.31), let $\xi' < \xi$ and consider the function

$$h_0' \equiv h_0(\theta - i\nu\xi', t - i\lambda\xi') .$$

By Cauchy's theorem we have:

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

Then, Parseval's identity yields

$$\sum |\hat{h}_{0_{(n,m)}}|^2 \, e^{2\xi'(n\cdot\nu+m\lambda)} \;=\; \int_{\mathbf{T}^{d+1}} |h_0'|^2 \, \frac{d\theta dt}{(2\pi)^{d+1}} \;\le\; \|h_0\|^2_{\xi,\mathcal{P}} \, .$$

Taking the supremum over $\xi' < \xi$ one obtains (10.31). From the maximum principle, Schwarz inequality, Assumption 2.11 and (10.31), it follows (dropping the index 0)

$$\begin{aligned}
|\, \partial^l_{\theta_j} D^{-1} h \,|_{\xi-\delta,\mathcal{P}} &= |\sum_{(n,m)\neq 0} \hat{h}_{(n,m)} \, \frac{n_j^l}{(\omega\cdot n+m)} \, e^{i(n\cdot\theta+mt)}|_{\xi-\delta,\mathcal{P}} \\
&= \sup_{\mathcal{P}} \sup_{(\nu,\lambda)\in[-1,1]^{d+1}} |\sum_{(n,m)\neq 0} \hat{h}_{(n,m)} \, \frac{n_j^l}{(\omega\cdot n+m)} \, e^{(n\cdot\nu+m\lambda)(\xi-\delta)}| \\
&\le \sup_{\mathcal{P}} \sum_{(n,m)\neq 0} |\hat{h}_{(n,m)}| \Big(\sum_{\lambda,\nu} e^{2(n\cdot\nu+m\lambda)\xi}\Big)^{1/2} e^{-\delta(|n|+|m|)} \, \frac{|n_j|^l}{|\omega\cdot n+m|} \qquad (10.32) \\
&\le \sigma_l(2\delta) \, \sup_{\mathcal{P}} \Big( \frac{1}{2^{d+1}} \sum_{(n,m)} |\hat{h}_{(n,m)}|^2 \sum_{\lambda,\nu} e^{2(n\cdot\nu+m\lambda)\xi} \Big)^{1/2} \\
&\le \sigma_l(2\delta) \, \|h\|_{\xi,\mathcal{P}} \;\le\; \sigma_l(2\delta) \, |h|_{\xi,\mathcal{P}} \, .
\end{aligned}$$

Now, let $h : \Delta_\xi \times \mathcal{P} \to \mathbb{C}^d$ and $l = 0$. Then, by (10.32)

$$\begin{aligned}
|D^{-1}h|_{\xi-\delta,\mathcal{P}} &\equiv \sum_i |D^{-1}h_i|_{\xi-\delta,\mathcal{P}} \\
&\le \sigma_0(2\delta) \sum_i |h_i|_{\xi,\mathcal{P}} \;\equiv\; \sigma_0(2\delta) \, |h|_{\xi,\mathcal{P}} \, .
\end{aligned}$$

If $l \ge 1$, then for $c_1 \in \mathbb{C}^d, \dots, c_l \in \mathbb{C}^d$,

$$\begin{aligned}
|\partial^l_\theta \, D^{-1}h|_{\xi-\delta,\mathcal{P}} &\equiv \sup_{|c_1|=1,\dots,|c_l|=1} |D^{-1} \, \partial^l_\theta h \; c_1 \dots c_l|_{\xi-\delta,\mathcal{P}} \\
&\le \sup_{|c_1|=1,\dots,|c_l|=1} |\partial^l_\theta \, D^{-1}h|_{\xi-\delta,\mathcal{P}} \; |c_1| \dots |c_l| \\
&\le \sup_{|c_1|=1,\dots,|c_l|=1} \sigma_l(2\delta) \; |h|_{\xi,\mathcal{P}} \; |c_1| \dots |c_l| \\
&= \sigma_l(2\delta) \; |h|_{\xi,\mathcal{P}} \, .
\end{aligned}$$

Finally, if $h : \Delta_\xi \times \mathcal{P} \to L^p(\mathbb{C}^d)$, applying again (10.32) one has

$$\begin{aligned}
|\partial^l_\theta \, D^{-1} \, h|_{\xi-\delta,\mathcal{P}} &= \sup_{|c_1|=1,\dots,|c_l|=1} |D^{-1} \, \partial^l_\theta h \; c_1 \dots c_l \;|_{\xi-\delta,\mathcal{P}} \\
&\le \sigma_l(2\delta) \; |h|_{\xi,\mathcal{P}} \, .
\end{aligned}$$

Now we want to show that if $(\omega, 1)$ verifies Assumption 2.11 then

$$\sigma_l(\rho) \;<\; K_l \gamma \delta^{-(\tau+l)} \, , \qquad K_l \equiv 2^{d+2-(\tau+l)} \sqrt{\Gamma(2(\tau+l)+1)} \, , \qquad (10.33)$$

where $\Gamma$ is the Euler's gamma function. Assume for $\omega$ the diophantine condition

$$|\omega\cdot n+m|^{-1} \;<\; \gamma |n|^\tau \, .$$



The solution $g$ of the equation $Dg = h$ is given by

$$g \;=\; D^{-1}h \;=\; \sum_{(n,m)\in\mathbf{Z}^{d+1}\setminus(0,0)} \frac{\hat{h}_{(n,m)} \; e^{i(n\cdot\theta+mt)}}{i(n\cdot\omega+m)} \; ;$$

therefore

$$\partial^l_{\theta_i} g(\theta,t) \;=\; \sum_{(n,m)\in\mathbf{Z}^{d+1}\setminus(0,0)} \frac{n_i^l}{n\cdot\omega+m} \; \hat{h}_{(n,m)} \; e^{i(n\cdot\theta+mt)} \, .$$

From the inequality (see [R3] p.180, formula (9.4)),

$$\sum_{n\in\mathbf{Z}^d, m\in\mathbf{Z}} |\hat{h}_{(n,m)}|^2 \; e^{2\xi(|n|+|m|)} \;<\; 2^{d+1} |h|^2_{\xi,\mathcal{P}} \, , \qquad (10.34)$$

the term $\partial^l_{\theta_i} g(\theta,t)$ can be estimated using Schwarz's inequality as

$$\begin{aligned}
|\partial^l_{\theta_i} g(\theta,t)| &\le \sum_{(n,m)} |\frac{n_i^l}{n\cdot\omega+m} \hat{h}_{(n,m)} \; e^{i(n\cdot\theta+mt)}| \\
&\le \sum_{(n,m)} |\frac{n_i^l}{n\cdot\omega+m}| e^{-\delta(|n|+|m|)} \; |\hat{h}_{(n,m)} \;| \; e^{\xi(|n|+|m|)} \\
&\le \sqrt{\sum_{(n,m)} |\hat{h}_{(n,m)} \;|^2 \; e^{2\xi(|n|+|m|)}} \; \sqrt{\sum_{(n,m)} |\frac{n_i^l}{n\cdot\omega+m}|^2 \; e^{-2\delta(|n|+|m|)}} \\
&\le 2^{\frac{d+1}{2}} \; \|h\| \; \sqrt{\Psi(\delta)} \;<\; 2^{\frac{d+1}{2}} \; |h| \; \sqrt{\Psi(\delta)} \, ,
\end{aligned}$$

where $\Psi(\delta) \equiv \sum_{(n,m)} |\frac{n_i^l}{n\cdot\omega+m}|^2 \; e^{-2\delta(|n|+|m|)}$. Let us estimate $\Psi(\delta)$ as follows.

$$\Psi(\delta) \le \sum_{k=1}^\infty \; \sum_{|n|+|m|=k} |\frac{n_i^l}{n\cdot\omega+m}|^2 \; e^{-2\delta k} \;\le\; \sum_{k=1}^\infty \Big( \sum_{|n|+|m|=k} |\frac{1}{n\cdot\omega+m}|^2 \Big) \; k^{2l} e^{-2\delta k} \, .$$

Finally defining $b_0 = 0$, $b_k = \sum_{0<|n|+|m|\le k} |\frac{1}{n\cdot\omega+m}|^2$, one has:

$$\begin{aligned}
\Psi(\delta) = \sum_{k=1}^\infty \, (b_k - b_{k-1}) k^{2l} \; e^{-2\delta k} \;&=\; \sum_{k=1}^\infty \; k^{2l} b_k e^{-2\delta k} - \sum_{k=1}^\infty (k+1)^{2l} b_k e^{-2\delta(k+1)} \\
&\le\; (1-e^{-2\delta}) \cdot 2^{d+3} \; \sum_{k=1}^\infty \frac{k^{2l} e^{-2\delta k}}{D_k^2} \, ,
\end{aligned}$$

where in the last inequality we used $b_k = \sum_{0<|n|+|m|\le k} |\frac{1}{n\cdot\omega+m}|^2 \;\le\; \frac{2^{d+3}}{D_k^2}$, where

$$D_k \;=\; \min_{0<|n|+|m|\le k} \; |n\cdot\omega+m|$$

(see the proof in [R2], [R4] suitably adapted to be valid with the actual choice of the norms). Since

$$D_k \ge \min_{0<|n|+|m|\le k} \; (\frac{1}{\gamma |n|^\tau}) = \frac{1}{\gamma k^\tau}$$

or $\frac{1}{D_k} \le \gamma k^\tau$, it follows

$$\Psi(\delta) \le (1-e^{-2\delta})\, 2^{d+3}\gamma^2 \sum_{k=1}^{\infty} k^{2l+2\tau}\, e^{-2\delta k}\,.$$

Let us estimate the sum as follows. Let $s = 2\tau$ and $\beta = 2\delta$:

$$\begin{aligned}\sum_{k=1}^{\infty} k^s e^{-\beta k} &= \sum_{k=1}^{\infty}\int_k^{k+1} k^s \beta e^{-\beta x}\,\frac{dx}{1-e^{-\beta}} \\ &= \frac{\beta}{1-e^{-\beta}}\sum_{k=1}^{\infty}\int_k^{k+1} k^s e^{-\beta x} dx \le \frac{\beta}{1-e^{-\beta}}\sum_{k=1}^{\infty}\int_k^{k+1} x^s e^{-\beta x} dx \\ &\le \frac{\beta^{-s}}{1-e^{-\beta}}\int_0^{\infty} y^s e^{-y} dy = \frac{\beta^{-s}}{1-e^{-\beta}}\,\Gamma(s+1)\,.\end{aligned}$$

Therefore,

$$\Psi(\delta) \le 2^{d+3-(2\tau+2l)}\gamma^2\delta^{-(2\tau+2l)}\,\Gamma(2(\tau+l)+1)$$

and finally

$$|\partial^l_{\theta,t} g(\theta,t)| \le 2^{d+2-(\tau+l)}\gamma\, |h|_{\xi,\mathcal{G}}\, \delta^{-(\tau+l)}\sqrt{\Gamma(2(\tau+l)+1)}\,. \ \square$$

## Appendix 4: Proof of Lemma 3.5

This appendix is devoted to the proof of the Lemma 3.5 of §3.4.

**Lemma:** *Let $\mathcal{T} \equiv \mathcal{M}^T \mathcal{L}_{yy}\mathcal{M}$ be the twist matrix of a non degenerate approximate solution $v$ and let $M$, $\overline{M}$, $L$, $\overline{L}$ denote upper bounds on (respectively) $|\mathcal{M}|_\xi$, $|\mathcal{M}^{-1}|_\xi$, $|\mathcal{L}_{yy}|_{\mathcal{D}\xi}$, $|\mathcal{L}_{yy}^{-1}|_{\mathcal{D}\xi}$. Then*

(i) $\overline{M}^{-2}\overline{L}^{-1} \le |\mathcal{T}|_0 \le |\mathcal{T}|_\xi \le M^2 L$

(ii) $M^{-2}L^{-1} \le |\mathcal{T}^{-1}|_0 \le |\mathcal{T}^{-1}|_\xi \le \overline{M}^2\overline{L}$

(iii) $|\langle\mathcal{T}^{-1}\rangle^{-1}|_0 \le |\mathcal{T}|_0 \qquad \left(\langle\cdot\rangle \equiv \int_{\mathbf{T}^{d+1}} \cdot \frac{d\theta dt}{(2\pi)^{d+1}}\right).$

**Proof:**

*(i)* By $|\mathcal{T}| = |\mathcal{T}^T|$, $|\mathcal{T}^{-1}| \ge |\mathcal{T}|^{-1}$ (since $1 = |\mathcal{T}\mathcal{T}^{-1}| \le |\mathcal{T}||\mathcal{T}^{-1}|$) and the positivity of the matrix $\mathcal{T}$ one has:

$$|\mathcal{T}| \le M^2 L$$

and

$$|\mathcal{T}|^{-1} \le |\mathcal{T}^{-1}| \le |(\mathcal{T}^T)^{-1}||(\mathcal{L}_{yy})^{-1}||\mathcal{T}^{-1}| \le \overline{M}^2\overline{L}$$

or

$$|\mathcal{T}| \ge \overline{M}^{-2}\overline{L}^{-1}\,.$$



*(ii)* The inequality $|\mathcal{T}^{-1}| \le \overline{M}^2\overline{L}$ has been proven in *(i)*. Moreover, $|\mathcal{T}| \le M^2 L$ implies

$$|\mathcal{T}^{-1}| \ge |\mathcal{T}|^{-1} \ge M^{-2}L^{-1}\,.$$

*(iii)* $\mathcal{T} > 0$ (i.e. $\mathcal{T} \ge 0$ and $\mathcal{T}$ invertible) implies $|\langle\mathcal{T}^{-1}\rangle^{-1}| \le \langle|\mathcal{T}|\rangle$.
$\mathcal{T} > 0$ follows from the following four general facts:
(*a*) $A = A^*$ (with $A^* = \overline{A}^T$) $\Rightarrow$ $A \le |A|$ ;
(*b*) $a \in \mathbb{R}_+$, $A \ge 0$, $A \le a$ $\Rightarrow$ $|A| \le a$ ;
(*c*) $a \in \mathbb{R}_+$, $A \ge 0$, $A \ge a$ $\Rightarrow$ $A^{-1} \le a^{-1}$ ;
(*d*) $A, B \ge 0$, $[A,B] \equiv AB - BA = 0$ $\Rightarrow$ $AB \ge 0$ .
Proof of (*a*), (*b*), (*c*), (*d*):
(*a*) $(Ax,x) \le |Ax||x| \le |A||x|^2 = (|A|x,x)$
(*b*) $A^* = A \Rightarrow |A| = \sup_{|x|=1}(Ax,x) \le a$
(*c*) $(Ax,x) \ge a(x,x) \Leftrightarrow (A^{1/2}x, A^{1/2}x) \ge a(x,x)$ ; setting $A^{1/2}x = y$ one has

$$a^{-1}(y,y) \ge (A^{-1/2}y, A^{-1/2}y) = (A^{-1}y,y)$$

(*d*) $[A,B] = 0 \Rightarrow [A, f(B)] = 0 \quad \forall f$ continuous; thus,

$$(ABx,x) = (AB^{1/2}B^{1/2}x,x) = (B^{1/2}AB^{1/2}x,x) = (AB^{1/2}x, B^{1/2}x) \ge 0\,.$$

By (*a*) $|\mathcal{T}| - \mathcal{T} \ge 0$ and since $|\mathcal{T}| - \mathcal{T}$ commutes with $\mathcal{T}^{-1}$, while by (*d*) $(|\mathcal{T}| - \mathcal{T})\mathcal{T}^{-1} \ge 0$ i.e. $|\mathcal{T}|\mathcal{T}^{-1} \ge \mathbb{1}$ and by averaging $\mathbb{1} \le \langle|\mathcal{T}|\rangle\langle\mathcal{T}^{-1}\rangle$ or, by (*c*), $\langle\mathcal{T}^{-1}\rangle^{-1} \le \langle|\mathcal{T}|\rangle$; finally (*b*) yields $\mathcal{T} > 0$. $\square$

## Appendix 5: Proof of Proposition 3.6

In this appendix we provide the details of the proof of Proposition 3.6 of §3.5. In particular we want to bound the error function

$$\varepsilon' \equiv q_1 + q_2 + q_3\,,$$

where $q_1$, $q_2$, $q_3$ have been defined in (2.47).
From (3.45) we have

$$|q_1|_{\xi'} \le |Dq_1^{(y)}|_{\xi'} + |q_1^{(x)}|_{\xi'}\,.$$

The term $|q_1^{(x)}|_{\xi'}$ can be bounded as in (3.53) by

$$|q_1^{(x)}|_{\xi'} \le \frac{L_3}{2}\,E^2\overline{L}^2 a^2(c+1)^2\,.$$

Now let

$$f = f(\omega+Dv, \theta+v, t) = \mathcal{L}_y^T(\omega+Dv, \theta+v, t)\,, \qquad f^+ = f(\omega+Dv+Dw, \theta+v+w, t)\,.$$

as in (3.54); then from (3.55)

$$
\begin{aligned}
|Dq_1^{(y)}|_{\xi'} &\equiv |D\Big[f^+ - f - f_y Dw - f_x w\Big]|_{\xi'} \\
&\equiv |\omega \cdot \partial_\theta(f^+ - f - f_y Dw - f_x w) + \partial_t(f^+ - f - f_y Dw - f_x w)|_{\xi'} \\
&\le |\partial_\theta(f^+ - f - f_y Dw - f_x w)\ \omega|_{\xi'} + |\partial_t(f^+ - f - f_y Dw - f_x w)|_{\xi'} \\
&\equiv A_1 + A_2 \ ,
\end{aligned}
\tag{10.35}
$$

where

$$
\begin{aligned}
A_1 &= |\partial_\theta(f^+ - f - f_y Dw - f_x w)\ \omega|_{\xi'} \\
&= |\partial_\theta\Big\{\int_0^1 (1-\beta)\Big[f_{yy} Dw Dw + f_{yx} Dw\ w + f_{xy} w Dw + f_{xx} ww\Big] d\beta\Big\}\ \omega|_{\xi'}
\end{aligned}
\tag{10.36}
$$

and

$$
A_2 = |\partial_t\Big\{\int_0^1 (1-\beta)\Big[f_{yy} Dw Dw + f_{yx} Dw\ w + f_{xy} w Dw + f_{xx} ww\Big] d\beta\Big\}|_{\xi'}
$$

(notice that the derivatives of $f$ are evaluated at $(\omega + Dv + \beta Dw, \theta + v + \beta w, t)$).

Let us start with the estimate of $A_1$. By (10.36) we have:

$$
\begin{aligned}
A_1 \le \Big\{\int_0^1 & (1-\beta)\Big\{\Big[|f_{yyy}|\ (|D\mathcal{M}| + \beta|Dw_\theta|) + |f_{yyx}|\ (|\mathcal{M}| + \beta|w_\theta|)\Big]|Dw|^2 \\
& + 2|f_{yy}||Dw_\theta||Dw| + 2\Big[|f_{yxy}|\ (|D\mathcal{M}| + \beta|Dw_\theta|) \\
& + |f_{yxx}|\ (|\mathcal{M}| + \beta|w_\theta|)\Big]|Dw||w| \\
& + 2|f_{yx}|\ (|Dw_\theta||w| + |Dw||w_\theta|) + \Big[|f_{xxy}|\ (|D\mathcal{M}| + \beta|Dw_\theta|) \\
& + |f_{xxx}|\ (|\mathcal{M}| + \beta|w_\theta|)\Big]|w|^2 + 2|f_{xx}|\ |w_\theta||w|\Big\} d\beta\Big\}\ |\omega| \\
\le \Big\{\frac{1}{2}\ & \Big[|f_{yyy}||D\mathcal{M}| + M|f_{yyx}|\Big]|Dw|^2 + |f_{yy}||Dw||Dw_\theta| + \Big[|f_{yxy}||D\mathcal{M}| \\
& + |f_{yxx}|M\Big]|w||Dw| + |f_{yx}|\Big[|w||Dw_\theta| + |w_\theta||Dw|\Big] \\
& + \frac{1}{2}\Big[|f_{xxy}||D\mathcal{M}| + |f_{xxx}|M\Big]|w|^2 + |f_{xx}||w||w_\theta| \\
& + \frac{1}{6}\Big[|f_{yyy}||Dw_\theta| + |f_{yyx}||w_\theta|\Big]|Dw|^2 + \frac{1}{3}\Big[|f_{yxy}||Dw_\theta| + |f_{yxx}||w_\theta|\Big]|w||Dw| \\
& + \frac{1}{6}\Big[|f_{xxy}||Dw_\theta| + |f_{xxx}||w_\theta|\Big]|w|^2\Big\}\ |\omega| \ .
\end{aligned}
$$

Finally, denoting by

$$
\begin{aligned}
\lambda_3' &\equiv \max\ \{\Omega^2|\mathcal{L}_{yyy}|,\ \Omega|\mathcal{L}_{yyx}|,\ |\mathcal{L}_{yxx}|\}\ , \\
\lambda_4 &\equiv \max\ \{\Omega^3|\mathcal{L}_{yyyy}|,\ \Omega^2|\mathcal{L}_{yyyx}|,\ \Omega^2|\mathcal{L}_{yyxy}|,\ \Omega|\mathcal{L}_{yyxx}|,\ \Omega|\mathcal{L}_{yxxy}|,\ |\mathcal{L}_{yxxx}|\}\ ,
\end{aligned}
$$



one has:

$$
\begin{aligned}
A_1 \ \le \ \Big\{& \frac{\lambda_4}{2} E^2\overline{L}^2 a^2 (c+1)^2\ \Big[(1+\delta^{-1})M + \frac{4}{3} E\overline{L} a \delta^{-2}(g + \frac{\delta}{4} + \frac{\delta^2}{4}\frac{s_1(\delta)}{s_0(\delta)}\frac{b_1}{b})\Big] \\
& + 4\lambda_3' E^2\overline{L}^2 a^2 \delta^{-2}\Big(g + \frac{\delta}{4} + \frac{\delta^2}{4}\frac{s_1(\delta)}{s_0(\delta)}\frac{b_1}{b}\Big)(c+1)\Big\}|\omega| \ .
\end{aligned}
$$

Analogously let us estimate the second term $A_2$ as follows:

$$
\begin{aligned}
A_2 \le \int_0^1 & (1-\beta)\Big\{\Big[|f_{yyy}|\ (|Dv_t| + \beta|Dw_t|) + |f_{yyx}|\ (|v_t| + \beta|w_t|)\Big]|Dw|^2 \\
& + 2|f_{yy}|\ |Dw_t||Dw| + 2\Big[|f_{yxy}|\ (|Dv_t| + \beta|Dw_t|) \\
& + |f_{yxx}|\ (|v_t| + \beta|w_t|)\Big]|Dw||w| \\
& + 2|f_{yx}|\ \Big[|Dw_t||w| + |Dw||w_t|\Big] + \Big[|f_{xxy}|\ (|Dv_t| + \beta|Dw_t|) \\
& + |f_{xxx}|\ (|v_t| + \beta|w_t|)\Big]|w|^2 + 2|f_{xx}|\ |w_t||w| \\
& + |f_{yyt}|\ |Dw|^2 + 2|f_{yxt}|\ |w||Dw| + |f_{xxt}|\ |w|^2\Big\} d\beta \\
\le \frac{1}{2}\ & \Big[|f_{yyy}||Dv_t| + |f_{yyx}||v_t|\Big]|Dw|^2 + |f_{yy}||Dw||Dw_t| \\
& + \Big[|f_{yxy}||Dv_t| + |f_{yxx}||v_t|\Big]|w||Dw| + |f_{yx}|\Big[|w||Dw_t| + |w_t||Dw|\Big] \\
& + \frac{1}{2}\Big[|f_{xxy}||Dv_t| + |f_{xxx}||v_t|\Big]|w|^2 + |f_{xx}||w||w_t| \\
& + \frac{1}{6}\Big[|f_{yyy}||Dw_t| + |f_{yyx}||w_t|\Big]|Dw|^2 \\
& + \frac{1}{3}\Big[|f_{yxy}||Dw_t| + |f_{yxx}||w_t|\Big]|w||Dw| \\
& + \frac{1}{6}\Big[|f_{xxy}||Dw_t| + |f_{xxx}||w_t|\Big]|w|^2 \\
& + \frac{1}{2}|f_{yyt}||Dw|^2 + |f_{yxt}||w||Dw| + \frac{1}{2}|f_{xxt}||w|^2 \ .
\end{aligned}
$$

We recall that if $S$ is an upper bound on $|v_t|_{\xi,\rho}$, then

$$
\begin{aligned}
|w_t|_{\xi'} &\le E\overline{L}a\ \ \Big(\delta^{-1} + \frac{s_1(\delta)}{s_0(\delta)}\frac{b_1}{b}\Big)\ , \\
|Dw|_{\xi'} &\le E\overline{L}ac\Omega\ , \\
|w|_{\xi'} &\le E\overline{L}a\ , \\
|Dv_t|_{\xi'} &\le S\delta^{-1}\Omega\ , \\
|Dw_t|_{\xi'} &\le 4E\overline{L}\delta^{-2}ag\Omega\ .
\end{aligned}
$$

Denoting by

$$
L_4' \ \equiv \ \max\{\Omega^2|\mathcal{L}_{yyyt}|,\ \Omega|\mathcal{L}_{yyxt}|,\ |\mathcal{L}_{yxxt}|\ \}\ ,
$$

one has

$$A_2 \le \frac{\lambda_4}{2}E^2\overline{L}^2a^2(c+1)^2\left[(1+\delta^{-1})S+\frac{4}{3}E\overline{L}a\delta^{-2}(g+\frac{\delta}{4}+\frac{\delta^2}{4}\frac{s_1(\delta)}{s_0(\delta)}\frac{b_1}{b})\right]$$
$$+4\lambda_3'E^2\overline{L}^2a^2\delta^{-2}\Big(g+\frac{\delta}{4}+\frac{\delta^2}{4}\frac{s_1(\delta)}{s_0(\delta)}\frac{b_1}{b}\Big)(c+1)$$
$$+\frac{L_4'}{2}E^2\overline{L}^2a^2(c+1)^2 .$$

Therefore

$$A_1+A_2\le\frac{E^2\overline{L}^2a^2}{2}(c+1)^2\Big\{\lambda_4(1+\delta^{-1})(S+M\Omega)+\frac{4}{3}\lambda_4\Omega E\overline{L}a\delta^{-2}\cdot$$
$$\cdot(g+\frac{\delta}{4}+\frac{\delta^2}{4}\frac{s_1(\delta)}{s_0(\delta)}\frac{b_1}{b})+L_4'\Big\}+4\lambda_3'\Omega E^2\overline{L}^2a^2\delta^{-2}(c+1)\cdot$$
$$\cdot\Big(g+\frac{\delta}{4}+\frac{\delta^2}{4}\frac{s_1(\delta)}{s_0(\delta)}\frac{b_1}{b}\Big) .$$

Let now

$$L_3'\equiv\max(\Omega\lambda_3') , \qquad L_4\equiv\max(\Omega\lambda_4) ;$$

then one has

$$|q_1|_{\xi'}\le\frac{E^2\overline{L}^2a^2}{2}(c+1)^2\Big[L_3+L_4(1+\delta^{-1})(S+M)+\frac{4}{3}L_4\,E\overline{L}a\delta^{-2}\Big(g+\frac{\delta}{4}$$
$$+\frac{\delta^2}{4}\frac{s_1(\delta)}{s_0(\delta)}\frac{b_1}{b}\Big)+L_4'\Big]+4L_3'E^2\overline{L}^2a^2\delta^{-2}(c+1)\Big(g+\frac{\delta}{4}+\frac{\delta^2}{4}\frac{s_1(\delta)}{s_0(\delta)}\frac{b_1}{b}\Big) .$$

The term $|q_2|_{\xi'}$ can be bounded as in (3.44) by

$$|q_2|_{\xi'}\ \le\ E^2\overline{L}\ \frac{a}{M\delta} .$$

Finally, from

$$|Dz|_{\xi'}\ \le\ \frac{E\overline{L}a'}{M} ,$$

one finds

$$|q_3|_{\xi'}\ \le|(\mathcal{M}^T)^{-1}|_{\xi'}\cdot|D^{-1}(\mathcal{M}^T\varepsilon_\theta-\varepsilon_\theta^T\mathcal{M})|_{\xi'}\cdot|Dz|_{\xi'}$$
$$\le 4E^2\overline{L}\ a'\delta^{-1}\overline{M}s_0(\delta) .$$

Collecting the above estimates, one has:

$$|\varepsilon|_{\xi-\delta}\ \equiv\ E^2\overline{L}a\ \Big\{\frac{\delta^{-1}}{M}+\frac{a}{2}(c+1)^2\overline{L}\ \Big[L_3+L_4(S+M)(1+\delta^{-1})+\frac{4}{3}E\overline{L}a\delta^{-2}L_4\cdot$$
$$\cdot\Big(g+\frac{\delta}{4}+\frac{\delta^2}{4}\frac{s_1(\delta)}{s_0(\delta)}\frac{b_1}{b}\Big)+L_4'\Big]+4L_3'\overline{L}\ a(c+1)\delta^{-2}\ \Big(g+\frac{\delta}{4}+\frac{\delta^2}{4}\frac{s_1(\delta)}{s_0(\delta)}\frac{b_1}{b}\Big)$$
$$+\ \chi_d\cdot 4\delta^{-1}\overline{M}s_0(\delta)\ \frac{a'}{a}\ \Big\} ,$$



where

$$\begin{cases}\chi_d=0 & d=1\\ \chi_d=1 & d\ge 2,\end{cases}$$

$$a\equiv(M\overline{M})^2s_0(\delta)^2b , \qquad b_1\equiv 1+(M\overline{M})^2L\overline{L}\ \frac{s_0(2\xi)}{s_0(\delta)} ,$$

$$b\equiv b_1+M\ \Big(\frac{s_0(\xi)}{s_0(\delta)}\Big)^2\ \Big[1+(M\overline{M})^2L\overline{L}\ \frac{s_0(2\xi)}{s_0(\xi)}\Big] , \qquad c\equiv\delta^{-1}+\frac{a'}{\Omega\, a} ,$$

$$a'\equiv(M\overline{M})^2s_0(2\delta)b_1' , \qquad b_1'\equiv 1+(M\overline{M})^2L\overline{L}\ \frac{s_0(2\xi)}{s_0(2\delta)} ,$$

$$a''\equiv(M\overline{M})^2s_0(\delta)b_1 , \qquad g\equiv 1+\frac{\delta}{4}\frac{a'}{a}+\frac{\delta}{4}s_1(\delta)\frac{a''}{a}+\frac{\delta}{2}\frac{a''}{a} .\ \square$$

## Appendix 6: Proof of (5.14)

In this appendix we want to prove the inequalities (5.14) of §5, i.e.

$$\begin{aligned}E_{i+1}\overline{L}\ &\le\ (E_i\overline{L})^2\beta_0{\gamma_0}^i\\ W_i\ &\le\ E_i\overline{L}\beta_1{\gamma_1}^i\\ W_{1i}\ &\le\ E_i\overline{L}\beta_2{\gamma_2}^i\\ |Dw^{(i)}|\ &\le\ E_i\overline{L}\beta_3{\gamma_3}^i\ \Omega\end{aligned} \tag{10.37}$$

with

$$\beta_0\ \equiv\ 8\cdot 208^2M^2(M\overline{M})^8(L\overline{L})^2(S+M)K_1K_0^3\lambda\gamma^4 2^{8\tau}\xi^{-4\tau-3} , \qquad \gamma_0\ \equiv\ 2^{4\tau+3}$$
$$\beta_1\ \equiv\ 13\cdot M(M\overline{M})^4L\overline{L}\ K_0^2\gamma^2 2^{4\tau+1}\xi^{-2\tau} , \qquad \gamma_1\ \equiv\ 2^{2\tau}$$
$$\beta_2\ \equiv\ 165\cdot M(M\overline{M})^4L\overline{L}\ K_1K_0\gamma^2 2^{4\tau}\xi^{-2\tau-1} , \qquad \gamma_2\ \equiv\ 2^{2\tau+1}$$
$$\beta_3\ \equiv\ 130\cdot M(M\overline{M})^4L\overline{L}\ K_0^2\gamma^2 2^{4\tau}\xi^{-2\tau-1} , \qquad \gamma_3\ \equiv\ 2^{2\tau+1} ,$$

where $\lambda\equiv\max(\overline{L}L_3,\overline{L}L_3',\overline{L}L_4,\overline{L}L_4',1)$.

Let us start to estimate $w^{(i)}$. To this end we need an upper bound on the two quantities $a_i$ and $b_i$, where

$$a_i\equiv(M_i\overline{M}_i)^2s_0(\delta_i)^2b_i ,$$

and

$$b_i\equiv b_{1i}+M_i\Big(\frac{s_0(\xi_i)}{s_0(\delta_i)}\Big)^2\ \Big[1+(M_i\overline{M}_i)^2L\overline{L}\frac{s_0(2\xi_i)}{s_0(\xi_i)}\Big] .$$
$$b_{1i}\equiv 1+(M_i\overline{M}_i)^2L\overline{L}\ \frac{s_0(2\xi_i)}{s_0(\delta_i)} .$$

From the definition of $M_i$, $\overline{M}_i$ we easily obtain

$$M_i\,\overline{M}_i\ \le\ 3M\overline{M} ;$$

moreover, from the estimates:

$$\begin{aligned}
\frac{s_0(2\xi_i)}{s_0(\delta_i)} &= (\frac{\delta_i}{2\xi_i})^\tau < (\frac{1}{8})^\tau < \frac{1}{8}\\
\Big(\frac{s_0(\xi_i)}{s_0(\delta_i)}\Big)^2 &= (\frac{1}{4^\tau})^2 < \frac{1}{16}\\
\frac{s_0(\xi_i)s_0(2\xi_i)}{s_0(\delta_i)^2} &= \frac{1}{2^\tau}\ (\frac{\delta_i}{\xi_i})^{2\tau} < (\frac{1}{32})^\tau\ ,
\end{aligned}$$

one has (notice that here we need to indicate the subscript $i$):

$$\begin{aligned}
b_i &= 1 + (M_i\overline{M}_i)^2 L\overline{L}\Big(\frac{s_0(2\xi_i)}{s_0(\delta_i)}\Big) + M_i\ \Big(\frac{s_0(\xi_i)}{s_0(\delta_i)}\Big)^2\ \Big[1 + (M_i\overline{M}_i)^2 L\overline{L}\frac{s_0(2\xi_i)}{s_0(\xi_i)}\Big]\\
&< 1 + 9(M\overline{M})^2 L\overline{L}\ (\frac{1}{8})^\tau + 2M\ (\frac{1}{4})^{2\tau}\ \Big[1 + 9(M\overline{M})^2 L\overline{L}\frac{1}{2^\tau}\Big]
\end{aligned}$$

and therefore

$$\begin{aligned}
a_i &= (M_i\overline{M}_i)^2 s_0(\delta_i)^2 b_i\\
&< 9\ (M\overline{M})^2 K_0^2\gamma^2\Big(\frac{2^{i+2}}{\xi}\Big)^{2\tau}\ \Big\{1 + 9(M\overline{M})^2 L\overline{L}\ (\frac{1}{8})^\tau + 2M(\frac{1}{4})^{2\tau}\\
&\qquad \cdot\Big[1 + \frac{9}{2^\tau}\ (M\overline{M})^2 L\overline{L}\Big]\Big\}\\
&< 9\ (M\overline{M})^4 K_0^2\gamma^2\ 2^{4\tau+2\tau i}\xi^{-2\tau}\ \Big\{1 + \frac{9}{8}L\overline{L} + \frac{2M}{16}\Big(1 + \frac{9}{2}L\overline{L}\Big)\Big\}\\
&< 26\ M(M\overline{M})^4(L\overline{L})K_0^2\gamma^2\ 2^{4\tau}2^{2\tau i}\xi^{-2\tau}\ .
\end{aligned}$$

Finally, we obtain

$$W_i \equiv E_i\overline{L}a_i \le\ E_i\overline{L}\beta_1{\gamma_1}^i\ .$$

with

$$\beta_1 \equiv 13\ M(M\overline{M})^4 L\overline{L}K_0^2\gamma^2\ 2^{4\tau+1}\xi^{-2\tau}\ ,\qquad \gamma_1 = 2^{2\tau}\ .$$

Next we estimate $w_\theta^{(i)}$. Since $\frac{b_{1i}}{b_i} \le 1$ and $\frac{K_0}{K_1} \le \frac{1}{\sqrt{3}}$, one has:

$$\begin{aligned}
W_{1i}\ &\le\ E_i\overline{L}a_i\ \Big(\delta_i^{-1} + \frac{s_1(\delta_i)}{s_0(\delta_i)}\Big)\\
&\le\ E_i\overline{L}\cdot 26M(M\overline{M})^4 K_0^2\gamma^2 2^{4\tau}2^{2\tau i}\xi^{-2\tau}(L\overline{L})\Big(\frac{2^{i+2}}{\xi} + \frac{K_1}{K_0}\frac{2^{i+2}}{\xi}\Big)\\
&\le\ E_i\overline{L}\cdot 165(M\overline{M})^4 MK_1K_0\gamma^2 2^{4\tau}L\overline{L}\ 2^{(2\tau+1)i}\xi^{-2\tau-1}\ ,
\end{aligned}$$

namely,

$$W_{i1}\ \le\ E_i\overline{L}\ \beta_2{\gamma_2}^i\ .$$

with

$$\beta_2 \equiv 165\cdot M(M\overline{M})^4 K_1K_0\gamma^2\ 2^{4\tau}\xi^{-2\tau-1}\ L\overline{L}\ ,\qquad \gamma_2 \equiv 2^{2\tau+1}\ .$$



From $\frac{a_i'}{a_i} < \frac{1}{4}$, we have:

$$\begin{aligned}
|Dw^{(i)}|\ &\le\ E_i\overline{L}a_i\Omega c_i\ \equiv\ E_i\overline{L}a_i\Omega\ \Big(\delta_i^{-1} + \frac{a_i'}{a_i\Omega}\Big)\\
&\le\ E_i\overline{L}\cdot 130M(M\overline{M})^4(L\overline{L})K_0^2\gamma^2\ 2^{4\tau}2^i2^{2\tau i}\xi^{-2\tau-1}\Omega\\
&=\ E_i\overline{L}\ \Omega\ \beta_3{\gamma_3}^i\ .
\end{aligned}$$

with

$$\beta_3 \equiv 130\ M(M\overline{M})^4(L\overline{L})K_0^2\gamma^2\ 2^{4\tau}\xi^{-2\tau-1}\ ,\qquad \gamma_3 \equiv 2^{2\tau+1}\ .$$

We finally come to the estimate of $E_{i+1}$. First we can bound $a_i$ from below as

$$\begin{aligned}
|a_i| &= 9\ (M\overline{M})^2 K_0^2\gamma^2\ \Big(\frac{2^{i+2}}{\xi}\Big)^{2\tau}\ \Big\{1 + \frac{9}{8^\tau}(M\overline{M})^2 L\overline{L}\\
&\qquad + \frac{2M}{4^{2\tau}}\ \Big(1 + \frac{9}{2^\tau}(M\overline{M})^2 L\overline{L}\Big)\Big\}\\
&>\ 132^2\cdot 2^{2\tau i}\xi^{-2\tau}(M\overline{M})^2\ .
\end{aligned}$$

Therefore $a_i$ is bounded from above and below as

$$132^2\cdot 2^{2\tau i}\xi^{-2\tau}(M\overline{M})^2\ < |a_i| <\ 26\cdot 2^{2\tau i}2^{4\tau}\xi^{-2\tau}K_0^2\gamma^2(L\overline{L})M(M\overline{M})^4\ .$$

Before proceding we need also the following estimates, which can be easily obtained from the definition of the various quantities:

$$\frac{b_1'}{b} < 2^\tau\ ,\qquad \frac{a'}{a} = \frac{s_0(2\delta)}{s_0(\delta)^2}\frac{b_1'}{b} < 1\ ,\qquad \frac{a''}{a} < \frac{\delta^\tau}{K_0\gamma}\ ,\qquad g < \frac{K_1}{K_0}\ .$$

and

$$\begin{aligned}
&\Omega\gamma \ge 1,\ \ \frac{b_1}{b} \le 1,\ \ \delta_i \equiv \frac{\xi}{2^{i+2}} < \frac{1}{4},\ \ \delta_i^\tau < \frac{1}{4}\\
&\gamma > 2,\qquad \tau \ge 1,\qquad K_0 \ge 11\cdot 2^{-\tau}\ .
\end{aligned}$$

Moreover from the definition of $K_0$ and $K_1$ (see (3.17)) we find

$$K_1 = \frac{K_0}{2}\sqrt{\frac{\Gamma(2\tau+3)}{\Gamma(2\tau+1)}}\ =\ \frac{K_0}{2}\ \sqrt{(2\tau+2)(2\tau+1)}\ ,\qquad K_1 \ge \sqrt{3}K_0\ .$$

Therefore, denoting by $A_i$ an upper bound on the norm of $a_i$ and by $\lambda \equiv \max(\overline{L}L_3, \overline{L}L_3', \overline{L}L_4, \overline{L}L_4', 1)$, we obtain

$$
\begin{aligned}
E_{i+1} &< E_i^2 \overline{L} A_i^2 \lambda \Big\{ \frac{\delta_i^{-1}}{A_i M_i} + \frac{4\overline{M}_i \delta_i^{-1}}{A_i} + \frac{1}{2}\Big(\delta_i^{-1} + \frac{s_0(2\delta_i)}{s_0(\delta_i)^2}\frac{b'_{1i}}{b_i \Omega} + 1\Big)^2 \\
&\quad \Big[1 + (S_i + M_i)(1 + \delta_i^{-1}) + \frac{4}{3} E_i \overline{L} A_i \delta_i^{-2}\Big(\frac{\delta_i}{4} + g_i + \frac{\delta_i^2}{4}\frac{s_1(\delta_i)}{s_0(\delta_i)}\frac{b_{1i}}{b_i}\Big) + 1\Big] \\
&\quad + 4\,\delta_i^{-2}\Big(\delta_i^{-1} + \frac{s_0(2\delta_i)}{s_0(\delta_i)^2}\frac{b'_{1i}}{b_i\Omega} + 1\Big)\Big(\frac{\delta_i}{4} + \frac{\delta_i^2}{4}\frac{s_1(\delta_i)}{s_0(\delta_i)}\frac{b_{1i}}{b_i} + g_i\Big)\Big\} \\
&< E_i^2 \overline{L} A_i^2 \lambda \Big\{ \frac{2^{i+2}}{132^2\, 2^{2\tau i} M_i (M_i \overline{M}_i)^2} + \frac{4\,\overline{M}_i 2^{i+2}}{132^2\, 2^{2\tau i}(M_i\overline{M}_i)^2} \\
&\quad + \frac{1}{2}\delta_i^{-3}\Big(1 + \frac{\xi}{K_0 \gamma \Omega\, 2^{i+2}} + \frac{\xi}{2^{i+2}}\Big)^2 \\
&\quad \Big[\frac{\xi}{2^i} + (S_i + M_i)(1 + \frac{\xi}{2^{i+2}}) + \frac{4}{3}\delta_i^{-1} E_i \overline{L} A_i \Big(\frac{\delta_i}{4} + \frac{K_1}{K_0} + \frac{\delta_i}{4}\frac{K_1}{K_0}\Big) + 1\Big] \\
&\quad + 4\delta_i^{-3}\Big(1 + \frac{\xi}{K_0\gamma\Omega 2^{i+2}} + \frac{\xi}{2^{i+2}}\Big)\Big(\frac{\delta_i}{4} + \frac{K_1}{K_0} + \frac{\delta_i}{4}\frac{K_1}{K_0}\Big)\Big\} \\
&< E_i^2 \overline{L} A_i^2 \lambda\, \frac{K_1}{K_0}\delta_i^{-3}(S+M)\Big\{\frac{4}{2^6 132^2\sqrt{3}} + \frac{16}{2^6 132^2 \sqrt{3}} + \frac{1}{2\sqrt{3}}(1 + \frac{1}{20} + \frac{1}{4})^2 \cdot \\
&\quad \cdot \Big[2 + 2(1 + \frac{1}{4}) + \frac{4}{3}E_i \overline{L} A_i \delta_i^{-1}(\frac{1}{16\sqrt{3}} + 1 + \frac{1}{16})\Big] \\
&\quad + 4(1 + \frac{1}{20} + \frac{1}{4})(\frac{1}{2^4\sqrt{3}} + 1 + \frac{1}{2^4})\Big\} \\
&< E_i^2 \overline{L} A_i^2 \lambda\, \frac{K_1}{K_0}\delta_i^{-3}(S+M)\Big[\frac{198}{25} + \frac{8}{11}E_i\overline{L}A_i\delta_i^{-1}\Big] \\
&< E_i^2 \overline{L} \cdot 26^2 M^2 (M\overline{M})^8 (L\overline{L})^2 (S+M) K_1 K_0^3 \lambda \gamma^4\, 2^{8\tau} 2^{4\tau i} \xi^{-4\tau-3} 2^{3i} 2^6 \\
&\quad \cdot \Big[\frac{198}{25} + \frac{8}{11}E_i \overline{L} A_i \delta_i^{-1}\Big].
\end{aligned}
$$

Defining

$$\beta_0 \equiv 8 \cdot 208^2 M^2 (M\overline{M})^8 (L\overline{L})^2 (S+M) K_1 K_0^3 \lambda \gamma^4\, 2^{8\tau} \xi^{-4\tau-3}\,, \qquad \gamma_0 \equiv 2^{4\tau+3}\,,$$

one has:

$$E_i \overline{L} \ < \ \frac{(E\overline{L}\beta_0\gamma_0)^{2^i}}{\beta_0 {\gamma_0}^{i+1}}$$

and by the hypothesis

$$E\overline{L}\beta_0\gamma_0 \ < \ 1\,,$$

one has

$$E_i\overline{L} \ < \ \frac{1}{\beta_0{\gamma_0}^{i+1}}\,.$$

Therefore,

$$
\begin{aligned}
E_{i+1} &< 208^2 \cdot 8M^2(M\overline{M})^8(L\overline{L})^2(S+M)K_1K_0^3\lambda\gamma^4\, 2^{8\tau}2^{(4\tau+3)i}\xi^{-4\tau-3}\,(E_i^2\overline{L}) \\
&< E_i^2\overline{L}\,\beta_0{\gamma_0}^i\,. \quad \square
\end{aligned}
$$



## Appendix 7: Up and Down of Real Numbers

Upper and lower bounds on the result of elementary operations can be obtained increasing or decreasing by one bit the last bit of the mantissa, with an eventual propagation of the carry.

In the Fortran function listed below it is shown how to obtain upper bounds on real numbers in G-floating representation.

The real number $r$, represented by 64-bits, is initially decomposed in 4 bytes (each one of 8 bits) labelled $kp(1)$, ..., $kp(4)$ by the Fortran "EQUIVALENCE" statement. Degenerate cases (i.e. bytes of all 0's or 1's) are treated properly.

```
Double precision function Up(r)
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
```

The lower bound of a number $s$ is obtained simply using the function Up as

$$\text{Down} = -\text{Up}(-s) .$$

## Appendix 8: Computation of the Diophantine Constant

In this appendix we prove that the golden mean

$$\omega = \omega_g \equiv \frac{\sqrt{5}-1}{2}$$

satisfies the diophantine inequality

$$|\omega - \frac{p}{q}| \geq \frac{1}{\gamma q^2} , \qquad \forall p, q \in \mathbb{Z},\ q \neq 0 , \tag{10.38}$$

with a constant

$$\gamma \equiv \frac{3+\sqrt{5}}{2} . \tag{10.39}$$

Let us review some properties of continued fractions (see [Kh]).

Let $\omega$ be a positive irrational number and let $[a_0; a_1, a_2, ...]$, $a_k \in \mathbf{N}$, its continued fraction expansion, namely

$$\omega \equiv a_0 + \frac{1}{a_1 + \frac{1}{a_2 + ...}} ;$$

let

$$\frac{p_k}{q_k} \equiv [a_0; a_1, ..., a_k] , \qquad r_k \equiv [a_k; a_{k+1}, ...] .$$

Then, the following relations hold (see [Kh]):

$$p_k = p_{k-1} a_k + p_{k-2} , \quad q_k = q_{k-1} a_k + q_{k-2} ,$$

for any $k \geq 1$, where $p_{-1} \equiv 1$, $q_{-1} \equiv 0$, $p_0 \equiv a_0$, $q_0 \equiv 1$;

$$p_{k-1} q_{k-2} - p_{k-2} q_{k-1} = (-1)^k \quad \forall k \geq 1 ; \quad \frac{p_{2k}}{q_{2k}} \nearrow \omega \swarrow \frac{p_{2k+1}}{q_{2k+1}} ; \tag{10.40}$$

$$\frac{1}{q_k(q_k + q_{k+1})} < |\omega - \frac{p_k}{q_k}| < \frac{1}{q_k q_{k+1}} ; \quad \omega = \frac{r_k p_{k-1} + p_{k-2}}{r_k q_{k-1} + q_{k-2}} . \tag{10.41}$$

**Lemma 1:** Let $\Phi : [1, \infty) \to [1, \infty)$ be a continuous non decreasing function. Then from the inequality

$$|\omega q_k - p_k| \geq \frac{1}{\Phi(q_k)} , \qquad \forall k \geq 0 ,$$



it follows

$$|\omega q - p| \geq \frac{1}{\Phi(q)} , \qquad \forall q \neq 0 .$$

**Proof:** If $\frac{p}{q} = \frac{p_k}{q_k}$ for some $k$ then there is nothing to prove. Hence, assume that $\frac{p}{q} \neq \frac{p_k}{q_k}$ $\forall k \geq 0$. Then three cases are possible:

$$(i)\ \frac{p}{q} < \frac{p_0}{q_0} \equiv a_0 ,$$
$$(ii)\ \frac{p}{q} > \frac{p_1}{q_1} ,$$
$$(iii)\ \frac{p}{q} \in I_k ,$$

where $I_k \equiv (\frac{p_{k+1}}{q_{k+1}}, \frac{p_{k-1}}{q_{k-1}})$ for $k$ odd and $I_k \equiv (\frac{p_{k-1}}{q_{k-1}}, \frac{p_{k+1}}{q_{k+1}})$ for $k$ even.

In case $(i)$:

$$|\omega q - p| \geq |\omega - \frac{p}{q}| > |\omega - a_0| = |\omega - \frac{p_0}{q_0}| \geq \frac{1}{\Phi(q_0)} = \frac{1}{\Phi(1)} \geq \frac{1}{\Phi(q)} .$$

In case $(ii)$:

$$|\frac{p}{q} - \omega| > |\frac{p}{q} - \frac{p_1}{q_1}| \geq \frac{1}{q q_1} \quad \Rightarrow \quad |p - \omega q| > \frac{1}{q_1} = \frac{1}{a_1} ;$$

since $|\omega - a_0| \leq \frac{1}{a_1}$, one has

$$|\omega q - p| > |\omega - p_0| = |\omega q_0 - p_0| \geq \frac{1}{\Phi(q_0)} \geq \frac{1}{\Phi(q)} .$$

In case $(iii)$, by (10.40):

$$\frac{1}{q q_{k-1}} \leq |\frac{p}{q} - \frac{p_{k-1}}{q_{k-1}}||\frac{p_{k+1}}{q_{k+1}} - \frac{p_{k-1}}{q_{k-1}}| < |\frac{p_k}{q_k} - \frac{p_{k-1}}{q_{k-1}}| = \frac{1}{q_k q_{k-1}} \quad \Rightarrow q > q_k .$$

Again by (10.40),

$$|\omega - \frac{p}{q}| > |\frac{p_{k+1}}{q_{k+1}} - \frac{p}{q}| \geq \frac{1}{q q_{k+1}} \qquad \Rightarrow |\omega q - p| \geq \frac{1}{q_{k+1}} ,$$

but, by (10.41) it is $|\omega q_k - p_k| \leq \frac{1}{q_{k+1}}$ and since $q > q_k$ Lemma 1 follows. □

**Lemma 2:** For all $k \geq 0$

$$|\omega - \frac{p_k}{q_k}| = \frac{1}{\sigma_k q_k^2}$$

with $\sigma_k \equiv r_{k+1} + \frac{q_{k-1}}{q_k}$.

**Proof:** By (10.41) and (10.40)

$$|\omega - \frac{p_k}{q_k}| \equiv |\frac{r_{k+1} p_k + p_{k-1}}{r_{k+1} q_k + q_{k-1}} - \frac{p_k}{q_k}| \equiv \frac{1}{q_k (r_{k+1} q_k + q_{k-1})} =$$
$$= \frac{1}{q_k^2} \frac{1}{(r_{k+1} + \frac{q_{k-1}}{q_k})} \equiv \frac{1}{q_k^2 \sigma_k} .$$

By Lemma 1 one has to check (10.38) for $(p, q) = (p_k, q_k)$ and by Lemma 2 we can take $\gamma = \sup_{k\geq 0} \sigma_k$. Since

$$\frac{\sqrt{5}-1}{2} = [0; 1, 1, 1, ...] \equiv [0; 1^\infty] \ ,$$

one finds

$$r_{k+1} = \frac{\sqrt{5}+1}{2} \qquad (\forall k \geq 0) \ .$$

Finally, from

$$\frac{q_{-1}}{q_0} = 0 \ , \quad \frac{q_0}{q_1} = 1 \ , \quad \frac{q_k}{q_{k+1}} < 1 \quad (\forall k \geq 1)$$

(10.38) and (10.39) follow. Notice that one may have better estimates using the identity

$$\frac{q_k}{q_{k-1}} = [a_k; a_{k-1}, ..., a_1] \ , \qquad \forall k \geq 1 \ . \ \square$$

## Appendix 9: Small-Divisor Series for Symplectic Maps

In this appendix we provide an upper bound, $s_l(\delta)$, on the small-divisor series,

$$[\sum_{n=1}^{\infty} (\frac{n^l}{\sin\frac{n\omega}{2}})^2 \ e^{-\delta n}]^{1/2} \ , \qquad l = 0, 1 \ , \tag{10.42}$$

arising in the theory of symplectic maps (cfr. §10).

We shall prove that, for any integers $N, l \geq 0$, the sum in (10.42) is bounded by:

$$s_l(\delta) \equiv [\sum_{n=1}^{N-1} (\frac{n^l}{\sin\frac{n\omega}{2}})^2 \ e^{-\delta n} \ + \ S_l^{(N)}]^{1/2} \ , \qquad N \in \mathbb{N} \ , \tag{10.43}$$

with

$$S_l^{(N)} \equiv (1 - e^{-\delta}) \ \frac{\pi^2 C^2}{12} \sum_{n=N}^{\infty} n^{2l+2} e^{-\delta n} \ ; \tag{10.44}$$

(for $N = 0$, the first sum in (10.43) is absent). For $l = 0, 1$ one can bound $S_l^{(N)}$ with:

$$\begin{aligned} S_0^{(N)} &\leq \frac{\pi^2 C^2}{4} (1 - e^{-\delta}) \ e^{\frac{\delta}{2}} e^{-\alpha(N-1)} \frac{1}{\alpha^3} \ [2 + (2N+1)\alpha + N^2\alpha^2] \ , \\ S_1^{(N)} &\leq \frac{\pi^2 C^2}{4} (1 - e^{-\delta}) \ e^{\frac{\delta}{2}} e^{-\alpha(N-1)} \frac{1}{\alpha^5} \ [24 + (24N+36)\alpha + (12N^2+24N+14)\alpha^2 \\ &\quad + (4N^3 + 6N^2 + 4N + 1)\alpha^3 + N^4\alpha^4] \ . \end{aligned} \tag{10.45}$$

where $\alpha \equiv \delta(1+\omega)$.



**Proof:** Let $b_n \equiv \sum_{N\leq k\leq n} \frac{1}{\sin^2(\frac{k\omega}{2})}$ for $n \geq N$ and $b_{N-1} = 0$; then, since $b_n - b_{n-1} = \frac{1}{\sin^2(\frac{n\omega}{2})}$, it follows that

$$\begin{aligned} \sum_{n=N}^{\infty} (\frac{n^l}{\sin(\frac{n\omega}{2})})^2 e^{-\delta n} &= \sum_{n=N}^{\infty} n^{2l} e^{-\delta n} (b_n - b_{n-1}) = \\ &= \sum_{n=N}^{\infty} n^{2l} e^{-\delta n} \ b_n \ - \ \sum_{n=N}^{\infty} (n+1)^{2l} e^{-\delta n} \ e^{-\delta} b_n \ \leq \\ &\leq (1 - e^{-\delta}) \ \sum_{n=N}^{\infty} n^{2l} e^{-\delta n} b_n \ \leq \\ &\leq (1 - e^{-\delta}) \ \frac{\pi^2 C^2}{12} \ \sum_{n=N}^{\infty} n^{2l+2} e^{-\delta n} \ , \end{aligned}$$

where in the last inequality we used Rüssmann's estimate (cfr [R2]):

$$\begin{aligned} |b_n| &= \sum_{N\leq k\leq n} \frac{1}{\sin^2(\frac{k\omega}{2})} \ \leq \ \sum_{1\leq k\leq n} \frac{1}{4 \min_{l\in\mathbb{Z}} |\frac{\omega}{2\pi} k - l|} \ \leq \\ &\leq \frac{\pi^2 C^2}{12} \ n^2 \ . \end{aligned}$$

Now using that for any $k \geq 0$

$$\sum_{n=N}^{\infty} n^k e^{-\delta n} = (-1)^k \ \frac{d^k}{d\delta^k} \frac{e^{-\delta N}}{1 - e^{-\delta}}$$

and the estimate

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

one obtains the claim. $\square$

## Appendix 10: KAM Statements for Symplectic Maps

Here we formulate the KAM algorithm and theorem for symplectic maps (see §10). The proofs are easily obtained by mimicking the arguments leading to §4 and §5 (and makes a good exercise); alternatively we refer the (tired) reader to [CC4] and [CC2].

### KAM Algorithm for the Maps in (10.1)

The notations are as in §10; the style as in §4. Let $v \equiv v^{(0)}$ be a non-degenerate approximate solution of (10.6), let $\varepsilon(\theta) \equiv \varepsilon^{(0)}(\theta)$ be the associated error-function (see (10.13)); let $v, \varepsilon$ be real-analytic in $\Delta_\xi^1$ (or possibly in $\Delta_\xi^1 \times \mathcal{P}$) and let $\{\delta_j\}$ be a sequence of positive numbers such that $\sum_{j\geq 0}^{\infty} \delta_j < \xi_0 \equiv \xi$. Let $v^{(j)}, \varepsilon^{(j)}$,... be the functions constructed by iteratively applying Lemma 10.3, provided, of course, one has the needed control on $(1 + v^{(j-1)})^{-1}$.

We define, now, the norm-parameters and the KAM algorithm. Let $\mathcal{T}^\xi \equiv \{x = \theta + v(\theta) \ | \ \theta \in \Delta_\xi^1\}$ and $\mathcal{D}_\rho^\xi \equiv \{x = x_0 + x_1 \ | \ x_0 \in \mathcal{T}^\xi, \ x_1 \in \mathbb{C}, |x_1| \leq \rho\}$. Let $j \geq 0$

and let

$$\mathcal{N}_j \equiv \{M_j, \overline{M}_j, V_j, V_{1j}, E_j, \rho_j, F_{2j}\} \tag{10.46}$$

be the set of positive numbers controlling the norms of $v^{(j)}$ on $\Delta^1_{\xi_j}$ and of $f$ on $\mathfrak{D}^{\xi}_{\rho_j}$ [i.e., $M_j \geq |1 + v_\theta^{(j)}|_{\xi_j} \equiv |1 + v_\theta^{(j)}|_j$, $\overline{M}_j \geq |(1 + v_\theta^{(j)})^{-1}|_{\xi_j}$, $V_j \geq |v^{(j)}|_{\xi_j}$, $V_{1j} \geq |v_\theta^{(j)}|_{\xi_j}$, $E_j \geq |\mathcal{E}(v^{(j)})|_{\xi_j}$, $F_{2j} \geq |f_{xx}|_{\mathfrak{D}^{\xi}_{\rho_j}} \equiv |f_{xx}|_j$ ]. Then, if $s_l(\delta_j)$ is an upper bound on the small-divisor series (see Appendix 9), the above norm-parameter can be defined as follows:

$$a_j \equiv (M_j \overline{M}_j s_0(\delta_j))^2 \{1 + (M_j \overline{M}_j)^2 \frac{s_0(2\xi_j)}{s_0(\delta_j)}\} ,$$

where $\xi_0 = \xi$ and (for $j \geq 1$) $\xi_j = \xi_{j-1} - \delta_{j-1}$,

$$W_j \equiv E_j a_j$$

and

$$W_{1j} \equiv E_j a_j \left(\frac{V_{1j}}{M_j} \delta_j^{-1} + \frac{s_1(\delta_j)}{s_0(\delta_j)}\right) .$$

Then one can take

$$M_{j+1} \equiv M_0 + \sum_{i=0}^{j} W_{1i} ,$$

$$\overline{M}_{j+1} \equiv \begin{cases} \overline{M}_j \cdot (1 - \overline{M}_j \sum_{i=0}^{j} W_{1i})^{-1} & \text{if } \sum_{i=0}^{j} W_{1i} < 1 \\ \infty & \text{if } \sum_{i=0}^{j} W_{1i} \geq 1 , \end{cases}$$

$$V_{j+1} \equiv V_0 + \sum_{i=0}^{j} W_i ,$$

$$V_{1(j+1)} \equiv V_{10} + \sum_{i=0}^{j} W_{1i}$$

and

$$E_{j+1} = (E_j)^2 \, a_j \left(\frac{a_j F_{2(j+1)}}{2} + \frac{\delta_j^{-1}}{M_j}\right) ,$$

where

$$F_2^{(j+1)} \equiv \sup_{\mathfrak{D}^{\xi}_{\rho_{j+1}}} | f_{xx} | , \qquad \rho_0 \equiv 0, \qquad \rho_{j+1} \equiv \sum_{i=0}^{j} W_i .$$

Finally, the smallness condition (i.e., the analog of (5.1) in Theorem 5.1 ), which has been used in [CC4] to obtain the results discussed in §10 (for rotation numbers with



$\tau = 1$) is:

$$154 \cdot 10^{13} \, C^5 M^2 (M\overline{M})^{21/2} \xi^{-8} \, F_2 \, E \leq 1 \tag{10.47}$$

where $F_2 \equiv \max\{1, \bar{F}_2\}$ , $\bar{F}_2 \equiv \sup_{\mathfrak{D}^{\xi}_r} |f_{xx}|$, $r = 1/67^4$.

Actually, the condition (10.47), which was deduced in [CC2] and used in [CC4], could be slightly improved using the techiques presented in this work.

## References


[A1] Arnold V.I., *Proof of a Theorem by A.N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian,* Russ. Math. Surveys **18**, 9 (1963)

[A2] Arnold V.I. (ed.), *Encyclopedia of Math. Sciences,* Dynamical System III, Springer-Verlag **3** (1988)

[AL] Aubry S., Le Daeron P.Y., *The discrete Frenkel-Kontorova model and its extensions I,* Physica **8D**, 381 (1983)

[BCCF] Berretti A., Celletti A., Chierchia L., Falcolini C., *Analytic properties of conjugacy of invariant curves,* J. of Stat. Physics **66**, 1613 (1992)

[BC] Berretti A., Chierchia L., *On the complex analytic structure of the golden invariant curve for the standard map,* Nonlinearity **3**, 39 (1990)

[B] Bost J.-B., *Tores invariants des systèmes dynamiques Hamiltoniens,* Séminaire Bourbaki **639**, 113 (1984-85)

[C1] Celletti A., *Analysis of resonances in the spin-orbit problem in Celestial Mechanics: The synchronous resonance (Part I),* J. of Appl. Math. and Phys. (ZAMP) **41**, 174 (1990)

[C2] Celletti A., *Analysis of resonances in the spin-orbit problem in Celestial Mechanics: Higher order resonances and some numerical experiments (Part II),* J. of Appl. Math. and Phys. (ZAMP) **41**, 453 (1990)

[CC1] Celletti A., Chierchia L., *Rigorous estimates for a computer-assisted KAM theory,* J. Math. Phys. **28**, 2078 (1987)

[CC2] Celletti A., Chierchia L., *Construction of analytic KAM surfaces and effective stability bounds,* Commun. Math. Phys. **118**, 119 (1988)

[CC3] Celletti A., Chierchia L., *A computer-assisted approach to small-divisors problems arising in hamiltonian mechanics,* IMA Volumes in Math. and its Applications, K.R. Meyer, D.S. Schmidt eds., **28**, 41 (1991)

[CC4] Celletti A., Chierchia L., *Invariant curves for area-preserving twist maps far from integrable,* to appear in J. of Stat. Physics

[CF] Celletti A., Falcolini C., *Capture probability of Mercury in spin-orbit resonances,* preprint CARR (1990)

[CFP] 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)

[CG] Celletti A., Giorgilli A., *On the numerical optimization of KAM estimates by classical perturbation theory,* J. Appl. Mathem. and Phys. (ZAMP) **39**, 743 (1988)

[ChF] Chierchia L., Falcolini C., *A non-linear programming problem arising in a KAM algorithm,* unpublished (1990)

[ChF2] Cierchia L., Falcolini C., *A direct proof of a theorem by Kolmogorov in Hamiltonian systems*, to appear in Annali Scuola Normale Superiore die Pisa Cl. Scienze.

[ChG] Cierchia L., Gallavotti G., *Drift and diffusion in phase space*, Ann. Inst. Henri Poincaré **60**, 1 (1994)

[C] Chirikov B.V., *A universal instability of many dimensional oscillator systems*, Physics Reports **52**, 263 (1979)

[D] Danby J.M.A., *Fundamentals of Celestial Mechanics*, Macmillan, New York (1962)

[Do] Douady R., *Une démonstration directe de léquivalence des théorèmes des tores invariants pour les difféomorphismes et les champs de vecteurs*, C.R. Acad. Sci. Paris **295**, 201 (1982)

[El] Eliasson H., *Absolutely convergent series expansions for quasiperiodic motions*, preprint, Univ. of Stockholm (1987)

[Es] Escande D.F., *Stochasticity in classical Hamiltonian systems: Universal aspects*, Physics Reports **121**, 165 (1985)

[ED] 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. Physics **26**, 257 (1981)

[G1] Gallavotti G., *The Elements of Mechanics*, Springer-Verlag, New York (1983)

[G2] Gallavotti G., *Perturbation theory for classical Hamiltonian systems*, in Scaling and Self-Similarity in Physics, ed. J. Fröhlich, PPh. **7**, Birkhauser, Boston (1984)

[GP] Goldreich P., Peale S., *Spin-orbit coupling in the solar system*, Astron. J. **71**, 425 (1966)

[Go] Goroff D., unpublished (1983)

[Gr] Greene J.M., *A method for determining a stochastic transition*, J. Math. Physics **20**, 1183 (1979)

[Ha] Hamilton R.S., *The inverse function theorem of Nash and Moser*, Bull. Amer. Math. Soc. **7**, 65 (1982)

[He] Henrard J., *Spin-orbit resonance and the adiabatic invariant*, in: "Resonances in the Motion of Planets, Satellites and Asteroids", S. Ferraz-Mello and W. Sessin eds., Sao Paulo, 19 (1985)

[Her1] Herman M., *Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations*, Pub. I.H.E.S. **49**, 5 (1979)

[Her2] Herman M., *Sur le courbes invariantes par le difféomorphismes de l'anneau*, Vol. 1 Astérisque **103-104** (1983)

[Her3] Herman M., *Recent results and some open questions on Siegel's linearization theorems of germs of complex analytic diffeomorphisms of* $\mathbf{C}^n$ *near a fixed point*, preprint (1987)

[Kh] Khintchine A., *Continued fractions*, Noordhoff Ltd. Groningen (1963)

[Ko] Kolmogorov A.N., *On the conservation of conditionally periodic motions under small perturbation of the Hamiltonian*, Dokl. Akad. Nauk. SSR **98**, 469 (1954)

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

[LR] De La Llave R., Rana D., *Accurate strategies for K.A.M. bounds and their implementation*, IMA Volumes in Math. and its Applications, K.R. Meyer, D.S. Schmidt eds., **28**, 127 (1991)



[MK] MacKay R.S., *Transition to chaos for area-preserving maps*, Lectures Notes in Physics **247**, 390 (1985)

[MKP] MacKay R.S., Percival I.C., *Converse KAM: Theory and practice*, Comm. Math. Phys. **98**, 469 (1985)

[Ma] Mather J.N., *Nonexistence of invariant circles*, Erg. theory and dynam. systems **4**, 301 (1984)

[Mo1] Moser J., *On invariant curves of area-preserving mappings of an annulus*, Nach. Akad. Wiss. Göttingen, Math. Phys. Kl. II **1**, 1 (1962)

[Mo2] Moser J., *A rapidly convergent iteration method and non-linear partial differential equations*, Ann. Scuola Norm. Sup. Pisa **20**, 265 (1966)

[Mo3] Moser J., *Convergent series expansions for quasi-periodic motions*, Math. Annalen **169**, 136 (1967)

[Mo4] Moser J., *Minimal solutions of variational problems on a torus*, Ann. Inst. Henri Poincaré **3**, 229 (1986)

[Mo5] Moser J., *Break-down of stability*, in Lect. Notes in Physics **247**, 492 (1986)

[Mo6] Moser J., *Monotone twist mappings and the calculus of variations*, Ergod. Th. & Dynam. Sys. **6**, 401 (1986)

[Pe] Percival I.C., *Variational principles for invariant tori and cantori, in nonlinear dynamics and the beam-beam interaction*, AIP Conference Proceedings **57**, ed. M. Month, J.C. Herrera 302 (1980)

[Po] Poincaré H., *Les Methodes Nouvelles de la Mechanique Celeste*, Gauthier-Villars, Paris (1892)

[R1] Rüssmann H., *On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus*, Lectures Notes in Physics, **38**, 598 (1975)

[R2] Rüssmann H., *On optimal estimates for the solutions of linear difference equations on the circle*, Celestial Mechanics **14**, 33 (1976)

[R3] Rüssmann H., *Konvergente Reihenentwicklungen in der Störungstheorie der Himmelsmechanik*, Selecta Mathematica V, Heidelberg Taschenbücher **201**, 93 (1978)

[R4] Rüssmann H., *Note on sums containing small divisors*, Comm. Pure and Appl. Math. **29**, 755 (1976)

[SZ] Salamon D., Zehnder E., *KAM theory in configuration space*, Comment. Math. Helvetici **64**, 84 (1989)

[S] Siegel C.L., *Iteration of analytic functions*, Ann. Math. **43**, 607 (1942)

[Vax] (no author listed) *Vax Architecture handbook*, Digital Equipment Corporation (1981)

[W] Wisdom J., Peale S.J., *The chaotic rotation of Hyperion*, Icarus **58**, 137 (1984)

[Z] Zehnder E., *Generalized implicit function theorems with applications to some small divisor problems*, Comm. Pure Appl. Math., I **28**, 91 (1975), II **29**, 49 (1976)