LINDSTEDT SERIES FOR LOWER DIMENSIONAL TORI

ÀNGEL JORBA, RAFAEL DE LA LLAVE AND MAORONG ZOU

1. Introduction

One of the first methods to compute quasi-periodic orbits (i. e. invariant tori with linear motions on them) was the Lindstedt method (see [14] Vol. 2) which produces an expansion of the quasi-periodic orbit with a fixed frequency in powers of a small parameter measuring the distance to integrability.

The convergence of these expansions is rather delicate because they involve small divisors. The convergence was finally established using KAM theory when the quasi-periodic solutions involve as many independent frequencies as degrees of freedom of the system (and the frequencies satisfy a Diophantine condition). A more recent development is the proof of KAM theorem using directly the compensations in the Lindstedt series (see [4], [6], [3] and the lectures of Gallavotti in these proceedings). When the number of independent frequencies is less than the number of degrees of freedom the situation is much less clear. On the one hand, the Lindstedt method can be carried out with only minor modifications. It is also possible to carry out a KAM theory. The somewhat puzzling observation is that the two of them involve different small divisors. The Lindstedt series involves only small divisors coming from the intrinsic frequencies on the torus, whereas the KAM theory needs to consider also frequencies of the oscillations of the variation equations.

In a first part of this lectures we review the calculation of Lindstedt series and show that, given nonresonance conditions on the torus, it is possible to solve the equations of invarince to any order in the perturbation parameter.

In a second part of the lectures, we will prove a *translated torus* theorem that may be of independent interest. This theorem states that near approximate solutions of certain equations satisfying certain non-degeneracy assumptions, we can find true solutions defined on a large set. When applied to the approximate solutions obtained using the Lindstedt series we can prove that, even if we cannot show that they converge, at least, they define a function which is analytic in a set which is quite close to being a disk (it suffices to exclude a countable set of balls with centers on the reals, contained in a wedge with vertex at zero and of arbitrarily small angle. The measure of the balls on the reals is zero to all orders on the distance to the origin). To our knowledge, the best results about analyticity domains of these Lindstedt series for lower dimensional tori that have been published those in [3] which established that they can be defined in sectors arbitrarily close to π.

We also point out that there are two basic approaches to KAM theory, one based on deriving functional equations for the objects we are interested in and another one based on making canonical transformations so that the desired objects become apparent.

In these lectures, we want to explore the approach of deriving functional equations for the objects we are interested in. Perhaps the main advantage of the transformation theory is that it can be used to yield global information (a modern exposition of the global information that can be obtained from the transformation method is [2]). Nevertheless, the methods based on functional equations seem to be better adapted to numerical work. Transformations are much harder to implement numerically (in particular, they involve a larger number of variables and the computational complexity increases greatly with the number of variables involved). If we discretize appropriately, a numerical method will produce a function that satisfies the required functional equation up to a small residual. Using constructive theorems of the type discussed here, we can show that if we find a function that has a sufficiently small residual, there is a true solution nearby (actually, it is even possible to write numerical methods – using e.g. interval arithmetic – that prove rigorous bounds on the approximation of the solution. For an implementation of this strategy for full dimensional tori, using a different constructive theorem than the one presented here, we refer to [15], [10]). As we show here, these translated curve theorems also can be used as the end game of other perturbation methods and yield a very efficient result. In this lecture, we will be concerned with Lindstedt perturbation theories that follow tori labeled by a frequency, but there are other more global perturbation theories [7] that produce many quasi-invariant tori giving a skeleton of the dynamics of the system. Many of the landmarks identified in these theories can be shown to exist and in positions close to the positions predicted by the theory.

We call attention to the fact that the Lindstedt series involve less small divisors than the KAM methods – be they based on transformation theory or in the study of the functional equation–. It seems an interesting problem to settle the question of whether the KAM conditions really do belong. In this respect, let us call attention to [13], where perturbation series were used to establish a theorem with more general small divisors than those required by KAM theory, and to the remarkable study of cancelations in Lindstedt series discovered in [4] (see also the lectures by Gallavotti and Falcolini in this volume and references therein).

2. A model map

To fix ideas, we will discuss mainly the so called Froeschlé map. In Hamiltonian form, it can be described by

$$F_\varepsilon(q,p) = (q + p + \varepsilon \nabla V(q) \bmod \mathbb{Z}^2,\ p + \varepsilon \nabla V(q)), \tag{1}$$

where $p = (p_1, p_2) \in \mathbb{R}^2$, $q = (q_1, q_2) \in \mathbb{R}^2/\mathbb{Z}^2 \equiv \mathbb{T}^2$ and $V : \mathbb{R}^2 \to \mathbb{R}$ satisfies $V(x+e) = V(x)$ whenever $e \in \mathbb{Z}^2$. The V most commonly used in numerical experiments is

$$V_{a,b,c} = \frac{a}{2\pi}\cos(2\pi q_1) + \frac{b}{2\pi}\cos(2\pi q_2) + \frac{c}{2\pi}\cos(2\pi(q_1+q_2)). \tag{2}$$

From the form of (2) we note that when $c = 0$ the map becomes two uncoupled copies of the standard map. When $a = b = 0$, it reduces to a standard map in the variables $(q_1+q_2), (p_1+p_2)$ and the identity in the variables $(q_1-q_2), (p_1-p_2)$. Note that a map of the form (1) is the same as a standard map except for the fact that the variables are two-dimensional. In particular, the same calculations used to show that the standard map is exact symplectic show that the Froeschlé map is exact symplectic (that is, $F_{\varepsilon *}(\sum_i p_1 dq_i) = \sum_i p_1 dq_i + dS_\varepsilon$) and reversible.

The Froeschlé map also admits a *"Lagrangian"* formulation. If we denote by x_n the q after n iterations, the Hamiltonian first order equation is equivalent to the second order difference equation

$$x_{n+1} + x_{n-1} - 2x_n - \varepsilon \nabla V(x_n) = 0. \tag{3}$$

For the Froeschlé map, the Lagrangian formulation is completely equivalent to the Hamiltonian one, since to pass from one to the other, one only needs to eliminate or introduce the p's. We will discuss the Lindstedt series in the Lagrangian formulation – but a Hamiltonian formulation is also available –. The KAM method we discuss does not use much geometry.

3. Lindstedt series for lower dimensional tori

The tori considered by the classical KAM theorem are two dimensional. Here, we will seek one dimensional tori. More precisely, we will look for a map $u_\varepsilon : \mathbb{T}^1 \to \mathbb{T}^2 \times \mathbb{R}^2$ in such a way that

$$F_\varepsilon(u_\varepsilon(\theta)) - u_\varepsilon(\theta + \omega) = 0. \tag{4}$$

In Lagrangian formulation, searching for invariant tori is the same as searching for functions $\ell_\varepsilon : \mathbb{R} \to \mathbb{R}^2$ in such a way that

$$\ell_\varepsilon(\theta + \omega) + \ell_\varepsilon(\theta - \omega) - 2\ell_\varepsilon(\theta) = \varepsilon \nabla V(\ell(\theta)), \quad \ell_\varepsilon(\theta + 1) = \ell_\varepsilon(\theta) + k, \tag{5}$$

where the second condition comes from imposing that ℓ_ε is a mapping from the circle to the torus. Rather than working with the function ℓ_ε, we find it more convenient to work with the periodic function $g_\varepsilon = \ell_\varepsilon - k\theta$, and study the equation

$$g_\varepsilon(\theta + \omega) + g_\varepsilon(\theta - \omega) - 2g_\varepsilon(\theta) = \varepsilon \nabla V(\theta k + g_\varepsilon(\theta)). \tag{6}$$

We point out that the solutions of these equivalent equations are never unique. If u_ε solves (4), then $\tilde{u}_\varepsilon$ defined by $\tilde{u}_\varepsilon(\theta) = u_\varepsilon(\theta + \alpha)$ is also a solution of the equation. Solutions that differ in this way describe the same invariant object, with only a different system of coordinates. They are, for our purposes, equivalent. The same considerations apply to the equivalent equations (5) or (6).

Lindstedt method consists in assuming that we can write

$$g_\varepsilon(\theta) = \sum_{n=0}^{\infty} g^n(\theta)\varepsilon^n. \tag{7}$$

If we equate the terms in ε^n after substituting (7) in (6) we obtain:

$$g^n(\theta + \omega) + g^n(\theta - \omega) - 2g^n(\theta) = R^n(\theta), \tag{8}$$

where R^n is an expression that involves only $g^0, \ldots, g^{n-1}$. Therefore, we can consider (8) as an equation to determine g^n when we know the coefficients of lower order. If this equation could be solved, then, we could compute all the coefficients of the expansion and, then, solve (6) in the sense of formal power series.

The theory of solvability of equations similar to (8) is very well known and can be found in all the books in KAM theory (see e.g. [17]).

Theorem 1 *Let ω be a Diophantine number (i.e. $|\omega - p/q| \geq K|q|^{-\nu}$). If R^n is an analytic function there exist a g^n solving (8) if and only if $\int_{\mathbb{T}^1} R(\theta)d\theta = 0$. In case that the solution exists, it is unique up to additive constants. Moreover, defining for $f(\theta) = \sum_j f_j e^{2\pi i j\theta}$ the norm $||f||_\delta = \sum_j |f_j| e^{|j|\delta}$, we have for the g^n with zero average $||g^n||_{\delta-\rho} \leq M(\nu, \rho)\frac{||R||_\delta}{K\rho^\nu}$.*

In view of Theorem 1, to show that there exists power series solutions, we just need to show that it is possible to adjust that the R^n has zero average. Note that the ε^0 equation becomes $g^0(\theta+\omega)+g^0(\theta-\omega)-2g^0(\theta)=0$. Hence, g^0 is a constant, according to Theorem 1.

The ε^1 equation becomes

$$g^1(\theta+\omega)+g^1(\theta-\omega)-2g^1(\theta)=\nabla V(\theta k+g^0).$$

The condition for the existence of solutions is $\int_0^1 \nabla V(\theta k+g^0)d\theta=0$. This is equivalent to the two conditions

$$\int_0^1 k\cdot\nabla V(\theta k+g^0)d\theta=0,\quad \int_0^1 k^\perp\cdot\nabla V(\theta k+g^0)d\theta=0, \tag{9}$$

where $k^\perp\in\mathbb{Z}^2$ is a vector perpendicular to k. The first of the two conditions (9) is always satisfied for all g^0 because $k\cdot\nabla V(\theta k+g^0)=\frac{d}{d\theta}V(\theta k+g^0)$. On the other hand, note that if we can find a g^0 solving (9), then $g^0+\beta k$ is also a solution. Hence, it suffices to seek $g^0=\alpha k^\perp$ so that the second condition in (9) is satisfied.

Such a solution can always be found because

$$\int_0^1 d\alpha \int_0^1 k^\perp\cdot\nabla V(\theta k+\alpha k^\perp)d\theta=\int_0^1 d\theta\int_0^1 d\alpha\,\frac{d}{d\alpha}V(\theta k+\alpha k^\perp)=0.$$

Noticing that if the integral of a periodic function vanishes it has to have at least two zeros, we conclude that the integrand with respect to α in the first integral has to vanish for two α's.

In the concrete form of the Froeschlé map (2) and for $k=(1,0)$, $k^\perp=(0,1)$ we can find exactly two solutions, $g^0=0(0,1)$, $g^0=\frac{1}{2}(0,1)$. For this particular example setting $\tilde{x}=x-g^0$, we have $\nabla(\tilde{x})=-\nabla V(-\tilde{x})$. This antisymmetry simplifies many of the calculations for this particular example.

Once that we have chosen g^0 in such a way that the compatibility conditions for g^1 are met, according to Theorem 1, we can determine g^1 up to an additive constant. This additive constant will, we expect, be determined in such a way that R^2 satisfies the compatibility conditions that allow to compute g^2 and so on. Indeed, this process can be carried out provided that the potential is not degenerate. A sufficient non-degeneracy condition is the following.

Lemma 1 *Let V satisfy $\int_0^1 k^\perp D^2V(\theta k)k^\perp d\theta\neq 0$, then, given $g^0,\dots,g^{n-1}$ we can find a unique $\alpha\in\mathbb{R}$ such that*

$$\int_0^1 R^n(g^0,g^1,\dots,g^{n-1}+\alpha k^\perp)d\theta=0.$$

Proof. Recall that R^n is the coefficient of ε^{n-1} in the expansion in ε of $\nabla V(\theta k + \sum_{i=0}^{n-1} g^i(\theta)\varepsilon^i + \alpha\varepsilon^{n-1}k^\perp)$. Expanding, we see that $R^n(\theta) = \alpha D^2V(\theta k)k^\perp + S^n(\theta)$ where S^n is an expression that involves only $g^0, \dots, g^{n-1}$.
Q.E.D

We have, therefore shown that there is a solution of (6) in the class of formal power series in ε whose coefficients are analytic functions in θ. A slightly weaker statement, which is the only one we will need, is the following:

Lemma 2 *Let V be an analytic non-degenerate (e.g. satisfying the hypothesis of Lemma 1). Then, given $\delta > 0$ sufficiently small, for every $N \in \mathbb{N}$, we can find a $g_\varepsilon^{[<N]} = \sum_{i<N} g^i(\theta)\varepsilon^i$ such that*

$$||g_\varepsilon^{[<N]}(\theta+\omega) + g_\varepsilon^{[<N]}(\theta-\omega) - 2g_\varepsilon^{[<N]}(\theta) - \varepsilon\nabla V(\theta k + g_\varepsilon^{[<N]}(\theta))||_\delta \leq C_{N,\delta}|\varepsilon|^N.$$

Of course, since the equations (6) and (4) are equivalent, there is a similar statement in terms of u_ε^N.

By performing the calculations outlined above with care it is possible to obtain explicit quantitative estimates of C_N. For example, for the explicit Froeschlé map of (2), it is possible to obtain $C_N \leq N^{aN}$. Even if this explicit bounds are certainly not enough to prove convergence, once the dependence on N is known, given an ε we can choose the N that gives the best bound for the remainder. For example, for the C_N above we obtain $\inf_N N^{aN}|\varepsilon|^N \approx \exp(-C|\varepsilon|^{-1/a})$.

Therefore, for ε sufficiently small, it is possible to produce functions that solve the equations up to an extremely small error. One can expect that if we start a Newton method in these approximate solutions, we would get exact solutions that are extremely close to those approximate solutions. This is indeed what we will do in the rest of the paper. Besides the usual analytical issues of making sense of a Newton method with unbounded derivatives, we face a formal difficulty that we discuss in the following section.

4. Reducibility

It seems that KAM theory can best study circles satisfying a certain extra property..

Definition 1 *We say that a $u : \mathbb{T}^1 \to \mathbb{R}^4$ is reducible when we can find $M : \mathbb{T}^1 \to \mathcal{M}_{4\times 4}$ and a $C \in \mathcal{M}_{4\times 4}$ in such a way that*

$$DF(u(\theta)) = M^{-1}(\theta+\omega)CM(\theta). \tag{10}$$

Of course, analogous definitions can be made in higher dimensions. We will not use them for the moment. Of course, in the above definition, the regularity required to the function M plays a role.

The concept of reducibility has a geometric meaning for invariant circles. To each point $x = u(\theta)$ of the circle we can associate a space E_x of infinitesimal displacements. The derivative $DF(x)$ is a map from E_x to $E_{F(x)}$. Reducibility is the same as making linear changes of variables in each of the spaces E_x (they may depend on x) in such a way that the derivative becomes a constant map. The fact that a torus is reducible makes it very convenient to study the effect of perturbations on it since there is a system of coordinates in which the infinitesimal perturbations are just constant.

Note also that the reduction is an analogue of Floquet theory to quasi-periodic motions rather than periodic. This makes it appear in many other contexts, e.g. Schroedinger equations with quesiperiodic potentials etc. We call attention to the lectures of Eliason in this volume for very remarkable recent progress as well as to the recent thesis of R. Krikorian.

A notable exception to the use of reducibility is the paper [1] and the papers inspired by it. This paper does not use reducibility but rather PDE methods – related to Frölich-Spencer methods in localization – We do not know of published papers that use an analogous method in finite dimensions!!.

If we take the u_ε that we have computed using Lindstedt method as in the previous section, we can write

$$DF_\varepsilon(u_\varepsilon(\theta)) = \begin{pmatrix} \mathrm{Id}_2 & \mathrm{Id}_2 \\ 0_2 & \mathrm{Id}_2 \end{pmatrix} + \varepsilon N_1(\theta) + \varepsilon^2 N_2(\theta) + \cdots, \tag{11}$$

where Id_2 denotes the 2 by 2 identity matrix and 0_2 denotes the 2 by 2 zero matrix. Each of the N_i is an analytic function in θ taking values in the set of 4 by 4 matrices and the whole expression is a formal power series in ε.

Our first task will be to show that an expression such as (11) can be reduced in the sense of power series expansions in ε with coefficients that are analytic in θ. That is, we will show that it is possible to find $M_i(\theta)$ analytic functions of θ taking values on 4 by 4 matrices and 4 by 4 matrices C_i in such a way that the formal power series expansions

$$\begin{aligned} M_\varepsilon(\theta) &= \mathrm{Id}_4 + \varepsilon M_1(\theta) + \varepsilon^2 M_2(\theta) + \cdots, \\ C_\varepsilon &= \begin{pmatrix} \mathrm{Id}_2 & \mathrm{Id}_2 \\ 0_2 & \mathrm{Id}_2 \end{pmatrix} + \varepsilon C_1 + \varepsilon^2 C_2 + \cdots, \end{aligned}$$

satisfy

$$M_\varepsilon(\theta + \omega) DF_\varepsilon(u_\varepsilon(\theta)) = C_\varepsilon M_\varepsilon(\theta), \tag{12}$$

in the sense of power series in ε. If we equate terms of order ε^n in (12) we obtain:

$$M_n(\theta + \omega) \begin{pmatrix} \mathrm{Id}_2 & \mathrm{Id}_2 \\ 0_2 & \mathrm{Id}_2 \end{pmatrix} = \begin{pmatrix} \mathrm{Id}_2 & \mathrm{Id}_2 \\ 0_2 & \mathrm{Id}_2 \end{pmatrix} M_n(\theta) + C_n + R_n, \tag{13}$$

where R_n is a 4 by 4 matrix that can be formed out of $M_1, M_2, \dots, M_{n-1}$, $C_1, C_2, \dots, C_{n-1}$, using just products and sums.

We will show that, if we assume that $M_1, M_2, \dots, M_{n-1}, C_1, C_2, \dots, C_{n-1}$ are known, then we can find M_n, C_n. It will follow by induction that all M_n can be found, that is, we can solve (12) in the sense of formal power series in ε.

That is, if we write $M_n(\theta) = \begin{pmatrix} a(\theta) & b(\theta) \\ c(\theta) & d(\theta) \end{pmatrix}$, $R_n(\theta) = \begin{pmatrix} s(\theta) & t(\theta) \\ u(\theta) & v(\theta) \end{pmatrix}$, (13) is equivalent to

$$\begin{aligned} a(\theta+\omega) &= a(\theta) + c(\theta) + c_{11} + s(\theta), \\ a(\theta+\omega) + b(\theta+\omega) &= b(\theta) + d(\theta) + c_{12} + t(\theta), \\ c(\theta+\omega) &= c(\theta) + c_{21} + u(\theta), \\ c(\theta+\omega) + d(\theta+\omega) &= d(\theta) + c_{22} + v(\theta). \end{aligned} \tag{14}$$

Using Theorem 1, we see that we can find a unique $c(\theta)$, c_{21} satisfying the third of the equations in (14) such that $\int c(\theta)\, d\theta = 0$. This c is analytic in θ. Once this c is known, the first and the fourth equations in (14) become equations for a, c_{11} and d, c_{22} respectively. The same argument as before shows that they have unique solutions that has zero average. Once a and d are known, we can solve the second equation for b, c_{12}.

Thus, as claimed, we can solve (12) in the sense of power series and have, therefore established:

Lemma 3 *Let V be an analytic non-degenerate potential. For every $\delta > 0$ sufficiently small, given any $N \in \mathbb{N}$, it is possible to find $u_\varepsilon^{[<N]}$, $M_\varepsilon^{[<N]}$, C_ε in such a way that*

$$||F_\varepsilon(u_\varepsilon^{[<N]}(\theta)) - u_\varepsilon^{[<N]}(\theta+\omega)||_\delta \le K_{N,\delta}|\varepsilon|^N,$$
$$\|M_\varepsilon^{[<N]}(\theta+\omega) DF_\varepsilon(u_\varepsilon^{[<N]}(\theta)) M_\varepsilon^{[<N]^{-1}}(\theta) - C_\varepsilon\|_\delta \le K_{N,\delta}|\varepsilon|^N.$$

In other words, we can not only solve approximately the invariance equation, but also reduce approximately the variation equation.

We remark that an elementary argument [14] §74, §79 shows that the eigenvalues of the matrix C_ε are 1, 1 – the eigenvectors correspond to the tangent direction to the tori and its symplectic conjugate – and the other two eigenvalues are a formal power series expansion in powers of $\varepsilon^{1/2}$. The leading terms are $1 \pm (a\varepsilon)^{1/2} + O(\varepsilon)$ where a is a real number that does not vanish given a non-degeneracy assumption (for example in (2) a is positive for one of the solutions and negative for the other). The corresponding eigenspaces form an angle $O(\varepsilon^{1/2})$. When $\varepsilon \in \mathbb{R}$, if $a\varepsilon > 0$, the two eigenvalues are of different size, one inside of the unit circle and the other outside. These are the so-called "*whiskered tori*". In the case that $a\varepsilon < 0$, both of them have to have modulus 1. These tori are called elliptic.

Note that at the level of perturbation theory, the eigenvalues are a function of $\varepsilon^{1/2}$, hence when ε goes around the origin, the stable eigendirection is changed into the unstable one and vice versa. The invariant tori themselves have trivial monodromy.

In summary, the Lindstedt series predict the existence of tori of lower dimension that are elliptic for real positive values of the parameter and whiskered for real negative values of the parameter. We will see that this prediction is correct, at least for a large set of values of the parameter.

5. Translated curve theorem

To motivate the translated curve theorem, let us discuss briefly the numerical resolution of (4) using a Newton method.

It is an easily justified calculation that, in the appropriate spaces, the derivative of (4) is the operator $\mathcal{A}$ defined by:

$$[\mathcal{A}\Delta_\varepsilon](\theta) = DF_\varepsilon(u_\varepsilon(\theta))\Delta_\varepsilon(\theta) - \Delta_\varepsilon(\theta+\omega).$$

For the Newton method we would need to invert the operator $\mathcal{A}$. Unfortunately, the operator $\mathcal{A}$ is not going to be invertible for a solution of (4). This is, of course, related to the fact that we discussed before, that the solutions of (4) are not unique and that, given a solution u_ε, the $\tilde{u}_\varepsilon$ defined by $\tilde{u}_\varepsilon(\theta+\alpha)$ is also a solution. If u_ε is a solution of (4), taking derivatives with respect to θ we obtain

$$DF_\varepsilon(u_\varepsilon(\theta))\partial_\theta u_\varepsilon(\theta) - \partial_\theta u_\varepsilon(\theta+\omega) = 0. \tag{15}$$

Moreover, if the matrix DF_ε is symplectic, we have $DF^t_\varepsilon(u_\varepsilon(\theta))JDF_\varepsilon(u_\varepsilon(\theta)) = J$ where J is the symplectic matrix, which also satisfies $J^2 = -1$. Then, we see that the vector $v_\varepsilon(\theta) \equiv J\partial_\theta u_\varepsilon(\theta)$ satisfies:

$$DF^t_\varepsilon(u_\varepsilon(\theta))v_\varepsilon(\theta+\omega) = v_\varepsilon(\theta).$$

This equation shows that v_ε should also be in the null space of $\mathcal{A}$.

Fortunately, we have only found a finite number (twice the dimension of the torus) of zero eigenvalues of $\mathcal{A}$. If those were the only ones, the infinite dimensional Newton method could be fixed using the same methods as in finite dimensions. First note that the zero eigenvalues corresponding to u_ε are related to the fact that the solutions of the equation (4) are indeed multiple, since we can change the origin of coordinates. This eigenvalue 0 of $\mathcal{A}$ can be eliminated by imposing a normalization that fixes the origin (for the problem of lower dimensional tori in the Froeschlé map, an adequate normalization could be $\int l_\varepsilon(\theta)\cdot k^\perp d\theta = 0$). The other zero eigenvalue is harder to eliminate. The best that we could hope is to add extra parameters to the equation so that they allow us to project out the component along this direction. In our case, this amounts to considering the equation in the unknowns u_ε, a_ε

$$F_\varepsilon(u_\varepsilon(\theta)) = u_\varepsilon(\theta+\omega) + a_\varepsilon, \tag{16}$$

where a_ε is constrained to belong to a finite dimensional space that has a non-trivial component along the eigenvalue that we want to eliminate.

For the Froeschlé map, a possibility would be to try to solve the equation

$$g_\varepsilon(\theta+\omega) + g_\varepsilon(\theta-\omega) - 2g_\varepsilon(\theta) = \varepsilon \nabla V(\theta k + g_\varepsilon(\theta)) + \alpha_\varepsilon k. \tag{17}$$

In the Hamiltonian formulation, this corresponds to adding a translation in the k direction along the actions. Note that such translation is the result of applying the symplectic matrix J to a translation in the k direction along the angles. A translation in the k direction along the angles, turns out to have a non-trivial component along the direction of reparameterizing the circles – this is very easy if the circles are perturbations of the unperturbed ones.

Trying to solve a modified equation does not seem so satisfactory since the geometric meaning of the equation (16) or (17) is not as interesting as that of our original equations. Nevertheless, it turns out that, for certain class of maps, the two equations are equivalent.

For example, (6) and (17) are equivalent. Indeed, if we take the dot product of (17) and $g_\varepsilon'(\theta)+k$ and we integrate over θ, using that $\int \nabla V(\theta k + g_\varepsilon(\theta)) \cdot (k + g_\varepsilon'(\theta)) d\theta = \int \frac{d}{d\theta} V(\theta k + g_\varepsilon(\theta)) d\theta = 0$ as well as integration by parts and changes of variables, we obtain that if (17) holds, then $\alpha_\varepsilon = 0$ (we will show more details later, as a preliminary to Lemma 4).

Notice that, if we did not have this cancelation, the result would indeed be false. For example, if the forcing was not the gradient of a periodic function – e.g. we took it to be a constant, so that the mappings were $(q,p) \mapsto (q+p, p+t_\varepsilon)$, we can see that, unless $t_\varepsilon = 0$, points do not come close to themselves so that this mapping does not posses invariant curves.

In the rest of these lectures we will formulate precisely and sketch the proof of the translated curve theorem as well as a *vanishing* theorem that shows that, for exact symplectic mappings the translations have to vanish.

Let us remark that, from the point of view of numerical implementations, the approach with the extra parameter is quite useful. The zero eigenvalues lead to ill conditioned matrices that are hard to deal with numerically – a bad pivot can wreck havoc in a Gaussian elimination even if it cancels afterwards. For very complicated systems that require software tools to write the programs, it is difficult to ensure that the derivative of the truncation is the truncation of the derivative. This damages some of the cancelations. Of course, when the calculation finishes, the vanishing theorem, assures us that the translation should be zero.

The translated curve theorem was formulated for two dimensional maps in [16] and proved using his “extra parameter method”. The proof we present here is based in another iteration scheme. The main difference is that we do not try to show that the translations vanish at each step. We also use the perturbation parameter as one of the relevant parameters that are used to keep the frequencies fixed, a technique already used in [8] and [9].

We will state it and prove it only for maps in 4 dimensions to simplify the notation, but it should be clear that it holds in greater generality.

Theorem 2 *Let ω be a Diophantine number, $\Omega \subset \mathbb{C}$ be a ball. Assume that we can find analytic functions of $\varepsilon \in \Omega$, u_ε,M_ε and c_ε, taking values respectively in the space of analytic functions $\mathbb{T}^1 \to \mathbb{R}^4$, $\mathbb{T}^1 \to \mathcal{M}_{4\times 4}$ and on $\mathcal{M}_{4\times 4}$ in such a way that:*

i) We have

$$||F_\varepsilon(u_\varepsilon(\theta)) - u_\varepsilon(\theta+\omega)||_{\delta_0} \le \beta, \quad ||M_\varepsilon(\theta+\omega)\, DF_\varepsilon(u_\varepsilon(\theta)) M_\varepsilon^{-1}(\theta) - c_\varepsilon||_{\delta_0} \le \beta.$$

ii) The eigenvalues λ_i of c_ε can be divided in those that are identically 1 and those that are never 1.

We denote by E^ε_Θ the spectral eigenspace of c_ε corresponding to eigenvalues 1 and by $E^\varepsilon_\mathcal{T}$ the eigenspace corresponding to derivatives along the angles.

iii) We assume that if λ^ε_i, λ^ε_j are not identically 1 and the eigenspace of λ^ε_i is not symplectically conjugate to that of λ^ε_j we have:

$$\left|\frac{d}{d\varepsilon}\lambda^\varepsilon_i\right| \ge a > 0; \quad \left|\ \left|\frac{d}{d\varepsilon}\lambda^\varepsilon_i\right| - \left|\frac{d}{d\varepsilon}\lambda^\varepsilon_j\right|\ \right| \ge a > 0; \quad \left|\frac{d^2}{d\varepsilon^2}\lambda^\varepsilon_i\right| \le b > 0; \qquad (18)$$

iv) We also assume that we can find a subspace $E^\varepsilon_{\tilde{\mathcal{C}}}$ in such a way that $E^\varepsilon_\Theta = E^\varepsilon_\mathcal{T} \oplus E^\varepsilon_{\tilde{\mathcal{C}}}$ and that the operator $\Gamma^\varepsilon \equiv \Pi^\varepsilon_\mathcal{T} c_\varepsilon : E^\varepsilon_{\tilde{\mathcal{C}}} \mapsto E^\varepsilon_\mathcal{T}$ has a bounded left inverse $(\Gamma^\varepsilon)^{-1}$ and that

$$||(\Gamma^\varepsilon)^{-1}|| \le b.$$

v) $\beta \le \beta^*(a, b, \omega, \delta_0)$ for some function β^* that will be made explicit during the proof.

Then, we can find a set $\mathcal{S} \subset \Omega$ in such a way that

a) $\mathcal{S} = \Omega - \cup_{i=1}^\infty B_{\varepsilon_i, r_i}$.

b) $|\Omega - \mathcal{S}| \le A^*(a, b, \omega, \delta)\beta^{1/4}$ (where A^* is also a function made explicit during the proof).

c) We can find analytic functions $\tilde{u}_\varepsilon$, $\tilde{M}_\varepsilon$, $\tilde{c}_\varepsilon$, $\tilde{A}_\varepsilon$ taking values in analytic functions $\mathbb{T}^1 \mapsto \mathbb{R}^4$, analytic functions $\mathbb{T}^1 \mapsto \mathcal{M}_{4\times 4}$ and into $\mathcal{M}_{4\times 4}$, $E^\varepsilon_{\tilde{\mathcal{C}}}$ and such that

$$F_\varepsilon(\tilde{u}_\varepsilon(\theta)) - \tilde{u}_\varepsilon(\theta+\omega) = M_\varepsilon(\theta)\tilde{A}_\varepsilon, \qquad \tilde{M}_\varepsilon(\theta+\omega)DF_\varepsilon(\tilde{u}_\varepsilon(\theta))\tilde{M}_\varepsilon^{-1}(\theta) = \tilde{c}_\varepsilon.$$

d) Moreover, for $\varepsilon \in \mathcal{S}$ we have

$$\max\{||\tilde{u}_\varepsilon - u_\varepsilon||_{\delta_0/2}, ||\tilde{M}_\varepsilon - M_\varepsilon||_{\delta_0/2}, ||\tilde{c}_\varepsilon - c_\varepsilon||, |tA_\varepsilon|\} \le A^*(a,b,\omega,\delta)\beta.$$

The meaning of the hypothesis is that we have an approximate invariant torus, which is approximately reducible (i)), that the spectra of the linearization changes with the perturbation (iii)) and with the coordinates symplectically conjugated to the motion on the torus (iv)). Then, if this approximation satisfies some explicit bounds – depending on the non-degeneracy conditions and the size of the functions – we conclude that there are exactly invariant tori for many parameters close by. Of course, if the system is quasi integrable, the tori of the integrable system satisfy the hypothesis. It is also possible to produce quasi-invariant quasi-reducible tori using a variety of other methods. For example, the output of numerical computations is function that satisfies the equations till very small error (truncation and roundoff). We note that to justify the validity of the numerical calculations we do not need to study the algorithm used to produce them. We only need to verify rigorously that the hypothesis of2 are verified. This is what is done in [15] [10]. One can also use other perturbation theories that, even if not convergent, produce quasi-invariant quasi-reducible tori. One such theory with a semi-global character is in [7]. The quantitative estimates in 2 and 3 also allow to use the series produced by the Lindstedt procedure as input to 2. Since the estimates of the error are smaller that all the powers of ϵ we obtain that there is an invariant curve defined for a set of ϵ depicted in Fig.1 and analytic there. Moreover, since the errors of the asymptotic expansion are zero to all orders in ϵ, using the conclusions in 2, we see that the function describing the invariant torus and the truncated Lindstedt series differ by amounts that are smaller than any power of ϵ. Hence, in the domain of definition, the Lindstedt series is the asymptotic expansion.

Note that for the Froeschlé map, since the directions whose eigenvalues are not 1 have to be symplectically conjugate. Then, the second part of iii) in Theorem 2 becomes empty. In systems with more frequencies it is not. The condition we have stated is not optimal, but is enough for our purposes.

Note that the tori are given by $\theta \mapsto (\theta k + g^0, \omega) + o(\varepsilon)$, the tangent to the torus is $(k,0) + o(\varepsilon)$ and the symplectic conjugate direction is $(0,k) + o(\varepsilon)$. The reduced matrix c_ε is $\begin{pmatrix} \mathrm{Id}_2 & \mathrm{Id}_2 \\ 0_2 & \mathrm{Id}_2 \end{pmatrix} + o(\varepsilon)$. The Γ^ε in hypothesis iv) is the upper right hand corner in this matrix.

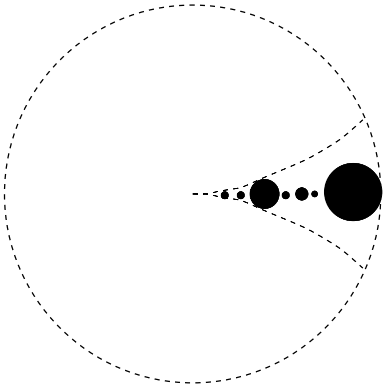

Figure 1. The set $\mathcal{S}$ applying Theorem 2 to the Lindstedt series.

Proof.

The proof will be done using a quadratically convergent procedure.

At the start of step n we will have u^n_ε, M^n_ε, c^n_ε and A^n_ε defined in $\mathcal{S}^n = \Omega - \cup_{i=1}^{N^n} B_{\sigma_i, r_i^n}$. Those of the above that are analytic functions in θ will be analytic in a set of the form $|\mathrm{Im}(\theta)| \le \delta_n$.

They will be approximate solutions to the equations that we want to solve. That is:

$$\begin{aligned} F_\varepsilon(u^n_\varepsilon(\theta)) - u^n_\varepsilon(\theta+\omega) &= R^n_\varepsilon(\theta) + A^n_\varepsilon, \\ M^n_\varepsilon(\theta+\omega) DF_\varepsilon(u^n_\varepsilon(\theta))(M^n_\varepsilon)^{-1}(\theta) - c^n_\varepsilon &= S^n_\varepsilon, \end{aligned} \tag{19}$$

where R^n_ε, S^n_ε are small in the sense of $||\cdot||_{\delta_n}$.

We will assume inductively that M^n_ε is in a fixed neighborhood of the identity, c^n_ε is in a fixed neighborhood of $\begin{pmatrix} \mathrm{Id}_2 & \mathrm{Id}_2 \\ 0_2 & \mathrm{Id}_2 \end{pmatrix}$ and that we have uniform bounds of the type assumed in *iii)* – with worse constants –.

The set $\mathcal{S}^n$ will also be assumed to be such that, for $|k| \le 2^n$, we have

$$\begin{aligned} |\lambda^{n,j}_\varepsilon - e^{2\pi i k\omega}| &\ge C|k|^{-\nu} ||R^n_\varepsilon||^{1/8}_{\delta_n}, \\ |(\lambda^{n,j}_\varepsilon) + e^{2\pi i k\omega}| &\ge C|k|^{-\nu} ||R^n_\varepsilon||^{1/8}_{\delta_n}, \\ |e^{2\pi i k\omega}\lambda^{n,i}_\varepsilon - \lambda^{n,j}_\varepsilon| &\ge C|k|^{-\nu} ||R^n_\varepsilon||^{1/8}_{\delta_n}. \end{aligned} \tag{20}$$

The goal of the step is to produce u^{n+1}_ε, M^{n+1}_ε, c^{n+1}_ε and A^{n+1}_ε defined in a slightly smaller set $\mathcal{S}^{n+1}$ and with slightly worse analyticity properties but which are much more approximate solutions to the equations. The set $\mathcal{S}^{n+1}$ will satisfy (20) for $|k| \le 2^{n+1}$ and for the eigenvalues $\lambda^{n+1,i}_\varepsilon$ corresponding to the matrix c^{n+1}_ε. We will also have to check that the inductive hypothesis about bounds can also be maintained.

To improve the solutions of the first equation of (19), we set $u_\varepsilon^{n+1} = u_\varepsilon^n + \Delta_\varepsilon^n$ and $A_\varepsilon^{n+1} = A_\varepsilon^n + B_\varepsilon^n$, where

$$DF_\varepsilon(u_\varepsilon^n(\theta))\Delta_\varepsilon^n(\theta) - \Delta_\varepsilon^n(\theta+\omega) = -R_\varepsilon^n + B_\varepsilon^n. \tag{21}$$

This equation can be justified heuristically as substituting $u_\varepsilon^n + \Delta_\varepsilon^n$ in the equation we want to solve and keeping only the linear terms in Δ_ε^n.

To solve (21), we use the second equation in (19), to transform (21) into

$$c_\varepsilon^n[M_\varepsilon^n\Delta_\varepsilon^n](\theta) - [M_\varepsilon^n\Delta_\varepsilon^n](\theta+\omega) = M_\varepsilon^n(\theta+\omega)[R_\varepsilon^n + B_\varepsilon^n] - S_\varepsilon^n(\theta)M_\varepsilon^n(\theta)\Delta_\varepsilon^n(\theta). \tag{22}$$

Rather than solving (22), we will solve the equation that results when we drop the last term (it can be argued that this is a reasonable approximation since this term contains the product of two small quantities S_ε^n and Δ_ε^n, we will justify later this heuristic procedure) and, moreover, we only consider the Fourier coefficients with $|k| \le 2^n$.

Equations for $\varphi_\varepsilon, B_\varepsilon$ of the form

$$c_\varepsilon\varphi(\theta) - \varphi_\varepsilon(\theta+\omega) = R_\varepsilon(\theta) + B_\varepsilon, \tag{23}$$

where R_ε and c_ε are given can be analyzed using Fourier series. If we take Fourier series, (23) becomes:

$$\begin{aligned}(c_\varepsilon - e^{2\pi ik\omega})\hat{\varphi}_{k,\varepsilon} &= \hat{R}_{k,\varepsilon}, \quad k \neq 0,\\ (c_\varepsilon - 1)\hat{\varphi}_{0,\varepsilon} &= \hat{R}_{0,\varepsilon} + B_\varepsilon.\end{aligned} \tag{24}$$

To obtain estimates for the solutions of (24), we recall that using Cramer's rule, if a finite-dimensional matrix satisfies $||A|| \le K$ we can bound $||A^{-1}|| \le C_K/|\det(A)|$ we solve the equation that results from (22) after dropping the indicated terms we have that $\det(c_\varepsilon - e^{2\pi ik\omega}) = (1 - e^{2\pi ik\omega})^2(\lambda^{n,1} - e^{2\pi ik\omega})(\lambda^{n,2} - e^{2\pi ik\omega})$. Therefore, given the bounds in (20), we obtain $|\det(c_\varepsilon - e^{2\pi ik\omega})| \ge C|k|^{-4\nu}||R_\varepsilon^n||_{\delta_n}^{1/4}$ where C is a uniform constant if c_ε remains uniformly bounded. Therefore,

$$|\hat{\varphi}_k| \le C|k|^{4\nu}||R_\varepsilon^n||_{\delta_n}^{-1/4}|\hat{R}_{\varepsilon,k}^n| \le C|k|^{4\nu}||R_\varepsilon^n||_{\delta_n}^{3/4}e^{-2\pi\delta_n|k|} \tag{25}$$

where the last inequality is just the Cauchy bounds.

For $k = 0$ we have to analyze the second equation in (24). Since c_ε has eigenvalues 1, the solutions may not exist or fail to be unique.

We can take as a guide, the fact that the equation we want to solve is very similar to

$$\left(\begin{pmatrix} \mathrm{Id}_2 & \mathrm{Id}_2 \\ 0_2 & \mathrm{Id}_2 \end{pmatrix} - 1\right)\begin{bmatrix} \varphi_1 \\ \varphi_2 \end{bmatrix} = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix} + \begin{bmatrix} 0 \\ B \end{bmatrix}.$$

Hence, to ensure that $\hat{\varphi}_{0,\varepsilon}$ exists, we need to choose B_ε in such a way that

$$\Pi_\Theta^{\varepsilon,n}(\overline{R_\varepsilon^n} + B_\varepsilon) = 0, \quad B_\varepsilon \in E_\Theta^{\varepsilon,n}.$$

Since $\Pi_\Theta^{\varepsilon,n}$ is a perturbation of the identity on $E_\Theta^{\varepsilon,n}$, we can find a unique B_ε and bound its size by a constant times $|\overline{R_\varepsilon^n}|$. A fortiori, by a constant times $||R_\varepsilon^n||_{\delta_n}$.

Once we have chosen B_ε in such a way, we note that the component of $\hat{\varphi}_{0,\varepsilon}$ along $E_\Theta^{\varepsilon,n}$ is arbitrary. We will choose it in such a way that the equation for the $E_{\mathcal{T}}^{\varepsilon,n}$ component becomes solvable. This can be done in a unique way because of assumption *iv*). Note that the equation for the $E_{\mathcal{T}}^{\varepsilon,n}$ component becomes

$$\Pi_{\mathcal{T}}^{\varepsilon,n} c_\varepsilon \Pi_\Theta^{\varepsilon,n} \hat{\varphi}_{0,\varepsilon} = \Pi_{\mathcal{T}}^{\varepsilon,n}(\overline{R_\varepsilon^n} + B_\varepsilon^n).$$

By assumption *iv*), $\Pi_{\mathcal{T}}^{\varepsilon,n} c_\varepsilon \Pi_\Theta^{\varepsilon,n}$ is invertible and, since the R.H.S. can be bounded by a constant $|\overline{R_\varepsilon^n}|$, the solution can be bounded in the same way that we bounded the other component.

Finally, we see that the equation for the the component of $\hat{\varphi}_{0,\varepsilon}$ along $E_{\mathcal{T}}^{\varepsilon,n}$ becomes completely undetermined. We just pick it to be zero.

Note that the fact that we can pick this component arbitrarily has a geometric meaning since we argued before that a constant along this direction corresponds just to changing the origin of the coordinate system on the torus.

If we use (25) and the estimates that we have for the zero-order term, we obtain that, for any $0 < \rho < \delta_n$

$$||\Delta_\varepsilon^n||_{\delta_n - \rho} \le C\rho^{-2\nu} ||R_\varepsilon^n||_{\delta_n}^{3/4}.$$

To estimate $||R^{n+1}||_{\delta_n - \rho}$ we observe that the only reasons why the procedure that we outlined above does not produce an exact solution are 1) we used a linear approximation, 2) we ignored one term in (22), 3) we truncated the R. Adding and subtracting appropriate terms, we can get to estimate each of the terms separately and, hence we get

$$\begin{aligned} ||R_\varepsilon^{n+1}||_{\delta_n - \rho} &\le K||\Delta_\varepsilon^n||_{\delta_n - \rho}^2 + ||\Delta^n||_{\delta_n - \rho} ||S_\varepsilon^n||_{\delta_n} + ||R_\varepsilon^n||_{\delta_n} e^{-\delta_n 2^n} \le \\ &\le K\rho^{-2\nu} ||R_\varepsilon^n||_{\delta_n} (||R_\varepsilon^n||_{\delta_n}^{1/2} + ||R_\varepsilon^n||_{\delta_n}^{-1/4} ||S_\varepsilon^n||_{\delta_n} + e^{-\delta_n 2^n}). \end{aligned}$$

We can also improve the reducibility equations. First we note that we do not want to reduce the matrix $DF_\varepsilon(u_\varepsilon^n)$ but rather $DF_\varepsilon(u_\varepsilon^{n+1})$. The difference between the two can be bounded as follows.

$$\begin{aligned} &M_\varepsilon^n(\theta+\omega) DF_\varepsilon(u_\varepsilon^{n+1}(\theta))(M_\varepsilon^n)^{-1}(\theta) - c_\varepsilon^n = \\ &\quad = S_\varepsilon^n + M_\varepsilon^n(\theta+\omega)[DF_\varepsilon(u_\varepsilon^{n+1}(\theta)) - DF_\varepsilon(u_\varepsilon^n(\theta))](M_\varepsilon^n)^{-1}(\theta) \equiv \tilde{S}_\varepsilon^n(\theta), \end{aligned}$$

and we have $||\tilde{S}^n_\varepsilon||_{\delta_n-\rho} \leq ||S^n_\varepsilon||_{\delta_n-\rho}+K||\Delta^n_\varepsilon||_{\delta_n-\rho} \leq ||S^n_\varepsilon||_{\delta_n-\rho}+K\rho^{-2\nu}||R^n_\varepsilon||^{3/4}_{\delta_n}$. If we write $M^{n+1}_\varepsilon = M^n_\varepsilon + \alpha^n_\varepsilon$, $c^{n+1}_\varepsilon = c^n_\varepsilon + D^n_\varepsilon$, substitute in (19) and keep only the terms that are linear in the small quantities, we are lead to the equation

$$\alpha^n_\varepsilon(\theta+\omega)c^n_\varepsilon - c^n_\varepsilon\alpha^n_\varepsilon(\theta) = -\tilde{S}^n_\varepsilon(\theta) + D^n_\varepsilon. \tag{26}$$

Again, this equation can be analyzed using Fourier series. Taking Fourier coefficients, we obtain:

$$\begin{aligned} e^{2\pi ik\omega}\hat{\alpha}^n_{\varepsilon,k}c^n_\varepsilon - c^n_\varepsilon\hat{\alpha}^n_{\varepsilon,k} &= \hat{\tilde{S}}^n_{\varepsilon,k}, \quad k \neq 0, \\ \hat{\alpha}^n_{\varepsilon,0}c^n_\varepsilon - \hat{\alpha}^n_{\varepsilon,0} &= -\hat{\tilde{S}}^n_{\varepsilon,0} + D^n_\varepsilon. \end{aligned} \tag{27}$$

We note that the equations for $k = 0$ are very easy to solve. We just set $D^n_\varepsilon = \hat{\tilde{S}}^n_{\varepsilon,0}$ and $\hat{\alpha}^n_{\varepsilon,0} = 0$. To solve the first of the two equations in (27), we note that the linear operator on $\mathcal{M}_{4\times 4}$ defined by $\alpha \mapsto \mu\alpha c - c\alpha$ has eigenvalues precisely $\mu\lambda^i - \lambda^j$, where $\{\lambda^i\}$ are the eigenvalues of c.

Then, (27) can be solved on the set $\mathcal{S}^n$ when $|k| \leq 2^n$ and we can bound the resulting Fourier coefficients in the same way that we bounded the other equation and we have

$$\begin{aligned} ||S^{n+1}_\varepsilon||_{\delta_n-\rho} &\leq K||\alpha^n_\varepsilon||^2_{\delta_n-\rho} + ||\tilde{S}^n_\varepsilon||_{\delta_n}e^{-\delta_n 2^n} \leq \\ &\leq K\rho^{-2\nu}(||S^n_\varepsilon||_{\delta_n} + ||R^n_\varepsilon||_{\delta_n})^2 + K(||S^n_\varepsilon||_{\delta_n} + K\rho^{2\nu}||R^n_\varepsilon||^{3/4}_{\delta_n})e^{-\delta_n 2^n}. \end{aligned}$$

The proof of convergence is standard in KAM theory. If we set $\rho_n = \delta_0 10^{-3}/n^2$ the bounds we have developed for R^n_ε, S^n_ε imply that provided that the initial residuals are small enough, the procedure can be iterated indefinitely and, moreover we have $||R^n_\varepsilon||_{\delta_n}$, $S^n_{\delta_n} \leq (A_0)^{\alpha^n}$ for some $0 < A_0 < 1$ and some $\alpha > 1$ and, similarly, the increments $||\Delta^n_\varepsilon||_{\delta_n}, ||\alpha^n_\varepsilon||_{\delta_n} \leq (A_0)^{\alpha^n}$. This means that we can pass to the limit and check that the final sum converges.

To finish the proof of Theorem 2, we just need to describe how the set $\mathcal{S}^{n+1}$ is produced. Note that in going form $\mathcal{S}^n$ to $\mathcal{S}^{n+1}$ we have to cope with two problems. One is that resonances change because c^n_ε changes. This can be controlled because $|c^{n+1}_\varepsilon - c^n_\varepsilon|$ is controlled by $||\tilde{S}^n_\varepsilon||_{\delta_n}$ hence, to exclude the resonances it suffices to enlarge the existing balls by an amount controlled by this quantity. The second problem we have to cope with is the fact that the number of resonances that have to be considered grows. In view of the conditions, of derivatives being bounded from below, we see that we have to exclude at most 3×2^n balls whose radius is bounded from above by $||R^n_\varepsilon||^{1/4}_{\delta_n}$. To control the amount of ε we have to give up, it is important to note that altering ε indeed changes the resonances. We have two types of resonances: $\lambda^\varepsilon_i - e^{2\pi ik\omega}$ that comes from solving (24) and

$e^{2\pi ik\omega}\lambda_i^\varepsilon - \lambda_j^\varepsilon$ that comes from solving (26). We also have to take care of the second condition in (20), whose meaning will become clear.

We can move from the first resonances by changing ε because $\frac{d}{d\varepsilon}\lambda_i^\varepsilon - e^{2\pi ik\omega} = \frac{d}{d\varepsilon}\lambda_i^\varepsilon$, which is under control because of the first bound in iii). Similarly, we can take care of the second condition in (20). For the second resonances, if λ_i^ε and λ_j^ε correspond to symplectically conjugated eigenvalues, we have $\lambda_j^\varepsilon = 1/\lambda_i^\varepsilon$ so $\frac{d}{d\varepsilon}\left(e^{2\pi ik\omega}\lambda_i^\varepsilon - \lambda_j^\varepsilon\right) = \left(\frac{d}{d\varepsilon}\lambda_i^\varepsilon\right)\left(e^{2\pi ik\omega} + (\lambda_i^\varepsilon)^{-2}\right)$. The first factor is bounded by the assumption on iii) and the second is bounded by the assumption on $\mathcal{S}$. If λ_i^εa and λ_j^ε do not correspond to symplectically conjugate eigenvectors, we have $\left|\frac{d}{d\varepsilon}\left(e^{2\pi ik\omega}\lambda_i^\varepsilon - \lambda_j^\varepsilon\right)\right| \geq \left|\ \left|\frac{d}{d\varepsilon}\lambda_i^\varepsilon\right| - \left|\frac{d}{d\varepsilon}\lambda_j^\varepsilon\right|\ \right|$ which is bounded from below by assumption iii).

We call attention to the fact that this procedure indicates that the balls we have to exclude to produce $\mathcal{S}$ have centers which are quite close to the resonant points of the unperturbed system. Similarly, the radius of a resonance of order $r \approx 2^n$ is $\approx (A_0)^{(1/4)\alpha^n} = (A_0)^{-Kr^a}$.

Q.E.D

We now argue that, under geometric assumptions (such as the map F_ε being exact symplectic) the translation has to vanish.

For the Froeschlé map, the argument is quite elementary and we give it first. If we choose the space along which to translate, the space $(0,k)$ in the Hamiltonian description, the Lagrangian version of the conclusions of Theorem 2 becomes

$$g_\varepsilon(\theta+\omega) + g_\varepsilon(\theta-\omega) - 2g_\varepsilon(\theta) = \varepsilon\nabla V(\theta k + g_\varepsilon(\theta) + ak,$$

for a real number a. If we multiply (dot product) this equation by $k + g_\varepsilon'(\theta)$ and integrate over θ and we use:

$$\begin{gathered}
\int k\cdot g_\varepsilon(\theta+\omega)\,d\theta = \int k\cdot g_\varepsilon(\theta-\omega)\,d\theta = \int k\cdot g_\varepsilon(\theta)\,d\theta,\\
\int g_\varepsilon'(\theta+\omega)\cdot g_\varepsilon(\theta)\,d\theta = -\int g_\varepsilon(\theta+\omega)\cdot g_\varepsilon'(\theta)\,d\theta = -\int g_\varepsilon(\theta)\cdot g_\varepsilon'(\theta-\omega)\,d\theta,\\
\int g_\varepsilon(\theta)\cdot g_\varepsilon'(\theta)\,d\theta = \tfrac{1}{2}\int \tfrac{d}{d\theta}|g_\varepsilon(\theta)|^2\,d\theta = 0,\\
\int \nabla V(\theta k + g_\varepsilon(\theta))\cdot(k + g_\varepsilon'(\theta))\,d\theta = \int \tfrac{d}{d\theta}V(\theta k + g_\varepsilon(\theta))\,d\theta = 0,
\end{gathered}$$

which can be readily proved using changes of variables, integration by parts or the fundamental theorem of calculus, we obtain

$$0 = a|k|^2.$$

This elementary calculation is a prototype of the following "*vanishing*" lemma.

Lemma 4 *Let γ be a symplectic form admitting a symplectic potential $\gamma = d\mu$. Let F be an exact symplectic mapping (i.e. $F_*\mu = \mu + dS$). Assume that for some parameterization u of $\mathbb{T}^n$ we have:*

$$F(u(\theta)) = u(\theta + \omega) + R(u(\theta + \omega)), \tag{28}$$

where for some $c > 0$ and for all $\lambda \in [0,1]$ we have:

$$\sum_i \gamma \left(\frac{\partial u}{\partial \theta_i} + \lambda \frac{\partial}{\partial \theta_i} R(u(\theta)), R(\theta) \right) \geq c|R(u(\theta))|. \tag{29}$$

Then, $R \equiv 0$.

The meaning of (28) is that the points in the torus get mapped into other points on the torus displaced by an amount R. The geometric meaning of (29) in Lemma 4 is that the translation by R is, roughly, in the direction of the symplectic conjugate to the tangent to the torus. Note that the condition is open in the sense that if it is verified by a parameterization u, field of displacements R, then it is verified by others which are close to them.

For the Froeschlé map, and the perturbed tori, if we take as the direction of the perturbations the k direction in the momentum and the tori are close to $u(\theta) = (\theta k + g^0, \omega)$ and $R(\theta) = (0, a(\theta)k)$ we see that $\gamma(\frac{\partial u}{\partial \theta} + \lambda R, R) = a(\theta)|k|^2$ so that the condition is easily verified for the perturbative tori of the Froeschlé map with the translations that we needed to prove the theorem.

Proof. Denote by $\hat{\theta}_i = (\theta_1, \cdots, \theta_{i-1}, \theta_{i+1}, \cdots, \theta_n)$ and by $\sigma_{i,\hat{\theta}_i}$ the loop defined by $\sigma_{i,\hat{\theta}_i}(\theta) = u(\theta_1, \cdots, \theta_{i-1}, \theta, \theta_{i+1}, \cdots, \theta_n)$. Since F is exact symplectic, we have

$$\int_{\sigma_{i,\hat{\theta}_i}} F_*\mu = \int_{\sigma_{i,\hat{\theta}_i}} \mu + dS = \int_{\sigma_{i,\hat{\theta}_i}} \mu.$$

On the other hand, if we compute using the definition,

$$\int_{\sigma_{i,\hat{\theta}_i}} F_*\mu = \int_{F(\sigma_{i,\hat{\theta}_i})} \mu = \int_{\tilde{\sigma}_{i,\hat{\theta}+\omega_i}} \mu = \int_{\sigma_{i,\hat{\theta}+\omega_i}} \mu + \int_{\Sigma_{i,\hat{\theta}_i}} \gamma,$$

where we have denoted by $\tilde{\sigma}_{i,\hat{\theta}_i}$ the loop defined by $\tilde{\sigma}_{i,\hat{\theta}_i}(\theta) = u(\theta_1, \cdots, \theta_{i-1}, \theta, \theta_{i+1}, \cdots, \theta_n) + R(u(\theta_1, \cdots, \theta_{i-1}, \theta, \theta_{i+1}, \cdots, \theta_n))$ and by $\Sigma_{i,\hat{\theta}_i}$ the surface whose parameterization is given by $\Sigma_{i,\hat{\theta}_i}(\theta, \lambda) = (1 - \lambda)\sigma_{i,\hat{\theta}_i}(\theta) + \lambda \tilde{\sigma}_{i,\hat{\theta}_i}(\theta)$. The crucial property of Σ is that $\partial \Sigma_{i,\hat{\theta}_i} = \sigma_{i,\hat{\theta}_i} - \tilde{\sigma}_{i,\hat{\theta}_i}$.

If we compute with this parameterization, we have:

$$\int_{\Sigma_{i,\hat{\theta}_i}} \gamma = \int_0^1 \int_0^1 \gamma \left(\frac{\partial}{\partial \theta} \Sigma_{i,\hat{\theta}_i}, \frac{\partial}{\partial \lambda} \Sigma_{i,\hat{\theta}_i} \right) = \int_0^1 \int_0^1 \gamma \left(\frac{\partial}{\partial \theta_i} u(\theta) + \lambda R(u(\theta)), R(u(\theta)) \right).$$

Summing over i and using the inequalities (29), we obtain that

$$\sum_i \int_{\sigma_{i,\hat{\theta}_i}} d\mu - \sum_i \int_{\sigma_{i,\hat{\theta}+\omega_i}} d\mu =$$

$$= \int_0^1 \int_0^1 \sum_i \gamma \left(\frac{\partial}{\partial \theta_i} u(\theta) + \lambda R(u(\theta)), R(u(\theta)) \right) \geq \int_0^1 \int_0^1 c|R(u(\theta + \omega))|. \tag{30}$$

If we now integrate (30) over all the values of $\hat{\theta}_i$, and note that the R.H.S. has zero integral by a change of variables, we obtain that $R \equiv 0$.
Q.E.D

We can apply Theorem 2 to the perturbation series constructed in Section 3. We take as a new variable $\sigma = \sqrt{\varepsilon}$. If we truncate the series up to order N, we see that in a sufficiently small neighborhood of 0, we can obtain remainders that are $O(|\sigma|^{2N})$ whereas the non-degeneracy constants remain uniformly bounded. Hence, applying Theorem 2 we obtain invariant tori that are analytic functions of σ as well as their reductions.

Nevertheless, we observe that the approximate tori themselves are analytic functions of ε. On the other hand, the approximate reduction has a non-trivial monodromy and changes the stable and unstable directions when we go around the origin in the complex ε plane. It is possible to show that the iterative procedure we just detailed does not change the monodromy. This is because the tori are not changed by the iterative procedure more than an amount $O(|\varepsilon|^N)$, but the same argument we have used shows that they are unique in these scales. Similarly, we see that the monodromy of the stable and unstable manifolds cannot collapse since the approximate stable and unstable manifolds differ by an amount that can be bounded from below by $K|\varepsilon|^{1/2}$ and the changes produced by the iterative procedure can be estimated from above by $|\varepsilon|^N$.

6. Some remarks about whiskered tori

The statement of Theorem 2 can be improved when the approximately invariant tori are whiskered. That is, when all the directions that are not corresponding to the directions tangent to the torus or their symplectic conjugate are either contracting or expanding.

In such a case, we do not need to assume approximate reducibility. It suffices to assume that the contracting and expanding directions form an angle bounded from below. In dimension 4, since the stable and unstable directions are one dimensional, this implies reducibility, but in higher dimensions, it is strictly more general.

The basic idea is that in this situation – which we will describe in more detail now – we can solve (21) without going through (22) and, hence, we do not need to consider (23). The second and third small divisors conditions in (20) were only used to achieve reducibility. Hence, they can be omitted from the assumptions.

We will only sketch the proof. Since this result is most interesting in dimensions at least 6, we will assume that this is the case. The notation does not change.

We will assume that for every θ we can find a splitting

$$F_\varepsilon(u^n_\varepsilon(\theta)) = \{\partial_\theta u^n_\varepsilon(\theta), J\partial_\theta u^n_\varepsilon(\theta)\} \oplus E^s_{n,\varepsilon,\theta} \oplus E^u_{n,\varepsilon,\theta},$$

and such that (to avoid cluttering the notation, we will suppress the indices n,ε that play little role)

$$\begin{aligned} DF(u(\theta))E^{s,u}_\theta &= E^{s,u}_{\theta+\omega}, \\ ||DF(u(\theta))|_{E^s_\theta}|| &\le \mu < 1, \\ ||(DF(u(\theta+\omega)))^{-1}|_{E^u_\theta}|| &\le \mu < 1. \end{aligned}$$

Considering the map $\theta \mapsto \theta + \omega$, $v \mapsto DF(u(\theta)v$ and extending slightly the theory of invariant foliations derived in [5] to include dependence on parameters and perturbations (this is done in [11]), we obtain that the splitting above depends analytically on the base point and on the parameter ε. If we now take projections over the invariant subspaces, and using the invariance properties of the splitting, (21) is equivalent to

$$\begin{aligned} DF(u(\theta))\Pi^s_\theta\Delta(\theta) - \Pi^s_{\theta+\omega} &= -\Pi^s_{\theta+\omega}R(\theta), \\ DF(u(\theta))\Pi^u_\theta\Delta(\theta) - \Pi^u_{\theta+\omega} &= -\Pi^u_{\theta+\omega}R(\theta), \\ DF(u(\theta))\Pi^c_\theta\Delta(\theta) - \Pi^c_{\theta+\omega} &= -\Pi^c_{\theta+\omega}R(\theta) + B. \end{aligned} \tag{31}$$

Using the contraction properties of DF on the spaces, we see that the first two equations in (31) admit the solution:

$$\begin{aligned} \Pi^s_\theta\Delta(\theta) &= \Pi^s_\theta R(\theta-\omega) + DF(u(\theta-\omega))\Pi^s_{\theta-\omega}R(\theta-2\omega) + \\ &\quad +DF(u(\theta-\omega))DF(u(\theta-2\omega)\Pi^s_{\theta-\omega}R(\theta-3\omega) + \cdots, \\ \Pi^u_\theta\Delta(\theta) &= -(DF(u(\theta)))^{-1}\Pi^u_{\theta+\omega}(\theta)R(\theta) - \\ &\quad -(DF(u(\theta)))^{-1}(DF(u(\theta+\omega)))^{-1}\Pi^u_{\theta+2\omega}(\theta)R(\theta+\omega) + \cdots. \end{aligned}$$

The uniform convergence of these series readily justifies that they are indeed solutions and the analytic dependence. As for the third equation of (31), we see that it becomes an equation of the type we considered in Theorem 1 up to high order errors.

Note that this construction does not use reducibility and, hence the KAM theory needs only the same conditions as the perturbation theory.

We note that this result can also be obtained using transformation theory (see [19]). We point out that this simplified proof when coupled with the estimates for the Lindstedt perturbation series can prove the analyticity of the tori in a domain that excludes an arbitrarily thin wedge around the real positive semi-axis. To obtain some of the elliptical tori, we need the more refined result given by Theorem 2.

7. Acknowledgments

R. de la Llave has been supported by NSF, TARP grants and a Centennial Fellowship from the AMS and a URI from UT. A. Jorba has been partially supported by the Spanish grant DGICYT PB94–0215, the EC grant ERBCHRXCT940460 and the Catalan CIRIT grant GRQ93–1135. The collaboration between the authors has been supported by the NATO grant CRG 950273 and the Catalan CIRIT grant BE94/III–513.

References

1. Craig, W. and Wayne, C. E. (1993) Newton's method and periodic solutions of nonlinear wave equatons. *Comp. Pure and Appl. Math.* **46**, pp. 1409–1498.
2. Delshams, A. and Gutiérrez, P. (1996) Effective stability and KAM Theorem, *J. of Diff. Equations*, **128**, pp. 415–490.
3. Chierchia, L. and Falcolini, C. (1995) A Direct Proof of a Theorem by Kolmogorov in Hamiltonian Systems, *Ann. Scuo. Norm. Pisa*, **4**, pp. 541–593.
4. Eliasson, L. H. (1996) Absolutely convergent expansions for quasi-periodic motions, *Math. Phys. Elec. Jour.* **2-4**, pp. 1-33.
5. Fenichel, N. (1973) Asymptotic stability with rate conditions. *Ind. Math. Jour.* **23** pp. 1109–1137.
6. Gallavotti, G. (1994) Twistless KAM theorem, quasi-flat intersections, and other cancelations in the perturbation series of certain integrable systems. A review. *Rev. Math. Phys.* **3** pp. 343-411.
7. Gutiérrez, P. (1995) Estabilitat efectiva i tors invariants de sistemes hamiltonians quasi-integrables, Ph. D. Thesis, Univ. de Barcelona.
8. Jorba, A. and Simó, C. (1992) On the reducibility of linear differential equations with quasiperiodic coefficients, *Jour. Diff. Eq.*, **98:1**, pp. 111–124
9. Jorba, A. and Simó, C. (1995) On quasiperiodic perturbations of elliptic equilibrium points, to appear in *SIAM J. on Math. Anal.*
10. Llave, R. and Rana D. (1991) Accurate strategies for K. A. M. bounds and their implementation, in *Computer Aided proofs in Analysis, K. Meyer, D. Schmidt ed.*, Springer Verlag.
11. Llave, R. and Wayne, C.E. (1988) Whiskered and low dimensional tori in nearly integrable Hamiltonian systems, *Preprint.*
12. R. Kirkorian (1996) Thése de doctorat, Ecole Ploytechnique.
13. Pöschel, J. (1986) Invariant manifolds of complex analytic mappings, Les Houches (XLIII), Vol. II, Ed. K. Osterwalder and R. Stora, North Holland, pp. 949–964.
14. Poincaré, H. (1899) *Les méthodes nouvelles de la mécanique céleste* Gauthier Villars, Paris
15. Rana, D. (1987) Proof of accurate upper and lower bounds to stability domains in small denominator problems, Ph. D. Thesis, Princeton Univ.
16. Rüssmann, H. (1976) On a new proof of Moser's twist mapping theorem, *Cel. Mech.*

14, pp. 19–31

17. Rüssmann, H. (1976) On optimal estimates for the solutions of linear difference equations on the circle, *Cel. Mech.*, **14**, pp. 33–38.
18. Treshev, D.V. (1991) A mechanism of the destruction of resonance tori in Hamiltonian systems, *Math. U.S.S.R. Sbor.* **68**, pp. 181-203
19. Zehnder, E. (1976) Generalized implicit function theorems with applications to some small divisor problems II, *Comm. Pure and Appl. Math.* **29** pp. 49–111