ON QUASIPERIODIC PERTURBATIONS OF ELLIPTIC EQUILIBRIUM POINTS*

ÀNGEL JORBA† AND CARLES SIMÓ‡

Abstract. This work focusses on quasiperiodic time-dependent perturbations of ordinary differential equations near elliptic equilibrium points. This means studying

$$\dot{x} = (A + \varepsilon Q(t, \varepsilon))x + \varepsilon g(t, \varepsilon) + h(x, t, \varepsilon),$$

where A is elliptic and h is $\mathcal{O}(x^2)$. It is shown that, under suitable hypothesis of analyticity, nonresonance and nondegeneracy with respect to ε, there exists a Cantorian set $\mathcal{E}$ such that for all $\varepsilon \in \mathcal{E}$ there exists a quasiperiodic solution such that it goes to zero when ε does. This quasiperiodic solution has the same set of basic frequencies as the perturbation. Moreover, the relative measure of the set $[0, \varepsilon_0] \setminus \mathcal{E}$ in $[0, \varepsilon_0]$ is exponentially small in ε_0. The case $g \equiv 0$, $h \equiv 0$ (quasiperiodic Floquet theorem) is also considered.

Finally, the Hamiltonian case is studied. In this situation, most of the invariant tori that are near the equilibrium point are not destroyed, but only slightly deformed and "shaken" in a quasiperiodic way. This quasiperiodic "shaking" has the same basic frequencies as the perturbation.

Key words. quasiperiodic perturbations, elliptic points, quasiperiodic solutions, small divisors, quasiperiodic Floquet theorem, KAM theory

AMS(MOS) subject classifications. 34C27, 34C50, 58F27, 58F30

1. Introduction. In this work we will consider autonomous differential equations under quasiperiodic time-dependent perturbations, near an elliptic equilibrium point. The kind of equations we shall deal with is

$$\dot{x} = (A + \varepsilon Q(t, \varepsilon))x + \varepsilon g(t, \varepsilon) + h(x, t, \varepsilon),$$

where A is assumed to be elliptic (that is, all the eigenvalues are purely imaginary and non zero), h is of second order in x and the system is autonomous when $\varepsilon = 0$. At this point we recall the definition of quasiperiodic function:

DEFINITION 1.1. *A function f is a quasiperiodic function with basic frequencies $\omega_1, \ldots, \omega_r$ if $f(t) = F(\theta_1, \ldots, \theta_r)$, where F is 2π periodic in all its arguments and $\theta_j = \omega_j t$ for $j = 1, \ldots, r$.*

We assume that the quasiperiodic functions appearing in our equations are analytical. For definiteness we give the following

DEFINITION 1.2. *A function f is analytic quasiperiodic on a strip of width ρ if it is quasiperiodic and F (see Definition 1.1) is analytical for $|\mathrm{Im}\ \theta_j| \le \rho$ for $j = 1, \ldots, r$. In this case we denote by $\|f\|_\rho$ the norm*

$$\sup\{|F(\theta_1, \ldots, \theta_r)| \ \text{ with } \ |\mathrm{Im}\ \theta_j| \le \rho,\ 1 \le j \le r\}.$$

This kind of equations appears in many problems. As an example, we can consider the equations of the motion near the Equilateral Libration points of the Earth-Moon system, including (quasiperiodic) perturbations coming from the noncircular motion

* This paper was written on October 28th, 1994.

† Departament de Matemàtica Aplicada I, ETSEIB, Universitat Politècnica de Catalunya, Diagonal 647, 08028 Barcelona, Spain (E-mail: jorba@ma1.upc.es).

‡ Departament de Matemàtica Aplicada i Anàlisi, Universitat de Barcelona, Gran Via 585, 08007 Barcelona, Spain (E-mail: carles@maia.ub.es).

of the Moon and the effect of the Sun (see [13], [5], [11], [12], [10] and [15]). In those works, some seminumerical methods have been applied to compute a quasiperiodic orbit replacing the equilateral relative equilibrium point (this means that, when the perturbation tends to zero, that quasiperiodic orbit tends to the libration point), but there is a lack of theoretical support to ensure that the methods used are really convergent, and the computed quasiperiodic orbit really exists. In Section 2 the existence of that dynamical equivalent is shown for a Cantorian set (of positive measure) of values of ε. Another problem related to this is the study of the stability of that quasiperiodic solution. In order to do this a kind of Floquet theory is available (see [16]), that now can be obtained as a result of the more general study presented here (see Section 2).

We also want to note that the Floquet theorem for the quasiperiodic case has already been considered in many papers. An approach similar to ours (based in KAM techniques) can be found in [3]. There, the reducibility to constant coefficients is studied for the case in which A is an hyperbolic matrix. For the case in which A is elliptic, they give some bounds on the measure of the set of matrices Q for which the system can be reduced to constant coefficients. The bounds on that measure, however, are not so good as the ones that can be derived from the work presented here.

Another approach to the reducibility of quasiperiodic linear equations can be found in [14]. The methods used there are not based in KAM techniques and the results can be applied to systems that are not close to constant coefficients. The main drawback is that the hypothesis used are quite restrictive,[1] and are very difficult to check in a practical example.

Finally, it is interesting to consider the Hamiltonian case. In Section 3 we show that most of the KAM tori of the autonomous system still persist when the quasiperiodic time-dependent perturbation is added.

Studies of this kind for the case of the one dimensional Schrödinger equation can be found in the literature (see, for instance [6], [18], [4] or [8]) and some of the methods and ideas used here are already contained in these papers. Note that, as the "unperturbed" problem is a harmonic oscillator, it is possible to obtain better results (see, e.g. [8]). Some of the ideas of the present paper could be already found in [17] and [7], although they deal with slightly different problems.

2. A dynamical equivalent to elliptic equilibrium points.

In this Section, we are going to focus on the equation

$$\dot{x} = (A + \varepsilon Q(t,\varepsilon))x + \varepsilon g(t,\varepsilon) + h(x,t,\varepsilon), \tag{1}$$

where the time-dependence is quasiperiodic with vector of basic frequencies $\omega = (\omega_1, \ldots, \omega_r)$ and analytic on a strip of width $\rho_0 > 0$. The reader should remind that h is of second order in x. We want to stress that the equations are not required to be Hamiltonian (the Hamiltonian case will be considered in Section 3).

2.1. The inductive scheme.

In order to study equation (1), let us perform some changes to simplify it. First of all, we shall try to eliminate the independent term $g(t)$ by means of quasiperiodic changes of variables. To do this, we shall need a scheme with quadratic convergence (otherwise the small divisors effect would make the method divergent). This kind of schemes are based in Newton method, that is, to linearize the problem in a known approximation of the solution, solve this linear problem and take this solution as a new (better) approximation to the solution we

[1] For instance, the system $\dot{x} = (A + \varepsilon Q(t,\varepsilon))x$, where A is elliptic and ε is small does not satisfy the required hypothesis.

are looking for. These algorithms can overcome the effect of the small divisors and ensure convergence on certain regions. To apply this method to our problem we have to consider the linearized problem (we take as initial guess the zero solution, and we linearize around this point):

$$\dot{x} = (A + \varepsilon Q(t, \varepsilon))x + \varepsilon g(t, \varepsilon).$$

We are looking for a quasiperiodic solution $\underline{x}(t, \varepsilon)$ with basic frequencies the ones of g and Q such that $\lim_{\varepsilon \to 0} \underline{x}(t, \varepsilon) = 0$. At this point we note that we do not need to know $\underline{x}(t, \varepsilon)$ exactly, because an approximation of order ε is enough. This is another property of Newton method: we do not need to know the Jacobian matrix exactly but some approximation of it, and it is enough that this approximation be of the order of the independent term we want to make zero. In our case, this can be done easily by considering the linear system

$$\dot{x} = Ax + \varepsilon g(t, \varepsilon). \tag{2}$$

Here we need a nonresonance condition. The usual one is

$$|(k, \omega)\sqrt{-1} - \lambda_i| > \frac{c}{|k|^{\gamma_0}}, \tag{3}$$

where λ_i are the eigenvalues of A and $|k| = |k_1| + \cdots + |k_r|$. Condition (3) as well as condition (7), an additional diophantine condition needed later, will be discussed in detail through Section 2.2.

Let us call $\underline{x}(t, \varepsilon)$ the solution of (2) which is quasiperiodic with respect to t (with basic frequencies the ones of g) and of order ε. The existence of that solution will be shown by Lemma 2.10. Now we can perform the change of variables $x = \underline{x}(t, \varepsilon) + y$ to equation (1) to obtain

$$\dot{y} = (A + \varepsilon Q_1(t, \varepsilon))y + \varepsilon^2 g_1(t, \varepsilon) + h_1(y, t, \varepsilon), \tag{4}$$

where, if $\varepsilon \neq 0$,

$$\begin{aligned} Q_1(t, \varepsilon) &= Q(t, \varepsilon) + \frac{1}{\varepsilon} D_x h(\underline{x}(t, \varepsilon), t, \varepsilon), \\ g_1(t, \varepsilon) &= \frac{1}{\varepsilon^2} h(\underline{x}(t, \varepsilon), t, \varepsilon) + \frac{1}{\varepsilon} Q(t, \varepsilon)\underline{x}(t, \varepsilon), \\ h_1(y, t, \varepsilon) &= h(\underline{x}(t, \varepsilon) + y, t, \varepsilon) - h(\underline{x}(t, \varepsilon), t, \varepsilon) - D_x h(\underline{x}(t, \varepsilon), t, \varepsilon)y. \end{aligned}$$

Note that this process can not be (successfully) iterated: now we need a solution of

$$\dot{y} = (A + \varepsilon Q_1(t, \varepsilon))y + \varepsilon^2 g_1(t, \varepsilon). \tag{5}$$

with an accuracy of order ε^2, and if we take the kind of approximation given by equation (2) (that is, dropping Q_1), we will have a divergent scheme. This is because in this way one obtains linear convergence in ε, which is overcomed by the effect of the small divisors.

To deal with this difficulty we perform a new change of variables to get something like $\varepsilon^2 Q_2$ instead of εQ_1. This can be done as follows: let us define the average of Q_1 as

$$\overline{Q}_1(\varepsilon) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} Q_1(t, \varepsilon)\, dt.$$

For the existence of the limit see, for instance, [9]. Consider now equation (5) after averaging with respect to t and some rearrangement,

$$\dot{y} = (\overline{A}(\varepsilon) + \varepsilon\widetilde{Q}_1(t,\varepsilon))y + \varepsilon^2 g_1(t,\varepsilon),$$

where $\widetilde{Q}_1(t,\varepsilon) = Q_1(t,\varepsilon) - \overline{Q}_1(\varepsilon)$, $\overline{A}(\varepsilon) = A + \varepsilon\overline{Q}_1(\varepsilon)$. Now we need to find a quasiperiodic solution of

$$\dot{P} = \overline{A}P - P\overline{A} + \widetilde{Q}_1, \tag{6}$$

with the same basic frequencies than $\widetilde{Q}_1$. This can be done if the eigenvalues of $\overline{A}$ satisfy a diophantine condition. The one used in [16] was

$$|(k,\omega)\sqrt{-1} - \overline{\lambda}_i + \overline{\lambda}_j| > \frac{c}{|k|^{\gamma_0}}. \tag{7}$$

Then, making the change of variables $y = (I + \varepsilon P)z$ (I denotes the identity matrix) to equation (4) (these changes of variables have already been considered in [3], [16] and [15]) we obtain the equation

$$\dot{z} = (\overline{A}(\varepsilon) + \varepsilon^2 Q_2(t,\varepsilon))z + \varepsilon^2 g_2(t,\varepsilon) + h_2(z,t,\varepsilon), \tag{8}$$

where $Q_2(t,\varepsilon) = (I + \varepsilon P(t,\varepsilon))^{-1}\widetilde{Q}_1 P(t,\varepsilon)$, $g_2(t,\varepsilon) = (I + \varepsilon P(t,\varepsilon))^{-1} g_1(t,\varepsilon)$ and $h_2(z,t,\varepsilon) = (I + \varepsilon P(t,\varepsilon))^{-1} h_1((I + \varepsilon P(t,\varepsilon))z, t, \varepsilon)$. Now, using $\dot{z} = \overline{A}z + \varepsilon^2 g_2(t)$ we are able to find an approximate solution of (8) with an accuracy of order ε^2 that allows to proceed with the Newton method. In this way, after n steps (each step is composed of the two changes of variables just explained above) the equation will look like

$$\dot{x}_n = (A_n(\varepsilon) + \varepsilon^{2^n} Q_n(t,\varepsilon))x_n + \varepsilon^{2^n} g_n(t,\varepsilon) + h_n(x_n,t,\varepsilon),$$

Then, if the norms of A_n, Q_n, g_n and h_n do not grow too fast with n, the scheme will be convergent to an equation like

$$\dot{y} = A_\infty(\varepsilon)y + h_\infty(y,t,\varepsilon),$$

That equation has the trivial solution $y = 0$ and this shows that, in the original system of equations, the origin is replaced by a quasiperiodic orbit whose basic frequencies are the ones of the perturbation. Note that we have also obtained the linearized flow (given by the "Floquet" matrix A_∞) around that quasiperiodic solution.

2.2. The resonances. We recall that the small divisors conditions needed at each step (to compute the changes of variables) are

$$|(k,\omega)\sqrt{-1} - \lambda_i| > \frac{c}{|k|^{\gamma_0}}, \quad |(k,\omega)\sqrt{-1} - \lambda_i + \lambda_j| > \frac{c}{|k|^{\gamma_0}}. \tag{9}$$

The first condition is needed to solve equations like (2) and the second one for equations like (6). Note that the eigenvalues λ_i are changed at each step of the process (because A is changed), and this implies that we do not know in advance if they will satisfy the diophantine conditions for all the steps.

To deal with this problem we need to have some control on the variation of the eigenvalues at each step. To explain the main idea, let us focus on the equation (2). As we are assuming that the eigenvalues of A verify the condition (3), at the first step we can solve the equation and proceed. In the second step, when we need to solve the same equation, we find that the matrix has been (slightly) changed: now it is $\overline{A}(\varepsilon) = A + \varepsilon\overline{Q}_1(\varepsilon)$. So, as the eigenvalues of $\overline{A}$ are different to the ones of A, we can not assure that they satisfy the condition (3).

To explain how to overcome this difficulty, let us denote by $\overline{\lambda}_i(\varepsilon)$, $i = 1, \dots, d$, the eigenvalues of the matrix $\overline{A}(\varepsilon)$. Let us write $\overline{\lambda}_i(\varepsilon)$ as

$$\overline{\lambda}_i(\varepsilon) = \lambda_i + \lambda_i^{(1)}\varepsilon + \lambda_i^{(2)}\varepsilon^2 + \cdots, \tag{10}$$

where λ_i is an eigenvalue of the unperturbed matrix A. If we look at $\overline{\lambda}_i(\varepsilon)$ as a function of ε, we can avoid the resonant values of $\overline{\lambda}_i(\varepsilon)$ by avoiding the corresponding values of ε. This implies that, to take out a (Cantor-like) set of resonant values of $\overline{\lambda}_i(\varepsilon)$ (this set is the usual union of small intervals centered in the values (k, ω)) is equivalent to take out the corresponding (by (10)) values of ε. In order to bound the measure of the "resonant" values of ε, we will ask that the relation (10) be Lipschitz from below with respect to ε. We also want to note that we need to take out values of ε at each step of the inductive process, so we need to have that condition at each step. Let us look at this: at first sight, it seems enough to ask for $\lambda_i^{(1)} \neq 0$, because this value is produced by the first averaging, so

1. it is left unchanged by all the others steps of the inductive procedure,
2. it can be computed easily at the beginning (it is a verifiable hypothesis).

The problem is that if we take out a Cantor-like set at each step, the dependence of $\overline{\lambda}_i(\varepsilon)$ of ε is not differentiable (because $\overline{\lambda}_i(\varepsilon)$ is defined only on a set with empty interior), and we do not even know if it is continuous. So, it is not obvious how to derive the Lipschitz condition that we need.

To deal with the latter difficulty, what we will do is to show explicitly that, at each step, the relation (10) is Lipschitz (note that the definition of Lipschitz holds perfectly on sets with empty interior). This will allow us to control the measure of the set of ε we are taking out.

Finally, we want to note that this technique has to be applied twice at each step: one for equations like (2) and one for equations like (6).

2.3. The measure of the resonant set. Another important point is to bound the measure of the set of values of ε too close to resonance. To do it, we will assume that ε belongs to an interval $[0, \varepsilon_0]$, being ε_0 small enough. What we are going to show is that it is possible to bound the measure of the set of resonant values of ε by a quantity exponentially small in ε_0. To simplify the discussion, we are going to focus again on the equation (2), so the corresponding small denominator is $(k, \omega)\sqrt{-1} - \lambda_i$.

The usual procedure is to use the bound given by (3), because it is good enough to produce convergence and to give a positive measure set of admissible frequencies (this has been done in [15] and [16]). Note that the size of the set of resonant values of λ is given by the bound of the small divisors.[2] This implies that we should try to select that value as small as possible. But, on the other hand, this value will appear in the denominators of the Fourier series. So, if it is too small, we will not be able to prove convergence.

[2] For instance, in the case of (3) the set of resonant values of λ is

$$\bigcup_{|k|\neq 0} B\left((k,\omega)\sqrt{-1}, \frac{c}{|k|^{\gamma_0}}\right),$$

where $B(a, r)$ denotes the ball (in the complex plane) centered in a with radius r.

The condition we have used is

$$|(k,\omega)\sqrt{-1}-\lambda_i| > \frac{c}{|k|^{\gamma_n}}e^{-\nu_n|k|} = D(k,n),$$

where γ_n has been taken equal to $\gamma_0 z^n$ $(1 < z < 2)$ and ν_n is $\nu_0/(n+1)^2$. Here, n denotes the actual step of the inductive process. To start the discussion of this expression, let us remark that the measure of the resonant set of λ at each step n is given by $\sum_{k\neq 0} 2D(k,n)$, $k \in \mathbb{Z}^r$, and the total measure is

$$\sum_{n\geq 0}\sum_{k\in\mathbb{Z}^r\setminus\{0\}} 2\frac{c}{|k|^{\gamma_n}}e^{-\nu_n|k|}. \tag{11}$$

So, a first condition we need is that those sums are convergent.

Before continuing with the discussion of $D(k,n)$, let us explain first where the exponentially small character (of the set of resonant values of ε) comes from. This will make clear (we hope) the reasons to choose an expression like $D(k,n)$.

As we have said before, the eigenvalues of the matrix A move at each step of the inductive process, in an amount of $\mathcal{O}(\varepsilon)$. Let us call $I_i(\varepsilon)$ the interval (with diameter $\mathcal{O}(\varepsilon)$) where the eigenvalue number i moves. This implies that, if the eigenvalues of the unperturbed matrix satisfy a condition like (3), the values (k,ω) are outside $I_i(\varepsilon)$ if $|k| < N(\varepsilon)$, for a suitable value $N(\varepsilon)$ (another way of saying this is that the values (k,ω) can not approach λ_i too fast, because of (3)). For that reason, we do not need to take out resonances with $|k| < N(\varepsilon)$, and this leads to the fact that, in (11), it is enough to start the sum in k when $|k| \geq N(\varepsilon)$. And this implies that, if the expression $D(k,n)$ decays exponentially with $|k|$, we will obtain something exponentially small with ε_0.

This is the reason of putting something like $\exp(-\nu|k|)$ in $D(k,n)$. As this value will appear in the denominators of the corresponding Fourier series, we will have the factor $\exp(\nu|k|)$ multiplying the coefficients of those series. This will produce a reduction of the analyticity strip of the series: the width will go from ρ to $\rho-\nu$. Of course, after a few steps, the functions will not be analytic. So all the inductive process will be over. To avoid that problem we have chosen ν depending on the actual step ($\nu_n = \nu_0/(n+1)^2$), in such a way that the total reduction on the analyticity strip remains bounded (of course, other selections of ν_n are possible, but they do not change the final result).

The next step is to realize that, with this selection of ν_n, the exponential goes to 1 when n goes to infinity. So we need to add some factor in front of the exponential, to ensure that the sum with respect to n is still exponentially small. For this reason we have added the factor $c/|k|^{\gamma_n}$. The selection $\gamma_n = \gamma_0 z^n$ is not the only one (one can use, for instance, $\gamma_n = \gamma_0 n^j$, for some j), but the results, with the present choice, seem to be better than for other choices. Finally, the value z has to be taken between 1 and 2. If it is taken equal to 2, then the divisor is too small and we are not able to guarantee convergence. This is seen clearer inside the proofs.

Finally, all this procedure is applied (at each step) in the same way for equation (6), using the same exponential bound for the denominators.

2.4. Some remarks. Before finishing this overview of the paper, it is interesting to remark the following: as the equations we are dealing with are not necessarily Hamiltonian, it is possible that, in some step of the inductive process, the eigenvalues of the matrix A leave the imaginary axis. In this case, we do not need to worry about resonances from this step onwards. As we can not know in advance if this is going to happen, we have considered during all the proof the worst case, that is, the eigenvalues are always on the imaginary axis. On the other hand, if the initial matrix A is partially elliptic and partially hyperbolic, the results are still valid. In the hyperbolic case they are, of course, much better: they hold for a full interval $[0, \varepsilon_0]$.

In some cases it is possible that at the first step of the inductive process the eigenvalues leave the imaginary axis (this is the general case, really). Theorem 2.4 ensures that this case can be detected averaging the original system and looking for the new equilibrium point of this autonomous system. The linearized equations around that point and the "Floquet" matrix (A_∞) of the quasiperiodic orbit differ in $\mathcal{O}(\varepsilon^2)$.

Another interesting point is to compare what we are doing here with the proof of the KAM Theorem. In the proof of the KAM Theorem (see, for instance, [1]) we use the action variables as a parameters, to avoid resonances. Here we use the eigenvalues of the matrix A but, as we can not move them directly, we move them by means of the single parameter ε. Note that the nondegeneracy condition of the KAM (non-zero Jacobian of the frequencies with respect to the actions) says, basically, that we can control the frequencies through the actions. Here, we want to control the eigenvalues by means of ε, so we ask for a suitable Lipschitz condition. As it is well known (see, for instance, [2]), the nondegeneracy condition of the KAM theorem can be relaxed to a second order condition. Here it is possible to do something similar (instead of asking for $\lambda_i^{(1)} \neq 0$ in (10), we can allow $\lambda_i^{(1)} = 0$ but asking for $\lambda_i^{(2)} \neq 0$, or even a higher order condition), and the estimates on the measure of the Cantorian set of ε are obtained in a similar way (in fact, the estimates can be even better). It is also remarkable that the scheme of the proof we are using is quite similar to the one of the KAM Theorem ([1]).

Finally, note that if the nonlinearity h and the independent term g of the initial equation (1) are both equal to zero, what we have is a Floquet theorem. Now, the result obtained is better than the one contained in [16]. There it was shown that the measure of the set of "resonant" $\varepsilon \in [0, \varepsilon_0]$ was $o(\varepsilon_0)$, and here is proved to be exponentially small with ε_0.

2.5. Theorems. From now on, if $x \in \mathbb{R}^n$ we denote by $||x||$ the sup norm of x. If A is a matrix, $||A||$ denotes the corresponding sup norm.

THEOREM 2.1. *Consider the differential equation*

$$\dot{x} = (A + \varepsilon Q(t, \varepsilon))x + \varepsilon g(t, \varepsilon) + h(x, t, \varepsilon), \tag{12}$$

where $Q(t, \varepsilon)$, $g(t, \varepsilon)$ and $h(x, t, \varepsilon)$ depend on time in a quasiperiodic way, with basic frequencies $(\omega_1, \ldots, \omega_r)^t$, $r \geq 2$, and $|\varepsilon| < \varepsilon_0$. We assume that A is a constant $d \times d$ matrix with d different eigenvalues λ_i and $\det A \neq 0$. Let us suppose that $h(x, t, \varepsilon)$ is analytic with respect to x on the ball $B_\tau(0)$, $h(0, t, \varepsilon) = 0$ and $D_x h(0, t, \varepsilon) = 0$. Moreover, we assume that

1. *Q, g and h are analytic with respect to t on a strip of width $\rho_0 > 0$, and they depend on ε in a Lipschitz way.*
2. *$||D_{xx} h(x, t, \varepsilon)|| \leq K$, where $||x|| \leq \tau$, $|\varepsilon| \leq \varepsilon_0$ and t belongs to the strip defined in 1.*
3. *The vector $(\lambda_1, \ldots, \lambda_d, \sqrt{-1}\omega_1, \ldots, \sqrt{-1}\omega_r)$ satisfies the nonresonance conditions*

$$|(k, \omega)\sqrt{-1} - \lambda_i| > \frac{2c}{|k|^{\gamma_0}}, \quad |(k, \omega)\sqrt{-1} - \lambda_i + \lambda_j| > \frac{2c}{|k|^{\gamma_0}}.$$

for all $1 \le i, j \le d$, $k \in \mathbb{Z}^r \setminus \{0\}$, $c > 0$ *and* $\gamma_0 \ge r-1$. *As usual,* $|k|$ *is taken as* $|k| = |k_1| + \cdots + |k_r|$.

4. *Let us denote by* $\underline{x}(t,\varepsilon)$ *the unique analytical quasiperiodic solution of* $\dot{x} = Ax + \varepsilon g(t,\varepsilon)$ *such that* $\lim_{\varepsilon\to 0} \underline{x}(t,\varepsilon) = 0$ *(the existence of that solution is shown by Lemma 2.10), and define*

$$\underline{A}(\varepsilon) = A + \varepsilon\overline{Q}(\varepsilon) + \overline{D_x h(\underline{x}(t,\varepsilon),t,\varepsilon)}.$$

Let $\lambda_j^0(\varepsilon)$, $j = 1,\ldots,d$ *be the eigenvalues of* $\underline{A}$. *We require the existence of* $\overline{\delta}$, $\delta > 0$ *such that*

$$\frac{\overline{\delta}}{2}|\varepsilon_1 - \varepsilon_2| > |\lambda_i^0(\varepsilon_1) - \lambda_j^0(\varepsilon_1) - (\lambda_i^0(\varepsilon_2) - \lambda_j^0(\varepsilon_2))| > 2\delta|\varepsilon_1 - \varepsilon_2| > 0,$$

$$\frac{\overline{\delta}}{2}|\varepsilon_1 - \varepsilon_2| > |\lambda_k^0(\varepsilon_1) - \lambda_k^0(\varepsilon_2)| > 2\delta|\varepsilon_1 - \varepsilon_2| > 0,$$

for all i, j, k *satisfying* $1 \le i < j \le d$, $1 \le k \le d$ *and provided that* $|\varepsilon_1|$ *and* $|\varepsilon_2|$ *are less than some small value* ε_0.

Then there exists a Cantorian set $\mathcal{E} \subset (0,\varepsilon_0)$ *with positive Lebesgue measure such that the equation (12) can be transformed, by means of a change of variables, into*

$$\dot{y} = A_\infty(\varepsilon)y + h_\infty(y,t,\varepsilon),$$

where A_∞ *is a constant matrix and* $h_\infty(y,t,\varepsilon)$ *is of second order in* y. *If* ε_0 *is small enough the relative measure of* $(0,\varepsilon_0)\setminus\mathcal{E}$ *in* $(0,\varepsilon_0)$ *is less than* $\exp(-c_1/\varepsilon_0^{c_2})$ *for* $c_1 > 0$ *and* $c_2 > 0$ *(independent of* ε_0*), where* c_2 *is any number such that* $c_2 < 1/\gamma_0$. *Furthermore the quasiperiodic change of variables that performs this transformation is analytic with respect to* t *and it has the same basic frequencies than* Q, g *and* h.

Remark 1. In the hypothesis 3 we use $2c$ instead of the usual c in the diophantine condition in order to simplify the notation inside the proofs.

Remark 2. During the proof of this Theorem, it will be supposed that $\rho_0 \ge 1 + \pi^2/6$. This condition can be achieved introducing a new time $\tau = st$, where

$$s = \max\left\{\frac{1+\frac{\pi^2}{6}}{\rho_0},\ 1\right\}.$$

This scaling may change the constant c and, therefore the set $\mathcal{E}$ is scaled by the same factor.

Remark 3. For fixed values of λ_i, $i = 1,\ldots,d$, $\lambda_i \ne \lambda_j$ if $i \ne j$, hypothesis 3 is not satisfied for any c only for a set of values of ω of zero measure if $\gamma_0 > r-1$.

Remark 4. If $r = 1$, that is, if the perturbation is periodic, no small divisors appear, if ε is small enough, and the results hold for all $\varepsilon \in (0,\varepsilon_0)$. The proof is classical. We shall assume, without explicit mention, that $r \ge 2$ in what follows.

Corollary 2.2. *Under the hypothesis of Theorem 2.1, there exists a Cantorian set* $\mathcal{E} \subset (0,\varepsilon_0)$ *with positive Lebesgue measure (and the complementary being exponentially small) such that the equation (12) has a quasiperiodic solution* $x_\varepsilon(t)$ *with basic frequencies* $(\omega_1,\ldots,\omega_r)$, *such that*

$$\lim_{\substack{\varepsilon\to 0\\ \varepsilon\in\mathcal{E}}} \|x_\varepsilon\| = 0.$$

Corollary 2.3. *(A Floquet theorem) Consider the linear differential equation*

$$\dot{x} = (A + \varepsilon Q(t, \varepsilon))x, \tag{13}$$

where $Q(t,\varepsilon)$ depends quasiperiodically on time, with basic frequencies $(\omega_1, \ldots, \omega_r)^t$, $r \geq 2$, and $|\varepsilon| < \varepsilon_0$. We assume that A is a constant $d \times d$ matrix with d different eigenvalues λ_i and $\det A \neq 0$. Moreover, we assume that

1. *Q is analytic with respect to t on a strip of width $\rho_0 > 0$, and it depends on ε in a Lipschitz way.*
2. *The vector $(\lambda_1, \ldots, \lambda_d, \sqrt{-1}\omega_1, \ldots, \sqrt{-1}\omega_r)$ satisfies the nonresonance condition*

$$|(k,\omega)\sqrt{-1} - \lambda_i + \lambda_j| > \frac{2c}{|k|^{\gamma_0}}.$$

   for all $1 \leq i,\, j \leq d$, $k \in \mathbb{Z}^r \setminus \{0\}$, $c > 0$ and $\gamma_0 \geq r-1$.
3. *Let us define*

$$\underline{A}(\varepsilon) = A + \varepsilon\overline{Q}(\varepsilon).$$

   Let $\lambda_j^0(\varepsilon)$, $j = 1, \ldots, d$ be the eigenvalues of $\underline{A}$. We require the existence of $\overline{\delta}$, $\delta > 0$ such that

$$\frac{\overline{\delta}}{2}|\varepsilon_1 - \varepsilon_2| > |\lambda_i^0(\varepsilon_1) - \lambda_j^0(\varepsilon_1) - (\lambda_i^0(\varepsilon_2) - \lambda_j^0(\varepsilon_2))| > 2\delta|\varepsilon_1 - \varepsilon_2| > 0,$$

   for all i, j, k satisfying $1 \leq i < j \leq d$, $1 \leq k \leq d$ and provided that $|\varepsilon_1|$ and $|\varepsilon_2|$ are less than some small value ε_0.

Then there exists a Cantorian set $\mathcal{E} \subset (0, \varepsilon_0)$ with positive Lebesgue measure such that the equation (13) can be reduced to a system with constant coefficients

$$\dot{y} = A_\infty(\varepsilon)y,$$

by means of a change of variables $x = (I + \varepsilon P(t,\varepsilon))y$, where I is the identity matrix and P is analytic and quasiperiodic with respect to t, having ω as a vector of basic frequencies. If ε_0 is small enough the relative measure of $(0,\varepsilon_0) \setminus \mathcal{E}$ in $(0, \varepsilon_0)$ is less than $\exp(-c_1/\varepsilon_0^{c_2})$ for $c_1 > 0$ and $c_2 > 0$ (independent of ε_0), where c_2 is any number such that $c_2 < 1/\gamma_0$.

Remark. This corollary is the result of taking $g \equiv h \equiv 0$ in Theorem 2.1. We have also weakened the nonresonance condition. This fact becomes clear by looking into the lines of the proof for that Theorem.

Theorem 2.4. *Let us consider the equation (12) and let us assume that all the hypothesis of Theorem 2.1 hold. Moreover, let us assume that the nonlinear part $h(x,t,\varepsilon)$ is of class $\mathcal{C}^2$ with respect to ε and $h(x,t,0) \equiv h(x)$. Then, if ε is sufficiently small, the averaged system*

$$\dot{y} = (A + \varepsilon\overline{Q})y + \varepsilon\overline{g} + \overline{h}(y, \varepsilon)$$

has an equilibrium point $x_0(\varepsilon)$ such that

1. $\lim_{\varepsilon \to 0} \|x_0(\varepsilon)\| = 0.$
2. *The matrix A_{x_0} of the linearized system around $x_0(\varepsilon)$ and the matrix A_∞ obtained in Theorem 2.1 satisfy $\|A_{x_0} - A_\infty\| = \mathcal{O}(\varepsilon^2)$.*

COROLLARY 2.5. *Let us define $\lambda_i^{x_0}$, $1 \le i \le d$, as the eigenvalues of the matrix A_{x_0} defined in Theorem 2.4. Then, under the hypothesis of Theorem 2.4, an equivalent version of the hypothesis 4 in Theorem 2.1 is obtained if λ_i^0 are replaced by $\lambda_i^{x_0}$.*

Proofs of the results above have been splitted in several parts in order to simplify the reading. Section 2.6 contains lemmas needed to show the convergence of the iterative scheme used to obtain Theorem 2.1. Section 2.7 presents the convergence proof. Up to this point we do not worry about the measure of the set of values of ε which has to be taken out. Section 2.8 includes the lemmas used to prove that the matrix A depend on ε in a Lipschitz way, at each step of the procedure. The lemma used to bound the measure of the Cantorian set where Theorem 2.1 holds is given in Section 2.9. Section 2.10 actually states the bounds for that measure, and finally, Section 2.11 is devoted to Theorem 2.4.

2.6. Convergence lemmas. In what follows, we will use that an analytic quasi-periodic function $f(t)$ on a strip of width ρ, having $\omega = (\omega_1, \ldots, \omega_r)$ as vector of basic frequencies, has Fourier coefficients defined by

$$f_k = \frac{1}{(2\pi)^r} \int_{\mathbb{T}^r} F(\theta_1, \ldots, \theta_r) e^{-(k,\theta)\sqrt{-1}}\, d\theta,$$

(F has been defined in Definition 1.1), such that f can be expanded as

$$f(t) = \sum_{k\in\mathbb{Z}^r} f_k e^{(k,\omega)\sqrt{-1}t},$$

For all t such that $|\mathrm{Im}\ t| < \rho/||\omega||$. Moreover, the analyticity of f implies that

$$|f_k| \le ||f||_\rho e^{-\rho|k|}.$$

LEMMA 2.6. *Let $\delta \in]0, 1]$, $\alpha \ge 1$. Let us define the function*

$$\chi(s) = \left(\frac{s-1}{e}\right)^{s-1} \sqrt{s-1}.$$

Then

$$\sum_{k\in\mathbb{Z}^r} |k|^\alpha e^{-\delta|k|} \le \frac{20r}{3\delta^{r+\alpha}} \chi(r+\alpha).$$

Proof. We shall use that $\#\{k \in \mathbb{Z}^r \ / \ |k| = m\} \le 2rm^{r-1}$. This is checked immediately for $m = 1$ or for $r \le 3$. Then one uses induction with respect to r for $m \ge 2$. Then we obtain

$$\sum_{k\in\mathbb{Z}^r} |k|^\alpha e^{-\delta|k|} \le 2r \sum_{m=0}^{\infty} m^{r+\alpha-1} e^{-\delta m} = (\triangle)$$

As the unique maximum of $g(x) = x^{r+\alpha-1}e^{-\delta x}$ is reached when $x = \frac{r+\alpha-1}{\delta}$, we can bound the sum above by this maximum plus the integral:

$$\begin{aligned}
(\triangle) \ &\le\ 2r\left[\left(\frac{r+\alpha-1}{\delta e}\right)^{r+\alpha-1} + \frac{1}{\delta^{r+\alpha}}\Gamma(r+\alpha)\right] = \\
&=\ \frac{2r}{\delta^{r+\alpha}}\left(\frac{r+\alpha-1}{e}\right)^{r+\alpha-1}\left[\delta + \left(\frac{e}{r+\alpha-1}\right)^{r+\alpha-1}\Gamma(r+\alpha)\right] < \\
&<\ \frac{2r}{\delta^{r+\alpha}}\left(\frac{r+\alpha-1}{e}\right)^{r+\alpha-1}\frac{10}{3}\sqrt{r+\alpha-1} = \frac{20r}{3}\frac{\chi(r+\alpha)}{\delta^{r+\alpha}}.
\end{aligned}$$

□

LEMMA 2.7. *Let $h : U \subset \mathbb{R}^d \to \mathbb{R}^d$ be a function of class $\mathcal{C}^2$ on a ball $B_\tau(0)$, satisfying that $h(0) = 0$, $D_x h(0) = 0$ and $\|D_{xx}h(x)\| \leq K$, where $x \in B_\tau(0)$. Then $\|h(x)\| \leq \frac{K}{2}\|x\|^2$ and $\|D_x h(x)\| \leq K\|x\|$.*

The proof follows from Taylor's formula.

LEMMA 2.8. *Let M be a diagonal matrix with d diferent nonzero eigenvalues μ_j, $j = 1, \ldots, d$, and $\alpha = \min\{\min_{i,j;i\neq j} |\mu_i - \mu_j|, \min_i |\mu_i|\}$. Let N be a matrix such that $(d+1)\|N\| < \alpha$. Let σ_j, $j = 1, \ldots, d$ be the eigenvalues of $M + N$, B a suitable matrix such that $B^{-1}(M + N)B = D = \mathrm{diag}(\sigma_j)$ with condition number $C(B)$. Then*

1. *$\beta = \min\{\min_{i,j;i\neq j} |\sigma_i - \sigma_j|, \min_i |\sigma_i|\} \geq \alpha - 2\|N\|$.*
2. *$C(B) \leq \frac{\alpha+(d-3)\|N\|}{\alpha-(d+1)\|N\|}$. In particular, if $\|N\| < \frac{\alpha}{3d-1}$ then $C(B) < 2$.*

Proof. It can be found in [16] or [15]. □

LEMMA 2.9. *Let A_0 be a $d \times d$ matrix such that $\mathrm{Spec}(A_0) = \{\lambda_1^0, \ldots, \lambda_d^0\}$, $|\lambda_i^0| > 2\mu$, $|\lambda_i^0 - \lambda_j^0| > 2\mu$, $i \neq j$, where $\mu > 0$. Let B_0 be a regular matrix such that $B_0^{-1}A_0B_0 = D_0 = diag(\lambda_1^0, \ldots, \lambda_d^0)$. Let us define $\beta_0 = \max\{\|B_0\|, \|B_0^{-1}\|\}$ and let α be a value such that $0 < \alpha < 2\mu/((3d-1)\beta_0^2)$. Then, if A verifies $\|A - A_0\| < \alpha$, the following conditions hold:*

1. *$\mathrm{Spec}(A) = \{\lambda_1, \ldots, \lambda_d\}$, and $|\lambda_i| > \mu$, $|\lambda_i - \lambda_j| > \mu$, $i \neq j$.*
2. *There exists a nonsingular matrix B such that $B^{-1}AB = \mathrm{diag}(\lambda_1, \ldots, \lambda_d)$ satisfying $\|B\| \leq \beta$ and $\|B^{-1}\| \leq \beta$, with $\beta = 2\beta_0$.*

Proof. Let A be a matrix, and we write $A = A_0 + (A - A_0)$. Then $B_0^{-1}AB_0 = D_0 + N$, where $N = B_0^{-1}(A - A_0)B_0$. Here we can apply Lemma 2.8 to obtain *1*, if $\|A - A_0\| \leq 2\mu/((d+1)\beta_o^2)$. Note that Lemma 2.8 states that the condition number of the matrix C that diagonalizes $D_0 + N$ is less than 2, provided that $\|A - A_0\| < 2\mu/((3d-1)\beta_0^2)$. In this case, the matrix that diagonalizes A can be obtained multiplying B_0 by C. Hence, its norm can be bounded by $2\beta_0$. □

In the next lemmas the parameters γ and ν are assumed to be positive.

LEMMA 2.10. *Let us consider the equation $\dot{x} = Ax + \varepsilon g(t)$, where A is a $d \times d$ matrix belonging to the ball $B_\alpha(A_0) \subset \mathcal{L}(\mathbb{R}^d, \mathbb{R}^d)$ with α as given by Lemma 2.9, $g(t) = (g_i(t))_{1\leq i\leq d}$ and $g_i(t)$ is an analytic quasiperiodic function on a strip of width ρ_1:*

$$g_i(t) = \sum_{k\in\mathbb{Z}^r} g_i^k e^{(k,\omega)\sqrt{-1}t}.$$

Let us assume that $|(k,\omega)\sqrt{-1} - \lambda_i| > \frac{c}{|k|^\gamma}e^{-\nu|k|}$ $\forall\lambda_i \in \mathrm{Spec}(A)$. Let ρ_2 such that $0 < \rho_2 < \rho_1 - \nu$ and $\delta = \rho_1 - \rho_2 - \nu \leq 1$. Then, there exists a unique quasiperiodic solution of $\dot{x} = Ax + \varepsilon g(t)$ having the same basic frequencies than g and satisfying

$$\|x\|_{\rho_2} \leq \varepsilon\|g\|_{\rho_1} L_1,$$

where $L_1 = 4\beta_0^2[\frac{1}{\mu} + \frac{20r}{3c}\frac{\chi(r+\gamma)}{\delta^{r+\gamma}}]$ and μ and β_0 are defined in Lemma 2.9.

Remark. In this Lemma as well as in the forthcoming ones, we consider A, Q, g, h depending also on ε (see Theorem 2.1) but, for simplicity, we do not write this explicitly.

Proof. Let B the matrix found in Lemma 2.9. Making the change of variables $x = By$ and defining $h(t) = B^{-1}g$ the equation becomes

$$\dot{y} = Dy + \varepsilon h(t).$$

As D is a diagonal matrix, we can handle this equation as d unidimensional equations, that can be solved easily: if $y = (y_i)_{1\le i\le d}$ and

$$y_i(t) = \sum_{k\in\mathbb{Z}^r} y_i^k e^{(k,\omega)\sqrt{-1}t},$$

the coefficients must be $y_i^k = \varepsilon \frac{h_i^k}{(k,\omega)\sqrt{-1}-\lambda_i}$, and they can be bounded by

$$|y_i^k| \le \begin{cases} \varepsilon \frac{\|h\|_{\rho_1}}{\mu} & \text{if } k = 0 \\ \varepsilon \|h\|_{\rho_1} \frac{|k|^\gamma}{c} e^{-(\rho_1-\nu)|k|} & \text{if } k \ne 0. \end{cases}$$

Now, we need to bound the norm $\|y\|_{\rho_2}$. Let t be a complex value such that $|\text{Im } \omega_i t| \le \rho_2$ (for all i). Then

$$|y_i(t)| \le \sum_{k\in\mathbb{Z}^r} |y_i^k|\, |e^{(k,\omega t)\sqrt{-1}}| \le \varepsilon \frac{\|h\|_{\rho_1}}{\mu} + \sum_{k\ne 0} \varepsilon \|h\|_{\rho_1} \frac{|k|^\gamma}{c} e^{-(\rho_1-\nu)|k|} e^{\rho_2|k|}.$$

Setting $\delta = \rho_1 - \rho_2 - \nu$, we can use Lemma 2.6 to bound the sum above:

$$|y_i(t)| \le \varepsilon \|h\|_{\rho_1} \left[\frac{1}{\mu} + \frac{20r\chi(r+\gamma)}{3c\delta^{r+\gamma}}\right].$$

As $\|h\|_{\rho_1} \le \|B^{-1}\| \|g\|_{\rho_1}$ and $\|x\|_{\rho_2} \le \|B\| \|y\|_{\rho_2}$ the result follows. □

LEMMA 2.11. *Let us consider the equation $\dot{P} = AP - PA + Q$, where $A \in B_\alpha(A_0)$, $Q = (q_{ij})$ where $q_{ij}(t)$ are analytic quasiperiodic functions on a strip of width ρ_1:*

$$q_{ij}(t) = \sum_{k\in\mathbb{Z}^r} q_{ij}^k e^{(k,\omega)\sqrt{-1}t}.$$

We also assume that Q has average equal to 0 and $|(k,\omega)\sqrt{-1} - \lambda_i + \lambda_j| > \frac{c}{|k|^\gamma} e^{-\nu|k|}$ $\forall \lambda_i \in \text{Spec}(A)$. Let ρ_2 such that $0 < \rho_2 < \rho_1 - \nu$ and $\delta = \rho_1 - \rho_2 - \nu \le 1$. Then, there exists a unique quasiperiodic solution of $\dot{P} = AP - PA + Q$ having the same basic frequencies than Q and satisfying

$$\|P\|_{\rho_2} \le \|Q\|_{\rho_1} L_2,$$

where $L_2 = 16\beta_0^2 \frac{20r\chi(r+\gamma)}{3c\delta^{r+\gamma}}$, and β_0 is defined in Lemma 2.9.

Proof. Let B be the matrix found in Lemma 2.9. Making the change of variables $P = BSB^{-1}$ and defining $R = B^{-1}QB$, the equation becomes

$$\dot{S} = DS - SD + R,$$

the matrix R having zero average. As D is a diagonal matrix, we can handle this equation as d^2 unidimensional equations, that can be solved easily: if $S = (s_{ij})$ and

$$s_{ij}(t) = \sum_{k\in\mathbb{Z}^r\setminus\{0\}} s_{ij}^k e^{(k,\omega)\sqrt{-1}t},$$

the coefficients must be $s_{ij}^k = \frac{r_{ij}^k}{(k,\omega)\sqrt{-1}-\lambda_i+\lambda_j}$, and they can be bounded by

$$|s_{ij}^k| \le \|r_{ij}\| \frac{|k|^\gamma}{c} e^{-(\rho_1-\nu)|k|}$$

Now, we need to bound the norm $||S||_{\rho_2}$. Let t be a complex value such that $|\text{Im }\omega_i t| \le \rho_2$ (for all i):

$$|s_{ij}(t)| \le \sum_{k\in\mathbb{Z}^r} |s_{ij}^k|\, |e^{(k,\omega t)\sqrt{-1}}| \le \sum_{k\ne 0} ||r_{ij}||_{\rho_1} \frac{|k|^\gamma}{c} e^{-(\rho_1-\nu)|k|} e^{\rho_2|k|}.$$

Now we can use Lemma 2.6, setting $\delta = \rho_1 - \rho_2 - \nu$, to bound the sum above:

$$|s_{ij}(t)| \le \frac{||r_{ij}||_{\rho_1}}{c}\left[\frac{20r\chi(r+\gamma)}{3\delta^{r+\gamma}}\right].$$

As $||P||_{\rho_2} \le ||B|| ||S||_{\rho_2} ||B^{-1}||$, we can use $||R||_{\rho_1} \le ||B^{-1}|| ||Q||_{\rho_1} ||B||$ to obtain the result. □

LEMMA 2.12. *Let us consider $\dot{x} = (A + \varepsilon Q(t))x + \varepsilon g(t) + h(x,t)$, where the time dependence is assumed to be analytic quasiperiodic on a strip of width ρ_1. We also assume that $h(x,t)$ is analytic with respect to x on the ball $B_\tau(0)$ and that satisfies $||D_{xx}h(x,t)||_{\rho_1} \le K$, $\forall x \in B_\tau(0)$. Moreover, $A \in B_\alpha(A_0)$ and $|(k,\omega)\sqrt{-1} - \lambda_i| > \frac{c}{|k|^\gamma} e^{-\nu|k|}$ $\forall \lambda_i \in \text{Spec}(A)$. Let ρ_2 such that $0 < \rho_2 < \rho_1 - \nu$ and $\delta = \rho_1 - \rho_2 - \nu \le 1$. Then, there exists a change of variables $x = y + \underline{x}(t)$ that transforms the initial equation in*

$$\dot{y} = (\overline{A} + \varepsilon\widetilde{Q}_1)y + \varepsilon^2 g_1(t) + h_1(x,t),$$

where $\widetilde{Q}$ has zero average and the following bounds hold:

1. *$||\widetilde{Q}_1||_{\rho_2} \le 2||Q||_{\rho_1} + 2KL_1||g||_{\rho_1}$, where L_1 was defined in Lemma 2.10.*
2. *$||g_1||_{\rho_2} \le \dfrac{KL_1^2||g||_{\rho_1}^2}{2} + L_1||Q||_{\rho_1}||g||_{\rho_1}$.*
3. *$||\overline{A}|| \le ||A|| + \varepsilon(||g||_{\rho_1}KL_1 + ||Q||_{\rho_1})$.*
4. *$||D_{yy}h_1(y,t)||_{\rho_2} \le K$.*
5. *$||\underline{x}||_{\rho_2} \le \varepsilon||g||_{\rho_1}L_1$,*

where $y \in B_{\tau_1}(0)$, $\tau_1 = \tau - ||\underline{x}||_{\rho_2}$ and ε is small enough.

Proof. Let $\underline{x}$ be such that $\dot{\underline{x}} = A\underline{x} + \varepsilon g$. In Lemma 2.10 we obtained

$$||\underline{x}||_{\rho_2} \le \varepsilon||g||_{\rho_1}L_1.$$

Making the change of variables $x = y + \underline{x}(t)$ we get

$$\dot{y} = (A + \varepsilon Q + D_x h(\underline{x}(t),t))y + h(\underline{x}(t),t) + \varepsilon Q\underline{x}(t) + h_1(y,t),$$

where $h_1(y,t) = h(\underline{x}(t) + y,t) - h(\underline{x}(t),t) - D_x h(\underline{x}(t),t)y$. Defining $Q_1 = Q + \frac{1}{\varepsilon}D_x h(\underline{x}(t),t)$ and $g_1 = \frac{1}{\varepsilon^2}h(\underline{x}(t),t) + \frac{1}{\varepsilon}Q\underline{x}(t)$ ($\varepsilon \ne 0$), the equation is then as follows:

$$\dot{y} = (A + \varepsilon Q_1(t))y + \varepsilon^2 g_1(t) + h_1(y,t).$$

To end up, the terms of this equation must be bounded. Let us start with Q_1. Using Lemma 2.7 we get

$$||Q_1||_{\rho_2} \le ||Q||_{\rho_2} + \frac{1}{\varepsilon}K||\underline{x}||_{\rho_2} \le ||Q||_{\rho_1} + ||g||_{\rho_1}KL_1.$$

Now, let us bound $||g_1||_{\rho_2}$, again by means of Lemma 2.7:

$$||g_1||_{\rho_2} \le \frac{1}{\varepsilon^2}\frac{K}{2}||\underline{x}||_{\rho_2}^2 + \frac{1}{\varepsilon}||Q||_{\rho_1}||\underline{x}||_{\rho_2} \le \frac{KL_1^2||g||_{\rho_1}^2}{2} + ||Q||_{\rho_1}||g||_{\rho_1}L_1.$$

Now is the turn of $D_{yy}h_1(y,t)$:

$$||D_{yy}h_1||_{\rho_2} = ||D_{xx}h(\underline{x}(t)+y,t)|| \leq K.$$

To do this we have to require $y \in B_{\tau_1}(0)$, where $\tau_1 = \tau - ||\underline{x}||_{\rho_2}$ (ε is supposed small enough). Now, using that $Q_1(t) = \overline{Q}_1 + \widetilde{Q}_1(t)$ and defining $\overline{A} = A + \varepsilon\overline{Q}_1$ we obtain

$$\dot{y} = (\overline{A} + \varepsilon\widetilde{Q}_1(t))y + \varepsilon^2 g_1(t) + h_1(y,t).$$

Finally,

$$||\overline{A}|| \leq ||A|| + \varepsilon||\overline{Q}_1||_{\rho_2},$$

and taking into account that $||\overline{Q}_1||_{\rho_2} \leq ||Q_1||_{\rho_2}$ and that $||\widetilde{Q}_1(t)||_{\rho_2} \leq 2||Q_1||_{\rho_2}$ the proof is finished. □

LEMMA 2.13. *Let us consider $\dot{x} = (A + \varepsilon Q(t))x + \varepsilon^2 g(t) + h(x,t)$, where the time dependence is assumed to be analytic quasiperiodic on a strip of width ρ_1 and Q has zero average. We also assume that $h(x,t)$ is analytic with respect to x on the ball $B_\tau(0)$ and that satisfies $||D_{xx}h(x,t)||_{\rho_1} \leq K$, $\forall x \in B_\tau(0)$. Moreover, $A \in B_\alpha(A_0)$ and $|(k,\omega)\sqrt{-1} - \lambda_i + \lambda_j| > \frac{c}{|k|^\gamma}e^{-\nu|k|}$ $\forall \lambda_i,\ \lambda_j \in \operatorname{Spec}(A)$. Let ρ_2 such that $0 < \rho_2 < \rho_1 - \nu$ and $\delta = \rho_1 - \rho_2 - \nu \leq 1$. Then, there exists a change of variables $x = (I + \varepsilon P(t))y$, where I is the identity $d \times d$ matrix and $P(t)$ is analytic quasiperiodic on a strip of width ρ_2, that transforms the initial equation in*

$$\dot{y} = (\overline{A} + \varepsilon^2\widetilde{Q}_1)y + \varepsilon^2 g_1(t) + h_1(y,t),$$

where $\widetilde{Q}_1$ has zero average and the following bounds hold:

1. $||\widetilde{Q}_1||_{\rho_2} \leq \dfrac{2||P||_{\rho_2}}{1 - \varepsilon||P||_{\rho_2}}||Q||_{\rho_1}$, *where $||P||_{\rho_2} \leq ||Q||_{\rho_1}L_2$ and L_2 was defined in Lemma 2.11.*
2. $||g_1||_{\rho_2} \leq \dfrac{1}{1 - \varepsilon||P||_{\rho_2}}||g||_{\rho_1}$.
3. $||D_{yy}h_1||_{\rho_2} \leq K\dfrac{(1 + \varepsilon||P||_{\rho_2})^2}{1 - \varepsilon||P||_{\rho_2}}$.
4. $||\overline{A}|| \leq ||A|| + \varepsilon^2\dfrac{||P||_{\rho_2}}{1 - \varepsilon||P||_{\rho_2}}||Q||_{\rho_1}$,

where $y \in B_{\tau_1}(0)$, $\tau_1 = \dfrac{\tau}{1 + \varepsilon||P||_{\rho_2}}$ and ε is small enough.

Proof. Using Lemma 2.11 we can solve $\dot{P} = AP - PA + Q$. The solution that we have found verifies

$$||P||_{\rho_2} \leq ||Q||_{\rho_1}L_2.$$

Now, by means of the change of variables $x = (I + \varepsilon P)y$ and introducing the notation $Q_1 = (I + \varepsilon P)^{-1}QP$, $g_1 = (I + \varepsilon P)^{-1}g$ and $h_1(y,t) = (I + \varepsilon P)^{-1}h((I + \varepsilon P)y,t)$ we obtain the following equation

$$\dot{y} = (A + \varepsilon^2 Q_1(t))y + \varepsilon^2 g_1(t) + h_1(y,t).$$

Now we are going to bound the terms of this equation. For this purpose we need the bound of $||P||_{\rho_2}$ provided by Lemma 2.11 and displayed above.

$$||Q_1||_{\rho_2} \leq \left(\sum_{i=0}^{\infty} \varepsilon^i ||P||_{\rho_2}^i\right) ||Q||_{\rho_2}||P||_{\rho_2} \leq \frac{||P||_{\rho_2}}{1 - \varepsilon||P||_{\rho_2}}||Q||_{\rho_1},$$

$$\|g_1\|_{\rho_2} \le \frac{1}{1-\varepsilon\|P\|_{\rho_2}} \|g\|_{\rho_1},$$

$$\|D_{yy}h_1\|_{\rho_2} \le \frac{1}{1-\varepsilon\|P\|_{\rho_2}} K\|I+\varepsilon P\|_{\rho_2}^2 \le K\frac{(1+\varepsilon\|P\|_{\rho_2})^2}{1-\varepsilon\|P\|_{\rho_2}}.$$

Of course, we require $y \in B_{\tau_1}(0)$, where $\tau_1 = \tau/(1+\varepsilon\|P\|_{\rho_2})$ and ε is small enough. To end up this, we rewrite the equation using $Q_1(t) = \overline{Q}_1 + \widetilde{Q}_1(t)$ and $\overline{A} = A + \varepsilon^2 \overline{Q}_1$ and we obtain

$$\dot{y} = (\overline{A} + \varepsilon^2 \widetilde{Q}_1)y + \varepsilon^2 g_1(t) + h_1(y,t),$$

and we only need to bound $\overline{A}$:

$$\|\overline{A}\| \le \|A\| + \varepsilon^2 \|Q_1\|_{\rho_2}.$$

□

Up to here, we have the main tools to carry out one step of the inductive process. Now we present a lemma which will be used to show the convergence.

LEMMA 2.14. *Let η_n be a sequence of real positive numbers such that*

$$\eta_{n+1} \le (\overline{\gamma} z^n)^{\overline{\gamma} z^n} \eta_n^2,$$

for all $n \ge 0$, where $\overline{\gamma} > 0$, $1 < z < 2$. Then

$$\eta_n \le \left[\left(\overline{\gamma} z^{\frac{z}{2-z}}\right)^{\frac{\overline{\gamma}}{2-z}} \eta_0 \right]^{2^n}.$$

Proof. Taking logarithms we have

$$\begin{aligned}
\log \eta_{n+1} &\le (\overline{\gamma} z^n)\log(\overline{\gamma} z^n) + 2\log\eta_n \le \\
&\le (\overline{\gamma} z^n)\log(\overline{\gamma} z^n) + 2\overline{\gamma} z^{n-1}\log(\overline{\gamma} z^{n-1}) + 4\log\eta_{n-1} \le \dots \le \\
&\le \overline{\gamma} \sum_{j=0}^{n} 2^j z^{n-j} (\log\overline{\gamma} + (n-j)\log z) + 2^{n+1}\log\eta_0 = \\
&= \overline{\gamma}\, 2^{n+1} \log\overline{\gamma} \sum_{l=0}^{n} \frac{z^l}{2^{l+1}} + \overline{\gamma}\, 2^{n+1}\log z \sum_{l=1}^{n} l\frac{z^l}{2^{l+1}} + 2^{n+1}\log\eta_0 \le \\
&\le \overline{\gamma}\, 2^{n+1} \frac{1}{2-z}\log\overline{\gamma} + \overline{\gamma}\, 2^{n+1}\frac{z}{(2-z)^2}\log z + 2^{n+1}\log\eta_0.
\end{aligned}$$

The result follows by exponentiation. □

LEMMA 2.15. *Let $\{a_n\}_n$ be a sequence of positive real numbers satisfying $a_n \in]0,1]$, $\prod_{n=0}^{\infty} a_n = a \in]0,1]$. Let $\{b_n\}_n$ be another sequence of positive real numbers satisfying $\sum_{n=0}^{\infty} b_n = b < +\infty$. Consider the new sequence $\{\tau_n\}_n$ defined by $\tau_{n+1} = a_n\tau_n - b_n$. Then the sequence $\{\tau_n\}_n$ converges to a limit value τ_∞ satisfying $\tau_\infty \ge a\tau_0 - b$.*

Proof. It is easy to see that

$$\tau_{n+1} = \left(\prod_{i=0}^{n} a_i\right)\tau_0 - \sum_{i=0}^{n-1}\left[\left(\prod_{j=i+1}^{n} a_j\right) b_i\right] - b_n.$$

As all the terms appearing in this expression converge, τ_n does. Moreover, using that

$$\prod_{j=i+1}^{n} a_j \leq 1$$

for all n, the result follows. □

2.7. Proof of Theorem 1 (part I). Here we are going to do the proof without worrying about resonances, and then (in the Part II) we shall take out the values of ε for which the proof fails.

First of all, let us denote by A_0 the initial matrix A (see Theorem 2.1) corresponding to the averaged linear part of the differential system. Let μ be a real value such that, if $\mathrm{Spec}(A_0) = \{\lambda_1^0, \ldots, \lambda_d^0\}$, then $|\lambda_i^0| > 2\mu$, $|\lambda_i^0 - \lambda_j^0| > 2\mu$ for all $i \neq j$. Then Lemma 2.9 can be applied to obtain values α and β such that all the matrices contained inside the ball $B_\alpha(A_0) = \{A \,/\, \|A - A_0\| < \alpha\}$ can be diagonalized. Moreover, the matrix B of the diagonalizing change of variables satisfies $\|B\| < \beta$, $\|B^{-1}\| < \beta$. During the proof we shall see that, if ε is small enough, all the matrices A_n that appear during the inductive process are inside that ball.

As we assume that the dependence of Q, g and h with respect to ε is Lipschitz, every time we compute some norm we mean, without explicit mention, that we look not only for the maximum with respect to t in the suitable strip, but also with respect to ε in the allowed range.

To begin the proof, we suppose that we have applied the method exposed before up to step n, and we are going to see that we can apply it again to get the $n+1$ step. In this way we shall obtain bounds for the quasiperiodic part at the n^{th} step and for the transformation at this step, and this allows us to prove the convergence.

We note that in the first step (that is, when the current data are the initial ones) the index n is equal to 0.

Now suppose that we are at n^{th} step. This means that the equation we have is

$$\dot{x}_n = (A_n(\varepsilon) + \varepsilon^{2^n} Q_n(t,\varepsilon))x_n + \varepsilon^{2^n} g_n(t,\varepsilon) + h_n(x_n,t,\varepsilon), \tag{14}$$

where A_n belongs to $B_\alpha(A_0)$, its eigenvalues λ_i verify the nonresonance condition

$$|(k,\omega)\sqrt{-1} - \lambda_i| > \frac{c}{|k|^{\gamma_n}} e^{-\nu_n |k|},$$

where $\gamma_n = \gamma_0 z^n$ $(1 < z < 2)$ and $\nu_n = \nu_0/(n+1)^2$, with $0 < \nu_0 < 1/4$. As we need to reduce the width of the analyticity strip of the quasiperiodic functions, we define $\rho_{n+1} = \rho_n - 1/(n+1)^2$ and $\sigma_n = \rho_n - 1/(2(n+1)^2)$, with $\rho_0 = 1 + \pi^2/6$. During the proof we shall see that the analyticity ball (with respect to x) of $h_n(x,t)$ has to be reduced at each step of the inductive process, and we shall found that, by selecting ε small enough, the limit radius of this ball is positive. Let us define τ_n as this radius at step n. Now we can apply Lemma 2.12 to transform equation (14) into

$$\dot{y}_n = (\widehat{A}_n(\varepsilon) + \varepsilon^{2^n} \widehat{Q}_n(t,\varepsilon))y_n + \varepsilon^{2^{n+1}} \widehat{g}_n(t,\varepsilon) + \widehat{h}_n(y_n,t,\varepsilon), \tag{15}$$

where the width of the analyticity strip has been reduced to σ_n. Now, assuming that the nonresonance condition

$$|(k,\omega)\sqrt{-1} - \lambda_i + \lambda_j| > \frac{c}{|k|^{\gamma_n}} e^{-\nu_n |k|},$$

holds for all λ_i, $\lambda_j \in \mathrm{Spec}(\widehat{A}_n(\varepsilon))$ we can apply Lemma 2.13 to equation (15) and to get

$$(16)\dot{x}_{n+1} = (A_{n+1}(\varepsilon) + \varepsilon^{2^{n+1}} Q_{n+1}(t,\varepsilon))x_{n+1} + \varepsilon^{2^{n+1}} g_{n+1}(t,\varepsilon) + h_{n+1}(x_{n+1}, t, \varepsilon).$$

Now, the width of the analyticity strip has been reduced to ρ_{n+1}. The next step of the proof is to obtain bounds of the terms appearing in equation (16) depending on the bounds of the terms of equation (14).

In what follows, $L_{1,n}$ and $L_{2,n}$ denote the values of L_1 and L_2 as introduced in Lemmas 2.10 and 2.11, when γ, ν and δ are replaced by γ_n, ν_n and $(\frac{1}{2} - \nu_0)/(n+1)^2$, respectively.

Using Lemma 2.13, and the condition $\varepsilon^{2^n} \|P_n\|_{\rho_{n+1}} \leq 1/2$ (see below) we get

$$\|Q_{n+1}\|_{\rho_{n+1}} \leq 4L_{2,n} \|\widehat{Q}_n\|^2_{\sigma_n}.$$

Here we need Lemma 2.12 to bound the expression above, but the bound provided by this Lemma has an "still unknown" term, that is, the bound of the second derivative of h_n. Let us call this value K_n. Note that it is "modified" at each step by Lemma 2.13. In order to bound it, we shall assume that ε is small enough to ensure that $\varepsilon^{2^n} \|P_n\|_{\rho_{n+1}}$ is less than 1/2. This implies that the value of ε will be reduced at each step, if necessary, to guarantee that condition. We will see that this condition is achieved from a certain step onwards, without modifying ε anymore. Therefore, we assume that $K_n \leq (9/2)^n K_0$ (when the convergence will be proved, we shall give a more realistic bound of K_n, that converges to a real number) and Lemma 2.12 states that

$$\|\widehat{Q}_n\|_{\sigma_n} \leq 2\|Q_n\|_{\rho_n} + 2K_n L_{1,n} \|g_n\|_{\rho_n},$$

Now we bound the norm of g_{n+1}:

$$\|g_{n+1}\|_{\rho_{n+1}} \leq 2\|\widehat{g}_n\|_{\sigma_n},$$

and from Lemma 2.12

$$\|g_{n+1}\|_{\rho_{n+1}} \leq K_n L^2_{1,n} \|g_n\|^2_{\rho_n} + 2L_{1,n} \|Q_n\|_{\rho_n} \|g_n\|_{\rho_n}.$$

For simplicity reasons, let us denote $\alpha_n = \|Q_n\|_{\rho_n}$ and $\beta_n = \|g_n\|_{\rho_n}$. This means that we have obtained the following bounds:

$$\begin{aligned} \alpha_{n+1} &\leq 16L_{2,n} \left(\alpha_n + \left(\frac{9}{2}\right)^n L_{1,n}\beta_n\right)^2, \\ \beta_{n+1} &\leq \left(\frac{9}{2}\right)^n L^2_{1,n}\beta^2_n + 2L_{1,n}\alpha_n\beta_n. \end{aligned}$$

To bound α_n and β_n we define $\eta_n = \max\{\alpha_n, \beta_n\}$. As $L_{2,n} < 4L_{1,n}$, after some rearranging we get

$$\begin{aligned} \alpha_{n+1} &\leq 64L_{1,n} \left(1 + \left(\frac{9}{2}\right)^n L_{1,n}\right)^2 \eta^2_n, \\ \beta_{n+1} &\leq \left(\left(\frac{9}{2}\right)^n L^2_{1,n} + 2L_{1,n}\right) \eta^2_n. \end{aligned}$$

As we can assume $c \leq 1$ without adding any additional constraint on the small divisors, we have $L_{1,n} > 1$. Hence

$$\eta_{n+1} < 128 \left(\frac{9}{2}\right)^n L_{1,n}^3 \eta_n^2.$$

It is immediate to check that there exists $\overline{\gamma}$ (depending on γ_0, r, β_0, c, ν_0 and z) such that

$$128 \left(\frac{9}{2}\right)^n L_{1,n}^3 < (\overline{\gamma} z^n)^{\overline{\gamma} z^n}, \quad \forall n \geq 0.$$

Using Lemma 2.14 we have $\eta_n \leq M_1^{2^n}$, where $M_1 = (\overline{\gamma} z^{\frac{z}{2-z}})^{\frac{\overline{\gamma}}{2-z}} \eta_0$. With this, we have proved that

$$\|Q_n\|_{\rho_n} \leq M_1^{2^n}, \quad \|g_n\|_{\rho_n} \leq M_1^{2^n}.$$

Note that this bound allows to ensure that, if $\varepsilon < \varepsilon_1 = M_1^{-1}$

$$\lim_{n\to\infty} \varepsilon^{2^n} \|Q_n\|_{\rho_n} = \lim_{n\to\infty} \varepsilon^{2^n} \|g_n\|_{\rho_n} = 0.$$

The next step is to bound $\|P_n\|_{\rho_{n+1}}$. For this purpose we use first Lemma 2.13 and then Lemma 2.12:

$$\|P_n\|_{\rho_{n+1}} \leq 2L_{2,n}(\|Q_n\|_{\rho_n} + K_n L_{1,n}\|g_n\|_{\rho_n}) \leq 4 \left(\frac{9}{2}\right)^n L_{1,n} L_{2,n} \eta_n.$$

Now, it is not difficult to prove that $L_{1,n} \leq M_2^{2^n}$, $L_{2,n} \leq M_2^{2^n}$, for a suitable constant M_2 (this is shown easily taking logarithms). Hence, we can derive

$$\|P_n\|_{\rho_{n+1}} \leq M_3^{2^n},$$

for a suitable constant M_3. This means that, if $\varepsilon < \varepsilon_1 = \min\{M_1^{-1}, M_3^{-1}\}$, we have

$$\lim_{n\to\infty} \varepsilon^{2^n} \|P_n\|_{\rho_{n+1}} = 0.$$

This allows to have the condition $\varepsilon^{2^n} \|P_n\|_{\rho_{n+1}} < 1/2$, without reducing the value of ε at each step. Now we are going to bound $\|\underline{x}_n\|_{\sigma_n}$:

$$\|\underline{x}_n\|_{\sigma_n} \leq \varepsilon^{2^n} L_{1,n} \|g_n\|_{\rho_n} < \varepsilon^{2^n} M_4^{2^n},$$

for a suitable M_4. When the changes of coordinates have been bounded, we can estimate the decrease of the radius τ_n of the ball where h_n is analytic with respect to x. It has been shown that

$$\tau_{n+1} = \frac{\widehat{\tau}_n}{1 + \varepsilon^{2^n} \|P_n\|_{\rho_{n+1}}} = \frac{1}{1 + \varepsilon^{2^n} \|P_n\|_{\rho_{n+1}}} \tau_n - \frac{\|\underline{x}_n\|_{\rho_n}}{1 + \varepsilon^{2^n} \|P_n\|_{\rho_{n+1}}}.$$

Now, we define

$$a_n = \frac{1}{1 + \varepsilon^{2^n} \|P_n\|_{\rho_{n+1}}}, \quad b_n = \frac{\|\underline{x}_n\|_{\rho_n}}{1 + \varepsilon^{2^n} \|P_n\|_{\rho_{n+1}}}.$$

It is easy to prove that $\prod_{n=0}^{\infty} a_n$ converges:

$$\left|\ln \prod_{n=0}^{N} a_n\right| \leq \sum_{n=0}^{N} |\ln a_n| < \sum_{n=0}^{N} \varepsilon^{2^n} \|P_n\|_{\rho_{n+1}} < \infty \quad \forall N \in \mathbb{N}.$$

As $\sum b_n$ is also convergent, we can apply Lemma 2.15 to get $\tau_\infty \geq a\tau_0 - b$, that is positive if ε is taken small enough.

Now, let us bound $\|A_n\|$:

$$\begin{aligned} \|A_{n+1}\| &\leq \|\widehat{A}_n\| + \varepsilon^{2^n} \frac{\|P_n\|_{\rho_{n+1}}}{1 - \varepsilon^{2^n}\|P_n\|_{\rho_{n+1}}} \\ &\leq \|\widehat{A}_n\| + \varepsilon^{2^n}\|Q_n\|_{\rho_n} + \varepsilon^{2^n} \frac{\|P_n\|_{\rho_{n+1}}}{1 - \varepsilon^{2^n}\|P_n\|_{\rho_{n+1}}}. \end{aligned}$$

Using the bounds found above we can write that

$$\|A_{n+1}\| \leq \|A_n\| + \kappa_n,$$

where $\kappa_n \leq \varepsilon^{2^n} M_5^{2^n}$ for a suitable M_5. As $\sum \kappa_n$ is convergent we can ensure that, if ε is selected small enough, the matrices A_n are always inside the ball $B_\alpha(A_0)$ defined before.

Consider now the value K_n. We have used above the pesimistic bound $K_n \leq (9/2)^n K_0$. Note that this bound does not allow to guarantee the convergence of the functions $h_n(x_n, t)$ to an analytic function $h_\infty(x_\infty, t)$ with respect to x. Now we can use a more accurate bound of that value to get this: From Lemma 2.13 we know that

$$K_{n+1} \leq K_n \frac{(1 + \varepsilon^{2^n}\|P_n\|_{\rho_{n+1}})^2}{1 - \varepsilon^{2^n}\|P_n\|_{\rho_{n+1}}},$$

and, by means of the inequality $1/(1-x) \leq 1 + 2x$ if $0 \leq x \leq 1/2$, we get

$$K_{n+1} \leq \left(1 + 2\varepsilon^{2^n}\|P_n\|_{\rho_{n+1}}\right)^3 K_n.$$

And, using the bounds of $\|P_n\|_{\rho_{n+1}}$ that we already know, it is easy to see that the (bound of the) value K_n converges.

Hence, we have obtained the convergence proof for all $|\varepsilon| < \varepsilon_0$, for a suitable ε_0, without taking into account the "bad set" of values of ε for which the diophantine conditions, at some step n, are not satisfied.

2.8. Lipschitz lemmas. In this Section we have included the lemmas needed to show that, at each step of the inductive process, the dependence on ε of the eigenvalues of the matrix A_n is Lipschitz.

LEMMA 2.16. *Let $f : [-\varepsilon, \varepsilon] \to \mathbb{C}$ be a Lipschitz function from above (with constant L) and from below (with constant l):*

$$|f(x) - f(y)| \leq L|x - y|, \quad |f(x) - f(y)| \geq l|x - y|.$$

Let $g : [-\varepsilon, \varepsilon] \to \mathbb{C}$ be another Lipschitz function from above with constant $\alpha < l$:

$$|g(x) - g(y)| \leq \alpha|x - y|.$$

Then, $h = f + g$ is Lipschitz from above with constant $L + \alpha$, and from below with constant $l - \alpha$:

$$|h(x) - h(y)| \leq (L + \alpha)|x - y|, \quad |h(x) - h(y)| \geq (l - \alpha)|x - y|.$$

Proof. The proof is elementary. □

Remark: From now on, all Lipschitz functions appearing in the text will be Lipschitz from above, unless otherwise stated. Moreover, sometimes we will use $\mathcal{L}(f)$ to denote the Lipschitz constant (always with respect to ε) of a Lipschitz function f. The set on which this constant is taken should be clear from the context. For instance, if $f(t, \varepsilon)$ is known to be defined for $|\mathrm{Im}\, t| \leq \rho$ and $\varepsilon \in E \subset \mathbb{R}$ and it is Lipschitz with respect to ε in E, then $|f(t, \varepsilon_2) - f(t, \varepsilon_1)| \leq \mathcal{L}(f)|\varepsilon_2 - \varepsilon_1|$ for all t, ε_1, ε_2 in the allowed domain.

In what follows, we shall denote by $\overline{\mathbb{N}}$ the set of non negative integers, that is, $\mathbb{N} \cup \{0\}$.

LEMMA 2.17. *Let us define*

$$f(z, \varepsilon) = \sum_{|k| \geq 2} a_k(\varepsilon) z^k, \qquad k \in \overline{\mathbb{N}}^d,$$

and assume that the sum is convergent $\forall z \in D = D_1 \times \cdots \times D_d \subset \mathbb{C}^d$, where D_j are fixed disks of $\mathbb{C}$. Moreover, we suppose that f depends on ε in a Lipschitz way, with Lipschitz constant L. Let us take $\widehat{D} \subset D$ such that $\widehat{D} = \widehat{D}_1 \times \cdots \times \widehat{D}_d$, and satisfying that $\mathrm{radius}(\widehat{D}_j) \leq \alpha\, \mathrm{radius}(D_j) = \alpha r_j$, $0 < \alpha < 1$. *Then, if $z \in \widehat{D}$, it holds*

1. $|f(z, \varepsilon_1) - f(z, \varepsilon_2)| \leq K_2(\alpha) L |\varepsilon_1 - \varepsilon_2| \alpha^2$.
2. $\|D_z f(z, \varepsilon_1) - D_z f(z, \varepsilon_2)\| \leq K_1(\alpha) L |\varepsilon_1 - \varepsilon_2| \alpha$,

where both $K_i(\alpha)$, $i = 1, 2$, defined for $\alpha < 1$, are continuous and increasing functions.

Proof. Let $\partial_0 D$ be $\partial D_1 \times \cdots \times \partial D_d$, where ∂ states for the boundary of the corresponding sets. As

$$a_k(\varepsilon) = \frac{1}{(2\pi\sqrt{-1})^d} \int_{\partial_0 D} \frac{f(z, \varepsilon)}{z_1^{k_1+1} \cdots z_d^{k_d+1}}\, dz_1 \ldots dz_d,$$

we have that

$$\begin{aligned} |a_k(\varepsilon_1) - a_k(\varepsilon_2)| &\leq \frac{1}{(2\pi)^d} \int_{\partial_0 D} \frac{|f(z, \varepsilon_1) - f(z, \varepsilon_2)|}{|z_1|^{k_1+1} \cdots |z_d|^{k_d+1}} |dz_1 \ldots dz_d| \leq \\ &\leq \frac{L|\varepsilon_1 - \varepsilon_2|}{(2\pi)^d} \int_{\partial_0 D} \frac{|dz_1| \cdots |dz_d|}{r_1^{k_1+1} \cdots r_d^{k_d+1}} = \frac{L}{r^k} |\varepsilon_1 - \varepsilon_2|. \end{aligned}$$

On the other hand,

$$\begin{aligned} |f(z, \varepsilon_1) - f(z, \varepsilon_2)| &\leq \sum_{|k| \geq 2} |a_k(\varepsilon_1) - a_k(\varepsilon_2)||z^k| \leq \\ &\leq L|\varepsilon_1 - \varepsilon_2| \sum_{|k| \geq 2} \frac{|z_1|^{k_1} \cdots |z_d|^{k_d}}{r_1^{k_1} \cdots r_d^{k_d}} = (\diamond). \end{aligned}$$

Now, using $z \in \widehat{D}$ (that is, $|z_j| \leq \alpha r_j$) and $\#\{k \in \overline{\mathbb{N}}^d \ / \ |k| = m\} \leq dm^{d-1}$ if $m \geq 1$ (that can be obtained by induction with respect to d), we obtain

$$(\diamond) \leq L|\varepsilon_1 - \varepsilon_2| \sum_{|k|\geq 2} \alpha^{|k|} \leq L|\varepsilon_1 - \varepsilon_2| \sum_{m=2}^{\infty} dm^{d-1}\alpha^m = K_2(\alpha)L|\varepsilon_1 - \varepsilon_2|\alpha^2,$$

where $K_2(\alpha) = d\sum_{m=2}^{\infty} m^{d-1}\alpha^{m-2}$. Finally, it is not difficult to see that $K_2(\alpha)$ is convergent if $|\alpha| < 1$. This completes the proof of 1.

As

$$\frac{\partial f}{\partial z_j}(z,\varepsilon) = \sum_{|k|\geq 2} k_j a_k(\varepsilon) z^{k-e_j},$$

we can proceed in the same way as before:

$$\begin{aligned}
\left|\frac{\partial f}{\partial z_j}(z,\varepsilon_1) \right.&- \left.\frac{\partial f}{\partial z_j}(z,\varepsilon_2)\right| \leq \sum_{|k|\geq 2} k_j |a_k(\varepsilon_1) - a_k(\varepsilon_2)||z|^{k-e_j} \leq \\
&\leq \sum_{|k|\geq 2} k_j \frac{L}{r^k}|\varepsilon_1 - \varepsilon_2||z|^{k-e_j} \leq \frac{L|\varepsilon_1 - \varepsilon_2|}{r_j} \sum_{|k|\geq 2} k_j \alpha^{k-e_j} \leq \\
&\leq \frac{L|\varepsilon_1 - \varepsilon_2|}{r_j} \sum_{m=2}^{\infty} dm^d \alpha^{m-1} \leq K_1(\alpha)L|\varepsilon_1 - \varepsilon_2|\alpha,
\end{aligned}$$

where $K_1(\alpha) = \frac{d}{\tau_\infty}\sum_{m=2}^{\infty} m^d\alpha^{m-2}$ and τ_∞ is a lower bound of the values r_j (see Section 2.7). Here, we note that $K_1(\alpha)$ is convergent if $|\alpha| < 1$. To complete the proof, we only need to take the sup norm of the vector of components $\frac{\partial f}{\partial z_j}(z,\varepsilon_1) - \frac{\partial f}{\partial z_j}(z,\varepsilon_2)$.
□

LEMMA 2.18. *Let us suppose that $P(t,\varepsilon)$ is a matrix depending on ε in a Lipschitz way, with constant L. If $\|P\| \leq 1/2$, then $(I + P(t,\varepsilon))^{-1}$ is Lipschitz with respect to ε with constant $4L$.*

Proof. It is known that

$$(I+P)^{-1} = I - P + P^2 - P^3 + \cdots,$$

and then, it is easy to see that

$$\begin{aligned}
\mathcal{L}((I+P)^{-1}) &\leq \mathcal{L}(P) + \mathcal{L}(P^2) + \mathcal{L}(P^3) + \cdots \leq \sum_{n=1}^{\infty} \left[n\|P\|^{n-1}L\right] = \\
&= \frac{1}{(1-\|P\|)^2}L \leq 4L.
\end{aligned}$$

□

LEMMA 2.19. *Let $q(t,\varepsilon)$ be an analytic quasiperiodic function on a strip of width ρ_1. We write*

$$q(t,\varepsilon) = \sum_{k\in\mathbb{Z}^r} q^k(\varepsilon)e^{(k,\omega)\sqrt{-1}t},$$

and we assume that all the coefficients $q^k(\varepsilon)$ depend on ε in a Lipschitz way, with constant L_k. Moreover, we suppose that $L_k \leq L|k|^\alpha e^{-\rho_1|k|}$ if $k \neq 0$, where L is a *positive constant. Let us take $\rho_2 \in]0, \rho_1[$. Then, if $q(t,\varepsilon)$ is restricted to a strip of width ρ_2, it depends on ε in a Lipschitz way with constant*

$$L' = L_0 + L\frac{20r}{3\delta^{r+\alpha}}\chi(r+\alpha),$$

where $L_0 = L_{k=0}$, $\delta = \rho_1 - \rho_2$ and χ was defined in Lemma 2.6.

Proof.

$$\begin{aligned} |q(t,\varepsilon_1) - q(t,\varepsilon_2)| &\leq \sum_{k\in\mathbb{Z}^r} |q^k(\varepsilon_1) - q^k(\varepsilon_2)||e^{(k,\omega)\sqrt{-1}t}| \leq \\ &\leq \left[L_0 + L\sum_{k\neq 0} |k|^\alpha e^{-\delta|k|} \right] |\varepsilon_1 - \varepsilon_2|. \end{aligned}$$

Here we can apply Lemma 2.6 to obtain the result wanted. □

LEMMA 2.20. *Let $q(t,\varepsilon)$ be an analytic quasiperiodic function on a strip of width ρ,*

$$q(t,\varepsilon) = \sum_{k\in\mathbb{Z}^r} q^k(\varepsilon)e^{(k,\omega)\sqrt{-1}t}.$$

Let us assume that $q(t,\varepsilon)$ depends on ε in a Lipschitz way, with constant L. Then, the coefficients $q^k(\varepsilon)$ depend on ε in a Lipschitz way,

$$|q^k(\varepsilon_1) - q^k(\varepsilon_2)| \leq L_k|\varepsilon_1 - \varepsilon_2|,$$

where $L_k = Le^{-\rho|k|}$.

Proof. Let us fix ε_1, ε_2 and define $p(t) = q(t,\varepsilon_1) - q(t,\varepsilon_2)$. As $\|p\|_\rho \leq L|\varepsilon_1 - \varepsilon_2|$, the Fourier coefficients of p satisfy

$$|p^k| \leq L|\varepsilon_1 - \varepsilon_2|e^{-\rho|k|},$$

and using that $|p^k| = |q^k(\varepsilon_1) - q^k(\varepsilon_2)|$ the result follows. □

LEMMA 2.21. *Let us define*

$$f(\varepsilon) = \frac{g(\varepsilon)}{\sigma - \lambda(\varepsilon)},$$

where $|\sigma - \lambda(\varepsilon)| \geq u$, and g and λ are Lipschitz functions with constants L_g and L_λ respectively. Then, f is Lipschitz with constant

$$L_f = \frac{L_g}{u} + \|g\|_\infty \frac{L_\lambda}{u^2}.$$

Proof. It is straightforward. □

LEMMA 2.22. *Let A_0 be a $d\times d$ matrix such that $\mathrm{Spec}(A_0) = \{\lambda_1^0, \ldots, \lambda_d^0\}$, $|\lambda_i^0| > 2\mu$, $|\lambda_i^0 - \lambda_j^0| > 2\mu$, $i \neq j$, where $\mu > 0$. Let $A(\varepsilon)$ be a matrix valued function such that $\|A(\varepsilon) - A_0\| < \alpha$ if $|\varepsilon| < \varepsilon_0$, and depending on ε in a Lipschitz way, with constant L_A. Let $B(\varepsilon)$ be the change of variables that diagonalizes $A(\varepsilon)$ (see Lemma 2.9). Then, there exist $\tau_1 = \tau_1(A_0,\alpha,\beta)$ and $\tau_2 = \tau_2(A_0,\alpha,\beta)$ such that*

$$\begin{aligned} \|B(\varepsilon_1) - B(\varepsilon_2)\| &\leq \tau_1 L_A|\varepsilon_1 - \varepsilon_2|, \\ \|B^{-1}(\varepsilon_1) - B^{-1}(\varepsilon_2)\| &\leq \tau_1 L_A|\varepsilon_1 - \varepsilon_2|, \\ |\lambda_j(\varepsilon_1) - \lambda_j(\varepsilon_2)| &\leq \tau_2 L_A|\varepsilon_1 - \varepsilon_2|, \end{aligned}$$

where $\lambda_j(\varepsilon)$ are the eigenvalues of $A(\varepsilon)$ and the definition of values α and β can be found in Lemma 2.9.

Proof. This result is essentially contained in [19], pages 66–67, but for an analytic dependence on ε. The result for a Lipschitz dependence on ε can be obtained as follows:

1. Let us consider the matrix A as a function of all its elements a_{ij}. This implies that, if the elements are close enough to the ones of A_0, the eigenvalues and eigenvectors depend on a_{ij} in an analytic way. Hence, in any compact inside the domain of analyticity they depend, also, in an analytic way.
2. The elements $a_{ij}(\varepsilon)$ of $A(\varepsilon)$ also depend on ε in Lipschitz way, with the same constant:

$$\begin{aligned}|a_{ij}(\varepsilon_1)-a_{ij}(\varepsilon_2)| &\le \max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}(\varepsilon_1)-a_{ij}(\varepsilon_2)| = \\ &= \|A(\varepsilon_1)-A(\varepsilon_2)\|\le L|\varepsilon_1-\varepsilon_2|.\end{aligned}$$

3. Finally we compose the Lipschitz dependence (of the eigenvalues and eigenvectors) on a_{ij} with the Lipschitz dependence of a_{ij} on ε.

□

Lemma 2.23. *Let us consider the equation*

$$\dot{x}=A(\varepsilon)x+g(t,\varepsilon),$$

under the same hypothesis as in Lemma 2.10. Let ρ_2 be such that $0<\rho_2<\rho_1-2\nu$ and $\delta=\rho_1-\rho_2-2\nu\le 1$. Moreover, we assume that $A(\varepsilon)$ and $g(t,\varepsilon)$ depend on ε in Lipschitz way with constants L_A and L_g, respectively. Then, the solution $x(t,\varepsilon)$ of this equation (see Lemma 2.10) depends on ε in a Lipschitz way, for t belonging to the strip of width ρ_2, with constant

$$L_x\le\frac{\chi(r+2\gamma)}{\delta^{r+2\gamma}}(E_1L_A\|g\|_{\rho_1}+E_2L_g),$$

where E_1 and E_2 are positive constants, that do not depend on the actual step of the inductive process of Section 2.7.

Proof. First of all, let us make the change of variables $x=B(\varepsilon)y$ (the matrix $B(\varepsilon)$ is given by Lemma 2.22) in order to diagonalize the matrix $A(\varepsilon)$. With this, the equation becomes

$$\dot{y}=D(\varepsilon)y+h(t,\varepsilon),$$

where $D(\varepsilon)$ is a diagonal matrix and $h(t,\varepsilon)=B^{-1}(\varepsilon)g(t,\varepsilon)$. Lemma 2.22 ensures that $L_D\equiv\mathcal{L}(D)\equiv\max_i L_{\lambda_i}=\tau_2L_A$ and $L_h\equiv\mathcal{L}(h)=\tau_1L_A\|g\|_{\rho_1}+\beta L_g$. Moreover, as

$$h(t,\varepsilon)=\sum_{k\in\mathbb{Z}^r}h^k(\varepsilon)e^{(k,\omega)\sqrt{-1}t},$$

we have, by Lemma 2.20, that $L_{h^k}\equiv\mathcal{L}(h^k)=L_he^{-\rho_1|k|}$.

As it has been shown in Lemma 2.10, the solution we are interested in is given by

$$y_i^k(\varepsilon)=\frac{h_i^k(\varepsilon)}{(k,\omega)\sqrt{-1}-\lambda_i(\varepsilon)}.$$

Now, let us compute $L_{y_i^k}\equiv\mathcal{L}(y_i^k)$. We distinguish two cases and, in both, we use Lemma 2.21:

1. Case $k=0$.

$$\begin{aligned} L_{y_i^0} &= \frac{L_{h^0}}{\mu} + |h^0| \frac{L_{\lambda_i}}{\mu^2} \le \frac{\tau_1 L_A \|g\|_{\rho_1} + \beta L_g}{\mu} + \|g\|_{\rho_1} \beta \frac{\tau_2 L_A}{\mu^2} = \\ &= \left(\frac{\tau_1}{\mu} + \frac{2\beta_0 \tau_2}{\mu^2} \right) L_A \|g\|_{\rho_1} + \frac{2\beta_0}{\mu} L_g \equiv C_1 L_A \|g\|_{\rho_1} + C_2 L_g, \end{aligned}$$

where μ has been defined in Lemma 2.9 and C_1 and C_2 do not depend on the step of the iterative process.

2. Case $k \neq 0$.

$$\begin{aligned} L_{y_i^k} &= \frac{L_{h^k}}{\frac{c}{|k|^\gamma} e^{-\nu|k|}} + |h^k| \frac{L_{\lambda_i}}{\left(\frac{c}{|k|^\gamma} e^{-\nu|k|} \right)^2} \le \\ &\le \frac{|k|^\gamma e^{\nu|k|}}{c} L_h e^{-\rho_1|k|} + \frac{|k|^{2\gamma} e^{2\nu|k|}}{c^2} \|h\|_{\rho_1} e^{-\rho_1} \tau_2 L_A \le \\ &\le |k|^{2\gamma} e^{-(\rho_1 - 2\nu)|k|} \left[\frac{\tau_1}{c} \|g\|_{\rho_1} L_A + \frac{\beta}{c} L_g + \frac{\beta \tau_2}{c^2} \|g\|_{\rho_1} L_A \right] \equiv \\ &\equiv |k|^{2\gamma} e^{-(\rho_1 - 2\nu)|k|} \left[C_3 \|g\|_{\rho_1} L_A + C_4 L_g \right], \end{aligned}$$

where now C_3 and C_4 do not depend on the step of the iterative process.

Now, we can apply this to bound the Lipschitz constant L_y corresponding to $y(t,\varepsilon) = \sum_k y_i^k(\varepsilon) e^{(k,\omega)\sqrt{-1}t}$. From Lemma 2.19 we obtain

$$L_y = C_1 L_A \|g\|_{\rho_1} + C_2 L_g + (C_3 L_A \|g\|_{\rho_1} + C_4 L_g) \frac{20r}{3\delta^{r+2\gamma}} \chi(r+2\gamma),$$

where $0 < \rho_2 < \rho_1 - 2\nu$ such that $\delta = \rho_1 - \rho_2 - 2\nu \le 1$. To simplify the following steps, we note that

$$L_y \le \frac{\chi(r+2\gamma)}{\delta^{r+2\gamma}} (C_5 L_A \|g\|_{\rho_1} + C_6 L_g),$$

for suitable constants C_5 and C_6, both independent on the actual step of the inductive process.

As $x = B(\varepsilon) y$, we have $L_x \equiv \mathcal{L}(x) \le \tau_1 L_A \|y\|_{\rho_2} + \beta L_y$, that allows (using the bound on $\|y\|_{\rho_2}$ given inside the proof of Lemma 2.10) to establish the following bound:

$$L_x \le \tau_1 L_A \left[\frac{1}{\mu} + \frac{20 r \chi(r+\gamma)}{3c\delta^{r+\gamma}} \right] \beta \|g\|_{\rho_1} + \beta \frac{\chi(r+2\gamma)}{\delta^{r+2\gamma}} (C_5 L_A \|g\|_{\rho_1} + C_6 L_g).$$

This can easily be rearranged to

$$L_x \le \frac{\chi(r+2\gamma)}{\delta^{r+2\gamma}} (E_1 L_A \|g\|_{\rho_1} + E_2 L_g),$$

where E_1 and E_2 are suitable constants, not depending on the actual step of the inductive process. □

2.9. Measure lemma. Here we give the basic lemma used to bound the measure of the resonances.

LEMMA 2.24. *Let $\omega \in \mathbb{R}^r$ and $v \in \sqrt{-1}\mathbb{R}$ such that*

$$|v - \sqrt{-1}(k,\omega)| \ge \frac{2c}{|k|^{\gamma_0}}.$$

for all $k \in \mathbb{Z}^r \setminus \{0\}$*, where* $c > 0$*,* $\gamma_0 > 0$*. We define the* n*-th resonant subset* $\mathcal{R}_\mu^{(n)} = \mathcal{R}_\mu^{(n)}(v)$ *as*

$$\mathcal{R}_\mu^{(n)} = \Big\{\varphi \in \sqrt{-1}\mathbb{R},\ |\varphi| < \mu \ / \ \exists k' \in \mathbb{Z}^r \setminus \{0\} \\ \text{such that } |\varphi + v - \sqrt{-1}(k',\omega)| < \frac{c}{|k'|^{\gamma_n}} e^{-\nu_n |k'|}\Big\},$$

where $\nu_n = \nu_0/(n+1)^2$*,* $0 < \nu_0 < 1/4$*,* $\gamma_n = \gamma_0 z^n$*,* $1 < z < 2$*,* $\gamma_0 \geq r-1$*. Let* $\mathcal{R}_\mu = \cup_{n\geq 0}\mathcal{R}_\mu^{(n)}$ *and* $\psi(\mu) = \frac{m(\mathcal{R}_\mu)}{2\mu}$*, where* m *denotes the Lebesgue measure. Then* $\psi(\mu) \leq \exp(-c_1/\mu^{c_2})$ *for some positive constants* c_1 *and* c_2*, with* $c_2 < 1/\gamma_0$*, provided* μ *is small enough.*

Proof. Let k' and φ such that

$$|\varphi + v - \sqrt{-1}(k',\omega)| < \frac{c}{|k'|^{\gamma_n}} e^{-\nu_n |k'|}.$$

As

$$|v - \sqrt{-1}(k',\omega)| \geq \frac{2c}{|k'|^{\gamma_0}},$$

one has

$$\mu > |\varphi| > \frac{2c}{|k'|^{\gamma_0}} - \frac{c}{|k'|^{\gamma_n}\exp(\nu_n|k'|)} > \frac{c}{|k'|^{\gamma_0}},$$

and, hence, $|k'| \geq \lceil (c/\mu)^{1/\gamma_0} \rceil \equiv M(\mu)$, where, for $\alpha \in \mathbb{R}$, $\lceil \alpha \rceil$ denotes the lowest integer greater than or equal to α. Now let us add for all $|k'| \geq M(\mu)$ and all $n \geq 0$ to have an upper bound on $m(\mathcal{R}_\mu)$. Adding for all $|k'| \geq M(\mu)$ and a fixed n we obtain

$$\sum_{|k'|\geq M(\mu)} \frac{2c}{|k'|^{\gamma_n}\exp(\nu_n|k'|)} 2c \sum_{j\geq M(\mu)} \frac{2rj^{r-1}}{j^{\gamma_n}} e^{-\nu_n j} < \\ < 4crM(\mu)^{r-1-\gamma_n}\frac{e^{-\nu_n M(\mu)}}{1-e^{-\nu_n}} < \frac{5cr}{\nu_0} M(\mu)^{r-1-\gamma_n}(n+1)^2 e^{-\nu_n M(\mu)},$$

because $r - 1 - \gamma_n \leq 0$, $1 - e^{-\alpha} > 0.8\alpha$ if $0 \leq \alpha \leq 1/4$. Adding for all n we have

$$m(\mathcal{R}_\mu) \leq \frac{5cr}{\nu_0} M(\mu)^{r-1} \sum_{n\geq 0} (n+1)^2 M(\mu)^{-\gamma_0 z^n} \exp\left(-\frac{\nu_0}{(n+1)^2} M(\mu)\right).$$

For our purposes a rough bound is enough. Let $n_* = \log(\nu_0 M(\mu))/\log z$. We split $\sum_{n\geq 0}$ as $\sum_{n=0}^{n_*} + \sum_{n>n_*}$ and assume μ small enough so that

$$\frac{(n+2)^2 M(\mu)^{-\gamma_0 z^{n+1}}}{(n+1)^2 M(\mu)^{-\gamma_0 z^n}} \leq \frac{1}{2}, \quad \forall n \geq 0.$$

Then

$$\sum_{n\geq 0} (n+1)^2 M(\mu)^{-\gamma_0 z^n} \exp\left(-\frac{\nu_0}{(n+1)^2} M(\mu)\right) \leq \\ \leq 2M(\mu)^{-\gamma_0} \exp\left(-\frac{\nu_0}{(n_*+1)^2} M(\mu)\right) + (n_*+1)^2 M(\mu)^{-\gamma_0 z^{n_*}}.$$

To finish the proof, after selecting any value of c_1 and $c_2 < 1/\gamma_0$, we want to show that each term is less than $\frac{1}{2}\exp(-c_1/\mu^{c_2})$. To this end we take logarithms. One has to prove

$$\log A - \log\left(\frac{c}{\mu}\right) - \frac{\nu_0}{\left[\frac{\log\left(\nu_0\left(\frac{c}{\mu}\right)^{\frac{1}{\gamma_0}}\right)}{\log z} + 1\right]^2}\left(\frac{c}{\mu}\right)^{\frac{1}{\gamma_0}} < -\log 4 - \frac{c_1}{\mu^{c_2}},$$

$$\log A + 2\log\left[\frac{\log\left(\nu_0\left(\frac{c}{\mu}\right)^{\frac{1}{\gamma_0}}\right)}{\log z} + 1\right] - \nu_0\left(\frac{c}{\mu}\right)^{\frac{1}{\gamma_0}}\log\left(\frac{c}{\mu}\right) < -\log 2 - \frac{c_1}{\mu^{c_2}},$$

where $A = \frac{5cr}{\nu_0}(\frac{c}{\mu})^{(r-1)/\gamma_0}$. Both inequalities are true if μ is small enough or, equivalently, for μ in a fixed range if c_1 is big enough. □

2.10. Proof of Theorem 1 (part II). Up to now, we have shown the convergence of the iterative scheme, provided that some nonresonance conditions hold at each step n (see Section 2.7). Now, our purpose is to show that all the matrices $A_n(\varepsilon)$ are Lipschitz (with respect to ε) from above and below, and their Lipschitz constants are bounded (from above and below respectively) by constants that do not depend on n. As we shall see later, this allows to take out a dense set (with small relative measure) of values of ε, for which the resonance conditions assumed during Section 2.7 might not hold at some step (i.e. for some n) of the proof.

To prove that $A_n(\varepsilon)$ is Lipschitz from below we shall proceed in the following way: as $A_0(\varepsilon)$ is Lipschitz from above and below (by hypothesis), it is enough to show that $A_n(\varepsilon)$ is Lipschitz from above and $\mathcal{L}(A_0) - \mathcal{L}(A_n) = \mathcal{O}(\varepsilon)$, since Lemma 2.16 implies that, if ε is sufficiently small, $A_n(\varepsilon)$ is also Lipschitz from below. For this reason, we are going to focus on Lipschitz constants from above, that, for simplicity, will be called Lipschitz constants. The notation used will be the following:

$$\begin{array}{lll} \mathcal{L}(A_n(\varepsilon)) = L_{A_n}, & \mathcal{L}(\widehat{A}_n(\varepsilon)) = L_{\widehat{A}_n}, & \mathcal{L}(\varepsilon^{2^n}Q_n(t,\varepsilon)) = L_{Q_n}, \\ \mathcal{L}(\varepsilon^{2^n}\widehat{Q}_n(t,\varepsilon)) = L_{\widehat{Q}_n}, & \mathcal{L}(\varepsilon^{2^n}g_n(t,\varepsilon)) = L_{g_n}, & \mathcal{L}(\varepsilon^{2^n}\widehat{g}_n(t,\varepsilon)) = L_{\widehat{g}_n}, \\ \mathcal{L}(\varepsilon^{2^n}P_n(t,\varepsilon)) = L_{P_n}, & \mathcal{L}(h_n(x,t,\varepsilon)) = L_{h_n}, & \mathcal{L}(\widehat{h}_n(x,t,\varepsilon)) = L_{\widehat{h}_n}. \end{array}$$

Our purpose now is to bound the Lipschitz constants of the equation terms at step $n+1$ as a function of the Lipschitz constants at step n. Let us assume that the scheme of Section 2.7 has been applied up to step n. Then, the following bounds can be established:

$$\begin{array}{rcl} \|\widehat{Q}_n\|_{\sigma_n} & \leq & N_1^{2^n}, \\ \|P_n\|_{\rho_{n+1}} & \leq & N_1^{2^n}, \\ \|\underline{x}_n\|_{\sigma_n} & \leq & (\varepsilon_0 N_1)^{2^n}, \\ \|\widehat{g}_n\|_{\sigma_n} & \leq & N_1^{2^{n+1}}, \end{array}$$

where N_1 is a positive constant and ε_0 was defined at the end of Section 2.7 and it is also assumed to satisfy $\varepsilon_0 N_1 < 1$. Proofs are not difficult, using the bounds given in Section 2.7.

We shall also use the bounds $\|Q_n\|_{\rho_n} \le N_1^{2^n}$ and $\|g_n\|_{\rho_n} \le N_1^{2^n}$ (see Section 2.7) when needed. The constant N_1 can be easily obtained from the constants M_i, $i = 1, \dots, 4$, introduced in Section 2.7.

Now let us bound $L_{Q_{n+1}}$. It is easy to obtain

$$\varepsilon^{2^{n+1}} Q_{n+1} = (I + \varepsilon^{2^n} P_n(t,\varepsilon))^{-1}(\varepsilon^{2^n}\widehat{Q}_n(t,\varepsilon))(\varepsilon^{2^n} P_n(t,\varepsilon)),$$

and this implies

$$\begin{aligned} L_{Q_{n+1}} \quad &\le \quad 4L_{P_n}\varepsilon_0^{2^n}\|\widehat{Q}_n\|_{\sigma_n}\varepsilon_0^{2^n}\|P_n\|_{\rho_{n+1}} + \|(I+\varepsilon^{2^n}P_n)^{-1}\|_{\rho_{n+1}} L_{\widehat{Q}_n}\varepsilon_0^{2^n}\|P_n\|_{\rho_{n+1}} + \\ &\quad + \|(I+\varepsilon^{2^n}P_n)^{-1}\|_{\rho_{n+1}}\varepsilon_0^{2^n}\|\widehat{Q}_n\|_{\sigma_n} L_{P_n}, \end{aligned}$$

Using that $\|(I+\varepsilon^{2^n}P_n)^{-1}\|_{\rho_{n+1}} \le 2$ (see Section 2.7) we obtain

$$\begin{aligned} L_{Q_{n+1}} \quad &\le \quad 4L_{P_n}(\varepsilon_0 N_1)^{2^n}(\varepsilon_0 N_1)^{2^n} + 2L_{\widehat{Q}_n}(\varepsilon_0 N_1)^{2^n} + 2(\varepsilon_0 N_1)^{2^n} L_{P_n} < \\ &< \quad 6(\varepsilon_0 N_1)^{2^n} L_{P_n} + 2(\varepsilon_0 N_1)^{2^n} L_{\widehat{Q}_n}. \end{aligned}$$

Now we are going to consider $L_{\underline{x}_n}$. From Lemma 2.23 is not difficult to obtain

$$L_{\underline{x}_n} \le \frac{\chi(r+2\gamma_n)}{\delta_n^{r+2\gamma_n}}(E_1 L_{A_n}\varepsilon_0^{2^n}\|g_n\|_{\rho_n} + E_2 L_{g_n}),$$

where δ_n is now $(\frac{1}{2} - 2\nu_0)/(n+1)^2$. Introducing $L_{3,n} = \frac{\chi(r+2\gamma_n)}{\delta_n^{r+2\gamma_n}}\max\{E_1,\, E_2\}$ we can write

$$L_{\underline{x}_n} \le L_{3,n}\left(L_{A_n}(\varepsilon_0 N_1)^{2^n} + L_{g_n}\right). \tag{17}$$

Let us consider now $L_{\widehat{Q}_n}$. We recall that $\widehat{Q}_n$ was defined as

$$\varepsilon^{2^n}\widehat{Q}_n = \varepsilon^{2^n} Q_n(t,\varepsilon) + D_x h_n(\underline{x}_n(t,\varepsilon), t, \varepsilon).$$

This allows to write

$$L_{\widehat{Q}_n} \le L_{Q_n} + \|D_{xx}h_n(\underline{x}_n(t,\varepsilon),t,\varepsilon)\| L_{\underline{x}_n} + \mathcal{L}(D_x h_n(\underline{x}, t, \varepsilon)).$$

Here we can use that $\|D_{xx}h_n(\underline{x}(t,\varepsilon),t,\varepsilon)\| \le K_\infty$ (see Section 2.7), Lemma 2.17 and (17) to write

$$L_{\widehat{Q}_n} \le L_{Q_n} + K_\infty L_{3,n}\left(L_{A_n}(\varepsilon_0 N_1)^{2^n} + L_{g_n}\right) + K_1(\alpha) L_{h_n}(\varepsilon_0 N_1)^{2^n},$$

where $\alpha = \|\underline{x}_n\|_{\sigma_n}/\tau_\infty$. Moreover, note that α goes to zero when $\|\underline{x}_n\|_{\sigma_n}$ does. This implies that, if ε is small enough, we can assume that $K_1(\alpha)$ is less than, for instance, $K_1(1/2)$.

Now we are going to focus on L_{P_n}. The definition of P_n was

$$\varepsilon^{2^n}\dot{P}_n(t,\varepsilon) = \widehat{A}_n(\varepsilon^{2^n}P_n(t,\varepsilon)) - (\varepsilon^{2^n}P_n(t,\varepsilon))\widehat{A}_n + \varepsilon^{2^n}\widehat{Q}_n(t,\varepsilon).$$

As this is a linear system of differential equations, we can apply a lemma which is essentially like Lemma 2.23 but for the actual system of equations and with new constants, $\overline{E}_1$, $\overline{E}_2$, to get

$$\begin{aligned} L_{P_n} \quad &\le \quad \frac{\chi(r+2\gamma_n)}{\delta_n^{r+2\gamma_n}}\left(\overline{E}_1\|\varepsilon_0^{2^n}\widehat{Q}_n\|_{\sigma_n} L_{A_n} + \overline{E}_2 L_{\widehat{Q}_n}\right) \le \\ &\le \quad \frac{\chi(r+2\gamma_n)}{\delta_n^{r+2\gamma_n}}\left(\overline{E}_1(\varepsilon_0 N_1)^{2^n} L_{\widehat{A}_n} + +\overline{E}_2 L_{\widehat{Q}_n}\right) \le L_{3,n}\left((\varepsilon_0 N_1)^{2^n} L_{\widehat{A}_n} + L_{\widehat{Q}_n}\right), \end{aligned}$$

where $L_{3,n}$ has been redefined as $L_{3,n} = \frac{\chi(r+2\gamma_n)}{\delta_n^{r+2\gamma_n}} \max\{E_1, E_2, \overline{E}_1, \overline{E}_2\}$.

Let us consider $L_{g_{n+1}}$. From

$$\varepsilon^{2^{n+1}} g_{n+1}(t,\varepsilon) = (I + \varepsilon^{2^n} P_n(t,\varepsilon))^{-1} (\varepsilon^{2^{n+1}} \widehat{g}_n(t,\varepsilon)),$$

follows

$$\begin{aligned} L_{g_{n+1}} &\leq 4L_{P_n} \varepsilon_0^{2^{n+1}} \|\widehat{g}_n\|_{\sigma_n} + \|(I + \varepsilon^{2^n} P_n(t,\varepsilon))^{-1}\|_{\rho_{n+1}} L_{\widehat{g}_n} \leq \\ &\leq 4L_{P_n} (\varepsilon_0 N_1)^{2^{n+1}} + 2L_{\widehat{g}_n}. \end{aligned}$$

Consider now $L_{\widehat{A}_n}$. As we have

$$\widehat{A}_n(\varepsilon) = A_n(\varepsilon) + \varepsilon^{2^n} \overline{\widehat{Q}}_n(\varepsilon),$$

it follows that

$$L_{\widehat{A}_n} = L_{A_n} + L_{\widehat{Q}_n}.$$

Moreover, as $A_{n+1} = \widehat{A}_n$ we also have

$$L_{A_{n+1}} = L_{\widehat{A}_n}.$$

Now let us bound $L_{\widehat{g}_n}$. Recall that

$$\varepsilon^{2^{n+1}} \widehat{g}_n(t,\varepsilon) = h_n(\underline{x}_n(t,\varepsilon), t, \varepsilon) + \varepsilon^{2^n} Q_n(t,\varepsilon) \underline{x}_n(t,\varepsilon),$$

that implies

$$\begin{aligned} L_{\widehat{g}_n} &\leq \|D_x h_n(\underline{x}_n(t,\varepsilon), t, \varepsilon)\|_{\sigma_n} L_{\underline{x}_n} + L_{h_n(\underline{x}_n, t, \varepsilon)} + \\ &\quad + L_{Q_n} \|\underline{x}_n(t,\varepsilon)\|_{\sigma_n} + \|\varepsilon^{2^n} Q_n(t,\varepsilon)\|_{\rho_n} L_{\underline{x}_n} \leq \\ &\leq [K_\infty \|\underline{x}_n(t,\varepsilon)\|_{\sigma_n} + \|\varepsilon^{2^n} Q_n(t,\varepsilon)\|_{\rho_n}] L_{\underline{x}_n} + \\ &\quad + K_2(1/2) \|\underline{x}_n(t,\varepsilon)\|_{\sigma_n}^2 L_{h_n} + \|\underline{x}_n(t,\varepsilon)\|_{\sigma_n} L_{Q_n}, \end{aligned}$$

where it has been used that $L_{h_n(\underline{x}_n, t, \varepsilon)} \leq K_2(\alpha) L_{h_n} \|\underline{x}_n\|^2$ and, as before, that $K_2(\alpha) \leq K_2(1/2)$ if ε_0 is small enough. On the other hand,

$$K_\infty \|\underline{x}_n(t,\varepsilon)\|_{\sigma_n} + \|\varepsilon^{2^n} Q_n(t,\varepsilon)\|_{\rho_n} \leq (K_\infty + 1)(\varepsilon_0 N_1)^{2^n}.$$

This implies

$$L_{\widehat{g}_n} \leq K_3 L_{3,n} (\varepsilon_0 N_1)^{2^n} (L_{A_n} + L_{g_n} + L_{h_n} + L_{Q_n}),$$

where $K_3 = \max\{K_\infty + 1,\ K_2(1/2)\}$ and we have used $\varepsilon_0 N_1 < 1$ and $L_{3,n} > 1$.

Let us follow with $L_{h_{n+1}}$. As

$$h_{n+1}(x_{n+1}, t, \varepsilon) = (I + \varepsilon^{2^n} P_n(t,\varepsilon))^{-1} \widehat{h}_n((I + \varepsilon^{2^n} P_n(t,\varepsilon)) x_{n+1}, t, \varepsilon),$$

we have

$$\begin{aligned} L_{h_{n+1}} &\leq 4L_{P_n} \|\widehat{h}_n((I + \varepsilon^{2^n} P_n) x_{n+1}, t, \varepsilon)\|_{\rho_{n+1}} + \\ &\quad + \|(I + \varepsilon^{2^n} P_n)^{-1}\|_{\rho_{n+1}} \|D_x \widehat{h}_n\|_{\rho_{n+1}} L_{P_n} + \|(I + \varepsilon^{2^n} P_n)^{-1}\|_{\rho_{n+1}} L_{\widehat{h}_n} \leq \\ &\leq 4L_{P_n} \frac{K_\infty}{2} \tau_\infty + 2K_\infty \tau_\infty L_{P_n} + 2L_{\widehat{h}_n}, \end{aligned}$$

that allows to write

$$L_{h_{n+1}} \leq 4K_\infty \tau_\infty L_{P_n} + 2L_{\widehat{h}_n}.$$

Finally, let us consider $L_{\widehat{h}_n}$. We recall

$$\widehat{h}_n(y_n, t, \varepsilon) = h_n(\underline{x}_n(t,\varepsilon) + y_n, t, \varepsilon) - h_n(\underline{x}_n(t,\varepsilon), t, \varepsilon) - D_x h_n(\underline{x}_n(t,\varepsilon), t, \varepsilon) y_n,$$

that implies

$$\begin{aligned} L_{\widehat{h}_n} &\leq \|D_x h_n(\underline{x}_n(t,\varepsilon) + y_n, t, \varepsilon)\| L_{\underline{x}_n} + L_{h_n} + \|D_x h_n(\underline{x}_n(t,\varepsilon), t, \varepsilon)\| L_{\underline{x}_n} + \\ &\quad + L_{h_n} + \|D_{xx} h_n(\underline{x}_n(t,\varepsilon), t, \varepsilon)\| \|y_n\| L_{\underline{x}_n} + \|y_n\| L_{D_x h_n} \leq \\ &\leq K_\infty \tau_\infty L_{\underline{x}_n} + L_{h_n} + K_\infty \tau_\infty L_{\underline{x}_n} + L_{h_n} + K_\infty \tau_\infty L_{\underline{x}_n} + \tau_\infty K_1(\alpha)\alpha L_{h_n}, \end{aligned}$$

and using that, if ε small enough, $\tau_\infty K_1(\alpha)\alpha \leq 1$ it can be obtained the following bound:

$$L_{\widehat{h}_n} \leq 3K_\infty \tau_\infty L_{\underline{x}_n} + 3L_{h_n}.$$

Up to now we have stated some bounds of the Lipschitz constants. Next step is to relate (in closed formulas) the bounds of step $n+1$ with bounds of step n.

Let us define $a_n = L_{A_n}$, $b_n = \max\{L_{Q_n}, L_{g_n}\}$, $c_n = L_{h_n}$ and let $e_n = (\varepsilon_0 N_1)^{2^n}$. Furthermore let $L_{4,n} = L_{3,n} \max\{K_\infty, 6, 2K_3, 6K_\infty \tau_\infty\}$. After some rearrangement one can write the bounds on the recurrences as

$$\begin{aligned} a_{n+1} &\leq a_n + b_n + L_{4,n}(a_n e_n + b_n) + K_1(\alpha) e_n c_n, \\ b_{n+1} &\leq 5L_{4,n}^2 e_n a_n + 8L_{4,n}^2 e_n b_n + (4K_1(\alpha)+1) L_{4,n} e_n c_n, \\ c_{n+1} &\leq 3L_{4,n}^2 e_n a_n + 4L_{4,n}^2 b_n + (6 + 2K_1(\alpha) L_{4,n}) c_n. \end{aligned} \tag{18}$$

Let $d_n = \max\{a_n, b_n, c_n\}$. It is immediate (recalling $e_n < 1$, $L_{4,n} > 1$) to obtain

$$d_{n+1} \leq R L_{4,n}^2 d_n,$$

where $R = 14 + 4K_1(\alpha)$. As before it is easy to obtain $RL_{4,n}^2 \leq M_6^{2^n}$ for some suitable M_6, independent on n. Therefore

$$d_n \leq \prod_{j=0}^{n} M_6^{2^j} d_0 < M_6^{2^{n+1}} d_0.$$

Going back to (18) we have

$$\begin{aligned} b_{n+1} \leq \frac{5}{14} M_6^{2^n} (\varepsilon_0 N_1)^{2^n} M_6^{2^{n+1}} d_0 + \frac{8}{14} M_6^{2^n} (\varepsilon_0 N_1)^{2^n} M_6^{2^{n+1}} d_0 + \\ + M_6^{2^n} (\varepsilon_0 N_1)^{2^n} M_6^{2^{n+1}} d_0 \leq 2d_0 \left(N_1 M_6^3\right)^{2^n} \varepsilon_0^{2^n} \leq (\varepsilon_0 S)^{2^n}, \end{aligned}$$

where S is a constant independent on n and ε_0. Taking $\varepsilon_0 < S^{-1}$ we have $b_n \to 0$ as $n \to 0$. Furthermore one obtains also

$$a_{n+1} - a_n < (\varepsilon_0 T)^{2^n},$$

for a suitable constant T, independent on n and ε_0. We also require $\varepsilon_0 < T^{-1}$. Then

$$a_n - a_0 < \sum_{j=0}^{n-1} (\varepsilon_0 T)^{2^j} < 2T\varepsilon_0,$$

for all n, provided $\varepsilon_0 < \min\{S^{-1}, \frac{1}{2}T^{-1}\}$.

This is the bound we were looking for: it shows that $A_n(\varepsilon)$ is a Lipschitz function of ε, and $\mathcal{L}(A_n) - \mathcal{L}(A_0) = \mathcal{O}(\varepsilon)$. This means, using Lemmas 2.16 and 2.22 that the eigenvalues λ_j^n and the differences $\lambda_{j_1}^n - \lambda_{j_2}^n$ are Lipschitz from above and from below if ε is small enough.

To end up the proof we are going to take into account the resonances. As we want to skip the posible resonances due to λ_j^n and $\lambda_{j_1}^n - \lambda_{j_2}^n$ at the n-th step, we have to apply Lemma 2.24 for each one of the eigenvalues and couples. This amounts to skip a measure at most d^2 times the one we skipped in the frequency space. To go back to the parameter space, that is, to ε, we use the Lipschitz constant from below. In this way we obtain the Cantorian $\mathcal{E}$ with the desired properties. □

2.11. Proof of Theorem 2. As $\det A \neq 0$, the Contraction Lemma ensures that, if ε is small enough, there exists a function $x_0(\varepsilon)$ such that

$$(A + \varepsilon\overline{Q})x_0(\varepsilon) + \varepsilon\overline{g} + \overline{h}(x_0(\varepsilon), \varepsilon) = 0,$$

and verifying $x_0(\varepsilon) = \mathcal{O}(\varepsilon)$. Let us define

$$A_{x_0} = A + \varepsilon\overline{Q} + D_x\overline{h}(x_0(\varepsilon), \varepsilon),$$

and let $\underline{x}(t, \varepsilon)$ be such that

$$\dot{\underline{x}} = A\underline{x} + \varepsilon g(t, \varepsilon). \tag{19}$$

(The existence of that solution was shown in Lemma 2.10 and we recall that is $\mathcal{O}(\varepsilon)$). The terms of order ε of the matrix A_∞ are provided by Lemma 2.12 at the first step of the inductive process. This modified matrix $\widehat{A}$ is

$$\widehat{A} = A + \varepsilon\overline{Q} + \overline{D_x h(\underline{x}, t, \varepsilon)}.$$

Then,

$$||A_{x_0} - \widehat{A}|| = ||D_x\overline{h}(x_0(\varepsilon), \varepsilon) - \overline{D_x h(\underline{x}(t, \varepsilon), t, \varepsilon)}||.$$

But

$$\overline{D_x h(\underline{x}(t, \varepsilon), t, \varepsilon)} = \overline{(C + \varepsilon R(t))\underline{x}(t)} + \mathcal{O}(\varepsilon^2) = C\overline{\underline{x}} + \mathcal{O}(\varepsilon^2),$$

where $C = \frac{1}{2}D_{xx}h(0, t, 0)$ is a constant matrix by hypothesis. Moreover, it is also easy to obtain that

$$D_x\overline{h}(x_0(\varepsilon), \varepsilon) = (C + \varepsilon\overline{R})x_0(\varepsilon) + \mathcal{O}(\varepsilon^2) = Cx_0(\varepsilon) + \mathcal{O}(\varepsilon^2)$$

We have obtained that

$$||A_{x_0} - \widehat{A}|| = C(x_0(\varepsilon) - \overline{\underline{x}}) + \mathcal{O}(\varepsilon^2).$$

Now, averaging (19) we get that $A\overline{\underline{x}} + \varepsilon\overline{g} = 0$, and using $Ax_0(\varepsilon) + \varepsilon\overline{g} = -(\varepsilon\overline{Q}x_0(\varepsilon) + \overline{h}(x_0(\varepsilon), \varepsilon)) = \mathcal{O}(\varepsilon^2)$, we obtain $||x_0(\varepsilon) - \overline{\underline{x}}|| = \mathcal{O}(\varepsilon^2)$, that ends the proof. □

3. The neighbourhood of an elliptic equilibrium point of a Hamiltonian system. Let us consider the Hamiltonian

$$H^0(p,q,t) = H_0(p) + H_1(p,q,t),$$

where $|H_1|$ is small and it depends on time in a quasiperiodic way, having $\varphi = (\varphi_1, \dots, \varphi_r)$ as a vector of basic frequencies. To obtain an autonomous system we define $q_2 = q$, $p_2 = p$ and $q_1 = \varphi t$. So, the Hamiltonian takes the form

$$H(p_1,p_2,q_1,q_2) = (\varphi, p_1) + H_0(p_2) + H_1(p_2,q_2,q_1),$$

where p_1 are the actions corresponding to q_1 (obviously, they are not relevant in this problem and have only been added to obtain a Hamiltonian form). We are interested in the invariant tori that the unperturbed system $H = H_0(p_2)$ had. Note that the KAM theorem (see [1]) can not be applied directly due to the degeneracy of this case.[3]

We have considered this case in Theorem 3.1, and we have found that the proof of the classical KAM theorem (see [1]) still works, because the perturbing frequencies are not modified in any step of the inductive process, and we only have to worry about the proper frequencies of the Hamiltonian, that can be controlled provided that the nondegeneracy condition

$$\det\left(\frac{\partial^2 H_0}{\partial (p_2)^2}\right) \neq 0,$$

holds. The result obtained is that there exist invariant tori near the origin for ε small enough. The frequencies of these tori are the ones of the unperturbed tori plus the ones of the perturbation. This can be described saying that the unperturbed tori are "quasiperiodically dancing" under the "rhythm" of the perturbation. The tori whose frequencies are in resonance with the ones of the perturbation are destroyed.

Finally, in case that the origin is not a fixed point of the perturbed Hamiltonian, we can reduce to this case performing a change of variables transforming the quasiperiodic orbit that replaces the equilibrium point (we recall that this orbit exists for a Cantorian set of values of ε) in a fixed point.

THEOREM 3.1. *Let us consider the Hamiltonian*

$$H(p_1,p_2,q_1,q_2) = (\varphi, p_1) + H_0(p_2) + \varepsilon H_1(p_2,q_1,q_2),$$

where q_1 are the angles of the perturbation, p_1 are the corresponding actions, q_2, p_2 are the angles and actions of the unperturbed system and $\varphi = (\varphi_1, \dots, \varphi_{n_1})$ is a constant vector of frequencies satisfying the nonresonance condition

$$| < k, \varphi > | > \frac{c}{|k|^\gamma}, \quad \forall k \in \mathbb{Z}^{n_1} \setminus \{0\}, \quad \gamma > n_1 - 1.$$

Let G^1 be a compact domain of $\mathbb{R}^{n_1}$, let G^2 be a compact domain of $\mathbb{R}^{n_2}$ and let G be $G^1 \times G^2$. Suppose now that this Hamiltonian function $H(p_1,p_2,q_1,q_2) = H(p,q)$ is analytic on the domain $F = \{(p,q)\,/\,p = (p_1,p_2) \in G, |\mathrm{Im}\, q| \le \rho\}$ and has period 2π with respect to the variables q. Let us assume that, in the domain F,

$$\det\left|\frac{\partial^2 H_0}{\partial (p_2)^2}\right| \neq 0.$$

[3] However, see the comments in [2], pp. 193–194, for a related result.

Then, if ε is small enough, the motion defined by the Hamiltonian equations

$$\begin{aligned} \dot{p}_1 &= -\frac{\partial H}{\partial q_1}, \quad \dot{q}_1 = \varphi, \\ \dot{p}_2 &= -\frac{\partial H}{\partial q_2}, \quad \dot{q}_2 = \frac{\partial H}{\partial p_2}, \end{aligned} \tag{20}$$

has the following properties:

1. *There exists a decomposition* $\mathrm{Re}\, F = F_1 + F_2$, *where* F_1 *is invariant and* F_2 *is small:* $\mathrm{mes}\, F_2 \leq \kappa_1(\varepsilon)\, \mathrm{mes}\, F$, *where* $\kappa_1(\varepsilon)$ *is* $o(\varepsilon)$.
2. F_1 *is composed of invariant* n*-dimensional analytic tori* I_ϕ, *defined parametrically by the equations*

   $$p = p_\phi + f_\phi(Q), \quad q = Q + g_\phi(Q),$$

   where f_ϕ, g_ϕ *are analytic functions of period* 2π *in the variables* Q, *and* ϕ *is a parameter determining the torus* I_ϕ. *In fact* ϕ *consists of all the frequencies, i.e., the ones of the external excitation and the proper frequencies:* $\phi = (\varphi_1, \ldots, \varphi_{n_1}, \omega_1, \ldots, \omega_{n_2})$.
3. *The invariant tori* I_ϕ *differ little from the tori* $p = p_\phi$*:*

   $$|f_\phi(Q)| < \kappa_2(\varepsilon), \quad |g_\phi(Q)| < \kappa_2(\varepsilon),$$

   where $\kappa_2(\varepsilon)$ *is* $o(\varepsilon)$.
4. *The motion (20) on the invariant torus* I_ϕ *is quasiperiodic with* n *frequencies* $\varphi_1, \ldots, \varphi_{n_1}$, $\omega_1, \ldots, \omega_{n_2}$ $(n = n_1 + n_2)$*:*

   $$\dot{Q} = \phi, \quad \omega = \left.\frac{\partial H_0}{\partial p_2}\right|_{p_\phi}.$$

3.1. Sketch of the proof of Theorem 3. The proof of this Theorem is essentially the same of the KAM Theorem contained in [1], and its technical details can be found inside [15]. Here we show the idea of that proof.

Let us define p and q as the vectors p_1, p_2 and q_1, q_2 respectively. Now, the Hamiltonian we have is

$$H^\varepsilon = (\varphi, p_1) + H_0(p_2) + \varepsilon \overline{H}_1(p_2) + \varepsilon \widetilde{H}_1(p_2, q), \tag{21}$$

and let us consider the generating function $S(P, q) = Pq + \sum_{k \neq 0} S^k(P_2) e^{(k,q)\sqrt{-1}}$. If we perform the canonical change of variables

$$\begin{aligned} p_1 &= P_1 + \varepsilon \frac{\partial S}{\partial q_1}, \\ p_2 &= P_2 + \varepsilon \frac{\partial S}{\partial q_2}, \\ Q_1 &= q_1, \\ Q_2 &= q_2 + \varepsilon \frac{\partial S}{\partial P_2}, \end{aligned}$$

to (21) we obtain

$$H^\varepsilon = (\varphi, P_1) + H_0(P_2) + \varepsilon \overline{H}_1(P_2) + \varepsilon F + \varepsilon^2 R(P_2, q),$$

where $F = (\varphi, S_{q_1}) + (\omega(P_2), S_{q_2}) + \widetilde{H}_1(P_2, q)$ and $\omega(p_2) = \frac{\partial H_0}{\partial p_2}(p_2)$. Let $\phi(P_2)$ be the vector $\varphi, \omega(P_2)$. We ask for the condition $F = 0$:

$$(\phi(P_2), S_q) + \widetilde{H}_1(P_2, q) = 0.$$

Now, using that $\widetilde{H}_1(P_2, q) = \sum_{k\neq 0} h_1^k(P_2) e^{(k,q)\sqrt{-1}}$, the coefficients of the Fourier expansion for the generating function S can be obtained easily:

$$S^k(P_2) = \frac{h_1^k}{(\phi(P_2), k)}\sqrt{-1}.$$

To ensure the convergence of that series, it is enough the usual nonresonance condition

$$|(\phi(P_2), k)| \geq \frac{c}{|k|^\gamma}, \tag{22}$$

that allows to prove the convergence in a smaller strip than the one on which $\widetilde{H}_1$ is analytic. With this the Hamiltonian takes the form

$$H^\varepsilon = (\varphi, P_1) + H_1^\varepsilon(P_2) + \varepsilon^2 \overline{H}_3(P_2) + \varepsilon^2 \widetilde{H}_3(P_2, Q).$$

This new Hamiltonian is very similar to (21), but with ε^2 instead of ε. Note that the difference between this proof and the one contained in [1] is the condition (22). Due to the fact that the first components of $\phi(P_2)$ are the ones of φ, *that are constant in all the inductive process*, we only have to worry about the last ones, $\omega(P_2)$. These components are different at each step of the process, but they can be controlled by the nondegeneracy condition

$$\det\left(\frac{\partial^2 H_0}{\partial (p_2)^2}\right) \neq 0.$$

The way to do that is exactly the same one showed in [1]. Note that, to get an rigorous proof of this theorem, we only need to copy the proof contained in [1], but adding the "parameter" φ. The unique difference is that now, the nonresonance condition is stronger, in the sense that we must eliminate a bigger set of (resonant) tori.

Acknowledgements. A. Jorba has been partially supported by the CIRIT grant BE94/III–513. He also wishes to thank the hospitality of R. de la Llave and of the Dept. of Mathematics of the University of Texas at Austin. C. Simó has been partially supported by the CICYT PB90–0580. Both authors thank useful comments by A. Neishtadt, the supports of CICYT grant ESP91–0403 and of the European Network grant ERBCHRXCT 940460.

REFERENCES

[1] Arnol'd V. I.: *Proof of a Theorem of A. N. Kolmogorov on the Invariance of Quasi-Periodic Motions Under Small Perturbations of the Hamiltonian*, Russ. Math. Surveys, 18:5 (1963), pp. 9–36.

[2] Arnol'd V. I., Kozlov V. V. and Neishtadt A. I.: *Dynamical Systems III*, Encyclopaedia of Mathematical Sciences, Springer-Verlag, Berlin, 1988.

[3] Bogoljubov N. N., Mitropoliski Ju. A. and Samoilenko A. M.: *Methods of Accelerated Convergence in Nonlinear Mechanics*, Springer-Verlag, New York, 1976.

[4] Chierchia L.: *Absolutely continuous spectra of quasiperiodic Schrödinger operators*, Journal of Math. Phys., 28 (1987), pp. 2891–2898.

[5] Díez C., Jorba A. and Simó C.: *A Dynamical Equivalent to the Equilateral Libration Points of the Earth-Moon System*, Celestial Mechanics, 50 (1991), pp. 13–29.

[6] Dinaburg E. I. and Sinai J. G.: *The One-dimensional Schrödinger Equation with Quasiperiodic Potential*, Funckcional. Anal. i Prilozen., 9 (1975), no. 4, pp. 8–21.

[7] Eliasson L. H.: *Perturbations of Stable Invariant Tori for Hamiltonian Systems*, Ann. Scuola Norm. Su. Pisa, 15 (1988), pp. 115–148.

[8] Eliasson L. H.: *Floquet Solutions for the 1-dimensional Quasi-periodic Schrödinger Equation*, Comm. Math. Phys., 146 (1992), no. 3, pp. 447–482.

[9] Fink A. M.: *Almost Periodic Differential Equations*, Lecture Notes in Mathematics 377, Springer-Verlag, Berlin, 1974.

[10] Gómez G., Jorba A., Masdemont J. and Simó C.: *A Quasiperiodic Solution as a Substitute of* L_4 *in the Earth-Moon System*, Proceedings of the 3rd International Symposium on Spacecraft Flight Dynamics, ESA Publications Division, ESTEC, Noordwijk, Holland (1991), pp. 35–41.

[11] Gómez G., Jorba A., Masdemont J. and Simó C.: *Study Refinement of Semi-Analytical Halo Orbit Theory*, ESOC Contract 8625/89/D/MD(SC), Final Report, 1991.

[12] Gómez G., Jorba A., Masdemont J. and Simó C.: *Study of Poincaré Maps for Orbits near Lagrangian Points*, ESOC Contract 9711/91/D/IM(SC), Final Report, 1993.

[13] Gómez G., Llibre J., Martínez R. and Simó C.: *Study on Orbits near the Triangular Libration Points in the perturbed Restricted Three-Body Problem*, ESOC Contract 6139/84/D/JS(SC), Final Report, 1987.

[14] Johnson R.A. and Sell G. R.: *Smoothness of Spectral Subbundles and Reducibility of Quasi-Periodic Linear Differential Systems*, Journal of Differential Equations, 41 (1981), pp. 262–288.

[15] Jorba A.: *On Quasiperiodic Perturbations of Ordinary Differential Equations*, Ph. D. Thesis, Universitat de Barcelona, 1991.

[16] Jorba A. and Simó C.: *On the Reducibility of Linear Differential Equations with Quasiperiodic Coefficients*, Journal of Differential Equations, 98:1 (1992), pp. 111–124.

[17] Moser J.: *Convergent Series Expansions for Quasiperiodic Motions*, Math. Annalen, 169 (1967), pp. 136–176.

[18] Moser J. and Pöschel J.: *On the Stationary Schrödinger Equation with a Quasiperiodic Potential*, Physica A, 124 (1984), pp. 535–542.

[19] Wilkinson J. H.: *The Algebraic Eigenvalue Problem*, Oxford Univ. Press, 1965.