Annali della
Scuola Normale Superiore di Pisa
Classe di Scienze

L. H. Eliasson

Perturbations of stable invariant tori for hamiltonian systems

Annali della Scuola Normale Superiore di Pisa, Classe di Scienze 4e série, tome 15, nº 1 (1988), p. 115-147

<http://www.numdam.org/item?id=ASNSP_1988_4_15_1_115_0>

Perturbations of Stable Invariant Tori for Hamiltonian Systems

L.H. ELIASSON

1. - Introduction

Let (X, α) be a symplectic manifold of dimension $2n + 2m$, and let $\underline{h} : X \to \mathbf{R}^{n+m}$ be an analytic mapping such that any two of its components commute for the Poisson bracket. We assume that $\underline{h}$ is proper with connected fibers, and that $D\underline{h}$ is of constant rank on each fiber. Then the fibers are tori of dimension 0 up to $n + m$.

Suppose that $\underline{h}^{-1}(c)$ is a torus of dimension n. Then, under fairly general conditions on $\underline{h}$, there exist a neighbourhood U of $\underline{h}^{-1}(c)$ and a diffeomorphism ϕ of U into $T^*(\mathbf{T}^n \times \mathbf{R}^m)$, $\mathbf{T} = \mathbf{R}/(2\pi)\mathbf{Z}$, such that

$$\phi_* \alpha = \sum_{1 \le j \le n} dx_j \wedge dy_j + \sum_{1 \le i \le m} dz_i \wedge dz_{i+m}$$

$$\phi(\underline{h}^{-1}(c)) = \{(x, y, z) \in \mathbf{T}^n \times \mathbf{R}^n \times \mathbf{R}^{2m} : y = z = 0\}$$

and

$$\underline{h} \circ \phi^{-1} \text{ is a function of } y \text{ and } y'$$

where $y'_i = \frac{1}{2}(z_i^2 + z_{i+m}^2)$, $1 \le i \le m$. The variables y and y' are *generalized action variables*. To y correspond the angle variables x, but y' are singular and do not possede any conjugate angle variables. (The case $m = 0$ is due to Arnold [1], the case $n = 0$ is a theorem of J. Vey [2], and the general result follows from these two cases).

An *integrable Hamiltonian system* $X = X_h$, with the components of $\underline{h}$ as integrals, has then a *Hamiltonian* h which is a function of y and y' only, and the action of this system leaves all tori $y =$ const., $y' =$ const. invariant. It is well known that the maximal tori, i.e. those of dimension $n + m$, in general persist under perturbations of the Hamiltonian [3,4]. In this paper, we will study the lower dimensional stable tori $y =$ const. $y' = 0$. Their perturbation theory

Pervenuto alla Redazione il 29 Maggio 1987 e in forma definitiva il 19 Aprile 1988.

involves a particular *small divisor problem* which we will describe and solve, and we will show that, under a certain non-degeneracy condition on h, many of these tori persist under perturbations.

The problem

We therefore consider a real analytic function of the form

$$h(y,z) = h_0(y) + \langle \Omega(y), y' \rangle + O^3(z)$$

defined for y, z in some open subset of $\mathbf{R}^n \times \mathbf{R}^{2m}$. h is the Hamiltonian of a system X_h in $T^*(\mathbf{T}^n \times \mathbf{R}^m)$ which, neglecting $O^3(z)$, is integrable with integrals y and y'. The *frequency map* of h is the map

$$y \longmapsto (\omega(y), \Omega(y)) \in \mathbf{R}^n \times \mathbf{R}^m, \text{ where } \omega(y) = Dh_0(y).$$

If we let $J = J(h, B)$ be the image of some open set B in $\mathbf{R}^n$ under this map, then, to any $(\omega, \Omega) \in J$, there corresponds a torus $y =$ const., $y' = 0$ which is invariant under X_h, and for which ω is the tangential frequency vector of the flow on this torus, and Ω is the normal frequency vector. (An invariant n-torus for a Hamiltonian system is said to have frequency vector (ω, Ω) if there exist symplectic coordinates in a neighbourhood of the torus in which the Hamiltonian can be written as $\langle \omega, y \rangle + \langle \Omega, y' \rangle$, neglecting higher order terms in y and z. These are normal coordinates in the sense of [6]). The problem is now what happens with such an invariant torus when we perturb the Hamiltonian system.

The Hamiltonian h is said to be *non-degenerate* if, for all y,

$$\det(D\omega(y)) \neq 0 \tag{1}$$

$$\langle l, \Omega(y) \rangle \neq \langle l, \omega(y)(D\omega(y))^{-1} D\Omega(y) \rangle \text{ for all } l \in \mathbf{Z}^m \backslash 0, \ |l| \leq 3. \tag{2}$$

(Here $\langle , \rangle$ is the usual scalar product and $|l| = |l_1| + \cdots + |l_m|$. Notice that matrices operate on vectors to the right). Since, by (1), $y \to \omega(y)$ is invertible, we can consider $\Omega(y)$ as a function of ω. The right hand side of (2) is then just the derivative of $\langle l, \Omega(y) \rangle$ in the direction ω.

Condition (1) is the standard non-degeneracy condition in the perturbation theory of invariant tori, and the second condition is needed since we are dealing with non-maximal tori. We briefly explain its role.

Let $\tau > n - 1$. We say that a frequency vector $(\omega, \Omega) \in \mathbf{R}^n \times \mathbf{R}^m$ satisfies the *Diophantine condition* $DC(K)$ if

$$|\langle k, \omega \rangle + \langle l, \Omega \rangle| \geq K^{-1}(|k| + |l|)^{-\tau} \text{ for all } (k, l) \in \mathbf{Z}^n \times \mathbf{Z}^m \backslash 0, \ |l| \leq 3. \tag{3}$$

(Since τ will be fixed throughout this paper, we will not make it explicit in the notations). That (ω, Ω) belongs to $DC = \bigcup DC(K)$ is the *sine qua non* condition for all perturbation theory of quasiperiodic motions. In fact, the scalar products $\langle k, \omega\rangle + \langle l, \Omega\rangle$ are the well known *small divisors* which appear as denominators in the Fourier expansion of the formal quasiperiodic solutions.

Let us also consider a *homogeneous* Diophantine condition. $\omega \in \mathbf{R}^n$ satisfies $DC_0(K)$ if

$$|\langle k, \omega\rangle| \geq K^{-1}|k|^{-\tau} \text{ for all } k \in \mathbf{Z}^n\backslash 0 \tag{4}$$

and $DC_0 = \bigcup DC_0(K)$.

Let B be an open set on which the frequency map is defined. Condition (1) now tells us that J is a graph over $\omega(B)$ and, as we shall see, that

$$\omega(B) \cap DC_0 \neq \emptyset.$$

And condition (2) implies not only that

$$J \cap DC \neq \emptyset$$

(this is assured also by other conditions) but something more. In fact, if $\omega \in DC_0$, then the whole line $\mathbf{R}\omega$ belongs to DC_0, and condition (2) implies that the projection of $J \cap DC$ on the first factor ω contains a dense subset of (a part of) $\mathbf{R}\omega$. This means that if $\omega(y_0) \in DC_0$ for some y_0, then $(\omega(y), \Omega(y)) \in DC$, if not for the same argument y_0, at least for an arbitrarily small 1-dimensional modification of y in the direction $\omega(y_0)$.

The particular constant K that can be used here depends on J, and it is better (i.e. smaller) the larger is B and the less degenerate is h.

Formulation of the result

THEOREM A. *Let $h(y,z) = h_0(y) + \langle\Omega(y), y'\rangle + O^3(z)$ be a non-degenerate real analytic function, defined in a neighbourhood $D_{\mathbf{R}}$ of $\{(x,y,z) : y = z = 0\}$ in $T^*(\mathbf{T}^n \times \mathbf{R}^m)$. Let B be a ball such that $\{(x,y,z) : \frac{1}{2}y \in B, z = 0\} \subset \underline{D}_{\mathbf{R}}$, and let D be a complex neighbourhood of $D_{\mathbf{R}}$. Then*

1) *there exists a constant $\mathbf{C}$, depending only on h and $D_{\mathbf{R}}$, such that $J(h,B) \cap DC(K) \neq \emptyset$ for $K \geq \mathbf{C}b^{-1}$, where b is the diameter of B;*

2) *there exists a constant $\mathbf{C}$, depending only on h and D, such that the following holds: for any $K > 0$, for any real analytic function f on D such that*

$$\sup_D |f - h| \leq \mathbf{C}\frac{1}{K^2}$$

and for any $(\omega(y_0), \Omega(y_0)) \in DC(K) \cap J(h,B)$, there exists an n-torus Λ, which is invariant under the Hamiltonian vector field X_f. X_f is linearizable *on* Λ, *with frequency vector* $(\tilde{\omega}, \tilde{\Omega})$ *satisfying* $\tilde{\omega} = t\omega(y_0)$, *for some scalar* t *such that*

$$|t-1| \leq \mathbf{C}K \ \sup_D |f-h|.$$

Moreover, Λ *is a graph* $\{(x, \iota(x)) : x \in \mathbf{T}^n\}$ *for some function* ι *and*

$$\sup_{x \in \mathbf{T}^n} |\iota(x) - (y_0, 0)| \leq \mathbf{C}K^2 \ \sup_D |f-h|.$$

This theorem establishes the existence of an invariant torus for X_f in a neighbourhood of $\{y = y_0, z = 0\}$. In order to have $(\tilde{\omega}, \tilde{\Omega}) = (\omega(y_0), \Omega(y_0))$ we need, as described in [5,6], more parameters than are available in our problem. We cannot even have $\tilde{\omega} = \omega(y_0)$. What we do achieve (and here condition (2) is essential) is that $\tilde{\omega}$ is parallel to $\omega(y_0)$. In fact, the lenght of the tangential frequency vector is an important parameter which we need in order to control the normal frequencies.

The theorem remains true under weaker assumptions. Indeed, one can replace $|l| \leq 3$ by $|l| \leq 2$ in (2) and (3).

It is also worth noticing that we have assumed nothing on that part of h which is cubic in z. In particular, it is possible that the perturbed system has no invariant tori of maximal dimension near Λ.

Theorem A is effective when f is of the form $h + \varepsilon\tilde{h}$, and h is non-degenerate. In this case it proves the existence of invariant tori for sufficiently small ε. Often, however, also h depends on ε, and then it is necessary to have a precise description of how the constant $\mathbf{C}$ depends on h. Our proof provides such a description in terms of the inequalities (1) and (2) and the supremum norm of h. As an immediate consequence of this description we get the following theorem.

THEOREM B. *Let* $h = h(w, z)$ *be defined and real analytic in a neighbourhood of the origin in* $\mathbf{R}^{2n} \times \mathbf{R}^{2m}$. *Let* $\sum dw_j \wedge dw_{j+n} + \sum dz_i \wedge dz_{i+m}$ *be the symplectic structure on* $\mathbf{R}^{2n} \times \mathbf{R}^{2m}$, *and let* $y_j = \frac{1}{2}(w_j^2 + w_{j+n}^2)$ *and* $y_i' = \frac{1}{2}(z_i^2 + z_{i+m}^2)$. *Assume that* $h = Dh = 0$ *at the origin, and that* h *satisfies the following two conditions*:

1) *the quadratic part* $h_2 = \langle \omega, y \rangle + \langle \Omega, y' \rangle$ *satisfies*

$$\langle \omega, k \rangle + \langle \Omega, l \rangle \neq 0, \ \ |k| + |l| \leq 4,$$

2) *the Birkhoff normal form* $h_2 + \frac{1}{2}\langle yM_1, y \rangle + \langle yM_2, y' \rangle + \frac{1}{2}\langle y'M_3, y' \rangle$ *is non-degenerate.*

Then the Hamiltonian flow of h *has* n*-dimensional invariant tori, close to the subspace* $z = 0$, *in any neighbourhood of the origin.*

About the proof

The proof proceeds by a quadratic iteration of Newton type (KAM-technique). We start with an Hamiltonian h on normal form, for example the form described above, and a small perturbation f. Let J be the image of the frequency map of h, and let $(\omega, \Omega) \in DC(K) \cap J$ be the unique image of some y_0, which we may assume to be 0. In a neighbourhood of the torus $y = 0, z = 0$ we construct a symplectic diffeomorphism Φ, close to the identity, such that $(h + f) \circ \Phi$ is a sum of a Hamiltonian h^+, on normal form, and a perturbation f^+ which is much smaller that f near the torus, but may become large far away. Moreover, f^+ becomes large with K.

Since there is an n-parameter family of n-dimensional tori, we have n parameters in our problem (by condition (1)). By an appropriate choice of $n-1$ of these parameters we can achieve that $h^+(y, y') = \langle\tilde{\omega}, y\rangle + \langle\tilde{\Omega}, y'\rangle$, neglecting non-linear terms in y and y', with $\tilde{\omega}$ parallel to ω. Since DC is dense in $\mathbf{R}^{n+m}$ (as can easily be shown) we can assume that $(\tilde{\omega}, \tilde{\Omega}) \in DC(K^+)$ for some K^+.

In order to proceed by an iteration we must control K^+ since it will influence the size of the perturbative term in the next step. Using the last parameter we can now start to vary the lenght of $\tilde{\omega}$. Let therefore $\tilde{\omega}_t = t\omega$, and let $\tilde{\Omega}_t$ be defined by the condition that $(\tilde{\omega}_t, \tilde{\Omega}_t) \in J^+$, where J^+ is the image of the frequency map of h^+. Since h^+ satisfies condition (2), $(\tilde{\omega}_t, \tilde{\Omega}_t) \in DC$ for almost all t (for which $\tilde{\Omega}_t$ is defined). Let K_t be the best admissible constant. It depends on t in the following way: the smaller one takes $|t-1|$ the larger will be K_t, and it is only for "not too small" values of $|t-1|$ that one can hope to have a K_t which is not too large (compared with K).

Now we got two contradictory requirements. In order to assure that, in the next step of the iteration, the perturbative term becomes sufficiently small we must keep f^+ small, i.e. we must stay "close to" the torus $y = 0, z = 0$. On the other hand, we must also control the admissible constant $K^+ = K_t$, which requires that t is "not too close to" 1, i.e. that we are "not too close to" the torus $y = 0, z = 0$.

In section II we give all the necessary details about the sets $DC(K)$ and $J \cap DC(K)$. Theorem A1) follows from proposition 1 proven there, but part 2) requires the more detailed description of proposition 2. In section III, where we formulate the usual "inductive lemma" (proposition 3), we show that one can control the size of both f^+ and K^+. In section IV, finally, we use this result as the inductive step in a standard rapid iteration process in order to prove the "stability" of certain Hamiltonians under perturbations (Theorem C.). Theorem A and B will then follow easily from this more general result.

We will formulate the proposition 3 and theorem C for a slightly different class of functions (*the normal forms*) than the ones introduced above. The particular normal form chosen is very much a matter of convenience, and the one we propose is certainly not the most general.

Literature

We want to finish by giving some references to previous works on the perturbation theory of quasi periodic motions.

Periodic solutions $(n = 1)$. This is a much simpler problem since there are no small divisors at all, and a perturbation theory for such solutions was constructed by Poincaré [7,16].

Maximal tori $(m = 0)$. In this case there are small divisors but also sufficiently many parameters available (under condition (1)) in order to restore the perturbed frequencies completely. This problem was solved by Kolmogorov and Arnold [3,4].

Next to maximal tori $(m = 1)$. For such solutions, a perturbation theory was constructed by Moser [5,6] under condition (1) and under the condition that

$$\det \begin{pmatrix} D\omega & D\Omega \\ \omega & \Omega \end{pmatrix} \neq 0.$$

(These two conditions are equivalent to (1) + (2), and our condition is indeed inspired by that of Moser).

In this case there are sufficiently many parameters available so that we can let the perturbed frequency vector be parallel to the unperturbed one.

Unstable invariant tori. In this case $\Omega_1, \cdots, \Omega_m$ are non real, and if the imaginary parts $\Im\Omega_1, \cdots, \Im\Omega_m$ avoid certain hyperplanes (finitely many) then DC reduces to DC_0. These tori have been studied by Moser and Graff [6,8,9]. (Actually, [8,9] deal with a more general situation).

The case of *stable invariant tori*, which is the object of this paper, poses more difficult small divisor problems. A perturbation theory using other non-degeneracy conditions has been described by Melnikov in two articles [10,11]. (We shall discuss his conditions in section II).

Acknowledgement

We want to thank L. Hörmander and J. Moser for their interest and criticism at various occasions during the preparation of this work, which was partly done at the Mittag-Leffler Institute.

II SMALL DIVISORS

In this section we study the intersection $J \cap DC$ defined in the introduction. In particular, we ask for which constants K the set $J \cap DC(K)$ is non void, and we explicitly determine K in terms of J. All results are of measure theoretical character, stating that these intersections are non void since they have positive Lebesgue measure. The results needed for the construction of invariant tori are formulated in proposition 1 and 2.

The set DC(K)

Let ω be a vector in $\mathbf{R}^n \backslash 0$, and let u be a real number. For any positive number d and any non-negative number N we define

$$\Gamma_N(d) = \Gamma_N(\omega, u, d) = \{k \in \mathbf{Z}^n : |\langle k, \omega \rangle + u| \le d, \ |k| \le N\}.$$

LEMMA 1. $\#\Gamma_N(d) \le C(N+1)^{n-1}(1 + d|\omega|^{-1})$, *where* C *is a constant that only depends on* n.

PROOF. By homogeneity, we can assume that $|\omega| = 1$. Now the balls of radius $\frac{1}{2}$ and centered at the points in $\Gamma_N(d)$ are disjoint and contained in the intersection of a slab of width $(d+1)$ with a ball of radius $N+1$, centered at the origin. Since this intersection has volume less than $C(N+1)^{n-1}(d+1)$, the estimate follows. ∎

LEMMA 2. *Let* B *be a ball in* $\mathbf{R}^{n+m}$ *of radius* b *and centered at* (ω_0, Ω_0). *Then*

1) $\text{meas}(\complement DC(K) \cap B) \le C b^{n+m-1} K^{-1}$;
2) $\text{meas}(\complement DC(K) \cap B \cap \{\omega = \omega_0\}) \le C b^{m-1} K^{-1}(b|\omega_0|^{-1} + 1)$, *if* $\omega_0 \in DC_0(K)$;

where C *is a constant only depending on* n, m *and* τ. ($\complement DC$ *denotes the complementary set of* DC, *and* meas *is the Lebesgue measure in* $\mathbf{R}^{n+m}$ *and* $\mathbf{R}^m$, *respectively*).

PROOF. Let $(k, l) \ne (0, 0)$ be given. The set of all $(\omega, \Omega) \in B$, such that

$$|\langle k, \omega \rangle + \langle l, \Omega \rangle| \le K^{-1}(|k| + |l|)^{-\tau}$$

is contained in the intersection of B with a slab of width less than $C_1 K^{-1}(|k| + |l|)^{-\tau-1}$, hence of measure less than $C_2 b^{n+m-1} K^{-1}(|k| + |l|)^{-\tau-1}$.
Since $\tau > n-1$, the series

$$\sum_{(k,l) \in \mathbf{Z}^{n+m} \backslash 0, \ |l| \le 3} (|k| + |l|)^{-\tau-1}$$

converges, and 1) follows.

The proof of 2) is slightly different. Let (k, l) be given as before. We can assume that $l \ne 0$ since $\omega_0 \in DC_0(K)$. Then the set X of all Ω in $B \cap \{\omega = \omega_0\}$, such that

$$|\langle k, \omega_0 \rangle + \langle l, \Omega \rangle| \le K^{-1}(|k| + |l|)^{-\tau}$$

is contained in the intersection of $B \cap \{\omega = \omega_0\}$ with a slab of width less than $C_3 K^{-1}(|k| + |l|)^{-\tau}$, hence of measure less than $C_4 b^{m-1} K^{-1}(|k| + |l|)^{-\tau}$.

Moreover, since

$$|\langle k, \omega_0\rangle + \langle l, \Omega_0\rangle| \leq |\langle k, \omega_0\rangle + \langle l, \Omega\rangle| + C_5 b$$

the set X is void unless

$$k \in G = \bigcup_{N \geq 0} G_N, \quad G_N = \Gamma_N(\omega_0, \langle l, \Omega_0\rangle, C_5(K^{-1} + b)).$$

Hence, the measure in 2) can be estimated by

$$C_6 b^{m-1} K^{-1} \sum_{k \in G} (|k| + 1)^{-\tau}$$

and, using Abel's summation formula, by

$$C_7 b^{m-1} K^{-1} \sum_{N \geq 0} (\# G_N)(N + 1)^{-\tau - 1}.$$

The result now follows from lemma 1, if we only observe that $K|\omega_0| \geq 1$ (since $\omega_0 \in DC_0(K)$). ∎

The conclusions of the lemma are the following: almost all (ω, Ω) in $\mathbf{R}^{n+m}$ satisfy DC; if ω_0 satisfies DC_0, then (ω_0, Ω) satisfies DC for almost all Ω in R^m. In both cases the admissible constant K depends on b^{-1}, and becomes very large when b is small. However, we shall see below that if B is centered at $(\omega_0, \Omega_0) \in DC$ ("a good frequency"), then we can get rid of this dependence on b^{-1}.

The set $J \cap DC(K)$

We now consider a C^1 frequency map $y \longmapsto (\omega(y), \Omega(y)) \in \mathbf{R}^{n+m}$, defined in a ball $B = B(b)$ in $\mathbf{R}^n$. ($B(r)$ denotes a ball with radius r and center at the origin). We assume that the map is *non-degenerate* in the sense that

(5) $$|(D\omega(y))^{-1}| \leq \mu_2 \text{ for all } y \in B$$

(6) $$|\langle l, \Omega(y) - \omega(y)(D\omega(y))^{-1} D\Omega(y)\rangle| \geq \mu_4^{-1} \text{ for all } y \in B, \ 0 < |l| \leq 3.$$

Let μ_1 and μ_3 satisfy

(7) $$\sup_B(|\omega| + |\Omega|) \leq \mu_3$$

(8) $$\sup_B(|D\omega| + |D\Omega|) \leq \mu_1.$$

It follows from this that

(9) $$min(\mu_1\mu_2, \mu_1\mu_2\mu_3\mu_4, \frac{\mu_2\mu_3}{b}) \geq C$$

where C is a positive constant that only depends on m and n.

PROPOSITION 1. *Under the assumptions* (5)-(8), *there exists a non void set G in B such that, for all $y \in G$,*

$$(\omega(y), \Omega(y)) \in DC(K) \text{ for any } K \geq C(\mu_1\mu_2\mu_3\mu_4)\frac{\mu_2}{b}$$

where C is a constant that only depends on m, n and τ.

PROOF. Suppose that there is a y_0 in $B(\frac{1}{2}b)$ such that $\omega(y_0) \in DC_0(K_0)$. Let $\omega_0 = \omega(y_0)$, and define $y(t)$ through the equation $\omega(y(t)) = (1+t)\omega_0$, $y(0) = y_0$. Then $y(t)$ is defined for $|t| < \frac{b}{2\mu_2\mu_3}$ since

$$|y(t) - y_0| = |\omega^{-1}((1+t)\omega_0) - \omega^{-1}(\omega_0)| \leq |t|\mu_2\mu_3.$$

Let now $0 < |l| \leq 3$ and consider

$$\langle k, \omega(y(t))\rangle + \langle l, \Omega(y(t))\rangle = (1+t)(\langle k, \omega_0\rangle + V(t)).$$

This equality defines the function $V = V_l$. Its derivative is

$$-(1+t)^{-2}\langle l, \Omega(y) - \omega(y)(D\omega(y))^{-1}D\Omega(y)\rangle \quad (y = y(t))$$

so we have

$$\frac{1}{C_2\mu_4} \leq |DV(t)| \leq C_2\nu, \ \nu = \mu_1\mu_2\mu_3,$$

for $|t| < \varepsilon = \frac{1}{2}\, min(\frac{b}{\mu_2\mu_3}, 1)$. (For the upper bound we have used (9)). This implies that the image of $]-\varepsilon, \varepsilon[$ under V contains a ball B_1 with radius $\varepsilon(C_3\mu_4)^{-1}$, and is contained in a ball B_2 with radius $C_3\varepsilon\nu$, both balls being centered at $\langle l, \Omega(y_0)\rangle$.

By lemma 2, it follows that the set X of all V in B_2, such that

$$|\langle k, \omega_0\rangle + V| \leq K^{-1}(|k| + |l|)^{-\tau} \text{ for some } k \in \mathbf{Z}^n$$

has measure less than

$$C_4K^{-1}(\varepsilon\nu|\omega_0|^{-1} + 1) \leq C_5K^{-1}(\varepsilon\nu K_0 + 1).$$

This measure is smaller than the volume of B_1, if

$$K \geq C_6(\mu_1\mu_2\mu_3\mu_4)\, max(\frac{\mu_2}{b}, K_0).$$

(Also here we have used (9)).

Hence, there exists a $t \in]-\varepsilon, \varepsilon[$, such that

$$|\langle k, \omega(y(t))\rangle + \langle l, \Omega(y(t))\rangle| \geq K^{-1}(|k| + |l|)^{-\tau} \text{ for all } k \in \mathbf{Z}^n, \ (k,l) \neq (0,0),$$

if K verifies the above inequality. It is clear that if we take C_6 sufficiently large, then this condition will be fulfilled for all $0 < |l| \leq 3$ and, hence, for all $|l| \leq 3$ simultaneously.

In order to conclude we need to establish the existence of a point $y_0 \in B(\frac{1}{2}b)$ such that $\omega(y_0) \in DC_0(K_0)$, where $K_0 = C_7\mu_2 b^{-1}$ for some constant C_7. But this follows from lemma 2, if we only observe that $\omega(B(\frac{1}{2}b))$ covers a ball of radius $\frac{b}{2\mu_2}$. This proves the proposition. ∎

So we see that the preimage G of $DC(K)$, under the frequency map, is non void if K is sufficiently large, as given by the proposition. Moreover, the proof gives a stronger statement than announced. The set G, in fact, is quite large since the projection of $\omega(G)$ on the unit sphere has positive measure. It is likely that $\omega(G)$ itself (and hence G) also has positive measure, but the argument does not suffice to establish this.

The set $DC(K)$ near good frequencies

The result of the preceeding paragraph shows that one can always find a Diophantine frequency vector in any (non-degenerate) n-parameter family. The admissible constant K depends (among other things) on the diameter b of the domain of definition of the family. When b is small this is rather bad, but we shall now show that one can get rid of this dependence on b if one is close to a given Diophantine frequency vector. We start with a lemma.

LEMMA 3. *Suppose* $\omega \in DC_0(K)$ *and*

$$|\langle k, \omega\rangle + u| \geq K^{-1}(|k| + 1)^{-\tau} \textit{ for all } k \in \mathbf{Z}^n, \ |k| + |u| \neq 0. \tag{10}$$

Then

$$\#\Gamma_N(d) \leq \begin{cases} C(N+1)^{n-1}(1+Kd)\, min((Kd)^{\frac{n-1}{\tau}}, 1) & \textit{if } Kd \geq (N+1)^{-\tau}; \\ 0 & \textit{otherwise}; \end{cases}$$

where the constant C *only depends on* n.

PROOF. By homogeneity, we may assume that $|\omega| = 1$. If now $k \in \Gamma_N(d)$, then

$$d \geq |\langle k, \omega\rangle + u| \geq K^{-1}(|k| + 1)^{-\tau}$$

which implies that $Kd \geq (N+1)^{-\tau}$. This proves the second part of the estimate. In order to prove the first part, we observe that this part follows from lemma 1 unless Kd is small as we now assume.

If k and l are different elements in $\Gamma_N(d)$, then

$$K^{-1}|k - l|^{-\tau} \leq |\langle k - l, \omega\rangle| \leq 2d$$

which implies that

$$(\langle k - l, k - l\rangle)^{\frac{1}{2}} \geq C_1|k - l| \geq C_1(2Kd)^{-\frac{1}{\tau}}.$$

Since $4d < C_1(2Kd)^{-\frac{1}{\tau}}$ if Kd is small, the distance between the projections of k and l, on the plane orthogonal to ω, must be greater that $\frac{1}{2}C_1(2Kd)^{-\frac{1}{\tau}}$. Since these projections are included in the ball with radius $(N + 1)$, centered at the origin, and since $(N + 1) \geq (Kd)^{-\frac{1}{\tau}}$, the lemma follows. ∎

We can also sharpen lemma 2 in a similar way if we assume that the ball B is centered at a Diophantine vector.

LEMMA 4. *Let B be a ball in $\mathbf{R}^{n+m}$, with radius b and center $(\omega_0, \Omega_0) \in DC(K_0)$. Then*

1) $\operatorname{meas}(\complement DC(K) \cap B) \leq Cb^{n+m}K_0K^{-1}$;

2) $\operatorname{meas}(\complement DC(K) \cap B \cap \{\omega = \omega_0\}) \leq Cb^m K_0 K^{-1}$;

where C is a constant that only depends on m, n and τ.

PROOF. We can assume that both K_0K^{-1} and K_0b are small, smaller than some C_1, say, which only depends on m, n and τ, and which is smaller than $\frac{1}{2}$. In fact, if $K_0K^{-1} \geq C_1$, then there is nothing to prove, and if $K_0b \geq C_1$, then lemma 4 follows from lemma 2.

Let now $(k, l) \neq (0, 0)$, $|l| \leq 3$, be given. The set X of all (ω, Ω) in B, such that

$$|\langle k, \omega\rangle + \langle l, \Omega\rangle| \leq K^{-1}(|k| + |l|)^{-\tau}$$

is of measure less than $C_2b^{n+m-1}K^{-1}(|k| + |l|)^{-\tau-1}$.

Moreover, X is void unless

$$K_0^{-1}(|k| + |l|)^{-\tau} \leq |\langle k, \omega_0\rangle + \langle l, \Omega_0\rangle| \leq (|k| + |l|)b + K^{-1}(|k| + |l|)^{-\tau}$$

which implies that

$$(|k| + |l|)^{\tau+1} \geq (2K_0b)^{-1} = (2N_0)^{\tau+1}$$

(where the last equation defines N_0) and this, in turn, implies that $|k| \geq N_0$, if we assume, as we may, that N_0 is sufficiently large. Moreover,

$$|\langle k, \omega_0 \rangle + \langle l, \Omega_0 \rangle| \leq (|k| + |l|)b + K^{-1}(|k| + |l|)^{-\tau} \leq 4|k|b$$

if KK_0^{-1} and $(K_0 b)^{-1}$, and hence $|k|$, are large as we have assumed.

Hence, the set X is void unless

$$k \in G = \bigcup_{N \geq N_0} G_N, \quad G_N = \Gamma_N(\omega_0, \langle l, \Omega_0 \rangle, 4Nb).$$

Then the measure in 1) can be estimated by

$$C_3 b^{n+m-1} K^{-1} \sum_{k \in G} (|k| + 1)^{-\tau-1}$$

which, by Abel's summation formula, can be estimated by

$$C_4 b^{n+m-1} K^{-1} \sum_{N \geq N_0} (\# G_N)(N+1)^{-\tau-2}.$$

It now follows, using lemma 3, that this sum is bounded by the estimate 1).

The proof of 2) is essentially the same. Since $\omega_0 \in DC_0(K)$ we can assume $l \neq 0$. The set X of all Ω in $B \cap \{\omega = \omega_0\}$, such that

$$|\langle k, \omega_0 \rangle + \langle l, \Omega \rangle| \leq K^{-1}(|k| + |l|)^{-\tau}$$

is of measure less than $C_5 b^{m-1} K^{-1}(|k| + |l|)^{-\tau}$. And one shows as above that this set is void unless

$$|k| \geq N_0 \text{ and } |\langle k, \omega_0 \rangle + \langle l, \Omega_0 \rangle| \leq 4b$$

where $(2N_0)^{-\tau} = 2K_0 b$, i.e. it is non void only if

$$k \in G = \bigcup_{N \geq N_0} G_N, \quad G_N = \Gamma_N(\omega_0, \langle l, \Omega_0 \rangle, 4b).$$

The estimate 2) now follows from lemma 3 in the same way as above (but easier). ∎

The set $J \cap DC(K)$ near good frequencies

We will now use lemma 4 to improve the estimate in proposition 1.

PROPOSITION 2. *Let* $y \longmapsto (\omega(y), \Omega(y))$ *be a* C^1 *family of frequencies defined on* $B(b) \subset \mathbf{R}^n$ *and satisfying* (5)-(8). *Assume that there exists* $(\omega_0, \Omega_0) \in DC(K_0)$, *such that*

$$|\omega(0) - \omega_0| + |\Omega(0) - \Omega_0| \leq \varepsilon\mu_3,$$

for some $\varepsilon \leq \frac{1}{2} min(\frac{b}{\mu_2 u_3}, 1)$. *Then, for any* $\delta \in]0,1[$, *there is a Cantor set* $\Lambda(\delta) \subset]-\varepsilon, \varepsilon[$, *of relative measure* $\geq (1 - \delta)$, *and, for each* $t \in \Lambda(\delta)$, *there is a* $y(t) \in B(b)$, *such that*

$$(\omega(y(t)), \Omega(y(t))) \in DC(K) \text{ for any } K \geq C(\mu_1\mu_2\mu_3\mu_4)K_0\delta^{-1}$$

where C *only depends on* m, n *and* τ. *Moreover,*

$$\omega(y(t)) = (1 + t)\omega_0$$

and

$$|y(t)| \leq 2\varepsilon\mu_2\mu_3 \text{ for } t \in \Lambda(\delta).$$

PROOF. The proof is merely a repetition of that of proposition 1, using 4 instead of lemma 2. In fact, there is a $y_0 \in B(\varepsilon\mu_2\mu_3)$ such that $\omega_0 = \omega(y_0) \in DC_0(K_0)$, and we can define $y(t)$ for all $|t| < \frac{b}{2\mu_2 u_3}$, as in proposition 1.

For $0 < |l| \leq 3$ we define $V = V_l$ as before. By 2) of lemma 4 it follows then that the set X of all $V \in B_2$ (a ball with radius $C_2\varepsilon\nu$), such that

$$|\langle k, \omega_0 \rangle + V| \leq K^{-1}(|k| + |l|)^{-\tau} \text{ for some } k \in \mathbf{Z}^n,$$

has measure less than $C_3\varepsilon\nu K_0 K^{-1}$. And this is less than $\delta\varepsilon(C_4\mu_4)^{-1}$, if

$$K \geq C_5(\mu_1\mu_2\mu_3\mu_4)\frac{K_0}{\delta}$$

for C_5 sufficiently large. Hence, if C_5 is large enough, then the set $\tilde{\Lambda}_l(\delta)$, which consists of all $t \in]-\varepsilon, \varepsilon[$ such that $V_l(t) \in X$, has measure less than $\delta\varepsilon(4m^2)^{-1}$. Since this is true for each l we can just let $\Lambda(\delta)$ be the complement of the union of all these sets $\tilde{\Lambda}_l(\delta)$.

This establishes the existence of the set $\Lambda(\delta)$. The requirements on $\omega(y(t))$ and on $|y(t)|$ are fulfilled by construction. ∎

REMARK. The results just proven are, of course, valid not only for $|l| \leq 3$, but for l belonging to any finite set Δ which contains 0. If Δ does not contain 0, the results are still true if we, to the assumptions in lemma 4 and proposition 2, add that $\omega_0 \in DC_0(K_0)$.

Comparison with other non-degeneracy conditions

Much of what we have done in this section is well known, and the measure theoretical techniques are standard. What is new is the formulation of the non-degeneracy condition (6), and the proof that if the homogeneous Diophantine condition is fulfilled for ω, then finitely many non-homogeneous conditions can be satisfied by varying a single parameter, namely the length of ω.

In [10] Melnikov describes a different approach to our problem, and in [11] yet another one in a more degenerate situation. The most interesting difference in our opinion is that he uses other non-degeneracy conditions, and we shall now compare them with (2) and its quantitative form (6).

Let $(\omega(y), \Omega(y)) \in \mathbf{R}^{n+m}$ be a family of frequency vectors which we assume to be real analytic in a ball $B = B(b) \subset \mathbf{R}^n$. As usual we assume that (1) holds, i.e.

$$D\omega(y) \text{ is invertible on } B$$

and we also assume that $(D\omega(y))^{-1}$ and $D\Omega(y)$ are bounded on B.

Consider the following condition, proposed by Melnikov in [11]:

$$\text{for all } l \in \mathbf{Z}^m,\ |l| \le 3, \text{ and for all } k \in \mathbf{Z}^n,\ (k,l) \ne 0,$$

$$k + l((D\omega)^{-1}D\Omega)^* \not\equiv 0 \text{ on } B$$

($*$ denotes transposition). This condition imposes, in fact, only finitely many restrictions on the family (since $(D\omega(y))^{-1}D\Omega(y)$ is assumed to be bounded), and it is therefore equivalent to the existence of an open (and dense) subset of B, which we still call B, such that

$$\text{for all } l \in \mathbf{Z}^m,\ |l| \le 3, \text{ and for all } k \in \mathbf{Z}^n,\ (k,l) \ne 0,$$

$$k + l((D\omega(y))^{-1}D\Omega(y))^* \ne 0 \text{ for all } y \in B. \tag{11}$$

(In the differentiable case this is still true for each l separately, but the subsets are no longer dense, so their non void intersection must be postulated explicitly).

That condition (11) is a good one, is easy to see. In order to show this, we formulate a quantitative version of (11):

$$|k + l((D\omega(y))^{-1}D\Omega(y))^*| \ge \frac{1}{\mu_4} \text{ for } y \in B,\ (k,l) \in \mathbf{Z}^{n+m}\backslash 0,\ |l| \le 3. \tag{12}$$

PROPOSITION. *Assume that* $(\omega(y), \Omega(y))$ *satisfies conditions* (5), (8) *and* (12). *Then the set*

$$\{y \in B : |\langle k, \omega(y)\rangle + \langle l, \Omega(y)\rangle|$$
$$\le K^{-1}(|k| + |l|)^{-\tau} \textit{ for some } (k,l) \in \mathbf{Z}^{n+m}\backslash 0,\ |l| \le 3\}$$

has measure

$$\leq Cb^n(\mu_1\mu_2\mu_4)K^{-1}\frac{\mu_2}{b}$$

where C *only depends on* m, n *and* τ.

PROOF. Let (k, l) be given and consider the function

$$F(y) = \langle k, \omega(y)\rangle + \langle l, \Omega(y)\rangle.$$

Then

$$DF(y) = (k + l((D\omega(y))^{-1}D\Omega(y))^*)(D\omega(y))^*$$

so $|DF(y)| \geq C_1(\mu_2\mu_4)^{-1}$ on B. On the other hand, $|DF(y)| \geq |k|(2\mu_2)^{-1}$ on B, if $|k| \geq C_2\mu_1\mu_2$.

Hence, the set

$$\{y \in B : |F(y)| \leq K^{-1}(|k| + |l|)^{-\tau}\}$$

has measure less than $C_3b^{n-1}K^{-1}(|k| + |l|)^{-\tau}a(k)$, where

$$a(k) = \begin{cases} \frac{2\mu_2}{|k|} & \text{if } |k| \geq C_2\mu_1\mu_2; \\ \mu_2\mu_4 & \text{otherwise.} \end{cases}$$

Then the proposition follows if we just observe that $\mu_1\mu_2$ and $\mu_1\mu_2\mu_4$ both are bounded from below by some positive constant. ■

So the intersection $J \cap DC$ is non void also under the condition (1)+(11). But (1)+(2), proposed in this article, permits us to first fix the tangential frequencies ω, and then find the normal frequencies Ω just by varying the length of ω. This is the reason why we get quasiperiodic solutions with a *fixed tangential frequency direction* in theorem A. Condition (2) can also be applied to certain linear eigenvalue problems, as studied by Sinai and Dinaburg (see [12] for references), where the tangential frequencies never gets perturbed and where we can use the eigenvalue to control the normal frequencies. In fact, the proof of proposition 1 shows that if $\omega_0 \in DC_0$, then any one-parameter family (ω_0, Ω_t) will intersect DC unless some of $\frac{d}{dt}\langle l, \Omega_t\rangle$, $|l| \leq 3$, is identically zero.

A case where condition (2) does not apply is "the small planet problem" for zero eccentricities and inclinations. By averaging over the fast variables (the mean anomalies) and taking the expansion of the averaged Hamiltonian up to order 2 in the eccentricities and inclinations, we get an integrable Hamiltonian whose frequency map verifies

$$\Omega(y) = \omega(y)(D\omega(y))^{-1}D\Omega(y) \text{ for all } y$$

i.e. condition (2) is fulfilled nowhere (see [13]). In this case Melnikovs condition (11) can perhaps be used to construct quasiperiodic solutions to this problem in celestial mechanics.

In [10] Melnikov proposes another condition:

$$\text{for all } l \in \mathbf{Z}^m,\ |l| \le 3, \text{ and for all } k \in \mathbf{Z}^n,\ (k,l) \ne 0,$$

$$\langle k, \omega \rangle + \langle l, \Omega \rangle \not\equiv 0 \text{ on } B. \tag{13}$$

Let $l \in \mathbf{Z}^m$, $|l| \le 3$, and suppose that (ω, Ω) satisfies $(13)_l$. If now

$$\langle l, \Omega(y) \rangle = \langle l, \omega(y)(D\omega(y))^{-1} D\Omega(y) \rangle \text{ for all } y \in B$$

then, for all k,

$$\langle k, \omega(y) \rangle + \langle l, \Omega(y) \rangle = \langle k + l((D\omega(y))^{-1} D\Omega(y))^*, \omega(y) \rangle \text{ for all } y \in B.$$

Hence, $(13)_l$ implies that

$$\langle l, \Omega - \omega(D\omega)^{-1} D\Omega \rangle \not\equiv 0 \text{ on } B$$

or

$$\text{for all } k \in \mathbf{Z}^n,\ (k,l) \ne 0,\ k + l((D\omega)^{-1} D\Omega)^* \not\equiv 0 \text{ on } B.$$

So the intersection $J \cap DC$ is non void also under condition (1)+(13). However, there seems to be no quantitative formulation of (13) as there are for (2) and (11) - namely (6) and (12).

Before ending this long remark we also want to mention a work of Pyartli [14]. He gives a generic condition which garanties that $(\omega(y), \Omega(y)) \in DC$ for almost all y (at least if $\tau > (n+m)^2 + n - m - 1$).

His condition is even valid for infinitely many l:s, much more than what we ask for, but it involves higher order derivatives in y which makes it harder to apply.

III INFINITESIMAL STABILITY OF NORMAL FORMS

On $T^*(\mathbf{T}^n \times \mathbf{R}^m)$ we introduce canonical coordinates $\{(x, y, z)\}$ so that the symplectic form becomes

$$\sum_{1 \le j \le n} dx_j \wedge dy_j + \sum_{1 \le i \le m} dz_i \wedge dz_{i+m}.$$

Corresponding to the Lagrangian foliation into tori $y_j = \text{const.}$, $y_i' = \frac{1}{2}(z_i^2 + z_{i+m}^2) = \text{const.}$, there is a linear action of the torus $\mathbf{T}^{n+m}$ on the symplectic space.

Let H be a real analytic function defined near the torus $\{(x,0,0)\}$. With H we associate its *mean value* $[H]$ under the action of $\mathbf{T}^{n+m}$. $[H]$ is a real analytic function in y and y'. We also define an *ordering* by writing $H = H_0 + H_1 + \cdots$, where

$$H_k(x,y,z) = \sum_{(\alpha,\beta)\in\mathbf{Z}^{n+2m},\ 2|\alpha|+|\beta|=k} \frac{1}{\alpha!\beta!}\frac{\partial^\alpha\partial^\beta}{\partial y^\alpha \partial z^\beta} H(x,0,0) y^\alpha\ z^\beta.$$

H is *homogeneous* if $H = H_k$ for some k, and H is *of order* k if $H_i = 0$ for $i < k$ and $H_k \neq 0$. We denote by $\hat{H}$ the function $H - (H_0 + H_1 + H_2 + H_3)$.

It is clear that the *Poisson bracket* $\{\ ,\ \}$, corresponding to the symplectic form, preserves this ordering.

We define two complex neighbourhoods of the torus $\{(x,0,0)\}$ through

$$D(r,s) = \{(x,y,z) : |\Im x| < r,\ |y| < s^2,\ |z| < s\},$$

and, for $\eta \in \mathbf{R}\backslash 0$,

$$(x,y,z) \in \eta D(r,s) \text{ if, and only if, } \eta^{-1}(x,y,z) \in D(r,s).$$

We finally let $|\ |_D$ denote the supremum norm over any neighbourhood D of the torus $\{(x,0,0)\}$.

The normal forms $\mathbf{NF}(r,s,\underline{\mu})\cap\mathbf{DC}(\omega,K)$

Let N be a real analytic function defined near the torus $\{(x,0,0)\}$, and let us write

$$[N_2] = \langle\omega, y\rangle + \langle\Omega, y'\rangle$$

$$[N_4] = \frac{1}{2}\langle yM_1, y\rangle + \langle yM_2, y'\rangle + \frac{1}{2}\langle y'M_3, y'\rangle,\ M_1^* = M_1.$$

Let $\underline{\mu} = (\mu_1,\mu_2,\mu_3,\mu_4) \in (\mathbf{R}^+)^4$. We say that $N \in \mathbf{NF}(r,s,\underline{\mu})$ if N is defined on $\bar{D}(r,s)$, and

$$N_1 = N_3 = 0 \tag{14}$$

$$N_0 = [N_0] \text{ and } N_2 = [N_2] \tag{15}$$

$$|M_1^{-1}| \leq \mu_2 \text{ and } |\langle l, \Omega - \omega M_1^{-1}M_2\rangle| \geq \frac{1}{\mu_4} \text{ for } 0 < |l| \leq 3 \tag{16}$$

$$|N_2|_{D(r,s')} \le \mu_3(s')^2 \text{ and } |\hat{N}|_{D(r,s')} \le \mu_1(s')^4, \ s' \le s. \tag{17}$$

The fact that N_0 is constant and $N_1 = 0$ implies that $\{(x,0,0)\}$ is invariant under the Hamiltonian flow of N, and that, by (15), the flow is linear on this torus. The assumption that Ω is independent of x is an important restriction, since no Floquet theory is available for quasiperiodic systems. Condition (16) is a sort of "twist condition" which tells us that $(\omega + yM_1, \Omega + yM_2)$ is a non-degenerate family of frequencies (i.e. satisfies conditions (5) and (6)), and that we therefore can apply the results of section II. The reason why we require N_3 to be 0 will be clear when we have stated the next condition.

We say that $N \in \mathbf{DC}(\omega, K)$ if

$$N_2 = \langle \omega, y\rangle + \langle \Omega, y'\rangle \text{ and } (\omega, \Omega) \in DC(K). \tag{18}$$

Now the assumption on N_3 appears to be no restriction at all. Any function satisfying (14), (15), (18), with the exception of $N_3 = 0$, can be symplectically transformed to one that satisfies (14), (15), (18) without exception. This is the Birkhoff normal form, which we have assumed for N in order to be able to formulate condition (16).

We notice that these parameters cannot be chosen independently, but that (16)-(18) implies that

$$min(\mu_1\mu_2, \mu_1\mu_2\mu_3\mu_4, K\mu_3) \ge C \tag{19}$$

where C is some constant that only depends on m, n and τ. For example, $I = M_1 M_1^{-1}$ implies the first inequality. We will frequently make use of these bounds in the proof of proposition 3.

It is possible to weaken the assumptions by requiring that (16) holds only for $|l| \le 2$. In that case it is not reasonable to assume that N_3 is 0, but only that it is independent of y. The result that we shall prove below is also true under this weaker assumptions, and the proof can easily be so modified, but we have refrained from doing this just to avoid too much technicalities.

The basic small divisor lemma

The following result is just a variant of a well known small divisor lemma which, in its sharpest form, is due to Rüssmann [15].

LEMMA 5. *Suppose $N_2 \in \mathbf{DC}(\omega, K)$ and let F be a real analytic function on $D(r,s)$, homogeneous of order $j \le 3$. Then there exists a unique real analytic function G, homogeneous of order j, such that*

$$\{G, N_2\} = F - [F], \ [G] = 0.$$

Moreover, this function satisfies

$$|G|_{D(r-\rho,s)} \leq CK\rho^{-\tau}|F|_{D(r,s)} \text{ for all } \rho < r$$

where C is a constant that only depends on m, n and τ.

PROOF. It is clearly sufficient to find a complex analytic function G. We may therefore introduce new variables $\hat{x}_j = x_j$, $\hat{y}_j = y_j$, $2z_i = (1+\sqrt{-1})(\hat{z}_i - \hat{z}_{i+m})$, $2z_{i+m} = (1-\sqrt{-1})(\hat{z}_i + \hat{z}_{i+m})$. This is a symplectic transformation, hence preserving the Poisson bracket, and it preserves $D(r,s)$. Moreover, $y_i' = \frac{1}{2}(z_i^2 + z_{i+m}^2)$ gets transformed into $-\sqrt{-1}(\hat{z}_i\hat{z}_{i+m})$, so we can therefore assume right away that $y_i' = -\sqrt{-1}(z_i z_{i+m})$, $i \leq m$.

We can also assume that F is independent of y, and, since the operator $G \to \{G, N_2\}$ preserves monomials in z (with coefficients that are functions of x), we can finally assume that F is such a monomial,

$$F(x,z) = (\sum_{k\in\mathbf{Z}^n} c(k)\exp(\sqrt{-1}\langle k,x\rangle))z^\beta$$

say, where $\beta = (\beta', \beta'') \in \mathbf{N}^{2m}$, with $|\beta| \leq 3$. By assumption, $|c(k)z^\beta| \leq C_1 e^{-r|k|}|F|_{D(r,s)}$, $|z| < s$.

Let

$$\gamma(k) = -\sqrt{-1}(\langle k,\omega\rangle + \langle \beta' - \beta'', \Omega\rangle)^{-1}$$

if $(k, \beta' - \beta'') \neq 0$, and $\gamma(k) = 0$ otherwise. Then clearly

$$G(x,z) = (\sum_{k\in\mathbf{Z}^n} \gamma(k)c(k)\exp(\sqrt{-1}\langle k,x\rangle))z^\beta.$$

And, by using Abel's summation formula, we find that $|G|_{D(r-\rho,s)}$ is less than

$$C_2(\sum_{k\in\mathbf{Z}^n} |\gamma(k)|e^{-|k|\rho})|F|_{D(r,s)} \leq C_3\rho(\sum_{N\geq 0}(\sum_{|k|\leq N} |\gamma(k)|e^{-N\rho})|F|_{D(r,s)}.$$

Consider now the function $\Gamma(d) = \#\Gamma_N(\omega, \langle\beta' - \beta'', \Omega\rangle, d)$. According to lemma 3 we have

$$d \geq \begin{cases} (\frac{\Gamma(d)}{C_4(N+1)^{n-1}})\frac{1}{K}, & \text{if } d \geq \frac{1}{K}; \\ (\frac{\Gamma(d)}{C_4(N+1)^{n-1}})^{\frac{\tau}{n-1}}\frac{1}{K}, & \text{if } d \leq \frac{1}{K}. \end{cases}$$

Let now $d_j = min\{d : \Gamma(d) = j\}$, defined for all positive integers j for which the set is non void. Let also $N_1 = C_4(N+1)^{n-1}$. Then

$$d_j \geq (\frac{j}{C_4(N+1)^{n-1}})^{\frac{\tau}{n-1}}\frac{1}{K}$$

if $j \leq N_1$, so that the sum

$$\sum_{|k|\leq N} |\gamma(k)| = \sum_{1\leq j\leq N_2} \frac{1}{d_j} \quad (N_2 = C_5(N+1)^n)$$

can be estimated by

$$C_6\Big(\sum_{1\leq j\leq N_1} K(N+1)^\tau j^{-\frac{\tau}{n-1}} + \sum_{N_1\leq j\leq N_2} K(N+1)^{n-1} j^{-1}\Big) \leq$$

$$C_7(K(N+1)^\tau + K(N+1)^{n-1}\log(N+1)) \leq C_8 K(N+1)^\tau$$

(since $\tau > n-1$). Finally, since

$$\sum_{N\geq 0} (N+1)^\tau e^{-N\rho} \leq C_9 \rho^{-\tau-1}$$

the lemma follows. ■

The inductive step

We shall now formulate the usual infinitesimal stability result which will serve as the inductive step in the iterative construction of the invariant tori.

PROPOSITION 3. *Let $N \in \mathbf{NF}(r, s, \underline{\mu}) \cap \mathbf{DC}(\omega, K)$. Let $r_+ < r$, $s_+ < s$, $1 < \nu < 2$ and let $\xi = (r - r_+)^{\tau+1}(K\mu_1)^{-1}$. We assume that r_+ and s_+ are so restricted so that $r - 1 < r_+, s < 2s_+$ and $\xi < s^2$.*

Then there exists a constant C, depending only on m, n and τ, such that the following holds: for all d and for all real analytic functions H, defined on $D = D(r, s)$ and such that

$$|H_0 + H_1|_D + \frac{\xi}{s^2}|H_2 + H_3|_D + \frac{\xi^2}{s^4}|\hat{H}|_D \leq d$$

$$\leq \frac{1}{C}\Big(\frac{(r-r_+)^{\tau+1}}{K\mu_1}\Big)^2\Big(1 - \frac{s_+}{s}\Big)\frac{\nu - 1}{(\mu_1\mu_2\mu_3\mu_4)(\mu_1\mu_2)\mu_2}$$

there exist a Cantor set $\Lambda \subset \mathbf{R}$ and, for all $t \in \Lambda$, a symplectic diffeomorphism

$$\Phi(t) : D^+ = D(r_+, s_+) \longrightarrow D(r, s)$$

such that

$$(N + H) \circ \Phi(t) = N^+ + H^+$$

where

$$N^+ \in \mathbf{NF}(r_+, s_+, \nu\underline{\mu}) \cap \mathbf{DC}((1+t)\omega, K^+) \text{ for some } K^+ < \infty$$

and

$$|H_0^+ + H_1^+|_{D^+} + \frac{\xi}{s_+^2}|H_2^+ + H_3^+|_{D^+} + \frac{\xi^2}{s_+^4}|\hat{H}^+|_{D^+}$$

$$\leq C(\frac{(r-r_+)^{\tau+1}}{K\mu_1})^{-2}(1-\frac{s_+}{s})^{-2}(\mu_1\mu_2)\mu_2 d^2.$$

Moreover,

$$\Lambda \subset \{t : |t| < \frac{K\mu_1}{(r-r_+)^{\tau+1}}\frac{d}{\mu_3}\}$$

and is of relative measure 1 *and*

$$(\Phi(t) - id) : D^+ \longrightarrow \eta D(r - r_+, s)$$

where $\eta = C(K\mu_1(r-r_+)^{-\tau-1})^2\mu_2 d$.

Finally, for all $0 < \delta < 1$, *the set*

$$\Lambda(\delta) = \{t \in \Lambda : N^+ \in \mathbf{DC}((1+t)\omega, K^+) \textit{ for any } K^+ \geq CK(\mu_1\mu_2\mu_3\mu_4)\frac{1}{\delta}\}$$

is of relative measure greater than $1 - \delta$ *in* Λ.

Proof of proposition 3

We shall first consider the case $s = 1$.

We want to construct functions $S = S_0 + S_1 + S_2 + S_3$ and $R = R_0 + R_2 + \hat{R}$ such that

$$(*) \qquad \{\langle\lambda, x\rangle + S, N\} = H - R, \ [S] = 0, \ [R_0 + R_2] = R_0 + R_2$$

where $\lambda = (\lambda_1, \cdots, \lambda_n)$ are n real parameters. If we let Φ_λ^t be the flow of the Hamiltonian vector field of $\tilde{S} = \langle\lambda, x\rangle + S$, then

$$(N + H) \circ \Phi_\lambda^1 = N^+ + \int_0^1 \{tH + (1-t)R, \tilde{S}\} \circ \Phi_\lambda^t dt, \ N^+ = N + R.$$

As the first step of the proof we shall, using lemma 5, solve equation (*) and estimate the solution.

The solution depends on λ. We shall then show that, for λ and d sufficiently small, N^+ belongs to $\mathbf{NF}(r_+, \frac{1}{2}, \nu\underline{\mu})$. This is the second step. In the third step we shall show that $\Phi_\lambda^1 : D^+ \to D$ for λ and d sufficiently small, and we shall estimate $(\Phi_\lambda^1 - id)$.

These thing can be done under the assumption that

$$d' \leq \frac{1}{C_2}(\frac{(r-r_+)^{\tau+1}}{K\mu_1})^2(1-s_+)\frac{\nu - 1}{(\mu_1\mu_2\mu_3\mu_4)\mu_2}$$

where

$$d' = d + |\lambda| \frac{(r - r_+)^{\tau+1}}{K}$$

and C_2 is some constant.

The fourth step consists in estimating H^+, an estimate that will depend on d' (and hence on $|\lambda|$).

In order to get H^+ as small as possible, we must restrict λ to a ball of radius $\sim K(r - r_+)^{-\tau-1} d$.

However, in the last step we shall show that there exists parameter values $\lambda = \lambda(t)$ in a ball of radius $\sim K(r - r_+)^{-\tau-1} d(\mu_1\mu_2)$, such that $\Phi(t) = \Phi^1_{\lambda(t)}$ has the required properties. (This larger value of λ will, of course, influence the estimate of H^+ which will be larger, but still sufficiently small). Now one sees readily that a choice of λ in this larger ball is consistent with the above restriction on d', if

$$d \leq \frac{1}{C_4} \left(\frac{(r - r_+)^{\tau+1}}{K\mu_1} \right)^2 (1 - s_+) \frac{\nu - 1}{(\mu_1\mu_2\mu_3\mu_4)(\mu_1\mu_2)\mu_2}$$

for some sufficiently large constant C_4.

Before proceeding to the details of the proof we introduce some more notation. Let

$$\rho = \frac{1}{6}(r - r_+), \ \ \sigma = \frac{1}{6}(1 - s_+), \ \ D_j = D(r - j\rho, 1 - j\sigma), \ \ 0 \leq j \leq 6.$$

We let $C_1, C_2, \cdots$ be an increasing sequence of constants which only depend on m, n and τ (and which will be determined by the expressions in which they first appear).

Solution of (*)

We first decompose (*) into its homogeneous components:

$$\{S_0, N_2\} = H_0 - R_0 - \langle \lambda, \omega \rangle, \ [S_0] = 0, \ [R_0] = R_0$$

$$\{S_1, N_2\} = H_1$$

$$\{S_2, N_2\} = H_2 - \{\langle \lambda, x \rangle + S_0, N_4\} - R_2, \ [S_2] = 0, \ [R_2] = R_2$$

$$\{S_3, N_2\} = H_3 - \{\langle \lambda, x \rangle + S_0, N_5\} - \{S_1, N_4\}$$

$$\hat{R} = \hat{H} - \{\langle \lambda, x \rangle + S_0, \hat{N} - N_4 - N_5\} - \{S_1, \hat{N} - N_4\} - \{S_2, \hat{N}\} - \{S_3, \hat{N}\}.$$

We let $R_0 = [H_0] - \langle \lambda, \omega \rangle$. Since $[H_1] = 0$, lemma 5 implies that

$$|S_0|_{D_1} \leq C_1 \frac{K}{\rho^\tau} |H_0|_D \text{ and } |S_1|_{D_1} \leq C_1 \frac{K}{\rho^\tau} |H_1|_D.$$

Using Cauchy estimates for the partial derivatives we find that

$$|H_2 + \{S_0, N_4\}|_{D_2} + |\{\langle \lambda, x \rangle, N_4\}|_{D_2} \leq C_2 \frac{K\mu_1}{\rho^{\tau+1}} d'.$$

Moreover,

$$[\{\langle\lambda, x\rangle, N_4\}] = [\langle\lambda, \frac{\partial N_4}{\partial y}\rangle] = \langle\lambda M_1, y\rangle + \langle\lambda M_2, y'\rangle$$

and, since R_2 shall be the mean value of the right hand side of the third equation, we get

$$R_2(y, y') = \langle P - \lambda M_1, y\rangle + \langle Q - \lambda M_2, y'\rangle, \text{ with } |P| + |Q| \le C_3 \frac{K\mu_1}{\rho^{\tau+1}} d. \tag{20}$$

(Notice that $|P| + |Q|$ is estimated by d and not by d'). Using lemma 5 again, we find that

$$\xi |R_2|_{D_3} \le C_4 d' \text{ and } \xi |S_2|_{D_3} \le C_4 \frac{K}{\rho^\tau} d'.$$

In the same way we can estimate S_3 and $\hat{R}$ by

$$\xi |S_3|_{D_4} \le C_5 \frac{K}{\rho^\tau} d' \text{ and } \xi^2 |\hat{R}|_{D_4} \le C_5 d' \frac{1}{\sigma}$$

where we have used $\rho < 1$ in the estimate of S_3, and $\rho, \xi < 1$ in that of $\hat{R}$. (Notice that σ never appears in the estimate of S or its derivatives since S is polynomial in y, z, and that $(1 - 3\sigma)^2 - (1 - 4\sigma)^2 \ge \frac{\sigma}{2}$. This explains why we get σ and not σ^2 in the denominator of the estimate of $\hat{R}$).

So we finally get

$$|S_0 + S_1|_{D_4} + \xi |S_2 + S_3|_{D_4} \le C_1 \frac{K}{\rho^\tau} d' \tag{21}$$

$$\xi |R_2|_{D_4} + \sigma \xi^2 |\hat{R}|_{D_4} \le C_1 d' \tag{22}$$

for some new constant which we denote by C_1 (and thereby forgetting all the other C_i:s).

$N^+ = N + \dot{R} \in \mathbf{NF}(r_+, s_+, \nu\underline{\mu})$.

We must verify conditions (14)-(17) with $\underline{\mu}$ replaced by $\nu\underline{\mu}$. Now (14) and (15) are automatic by construction, and (17) follows easily from (22) if i) $d' \le C_2^{-1}(\nu - 1)\mu_3 \xi$ and ii) $d' \le C_2^{-1}(\nu - 1)\mu_1 \xi^2 \sigma$.

So we only need to consider (16) in some detail.

If we write

$$[N_4 + R_4] = \frac{1}{2}\langle y\tilde{M}_1, y\rangle + \langle y\tilde{M}_2, y'\rangle + \frac{1}{2}\langle y'\tilde{M}_3, y'\rangle$$

then we clearly have

$$|\tilde{M}_1 - M_1| + |\tilde{M}_2 - M_2| \le C_3 |R_4|_{D_4}.$$

Since $\tilde{M}_1^{-1} = M_1^{-1}(I + (\tilde{M}_1 - M_1)M_1^{-1})^{-1}$, we find that

$$|\tilde{M}_1^{-1}| \leq \mu_2(1 + C_4\mu_2|R_4|_{D_4})$$

which is less than $\nu\mu_2$ if iii) $d' \leq C_5^{-1}\mu_2^{-1}(\nu - 1)\xi^2\sigma$. Finally, if we let $\omega(\lambda) = \omega + P - \lambda M_1$ and $\Omega(\lambda) = \Omega + Q - \lambda M_2$, and write

$$\begin{gathered}\Omega(\lambda) - \omega(\lambda)\tilde{M}_1^{-1}\tilde{M}_2 = \\ (\Omega - \omega M_1^{-1}M_2) + (Q - \lambda M_2 - (P - \lambda M_1)\tilde{M}_1^{-1}\tilde{M}_2) \\ +\omega\tilde{M}_1^{-1}((\tilde{M}_1 - M_1)M_1^{-1}M_2 + (M_2 - \tilde{M}_2))\end{gathered}$$

we find that

$$|\langle l, \Omega(\lambda) - \omega(\lambda)\tilde{M}_1^{-1}\tilde{M}_2\rangle| \geq \frac{1}{\mu_4} - C_6(\mu_1\mu_2|R_2|_{D_4} + \mu_1\mu_2\mu_2\mu_3|R_4|_{D_4})$$

which is $\geq (\nu\mu_4)^{-1}$, if iv) $d' \leq C_7^{-1}(\nu - 1)(\mu_1\mu_2\mu_3\mu_4\mu_2)^{-1}\xi^2\sigma$ and v) $d' \leq C_7^{-1}(\nu - 1)(\mu_1\mu_2\mu_4)^{-1}\xi$.

All these conditions i)-v) on d' are fulfilled if

$$d' \leq \frac{1}{C_2}\frac{\nu - 1}{\mu_1\mu_2\mu_3\mu_4\mu_2}\xi^2(1 - s_+) \tag{23}$$

for some new constant C_2. (Here we have used (19)).

Estimate of Φ_λ^1

We get from (21) that

$$|S|_{D_4} \leq C_1\xi^{-2}\rho\mu_1^{-1}d' = C_1\rho\eta'$$

where the last equality defines η'. Then

$$|\lambda + \frac{\partial S}{\partial x}|_{D_5} \leq C_3\eta', \ |\frac{\partial S}{\partial y}|_{D_5} + |\frac{\partial S}{\partial z}|_{D_5} \leq C_3\rho\eta'$$

which implies that

$$\Phi_\lambda^t - id : D_6 \longrightarrow (C_4\eta')D(\rho, 1), \quad |t| \leq 1$$

and

$$\Phi_\lambda^1 : D_6 \longrightarrow D_5$$

if $d' \leq C_5^{-1}\mu_1\xi^2\sigma$, for some constant C_5. And this condition is implied by (23), if C_2 is sufficiently large.

Estimate of H^+

Let $F_t = tH + (1-t)R$, $|t| \leq 1$. Then $H^+ = \int_0^1 \{F_t, \tilde{S}\} \circ \Phi_\lambda^t \, dt$, and we get an estimate of $|H^+|_{D^6}$ from

$$|\{F_t, \tilde{S}\}|_{D_5} \leq C_3 \xi^{-4} \mu_1^{-1} (d')^2 \sigma^{-2}.$$

($C_3, C_4, \cdots$ are now new constants).

But one can do better. If we write $G = \{F_t, \tilde{S}\}$ and estimate each homogeneous component separately we find

$$|G_0 + G_1|_{D_5} + \xi |G_2 + G_3|_{D_5} + \xi^2 |\hat{G}|_{D_5} \leq C_4 \xi^{-2} \mu_1^{-1} (d')^2 \sigma^{-2}$$

which is the estimate we want. However, we need this estimate for $\{F_t, \tilde{S}\} \circ \Phi_\lambda^t$, and this does not follow immediately, since composition by Φ_λ^t does not preserve the ordering.

We therefore let $F = F_{t_0}$, for some fixed t_0 between 0 and 1, and consider the analytic function $G \circ \Phi_\lambda^t = \{F, \tilde{S}\} \circ \Phi_\lambda^t$, defined in a neighbourhood of the closed unit disc in the complex t-plane. Also each homogeneous component of $G \circ \Phi_\lambda^t$ is such an analytic function. So let us consider the component of order j. It has a power series expansion

$$\sum_{0 \leq k \leq \infty} a_k t^k, \qquad a_0 = G_j \text{ and } a_k = \frac{1}{k}(\{a_{k-1}, \tilde{S}\})_j$$

convergent for $|t| \leq 1$. Since F and S are defined in D_5, each coefficient is defined there, and we get Cauchy estimates for the derivatives in some smaller domain.

Let now ρ' and σ' be such that $k\rho' \leq \rho$ and $k\sigma' \leq \sigma$. Then

$$|a_k|_{D(r-5\rho-k\rho', 1-5\sigma-k\sigma')}$$
$$\leq C_5 \frac{1}{k}\left(\frac{|\lambda|}{\sigma'} + \frac{1}{\sigma' k\rho'}|S|_{D_5}\right)|a_{k-1}|_{D(r-5\rho-(k-1)\rho', 1-5\sigma'-(k-1)\sigma')}$$

which, by induction, is less than

$$\frac{1}{(k!)^2}\left(C_6 \frac{1}{\rho'\sigma'}(|\lambda| k\rho' + |S|_{D_5})\right)^k |a_0|_{D_5}.$$

If we now set $\rho' = \frac{\rho}{k}$ and $\sigma' = \frac{\sigma}{k}$, then we find that

$$|a_k|_{D_6} \leq \left(\frac{k^k}{k!}\right)^2 \left(C_7 \frac{1}{\rho\sigma}(|\lambda|\rho + |S|_{D_5})\right)^k |G_j|_{D_5} \leq 2^{-k} |G_j|_{D_5}$$

if $d' \leq C_8^{-1} \mu_1 \xi^2 \sigma$ for some constant C_8, i.e. if d' fulfills condition (23).

Determination of λ

We are therefore forced to assume that d' fulfills (23), and this imposes a condition on the domain $\{\lambda : |\lambda| \leq b\}$ over which λ is allowed to vary:

$$b < \frac{1}{C_2} \frac{\nu - 1}{(\mu_1\mu_2\mu_3\mu_4)(\mu_1\mu_2)} \xi(1 - s_+).$$

Let now $\omega(\lambda) = \omega + P - \lambda M_1$ and $\Omega(\lambda) = \Omega + Q - \lambda M_2$, and let $|\lambda| \leq b$. This family satisfies conditions (5)-(8) of section II for the parameters $(\mu_1', \mu_2', \mu_3', \mu_4')$, where

$$\mu_i' = C_3\mu_i, \quad i \leq 4$$

for some constant C_3 (as usual, we have only keept C_1 and C_2 from above). In fact (5) and (8) are trivial, and (7) follows from (20) (under condition (23)). The verification of (6) is done in the part "$N^+ = N + R \in \mathbf{NF}(r_+, s_+, \nu\underline{\mu})$", if we there let $\tilde{M}_1$ and $\tilde{M}_2$ be equal to M_1 and M_2.

Let now ε be defined by

$$\varepsilon = \frac{K\mu_1}{(r - r_+)^{\tau+1}} \frac{d}{\mu_3}.$$

Then, from (20)+(23),

$$|P| + |Q| \leq C_1 \frac{K\mu_1}{\rho^{\tau+1}} d \leq \varepsilon\mu_3'$$

if $C_3 > C_1 6^{\tau+1}$ as we may assume.

In order to apply proposition 2 we must assure that

$$\varepsilon \leq \frac{1}{2} min(\frac{b}{\mu_2'\mu_3'}, 1).$$

The second of these two conditions is fulfilled by (23) (if C_2 is sufficiently large), and the first one requires

$$b \geq 2C_3^2 \xi^{-1} \mu_2 d.$$

Hence, if we define b by the right hand side of this inequality, we can apply proposition 2 which gives the existence of Λ and $\Lambda(\delta)$.

Finally, in order for (23) to be fulfilled for this choice of b, we need

$$d \leq \frac{1}{C_4} \frac{\nu - 1}{(\mu_1\mu_2\mu_3\mu_4)(\mu_1\mu_2)\mu_2} \xi^2(1 - s_+). \tag{24}$$

This proves the case $s = 1$.

The case of an arbitrary s follows by scaling the variables, i.e. by first doing the change of variables

$$x \longmapsto x, \; y \longmapsto \frac{y}{s^2}, \; z \longmapsto \frac{z}{s}$$

and then dividing $N + H$ by s^2. Under this transformation, the constants will change in the following way:

$$d \longmapsto \frac{d}{s^2}, \ \mu_1 \longmapsto s^2\mu_1, \ \mu_2 \longmapsto \frac{\mu_2}{s^2}, \ \mu_3 \longmapsto \mu_3, \ \mu_4 \longmapsto \mu_4, \ K \longmapsto K.$$

And this reduces the problem to the case $s = 1$. This concludes the proof of proposition 3. ∎

IV STABILITY OF NORMAL FORMS

The main theorem

THEOREM C. *Let* $N \in \mathbf{NF}(r, s, \underline{\mu}) \cap \mathbf{DC}(\omega, K)$.

Then there exists a constant C, *depending only on* m, n *and* τ, *such that the following holds: for any* $0 < \delta < 1$, *for any* d *and for any real analytic function* H *such that*

$$|H|_{D(r,s)} \leq d \leq \frac{1}{C} \frac{1}{K^2(\mu_1\mu_2\mu_3\mu_4)^2(\mu_1\mu_2)^2\mu_1} \min(r^{\tau+1}, K\mu_1 s^2, 1)^2\delta^2$$

there exist a 1*-dimensional Cantor set* Λ_δ *and, for any* $t \in \Lambda_\delta$, *a symplectic diffeomorphism*

$$\Psi(t) : D(\frac{r}{4}, \frac{s}{4}) \longrightarrow D(r, s)$$

such that

$$(N + H) \circ \Psi(t) \in \mathbf{NF}(\frac{r}{4}, \frac{s}{4}, 2\underline{\mu}) \cap \mathbf{DC}(t\omega, \infty).$$

Moreover,

$$\Lambda_\delta \subset \{t \in \mathbf{R} : |t - 1| \leq K\mu_1 \frac{1}{\mu_3 \min(r^{\tau+1}, K\mu_1 s^2, 1)} d\}$$

and is of relative measure $\geq (1 - \delta)$, *and*

$$(\Psi(t) - id) : D(\frac{r}{4}, \frac{s}{4}) \longrightarrow CK^2\mu_1(\mu_1\mu_2)\frac{d}{\min(r^{\tau+1}, K\mu_1 s^2, 1)^2} D(r, s).$$

So the system $N + H$ has an invariant n-torus with frequency vector (ω_t, Ω_t). This vector is, of course, close to the unperturbed one, and $\omega_t = t\omega$, but we loose control on the arithmetical properties of Ω_t. In particular, we don't know if (ω_t, Ω_t) belongs to DC. It is likely, however, that (eventually under some what stronger smallness assumptions) Ω_t is regular, C^1 in the sense of Whitney say, as a function of t, and that there therefore exists a subset $\Lambda' \subset \Lambda$ (of full measure) such that $(\omega_t, \Omega_t) \in DC$ for all $t \in \Lambda'$.

Proof of theorem C

We will first consider the case $s = 1$. The proof will be done by an iteration process, using proposition 3 as inductive step. We start by proving that there exists a non void set Λ_δ with the stated properties and after this we shall prove the statement about the relative measure of this set.

We shall first define a sequence of parameters. Let C_0 be the constant whose existence is affirmed in proposition 3.

Definition of parameters

Let a be defined by $a^{\tau+1} = min(r^{\tau+1}, K\mu_1, 1)$, and define for $k \geq 1$:

$$r_k = r - (1 - 2^{-k+1})\frac{a}{2}, \; s_k = (1 + 2^{-k+1})\frac{1}{2}$$

$$\nu_k = 2^{(\frac{1}{2})^k}, \; \underline{\mu}^{k+1} = \nu_k \underline{\mu}^k = \nu_k(\mu_1^k, \mu_2^k, \mu_3^k, \mu_4^k), \; \underline{\mu}^1 = \underline{\mu}$$

(we shall avoid confusion by always using parenthésis when taking powers of the μ:s)

$$K_{k+1} = C_0 \alpha^k (\mu_1^k \mu_2^k \mu_3^k \mu_4^k) K_k \delta^{-1}, \; K_1 = K,$$

where α is some fixed number greater than $4^{\tau+1}3$.

Then it follows, by the choice of a, that

$$r_k - 1 < r_{k+1} < r_k, \; \frac{s_k}{2} < s_{k+1}, \; \xi_k = \frac{(r_k - r_{k+1})^{\tau+1}}{K_k \mu_1^k} \leq s_k^2.$$

Let

$$d_{k+1} = C_0(\mu_1^k \mu_2^k)\mu_2^k \xi_k^{-2} d_k^2 (1 - \frac{s_{k+1}}{s_k})^{-2}, \; d_1 = d$$

$$\eta_k = C_0 \mu_2^k \xi_k^{-2} d_k.$$

Now it is straightforward to verify that, if C is sufficiently large (depending on α), then

$$d_k \leq \frac{1}{C_0} \frac{(\nu_k - 1)}{(\mu_1^k \mu_2^k \mu_3^k \mu_4^k)(\mu_1^k \mu_2^k)\mu_2^k} \xi_k^2 (1 - \frac{s_{k+1}}{s_k})$$

hence $d_k \xi_k^{-2} \to 0$, and

$$\sum_{1 \leq j \leq \infty} \eta_j \leq C \frac{K^2 \mu_1 (\mu_1 \mu_2)}{a^{2\tau+2}} d.$$

(P_k) *implies the theorem*

Let $\gamma = K\mu_1\mu_3^{-1}a^{-\tau-1}d$. For $k \geq 2$, consider the statement:

(P_k) there exists a Cantor set $\Lambda^k \subset]1-\gamma, 1+\gamma[$ and, for each $t \in \Lambda^k$, there exists a symplectic diffeomorphism

$$\Psi_k(t) : D(r_k, s_k) \longrightarrow D(r, 1)$$

such that $(N+H) \circ \Psi_k(t) = N^k + H^k$, with

$$N^k \in \mathbf{NF}(r_k, s_k, \underline{\mu}^k) \cap \mathbf{DC}(t\omega, K_k)$$

and

$$|H_0^k + H_1^k|_{D(r_k, s_k)} + \frac{\xi_k}{s_k^2}|H_2^k + H_3^k|_{D(r_k, s_k)} + \frac{\xi_k^2}{s_k^4}|\hat{H}^k|_{D(r_k, s_k)} \leq d_k.$$

Moreover,

$$\Lambda^k \text{ is of measure } \geq \prod_{1 \leq j \leq k-1} (1 - \frac{4^{\tau+1}\delta}{\alpha^j})2\gamma$$

and

$$(\Psi_k(t) - id) : D(r_k, s_k) \longrightarrow \sum_{1 \leq j \leq k-1} \eta_j D(r, 1).$$

Suppose now that (P_k) is true for all $k \geq 2$. Then, if we define

$$\Lambda_\delta = \bigcap_{j=2}^{\infty} \bigcup_{k=j}^{\infty} \Lambda^k$$

we find

$$\text{meas}(\Lambda_\delta) \geq \sum_{1 \leq j < \infty} (1 - \frac{4^{\tau+1}\delta}{\alpha^j})2\gamma \geq (1-\delta)2\gamma$$

by the choice of α.

Let now $t \in \Lambda_\delta$. Then there exists an increasing sequence $k_j \nearrow \infty$, such that $t \in \Lambda^{k_j}$ for all j. By (P_{k_j}), there is a diffeomorphism $\Psi_j : D(r_{k_j}, s_{k_j}) \to D(r, 1)$, such that $N^{k_j} + H^{k_j} = (N+H) \circ \Psi_j$ fulfills the statement of (P_{k_j}). Since the Ψ_j:s are uniformly bounded on $D(\frac{r}{2}, \frac{1}{2})$, we may assume (eventually by taking a subsequence) that $\{\Psi_j\}$ converges to an analytic function Ψ on $D(\frac{r}{4}, \frac{1}{4})$. Since $\{H^{k_j}\}$ converges uniformly to 0 on $D(\frac{r}{2}, \frac{1}{2})$, it follows that N^{k_j} converges uniformly to $(N+H) \circ \Psi$.

This proves the theorem $(s = 1)$ modulo (P_k).

Proof of (P_k)

If we define γ_{k+1} by $K_k \mu_1^k (\mu_3^k)^{-1} (r_k - r_{k+1})^{-\tau-1} d_k$, we see that (P_2) follows from proposition 3 (with δ replaced by $\delta\alpha^{-1}$) since $\gamma_2 = 4^{\tau+1}\gamma$.

Suppose now that (P_k) is true, and take $t \in \Lambda^k$. By proposition 3 there exist a subset $\Lambda_t(\delta\alpha^{-k})$ of $]-\gamma_{k+1}, \gamma_{k+1}[$ and, for all $\sigma \in \Lambda_t(\delta\alpha^{-k})$, a diffeomorphism

$$\Phi(\sigma) : D(r_{k+1}, s_{k+1}) \longrightarrow D(r_k, s_k)$$

such that $((N + H) \circ \Psi(t)) \circ \Phi(\sigma) = N^{k+1} + H^{k+1}$, with N^{k+1} and H^{k+1} (both depending on σ and t) satisfying the statement of (P_{k+1}). In particular,

$$N^{k+1} \in \mathbf{DC}((1 + \sigma)t\omega, K_{k+1}).$$

So we let Λ^{k+1} consist of all such products $(1 + \sigma)t$, with $t \in \Lambda^k$ and $\sigma \in \Lambda_t(\delta\alpha^{-k})$, and we let $\Psi((1 + \sigma)t) = \Psi(t) \circ \Phi((1 + \sigma))$. Then all properties of (P_{k+1}) are immediately fulfilled, except the one concerning the measure of Λ^{k+1} to whose proof we now turn.

Choose β such that $\gamma\beta^{-1}$ is an integer and $\frac{1}{8}\gamma_{k+1} \leq \beta \leq \frac{1}{4}\gamma_{k+1}$, $k \geq 2$. (Such an integer exists if γ_{k+1} is smaller than γ, which is the case if C is sufficiently large). Then we can cover $]1 - \gamma, 1 + \gamma[$ by $\gamma\beta^{-1}$ many disjoint intervals of length 2β. Obviously, the number of all such intervals that intersects Λ^k is greater than $\text{meas}(\Lambda^k)(2\beta)^{-1}$. Let I be such an interval and let $t \in I \cap \Lambda^k$.

Let now $J = \{r : (1 + r)t \subset I\}$. Then J is an interval containing 0 and of length less than 4β (since $\gamma \leq \frac{1}{2}$, if C is sufficiently large, and therefore $t > \frac{1}{2}$). Moreover, J is included in $] - \gamma_{k+1}, \gamma_{k+1}[$ since $\gamma_{k+1} \geq 4\beta$.

Moreover, by proposition 3, we know that $\Lambda_t(\delta\alpha^{-k}) \subset] - \gamma_{k+1}, \gamma_{k+1}[$, and have measure larger than $(1 - \delta\alpha^{-k})2\gamma_{k+1}$, i.e. the complementary set have measure less than

$$\frac{\delta}{\alpha^k}2\gamma_{k+1} \leq 8\frac{\delta}{\alpha^k}2\beta.$$

Since $t \leq 2$, it follows that the set $(\complement\Lambda^{k+1}) \cap I$ have measure less than $16\delta\alpha^{-k}2\beta$, i.e.

$$\text{meas}(\Lambda^{k+1} \cap I) \geq (1 - 16\frac{\delta}{\alpha^k})2\beta.$$

Hence,

$$\text{meas}(\Lambda^{k+1}) \geq (1 - 16\frac{\delta}{\alpha^k})2\beta \ \text{meas}(\Lambda^k)\frac{1}{2\beta} \geq (1 - 4^{\tau+1}\frac{\delta}{\alpha^k}) \ \text{meas}(\Lambda^k).$$

This completes the proof of (P_{k+1}) and, hence, of the theorem when $s = 1$.

The case of an arbitrary s reduces to what we already have done, in the same way as in proposition 3. That is, by scaling the variables through

$$x \longmapsto x, \ y \longmapsto \frac{y}{s^2}, \ z \longmapsto \frac{z}{s}$$

and then dividing $N + H$ by s^2. Then the proof of the theorem is immediate. ∎

Proof of Theorem A.

Consider h given in theorem A, and let D be the domain

$$|\Im x| < r, \ |\Re y| < s_1, \ |\Im y| < s^2, \ |z| < s \tag{25}$$

where we assume that $s_1 \geq 2s^2$. Let B be the ball with center at 0 and radius $\frac{s_1}{2}$ in $\mathbf{R}^n$, and let $\omega(y) = Dh_0(y)$. Let finally μ_1 and μ_3 be such that

$$|\frac{\partial h}{\partial y}|_D + |\frac{\partial^2 h}{\partial z^2}|_D \leq \mu_3 \tag{26}$$

$$|\frac{\partial^3 h}{\partial z^3}|_D + |\frac{\partial^2 h}{\partial y^2}|_D + |\frac{\partial^3 h}{\partial y \partial z^2}|_D + |\frac{\partial^4 h}{\partial z^4}|_D \leq \mu_1. \tag{27}$$

By assumption, there exist μ_2, μ_4 such that (5)-(8) are fulfilled for $(\omega(y), \Omega(y))$. Proposition 1 then gives that

$$J(h, B) \cap DC(K) \neq \emptyset \text{ for } K \geq C(\mu_1\mu_2\mu_3\mu_4)\frac{\mu_2}{s_1}$$

where C only depends on m, n and τ. This proves part 1) of the theorem.

Let now $K > 0$ be given and assume that there exists a $y_0 \in B$, such that $(\omega, \Omega) = (\omega(y_0), \Omega(y_0)) \in J(h, B) \cap DC(K)$. By replacing y by $y - y_0$ we can assume that $y_0 = 0$. We now restrict h to $D(r, s)$.

By using a symplectic transformation which preserves $y = z = 0$ we can transform h to a function, still called h such that $\frac{\partial^3 h}{\partial z^3}(0,0) = 0$ and such that (26) and (27) are fulfilled (for some other constants μ_1, μ_3 and domain D).

Hence, we can assume that $\frac{\partial^3 h}{\partial z^3}(0,0) = 0$, and we can apply theorem C if

$$|f - h|_{D(r,s)} \leq d \leq \frac{1}{C}\frac{1}{K^2(\mu_1\mu_2\mu_3\mu_4)^2(\mu_1\mu_2)^2\mu_1} \ min(r^{\tau+1}, K\mu_1 s^2, 1)^2\delta^2. \tag{28}$$

This proves theorem A. ∎

Notice that (28) gives a quite explicit expression for the constant **C** in theorem A. If

$$D \text{ is the domain (25) with } s_1 \geq 2s^2$$

and

$$h \text{ verifies (5), (6), (26), (27) and } \frac{\partial^3 h}{\partial z^3}(0,0) = 0$$

then the existence of n-dimensional invariant tori is garanteed by the condition

$$|f - h|_D \leq \frac{1}{C}\frac{1}{(\mu_1\mu_2\mu_3\mu_4)^4(\mu_1\mu_2)^3\mu_2} \ min(r^{\tau+1}s_1, s^2)^2\delta^2 \tag{29}$$

where C only depends on m, n and τ.

Proof of theorem B

Since $\langle k, \omega\rangle + \langle l, \Omega\rangle \neq 0$ for $0 < |k| + |l| \leq 4$, there exists a symplectic diffeomorphism Φ such that $h \circ \Phi$ is on Birkhoff normal form up to order 4 (see [16]). We can therefore assume that h is of the form

$$h = h_2 + h_4 + \cdots, \text{ with } h_2 = \langle \omega, y\rangle + \langle \Omega, y'\rangle \text{ and}$$

$$h_4 = \frac{1}{2}\langle yM_1, y\rangle + \langle yM_2, y'\rangle + \frac{1}{2}\langle y'M_3, y'\rangle.$$

By assumption,

$$\det M_1 \neq 0 \text{ and } \langle l, \Omega - \omega M_1^{-1}M_2\rangle \neq 0 \text{ for } 0 < |l| \leq 3.$$

If we stretch the variables by a factor $\sqrt{\varepsilon}$ and divide by ε, then we get

$$h = h_2 + \varepsilon h_4 + O(\varepsilon^{\frac{3}{2}}),$$

defined for $|w| < 1, |z| < 1, 0 \leq \varepsilon < \varepsilon_0$.

By introducing symplectic polar coordinates

$$w_j = \sqrt{2(\frac{1}{4} - y_j)}\cos x_j, \quad w_{j+k} = \sqrt{2(\frac{1}{4} - y_j)}\sin x_j$$

we get, up to a constant factor,

$$h = \langle \omega, y\rangle + \langle \Omega, y'\rangle + \varepsilon h_4 + O(\varepsilon^{\frac{3}{2}}),$$

defined for $|y| < \frac{1}{4}$, $x \in \mathbf{T}^n$, $|z| < 1, 0 \leq \varepsilon < \varepsilon_0$. Then we can apply theorem A with constants $(\varepsilon\mu_1, \frac{1}{\varepsilon}\mu_2, \mu_3, \mu_4)$, where the μ:s are independent of ε.

If we now use the smallness condition (29), theorem B follows since

$$\frac{1}{(\varepsilon\mu_1\frac{1}{\varepsilon}\mu_2)^7(\mu_3\mu_4)^4\frac{1}{\varepsilon}\mu_2}$$

is of size $O(\varepsilon)$ which is larger than $O(\varepsilon^{\frac{3}{2}})$. This proves theorem B. ∎

REFERENCES

[1] V.I. Arnold, *Mathematical methods of classical mechanics*, Springer (1978).

[2] J. Vey, *Sur certains systèmes dynamiques séparables*, American Journal of Mathematics, **100** (1978), pp. 591-614.

[3] A.N. Kolmogorov, *The general theory of dynamical systems and classical mechanics*, Proc. of the 1954 Intern. Congr. of Math., in R. Abraham, J.E. Marsden, Foundations of mechanics, Benjamin (1978).

[4] V.I. ARNOLD, *Proof of a theorem of A.N. Kolmogorov on the invariance of quasiperiodic motions under small perturbations of the Hamiltonian*, Russian Mathematical Surveys, **18** (1962), No. 5, pp. 9-36.
[5] J. MOSER, *On the theory of quasiperiodic motions*, Siam Review **8** (1966), pp. 145-172.
[6] J. MOSER, *Convergent series expansions for quasiperiodic motions*, Mathematische Annalen, **169** (1967), pp. 136-176.
[7] H. POINCARÉ, *Les méthodes nouvelles de la mécanique céleste*, vol. I, Paris (1892).
[8] S. GRAFF, *On the conservation of hyperbolic invariant tori for Hamiltonian systems*, Journal of Differential Equations, **15** (1974), pp. 1-69.
[9] E. ZEHNDER, *Generalized implicit function theorems with application to some small divisor problems II*, Communications on Pure and Applied Mathematics, **29** (1976), pp. 49-111.
[10] V.K. MELNIKOV, *On some cases of conservation of conditionally periodic motions under a small change of the Hamiltonian function*, Soviet Mathematics Doklady, **6** (1965), pp. 1592-1596.
[11] V.K. MELNIKOV, *A family of conditionally periodic solutions of a Hamiltonian system*, Soviet Mathematics Doklady, **9** (1968), No. 4, pp. 882-885.
[12] J. MOSER - J. PÖSCHEL, *An extension of a result by Dinaburg and Sinai on quasiperiodic potentials*, Commentarii Mathematici Helvetici, **59** (1984), pp. 39-85.
[13] C.L. CHARLIER, *Die Mechanik des Himmels*, vol. 2, Leipzig (1907).
[14] A.S. PYARTLI, *Diophantine approximation on submanifolds of euclidean space*, Functional Analysis and its Applications, **3** (1969), pp. 303-306.
[15] H. RÜSSMANN, *On optimal estimates for the solutions of linear partial differential equations of the first order with constant coefficients on the torus*, Springer, Lecture Notes in Physics, Vol. 38 (1975), pp. 598-624.
[16] C.L. SIEGEL - J. MOSER, *Lectures on celestial mechanics*, Springer (1971).

Department of Mathematics
University of Stockholm
Box 6701
S-113 85 Stockholm, Sweden