A PROOF OF NEKHOROSHEV'S THEOREM FOR THE STABILITY TIMES IN NEARLY INTEGRABLE HAMILTONIAN SYSTEMS

GIANCARLO BENETTIN
Dipartimento di Fisica dell'Università, Padova, Gruppo Nazionale di Struttura della Materia del C.N.R - Unità di Padova, Via Marzolo 8, Padova, Italia

LUIGI GALGANI
Dipartimento di Matematica dell'Università, Via Saldini 50, Milano, Italia

and

ANTONIO GIORGILLI
Dipartimento di Fisica dell'Università, Via Celoria 16, Milano, Italia

(Received 5 December, 1984; accepted 11 July, 1985)

Abstract. In the present paper we give a proof of Nekhoroshev's theorem, which is concerned with an exponential estimate for the stability times in nearly integrable Hamiltonian systems. At variance with the already published proof, which refers to the case of an unperturbed Hamiltonian having the generic property of steepness, we consider here the particular case of a convex unperturbed Hamiltonian. The corresponding simplification in the proof might be convenient for an introduction to the subject.

1. Introduction

Classical perturbation theory is concerned with nearly integrable Hamiltonians of the form

\[ H(p, q) = h(p) + \epsilon f(p, q), \]

where \( p = (p_1, \ldots, p_n) \in \mathcal{G} \subset \mathbb{R}^n \) and \( q = (q_1, \ldots, q_n) \in \mathbb{T}^n \) are action-angle variables, \( \mathcal{G} \) denoting an open bounded domain of \( \mathbb{R}^n \) and \( \mathbb{T}^n \) the \( n \)-dimensional torus; \( H \) is assumed to be smooth (typically analytic) in its domain of definition \( \mathcal{D} = \mathcal{G} \times \mathbb{T}^n \).

For vanishing \( \epsilon \) the motion is trivial, and in particular the actions \( p_1, \ldots, p_n \) are constant; for \( \epsilon \neq 0 \), instead, the only result that can be easily read directly by inspection from Hamiltonian (1) is the very poor a priori estimate \( |p_j(t) - p_j(0)| = O(\epsilon t), (j = 1, \ldots, n) \). Generally speaking, the purpose of classical perturbation theory is to go beyond

such elementary result, providing a deeper insight into the Hamiltonian system (1), in order to obtain better estimates for the time evolution of the "slow variables" \( p_j \).

The basic tool of classical Hamiltonian perturbation theory are canonical transformations \( \psi : D' \to D \) near the identity, where both the domain of definition \( D' = G' \times T^n \) and its image \( \psi(D') \) are required to be, in some sense, near the original domain of definition \( D \). The requirement for the correspondingly transformed Hamiltonian \( H' = H \circ \psi \) is that the behavior of the system be more transparent. In fact, let us consider a canonical transformation near the identity \( (p, q) = \psi(p', q') \) for which one has \( |p_j - p_j'| \leq B(\epsilon) \), \( B \) being infinitesimal with \( \epsilon \). If one could obtain, in the new variables, a \( q \)-independent Hamiltonian

\[
H'(p', q') = h'(p'),
\]

then, in the old variables, every motion \( (p(t), q(t)) \) with initial datum in \( \psi(D') \) would satisfy the beautiful estimate

\[
|p_j(t) - p_j(0)| \leq 2B(\epsilon)
\]

for any \( t \). Unfortunately, as stressed by Poincaré\(^1\), this program in general fails, because of the well known difficulty related to the density of resonances in phase space, which manifests itself analytically as the "problem of the small divisors".

Two powerful ways out, however, were found, within quite general hypotheses on the Hamiltonian. The first is the one provided by the celebrated Kolmogorov Arnold Moser (KAM) theorem\(^2\), which, following a recent formulation\(^3\), can be stated as guaranteeing form (2) of the Hamiltonian in a quite peculiar domain, which is close to \( D \) in measure, but is topologically quite different, being the complement of an open dense subset of \( D \) (see also ref. 4 for a proof along the original Kolmogorov's method). The second more recent one, due to Nekhoroshev\(^5\-\(^6\), is concerned instead with estimates in a "good domain" \( D' \), but this can be obtained by renouncing to have estimates for all times. In fact one can guarantee

\[
|p_j(t) - p_j(0)| \leq P \epsilon^b
\]

for times even larger than any power of \( \frac{1}{\epsilon} \), precisely for \( t \in (0, T) \) with

\[
T = T \left( \frac{1}{\epsilon} \right)^{\frac{1}{2}} e^{(\frac{1}{\epsilon})^a}
\]

where \( P \) and \( T \) are dimensional constants and \( a, b \) are positive constants; the only trivial restriction on the initial datum is that the distance between \( p(0) \) and the border of the original domain \( D \) be not less than \( P \epsilon^b \) (in the original paper of Nekhoroshev one finds 1 instead of \( \frac{1}{2} \) as the power of \( \frac{1}{\epsilon} \) in (5)).

Now, at variance with KAM theorem, the theorem of Nekhoroshev is not much known, and in addition a relevant part of its proof, called by that author the "analytic part", is published only in russian; on the other hand such theorem is certainly physically relevant, possibly even more than KAM theorem\(^7\). So it seemed to us to be useful to present here a complete proof of it. For what concerns the conditions required by the theorem, they are essentially smoothness (in fact analyticity) of \( H \), smallness of \( \epsilon \), and a geometrical condition on the unperturbed Hamiltonian \( h \) firstly considered by Glimm\(^8\).
and called "steepness", which is in a sense a generalization of convexity. As shown by Nekhoroshev, such condition is generic. In the present paper, in order to give a complete proof showing the essence of the problem without entering the quite complicated details related to the conditions of steepness, we decided to limit ourselves to the special but quite representative case of convex unperturbed Hamiltonians $h$. In the same spirit, for the numerous estimates entering the theorem, attention was payed not to obtain optimal results, but to give simple explicit expressions.

In section 2, having preliminarily fixed the notations, we state precisely the theorem. Section 3 is devoted to the analytic part of the proof, and section 4 to the geometrical part, while the conclusion of the proof is given in section 5. Section 6 contains the proof of the analytic lemma, used in section 2. This is the analogue of the corresponding lemma published in the second of Nekhoroshev's papers\[^{[6]}\]. Some minor details are deferred to appendices A, B and C.

2. Preliminaries and statement of the theorem

Let us fix some notations. First of all, for vectors $v \in \mathbb{C}^n$ we will consider the norm $\|v\| = \max_i |v_i|$. Instead, in the case of vectors $k \in \mathbb{Z}^n$ of integer components we will usually consider the norm $|k| = \sum_i |k_i|$. Both for real and complex vectors $v, w$ we denote $v \cdot w = \sum_i v_i w_i$. As usual, a linear subspace $\mathcal{M}$ of $\mathbb{Z}^n$ will be called a module.

As usual in Hamiltonian perturbation theory, one deals with action-angle variables $(p, q)$ defined in a domain $\mathcal{G} \times \mathbb{T}^n$, where $\mathcal{G}$ (space of actions) is a domain of $\mathbb{R}^n$ and $\mathbb{T}^n$ (space of angles) is the $n$-dimensional torus. As in the proof of KAM theorem in the analytic framework, use will be made here too of the standard Cauchy inequalities for estimating the derivatives of an analytic function, which are recalled in Appendix A. Thus, for a generic domain $G \times \mathbb{T}^n$ with $G \subset \mathcal{G}$, we need consider complex extensions of the form

$$D \equiv F(G, \rho, \sigma) \equiv \{(p, q); \text{Re } p \in G, \|\text{Im } p\| \leq \rho; \text{Re } q \in \mathbb{T}^n, \|\text{Im } q\| \leq \sigma\},$$  \hspace{1cm} (6)

where $\rho$ and $\sigma$ are positive parameters. For an analytic function $f$ defined on a domain $D$ as above, the norm $\|f\|_D$ is then defined by

$$\|f\|_D = \sup_{(p, q) \in D} |f(p, q)|.$$  \hspace{1cm} (7)

In addition, for any domain $G \subset \mathcal{G}$ we define as usual $G - \delta$ as the set of points which are contained in $G$ together with a $\delta$-neighborhood, and $G + \delta$ as the union of the $\delta$-neighborhoods of the points of $G$. In an analogous way, for a complex domain $D = F(G, \rho, \sigma)$ defined as in (6), for given positive $\delta < \rho$ and $\xi < \sigma$ we denote $D - (\delta, \xi) = F(G - \delta, \rho - \delta, \sigma - \xi)$.

We can now come to the statement of the
Theorem. Let \( H(p, q) = h(p) + f(p, q) \) be analytic in the domain \( D = F(\mathcal{G}, \rho, \sigma) \) defined by (6) with \( \mathcal{G} \) bounded, open and convex, and \( \rho \) and \( \sigma \) positive, and \( n > 1 \). Consider the matrix \( A(p) \) defined by \( A_{ij} = \frac{\partial^2 h}{\partial p_i \partial p_j} \), and introduce four parameters \( E, M, m, \epsilon \) by

\[
\|H\|_D \leq E
\]
\[
\|Av\|_D \leq M\|v\| \text{ for all } v \in C^n
\]
\[
|A(p)v \cdot v| \geq m\nu \cdot v \text{ for all } p \in \mathcal{G} \text{ and all } v \in \mathbb{R}^n
\]
\[
\|f\|_D \leq \epsilon E.
\]

Suppose

\[
\epsilon < \epsilon_0 = \min \left( \left( \frac{M\rho^2}{E} \right)^2, \left( \frac{\sigma}{8n} \right)^c \left( \frac{m}{M} \right)^{8n} \right)
\]

with

\[
c = 4(n^2 + 2n + 2);
\]

assume also (which is not restrictive)

\[
\sigma \leq 1, \quad \rho \leq \left( \frac{E}{M} \right)^{\frac{1}{2}}.
\]

Then one has

\[
\|p(t) - p(0)\| \leq \Delta \quad \text{for any} \quad t \in [0, T],
\]

for all real motions \((p(t), q(t))\) with \( p(0) \in \mathcal{G} - 2\Delta \), where

\[
\Delta = 2(n + 1)\rho \epsilon^\frac{1}{c},
\]

\[
T = \mathcal{T} \left( \frac{1}{\epsilon} \right)^{\frac{1}{c}} e\left( \frac{1}{c} \right)^{\frac{1}{c}};
\]

\(\mathcal{P}\) and \(\mathcal{T}\) being dimensional constants defined by

\[
\mathcal{P} = \left( \frac{ME}{m} \right)^{\frac{1}{2}}, \quad \mathcal{T} = \left( \frac{1}{EM} \right)^{\frac{1}{2}}.
\]

One might notice that relation (10) expresses a sort of uniform convexity for the unperturbed Hamiltonian \(h\); at variance with the much more general steepness condition of Nekhoroshev, which requires to consider \(3n - 2\) parameters, the former just needs one parameter \(m\). One can see that, in the statement of the theorem, \(m\) appears only in the expression of the estimate \(\Delta\) for the allowed variation of the action; this fact reflects the circumstance that \(m\) is used only in the geometric part of the proof. One might also notice that the above convexity condition is stronger than the condition \(|\text{det} A(p)| \geq m > 0\) usually assumed in KAM theorem; however, as is shown in Nekhoroshev's paper, the more general steepness condition is independent of the latter one. For what concerns the parameter \(\epsilon\) measuring the size of the perturbation \(f\), one might point out the difference with respect to the more special way it entered the problem in (1). Finally, we considered to be of interest to give all formulae with the appropriate dimensional constants; in fact,
as was pointed out particularly by Gallavotti in connection with KAM theorem\[^3\], this turns out to be useful also in handling formulae for the proof of the theorem.

3. The analytic part of the proof

In perturbation theory, starting with an Hamiltonian of the form \( H(p, q) = h(p) + f(p, q) \), one usually looks for a canonical transformation such that the new Hamiltonian \( H' \) take the approximate “normal form” \( H'(p, q) = h(p) + Z(p) + R(p, q) \) in which the angles \( q \) have been eliminated up to a “small” remainder \( R \). In doing that, as already recalled, one meets the small divisors \( \omega(p) \cdot k \) (where \( \omega = \frac{\partial h}{\partial p} \) are the unperturbed frequencies, and \( k \in \mathbb{Z}^n \)), which in general vanish in a dense subset of the domain of definition \( \mathcal{G} \) of \( h \). However, as pointed out by Arnold, for analytic Hamiltonians it is sufficient to take care only of a finite number of small denominators, say those with \( |k| = |k_1| + \cdots + |k_n| \leq K \) for a given (large) \( K \), so that one can control the problem by only guaranteeing a condition of the type \( |\omega(p) \cdot k| \geq \alpha \) for \( |k| \leq K \). In such a way one has a description adapted to the “nonresonant region” of phase space. On the other hand, it is well known that even in the resonant regions it is possible to give a suitable normalization; indeed, if in a certain domain one can guarantee \( |\omega(p) \cdot k| \geq \alpha \) for \( |k| \leq K \) and \( k \notin \mathcal{M} \), where \( \mathcal{M} \) is a given subspace (“resonant module”) of \( \mathbb{Z}^n \), then, in a corresponding region of phase space one can put the Hamiltonian into the approximate normal form adapted to it, namely \( H'(p, q) = h(p) + Z(p, q) + R(p, q) \), where \( Z(p, q) = \sum_{k \in \mathcal{M}} z_k(p) e^{ik \cdot q} \) and the remainder \( R \) is “small”. In particular, for \( \mathcal{M} = 0 \), one recovers the nonresonant normal form. For an elementary illustration of such aspects of perturbation theory, see for example ref. 9.

The possibility of performing such a normalization with a remainder exponentially decreasing with \( K \), if \( K \) is sufficiently large and the size \( \epsilon \) of the perturbation is sufficiently small, is guaranteed by the following

**Analytic Lemma.** Let \( H(p, q) = h(p) + f(p, q) \) be analytic in the domain \( D = F(G, \rho, \sigma) \) defined as in (6) with \( G \subset \mathbb{R}^n \) compact and \( \rho \) and \( \sigma \) positive, and introduce positive numbers \( E, \epsilon \) by

\[
\| H \|_D \leq E
\]

\[
\| f \|_D \leq \epsilon E
\]

With reference to a given module \( \mathcal{M} \subset \mathbb{Z}^n \) and to positive numbers \( \alpha, K \) assume, everywhere in \( D \),

\[
|\omega(p) \cdot k| \geq \frac{\alpha}{2} \text{ for all } k \notin \mathcal{M}, \; |k| < K
\]

where \( \omega = \frac{\partial h}{\partial p} \). Assume also that the quantities \( \rho, \sigma, E, \epsilon, \alpha, K \), together with a positive \( \delta < \rho \) satisfy the inequalities

\[
\sigma \leq 1 \; , \; \rho \leq \frac{E}{\alpha}
\]

\[
8n\sigma^{-1}K^{-\frac{1}{4}} < 1
\]
\[ B \left( \frac{E}{\delta \alpha} \right)^2 \sigma^{-2n-2} K^{n+2} \epsilon < 1 \]  

(24)

with

\[ B = 2^{6n+16} (n+2)^{2n+4} \]  

(25)

Then there exists an analytic canonical transformation \( \psi = (\psi^p, \psi^q) : D' \to D \) where \( D' = F(G', \rho - \delta, \sigma/2) \) with \( G' \) compact such that

\[ G - \delta \subset G' \subset G , \quad G - \delta \subset \psi^p(G') \subset G , \]  

(26)

which is near the identity \( I = (I^p, I^q) \), and in particular satisfies

\[ \| \psi^p - I^p \|_{D'} < \frac{\delta}{2} . \]  

(27)

In \( D' \) the new Hamiltonian \( H' = H \circ \psi \) takes the form

\[ H'(p, q) = h(p) + Z'(p, q) + R'(p, q) \]  

(28)

where

\[ Z'(p, q) = \sum_{k \in \mathcal{M}} z'_k(p) e^{ik \cdot q} , \]  

(29)

and \( z'_k \) are suitable Fourier coefficients, while for the remainder \( R' \) one has the exponential estimate

\[ \| R'\|_{D'} < 16 \epsilon E e^{-K^{\frac{1}{4}}} . \]  

(30)

The proof is deferred to section 6. The use of the normal form (28) is the following. Due to the Hamiltonian character of the equations of motion, from (28) one immediately obtains 

\[ \dot{p} = \sum_{k \in \mathcal{M}} k \omega_k(p, q) + \frac{\partial R'}{\partial q} \] \( \omega_k(p, q) = iz'_k(p) e^{ik \cdot q} \), so that one can have a fast drift only along the hyperplane \( \Pi_{\mathcal{M}}(p(0)) \) through \( p(0) \) parallel to \( \mathcal{M} \) and of the same dimensionality as \( \mathcal{M} \), while the deviation from such plane, due to the remainder, should in any case be small. In fact one then easily proves the following

Main Lemma. With the same notations of the theorem, let \( H(p, q) = h(p) + f(p, q) \) be analytic in the domain \( D = F(G, \rho, \sigma) \) and let \( E, M, \epsilon, \sigma \) be defined as there. Suppose \( E, M, \epsilon, \sigma \), together with given positive numbers \( K, \alpha \), satisfy the inequalities

\[ \sigma \leq 1 , \quad \rho \leq \frac{E}{\alpha} \]  

(31)

\[ \frac{E}{M \rho^2 \epsilon^{\frac{3}{2}}} < 1 \]  

(32)

\[ 8n \sigma^{-1} K^{-\frac{1}{4}} < 1 \]  

(33)

\[ B \frac{EM}{\alpha^2} \sigma^{-2n-2} K^{n+2} \epsilon^{\frac{1}{2}} < 1 \]  

(34)
A PROOF OF NEKHOROSHEV'S THEOREM

with \( B \) given by (25). Denote

\[
d = 2 \left( \frac{E}{M} \right)^{\frac{1}{2}} \epsilon^\frac{1}{4} \tag{35}\]

and consider any set \( \Gamma \subset G - 2d \) such that, with reference to a given module \( \mathcal{M} \subset \mathbb{Z}^n \), one has, everywhere in \( \Gamma \),

\[
|\omega(p) \cdot k| \geq \alpha \quad \text{for all} \quad k \notin \mathcal{M}, \quad |k| \leq K. \tag{36}\]

Finally, denote by \( \tau^+ \) and \( \tau^- \), \( \tau^- < 0 < \tau^+ \), the (possibly infinite) times of escape of \( p(t) \) from \( \Gamma \), at positive and negative times respectively. Then for any real motion \( (p(t), q(t)) \) with \( p(0) \in \Gamma \) one has

\[
\text{dist} [p(t), \Pi_\mathcal{M}(p(0))] < d \tag{37}\]

for \( t \in [-\min(T, -\tau^-), \min(T, \tau^+)] \), where \( \Pi_\mathcal{M}(p(0)) \) is the hyperplane through \( p(0) \) parallel to \( \mathcal{M} \) and of the same dimensionality as \( \mathcal{M} \), and

\[
T = \left( \frac{1}{EM} \right)^{\frac{1}{2}} \left( \frac{1}{\epsilon} \right)^{\frac{1}{2}} e^{K^\frac{1}{4}}. \tag{38}\]

Proof. Take \( \delta = \frac{d}{2} = (\frac{E}{M})^{\frac{1}{2}} \epsilon^\frac{1}{4} \); by (32) it is \( \delta < \rho \). Consider then the domain \( G = \overline{G + \delta} \) (where the overline denotes closure) and the corresponding domain \( D = F(G, \delta, \sigma) \) with \( F \) defined as in (6). Now, all hypotheses of the analytic lemma are easily seen to be satisfied. Indeed, (31) and (33) coincide with (22) and (23), respectively; and (34), together with the definition of \( \delta \), gives (24). There remains to show that (21) is also satisfied, and this is seen as follows. For any \( p \in G \) there exists \( \hat{p} \in \Gamma \) satisfying \( \|p - \hat{p}\| < 2\delta = d \), and consequently, from the very definition (9) of \( \mathcal{M} \), also satisfying \( \|\omega(p) - \omega(\hat{p})\| \leq Md \); moreover, from (34) one has in particular \( MdK < \frac{\alpha}{2} \), so that from (36) one finally obtains

\[
|\omega(p) \cdot k| \geq |\omega(\hat{p}) \cdot k| - |(\omega(p) - \omega(\hat{p})) \cdot k| \geq |\alpha - MdK| > \frac{\alpha}{2}
\]

for any \( k \notin \mathcal{M}, \quad |k| \leq K \).

One can then apply the analytic lemma in the domain \( D = F(G, \delta, \sigma) \) and consider the motion \( (p'(t), q'(t)) = \psi^{-1}(p(t), q(t)) \), which, by (26), is certainly defined for \( p(t) \in \Gamma \), i.e. for \( t \in [\tau^-, \tau^+] \). This motion is solution of the Hamilton equations corresponding to Hamiltonian (28). Thus one can write

\[
p'(t) = p'(0) + \int_0^t \frac{\partial H'}{\partial q} (p'(t'), q'(t')) \, dt'
\]

i.e.

\[
p'(t) = \tilde{p}(t) + \int_0^t \frac{\partial R}{\partial q} (p'(t'), q'(t')) \, dt'
\]

with

\[
\tilde{p}(t) = p'(0) + \int_0^t \frac{\partial Z'}{\partial q} (p'(t'), q'(t')) \, dt'.
\]
In particular, by the form (29) of $Z'$, one has $\tilde{p}(t) \in \Pi_M(p'(0))$ for any $t \in [\tau^-, \tau^+]$, while, from the estimate (30) of the remainder and Cauchy inequality together with the definition (35) of $d$, for any $t \in [-\min(T, -\tau^-), \min(T, \tau^+)]$ one has $\|p'(t) - \tilde{p}(t)\| \leq 32\sigma^{-1}e^{\frac{t}{4}}d \leq \frac{d}{4}$, where (34) was also used. On the other hand, from the property (27) of the canonical transformation one has $\|p(t) - p'(t)\| \leq \frac{d}{4}$ as well as $\text{dist} [\tilde{p}(t), \Pi_M(p(0))] \leq \frac{d}{4}$, so that one finally gets

$$\text{dist} [p(t), \Pi_M(p(0))] \leq \|p(t) - p'(t)\| + \|p'(t) - \tilde{p}(t)\| + \text{dist} [\tilde{p}(t), \Pi_M(p(0))]$$

$$\leq \frac{d}{2} + \frac{d}{4} + \frac{d}{4} = d$$

for $t \in [-\min(T, -\tau^-), \min(T, \tau^+)]$, as claimed.

The meaning of the main lemma is essentially the following. Consider the “cylinder” $\Pi_M(p(0)) + d$, i.e. the “cylinder” with “axis” $\Pi_M(p(0))$ and “radius” $d$, and let $C$ be its intersection with the set $\Gamma$. We then have shown that $p(t)$ is constrained to remain inside $C$ up to time $T$, unless at $\tau^+ < T$ it leaves $\Gamma$; but in such case it can go out of $\Gamma$ only through one of the “bases” of the cylinder. Similarly, $p(t)$ can enter $C$ only through a basis.

By means of this lemma we achieve a good control on the projections of the motions $p(t)$ “transversal” to the planes $\Pi_M(p(0))$. Nothing can be said instead, for the moment, on the motion (fast drift) along the axis of the cylinder. Such “cylinders of fast drift” play a relevant role in the geometric part of Nekhoroshev’s theorem.

4. The geometric part of the proof

A. Geography of resonances

In order to apply the main lemma to our Hamiltonian system, we need to divide the real part $\mathcal{G}$ of the action space into regions having well defined resonance properties.

First of all, we are interested in resonances only up to a maximal order $K$; thus, from now on, we will consider only those moduli $M \subset \mathbb{Z}^n$, $\dim M = r$, which admit a basis $(k^{(1)}, \ldots, k^{(r)})$ satisfying $|k^{(j)}| \leq K$ for $j \leq r$ (the 0-dimensional module $M = 0$ is also included). Such a basis will be called a $K$-basis, and the corresponding module will also be called a $K$-module. In addition to $K$, our geometrical construction will be characterized by $n+2$ more parameters, precisely a positive parameter $d$, to be later identified with the $d$ appearing in the main lemma, and a sequence of $n+1$ positive numbers $\alpha_0 < \alpha_1 < \ldots < \alpha_n$, to be used in defining resonance relations with moduli of different dimensionality.

Making reference to these parameters, we then define, for each $K$-module $M \subset \mathbb{Z}^n$ as above, with $\dim M = r$:

i) The resonant manifold

$$\mathcal{R}_M = \{ p \in \mathcal{G} - 2d ; \quad \omega(p) \cdot k = 0 \text{ for all } k \in M \} ; \quad (39)$$

© Kluwer Academic Publishers • Provided by the NASA Astrophysics Data System
ii) The resonant zone

\[ Z_M = \left\{ p \in \mathcal{G} - 2d; \left| \omega(p) \cdot k^{(j)} \right| < \alpha_r \right\} \quad (j = 1, \ldots, r) \]  

(40)

where \( k^{(1)}, \ldots, k^{(r)} \) is a \( K \)-basis.

Notice that, for \( r = n \), the resonant manifold reduces to a point. Instead, for \( r = 0 \), both \( \mathcal{R}_0 \) and \( Z_0 \) are defined as coinciding with \( \mathcal{G} - 2d \).

Denote by \( Z_r^* \), \( 0 \leq r \leq n \), the union of all resonant zones with the same \( r \), i.e.

\[ Z_r^* = \bigcup_{\dim M = r} Z_M \]  

(41)

and extend the definition by setting \( Z_{n+1}^* = \emptyset \), the empty set, (while we already have \( Z_0^* = \mathcal{G} - 2d \)). We can then define, for each \( K \)-module \( M \subset \mathbb{Z}^n \) with \( \dim M = r \),

iii) The resonant block

\[ B_M = Z_M \setminus Z_{r+1}^* \]  

(42)

The dimension \( r \) of \( M \) will also be called the multiplicity of the corresponding resonant manifold, zone or block.

Let us stress a few basic properties of this construction:

a) The blocks constitute a covering of \( \mathcal{G} - 2d \) (although not a partition, because superpositions are possible). Of course, from a given multiplicity on, they could also be empty.

b) In each block of multiplicity \( r \), there are exactly \( r \) resonances “within \( \alpha_r \)”, while, by definition, \( r + 1 \) resonances are excluded even within \( \alpha_{r+1} > \alpha_r \). Notice however that \( r'' \) resonances with \( r'' > r + 1 \) are not excluded.

c) If \( p \) does not belong to any zone \( Z_M \) of multiplicity \( r \), then there exists a block \( B_{M'} \) of multiplicity \( r' < r \) containing \( p \), as can be easily seen by induction. Notice however that this does not exclude the existence of a block \( B_{M''} \) of multiplicity \( r'' > r \) containing it.

Let us proceed, and complete our geometrical construction. To each \( p \) belonging to a given block \( B_M \) we associate

iv) the “cylinder” of “radius” \( d \)

\[ C_{M,d}(p) = (\Pi_M(p) + d) \cap Z_M \]  

(43)

where \( \Pi_M(p) \) is the hyperplane through \( p \) parallel to \( M \) and of the same dimensionality. The set

\[ \Lambda_M(p) = \Pi_M(p) \cap Z_M \]  

(44)

can be considered to be the cylinder axis, while the intersection of \( C_{M,d}(p) \) with the border of \( Z_M \) gives the cylinder bases. Let us stress that such bases do not belong to \( Z_M \), which was taken to be open.

Finally, to each block \( B_M \) we associate the corresponding
v) Extended block

\[ B^{\text{ext}}_{M,t} = \bigcup_{p \in B_M} C_{M,t}(p) \ . \quad (45) \]

As a final comment, we notice that our terminology was a little changed with respect to that of Nekhoroshev; indeed his disks are here called cylinders, and his lateral surfaces are here called bases.

\[ B^{\text{ext}}_{M,t} \]

B. The condition of nonoverlapping of resonances

We will apply the main lemma choosing as basic set \( \Gamma \) one of the extended blocks \( B^{\text{ext}}_{M,t} \). A condition required for the applicability of the lemma is (36), which in our new language reads

\[ |\omega(p) \cdot k| \geq \alpha_r \text{ for all } p \in B^{\text{ext}}_{M,t} \text{ and all } k \notin M , \ |k| \leq K \ . \quad (46) \]

This is called "condition of nonoverlapping of resonances", because of its geometrical meaning: in fact, it states that there is no intersection between an extended block \( B^{\text{ext}}_{M,t} \) and any resonant zone \( Z_{M'} \) of the same multiplicity, with \( M' \neq M \).

Once this condition is satisfied, quite relevant consequences can be drawn. Indeed, in such case, if one takes \( p \) belonging to \( B_M \) of multiplicity \( r \), then the bases of the cylinder \( C_{M,t}(p) \) do not belong to any resonant zone of the same multiplicity. As a consequence, as already remarked, such bases do belong to a block \( B_{M'} \) of multiplicity \( r' < r \), unless they are outside \( G - 2d \), and this will be a crucial point in the proof of the following

Geometric lemma. With the same notations of the theorem, let \( H(p,q) = h(p) + f(p,q) \) be analytic in the domain \( D = F(G,\rho,\sigma) \), let \( E,M,m,\epsilon \) be defined by (8) -(11) , and denote

\[ d = 2 \left( \frac{E}{M} \right)^{\frac{1}{2}} \epsilon^{\frac{1}{2}} \ . \quad (47) \]

Take positive parameters \( K \) and \( \alpha_0 < \alpha_1 < \cdots < \alpha_n \). For any extended block \( B^{\text{ext}}_{M,t} \) of multiplicity \( r = 0,1,\ldots,n \) defined as in section 3A with \( \omega = \frac{\partial h}{\partial p} \), assume the condition of nonoverlapping of resonances, namely

\[ |\omega(p) \cdot k| \geq \alpha_r \text{ for all } p \in B^{\text{ext}}_{M,t} \text{ and all } k \in M , |k| < K . \quad (48) \]

Assume furthermore the conditions entering the main lemma, namely

\[ \sigma \leq 1 \ , \ \rho \leq \frac{E}{\alpha_r} \quad (49) \]

\[ \frac{E}{M\rho^2} \epsilon^{\frac{1}{2}} < 1 \quad (50) \]

\[ 8n\sigma^{-1}K^{-\frac{1}{2}} < 1 \quad (51) \]

\[ B^{EM}_{\alpha_r^2} \sigma^{-2n-2} K^{n+2} \epsilon^{\frac{1}{2}} < 1 \ , \ (r = 0,1,\ldots,n) \ , \quad (52) \]
where the constant $B$ is defined by (25). Consider any real motion $(p(t), q(t))$ of Hamiltonian $H$, for any $t \in [0, \min(T, \tau_0)]$, where $\tau_0$ is the escape time of $p(t)$ from $\mathcal{G} - 2d$ and $T$ is given by (38), namely

$$T = \left( \frac{1}{EM} \right)^{\frac{1}{3}} \left( \frac{1}{\epsilon} \right)^{\frac{1}{3}} e^{K \frac{t}{4}}. \quad (53)$$

Then $p(t)$ is confined to a cylinder $C_{M^*}d(p(t^*))$, with $t^* \in [0, \min(T, \tau_0)]$, and $p(t^*) \in \mathcal{B}_{M^*}$ of multiplicity $r^*$, minimal among those of the blocks visited by $p(t)$ during the above time interval.

**Remark.** The cylinder $C_{M^*}d(p(t^*))$ coincides with the cylinder $C_{M,d}(p(0))$ such that $p(0) \in \mathcal{B}_{M}$, unless $p(0)$ is so close to a basis of the latter cylinder that $p(0)$, although not belonging to $\mathcal{B}_{M^*}$, belongs nevertheless to $\mathcal{B}_{M^*}^{ext,d}$.

**Proof.** Suppose that, for $t \in [0, \min(T, \tau_0)]$, $(p(t), q(t))$ is a motion of our Hamiltonian system, $\tau_0$ being the (positive, possibly infinite) escape time of $p(t)$ from $\mathcal{G} - 2d$. Suppose $p(t)$ visits several blocks; let $r^*$ be their minimal multiplicity, and $t^*$ be such that $p(t^*) \in \mathcal{B}_{M^*}$ of multiplicity $r^*$. By the main lemma (using $t^*$ as the origin of times) one immediately obtains that $p(t)$ cannot enter or leave $C_{M^*}d(p(t^*))$ for any $t \in [0, \min(T, \tau_0)]$: indeed, in such case it should enter or leave the cylinder through a basis, thus coming from, or arriving to, a block of multiplicity less than $r^*$, which is a contradiction.

From this lemma one trivially obtains the following

**Corollary.** In the same conditions of the geometric lemma, up to the same time, for all $p(0) \in \mathcal{B}_{M}$ one has

$$||p(t) - p(0)|| < \text{diam}(\Lambda_M(p(0))) + d \quad (54)$$

where $\Lambda_M$ is defined by (44) and diam denotes the usual diameter of a set.

**C. Using convexity**

In this subsection we use the convexity of $h(p)$ in order to estimate the diameters of the sets $\Lambda_M(p)$ defined in (44). In fact, we prove here the following

**Lemma on diameter of $\Lambda_M(p)$.** For a given $r$-dimensional module $M$ admitting a $K$-basis, let the zone $Z_M$ and the set (cylinder axis) $\Lambda_M(p)$, $p \in Z_M$, be defined as in (40), (44) with $\omega = \frac{\partial h}{\partial p}$ and a given $\alpha_r$. Assume furthermore $h$ satisfies the property of uniform convexity

$$|A(p)v \cdot v| \geq mv \cdot v \text{ for all } v \in \mathbb{R}^n \text{ and all } p \in Z_M \ , \quad (55)$$

where $A = \left( \frac{\partial^2 h}{\partial p_i \partial p_j} \right)$. Then, for any $p \in Z_M$ one has

$$\text{diam} \Lambda_M(p) \leq d_r \ , \quad (56)$$

where

$$d_r = \frac{2r}{m} K^{r-1} \alpha_r \ . \quad (57)$$
The proof is based on the following

Technical lemma 1. Let \(k^{(1)}, \ldots, k^{(r)}\) be linearly independent vectors of \(\mathbb{Z}^n\), with \(|k^{(j)}| \leq K, (j = 1, \ldots, r)\), and let \(w \in \mathbb{R}^n\) be any linear combination of \(k^{(1)}, \ldots, k^{(r)}\) satisfying

\[
|w \cdot k^{(j)}| \leq \alpha, \quad (j = 1, \ldots, r).
\]

Then one has

\[
\|w\|_e < rK^{r-1}\alpha,
\]

where \(\| \cdots \|_e\) denotes the Euclidean norm in \(\mathbb{R}^n\), \(\|w\|_e = (w \cdot w)^{\frac{1}{2}}\).

Deferring to Appendix B the proof of this lemma, we can now deduce from it the

Proof of the lemma on diameters of \(\Lambda_M\). Take any pair of points \(p', p''\) belonging to \(\Lambda_M(p)\) and so in particular to \(Z_M\); thus one has

\[
\left|\omega(p'') - \omega(p')\right| \cdot k^{(j)} < 2\alpha_r \quad (j = 1, \ldots, r)
\]

for a \(K\)-basis \(k^{(1)}, \ldots, k^{(r)}\) of \(M\). So one has also, by the mean value theorem and the definition of \(A\),

\[
\left|A^*(p'' - p') \cdot k^{(j)}\right| < 2\alpha_r, \quad (j = 1, \ldots, r),
\]

with \(A^* = A(p^*)\) for a suitable \(p^*\) belonging to the segment joining \(p'\) and \(p''\), which by the convexity of \(\mathcal{G}\) is in \(\mathcal{G}\). Denoting by \(w\) the parallel component of \(A^*(p'' - p')\) on \(M\), we have then \(|w \cdot k^{(j)}| < 2\alpha_r, \quad (j = 1, \ldots, r)\), and thus, by Technical lemma 1,

\[
\|w\|_e < 2rK^{r-1}\alpha_r.
\]

On the other hand, by convexity, one has

\[
m(\|p' - p''\|_e)^2 \leq |A^*(p' - p'') \cdot (p' - p'')| = |w \cdot (p' - p'')| \leq \|w\|_e\|p' - p''\|_e
\]

so that, using (60), one obtains

\[
\|p' - p''\| \leq \|p' - p''\|_e \leq \frac{2r}{m}K^{r-1}\alpha_r,
\]

i.e. (57).

Let us add some comments. The estimate of the diameters of the sets \(\Lambda_M(p)\) relies essentially on a geometric property, namely the transversality of \(\mathcal{R}_M\) and \(\Pi_M\). Indeed, in case of transversality, the diameter of \(\Lambda_M(p)\) is essentially, up to a factor, the width of the resonant zone, which is small with \(\alpha\); instead, in case of non transversality, the diameter of \(\Lambda_M(p)\) could even be the whole "length" of such zone, which is not small. In fact, as shown by Nekhoroshev, transversality can be replaced by a much weaker requirement (which of course leads to rougher estimates), namely steepness, according to which the contact between the resonance manifold \(\mathcal{R}_M\) and a plane \(\Pi_M\) be of finite order. In the case of convexity of \(h\) considered here, it is immediate to see that transversality follows, i.e.
A PROOF OF NEKHOROSHEV'S THEOREM

that any (nonvanishing) vector tangent to $\mathcal{R}_M$ does not belong to $\Pi_M$. Indeed, consider any vector $v = \sum a_j k^{(j)}$ of $\Pi_M$. Now, tangency to $\mathcal{R}_M$ requires $u_j \cdot v = 0 \ (j = 1, \ldots, r)$, where $u_j = \partial \omega(p) \cdot k^{(j)} = Ak^{(j)}$, and so also $0 = \sum a_j u_j \cdot v = Av \cdot v$, from which $v = 0$ follows by convexity.

5. Proof of the theorem

In order to prove the theorem, we have now to make a choice for the still free parameters $K, \alpha_0, \alpha_1, \ldots, \alpha_n$ in such a way that, together with the parameters $E, M, m, \epsilon, \rho, \sigma$ characterizing the Hamiltonian, they satisfy the conditions (48) - (52) of the geometric lemma; after this, it will be sufficient to verify that the corollary to the geometric lemma, taking into account the estimates (56) - (57), implies the thesis of the theorem.

Let us first show that the condition (48) of nonoverlapping of resonances is satisfied if one makes for $\alpha_r$ the choice

$$\alpha_r = r! \left(\frac{2M}{m} \right)^r K^{\frac{r}{2}} r^{r-1} \alpha_0 \quad (r = 1, \ldots, n) \quad (61)$$

with

$$\alpha_0 \geq 2 (EM)^{\frac{1}{2}} K \epsilon^\frac{1}{4} \quad (62).$$

Indeed, concerning condition (48), one can remark that, given $p \in B^p_{M,d}$ of multiplicity $r$, by definition of extended block there exists $p' \in B_M$ such that $p \in C_{M,d}(p')$ and so $||p' - p|| < d_r + d$ where $d_r$ is given by (57) and $d$ is given by (47). Then, for any $k \notin \mathcal{M}, |k| \leq K$, one can write

$$|\omega(p) \cdot k| \geq |(|\omega(p') \cdot k| - |(p - \omega(p')) \cdot k|)| \geq |\alpha_{r+1} - M(d_r + d)K| \quad (63),$$

where use has been made of the very definition of $M$ and of $|\omega(p') \cdot k| \geq \alpha_{r+1}$, which follows from the remark b) below the definition of blocks. Thus (48) is satisfied if one assumes

$$\alpha_{r+1} \geq MK(d_r + d) + \alpha_r \quad (64).$$

On the other hand, recalling the expressions (57) of $d_r$ and (47) of $d$, one sees that (62) guarantees $MKd \leq \alpha_0 \leq \alpha_r, \ (r = 1, \ldots, n)$, and so also

$$MK(d_r + d) \leq \alpha_r \left(\frac{2rM}{m} K^{r} + 1 \right) \quad (65).$$

Thus (48) is satisfied if one assumes (62) and defines recursively $\alpha_r$ by

$$\alpha_{r+1} = 2(r + 1) \frac{M}{m} K^r \alpha_r \quad (r = 0, \ldots, n-1),$$

namely by (61). So the statement made above is proven. However, one can notice that (52) requires

$$\alpha_0 \geq \frac{BME^\frac{1}{2}}{\sigma^{n+1}} K^{\frac{n+3}{2}} \epsilon^\frac{1}{4} \quad (66)$$

which is stronger than (62); thus, in order to give a lower bound to the width of the resonant zones, we will take (66) instead of (62).
On the other hand, we will also impose a higher bound to \( \alpha_0 \), because, in virtue of the relation (57) between \( d_r \) and \( \alpha_r \), this will give a higher bound to \( d_r \), and so also to the maximal allowed variation \( \Delta \) of the actions, which enters the statement of the theorem and should be kept as small as possible. To this end, a convenient requirement is that \( d_n \) be of the order of \( K^{-1} \), which can be obtained by imposing

\[
\alpha_0 \leq (EM)^{\frac{1}{2}} K^{-n},
\]

i.e., in virtue of (61),

\[
\alpha_0 \leq \frac{(EM)^{\frac{1}{2}}}{n!} \left( \frac{m}{2M} \right)^{n} K^{\frac{n(n+1)}{2}}.
\]

So \( \alpha_0 \) can be arbitrarily chosen in the interval

\[
\frac{BEM^{\frac{1}{2}}}{\sigma^{n+1}} K^{\frac{n+2}{2}} \epsilon^4 \leq \alpha_0 \leq \frac{(EM)^{\frac{1}{2}}}{n!} \left( \frac{m}{2M} \right)^n K^{\frac{n(n+1)}{2}},
\]

and this requires first of all the consistency condition that the l.h.s. be smaller than the r.h.s., i.e.

\[
\epsilon^4 \leq \frac{\sigma^{n+1}}{B^{\frac{1}{2}} n!} \left( \frac{m}{2M} \right)^n K^{\frac{n^2+2n+2}{2}}.
\]

It is then quite natural to set

\[
K = \epsilon^{-\frac{1}{c}}
\]

with

\[
\epsilon = 4 \left( n^2 + 2n + 2 \right),
\]

so that (68) is satisfied if one assumes

\[
\epsilon \leq \left[ \frac{\sigma^{n+1}}{B^{\frac{1}{2}} n!} \left( \frac{m}{2M} \right)^n \right]^{\frac{1}{8}}.
\]

Moreover, condition (50) and the first of conditions (49) are explicitly assumed in the theorem, while the second of (49), namely \( \rho \leq \frac{E}{\alpha_n} \), with \( \alpha_n \) estimated by (65), is trivially seen to be weaker than the condition \( \rho \leq \left( \frac{E}{M} \right)^{1/2} \) assumed in the theorem. Finally, there only remains to satisfy condition (51), which now reads

\[
\epsilon < \left( \frac{a}{8n} \right)^{4c}.
\]

On the other hand, one immediately checks that both (71) and (72) are satisfied if one assumes

\[
\epsilon < \left( \frac{a}{8n} \right)^{4c} \left( \frac{m}{M} \right)^{8n},
\]

which is the condition occurring in the theorem.

So we have shown that the hypotheses of the theorem imply the hypotheses of the geometric lemma. Thus, for any \( t \in [0, \min(T, \tau_0)] \), where \( T \) is given by (53) and \( \tau_0 \) is the escape time from \( \mathcal{G} - 2d \), one has

\[
\| p(t) - p(0) \| \leq d_r + d,
\]

© Kluwer Academic Publishers • Provided by the NASA Astrophysics Data System
with \( d_r \) given by (57) and \( d \) given by (47). The worst case corresponds to \( r = n \). Recalling the estimate (65) for \( \alpha_n \) one has

\[
d_n \leq \frac{2n}{m} (EM)^{\frac{1}{2}} \epsilon^4,
\]

so that one gets \( \| p(t) - p(0) \| \leq \Delta \) with \( \Delta \) estimated by (16). This is true for any initial datum \( p(0) \) belonging to any block \( B_{\mathcal{M}} \), and thus to the union of all blocks, namely \( \mathcal{G} - 2d \). Now if one further restricts the initial datum to \( \mathcal{G} - 2\Delta \subset \mathcal{G} - 2d - \Delta \), one is also insured that for the escape time \( \tau_0 \) one has \( \tau_0 > T \), so that the theorem is proven.

6. Proof of the analytic lemma

The present section is divided into parts A and B. In part A we state and prove an iterative lemma, which is then used in part B to prove the analytic lemma.

A. The iterative lemma.

Iterative lemma. Consider an Hamiltonian \( H(p, q) = h(p) + f(p, q) \) analytic in a domain \( D = F(G, \rho, \sigma) \) defined as in (6) with \( G \subset \mathbb{R}^n \) compact and \( \rho \) and \( \sigma \) positive. With reference to a modulus \( \mathcal{M} \) and to positive numbers \( K, \alpha \), and with \( \omega = \frac{\partial h}{\partial p} \), assume, everywhere in \( D \),

\[
|\omega(p) \cdot k| \geq \frac{\alpha}{2} \quad \text{for all} \quad k \not\in \mathcal{M}, |k| < K .
\]

Consider the decomposition

\[
H(p, q) = h(p) + Z(p, q) + R(p, q)
\]

(75)

where

\[
Z(p, q) = \sum_{k \in \mathcal{M}} z_k(p)e^{ik \cdot q} ,
\]

(76)

\[
R(p, q) = \sum_{k \not\in \mathcal{M}} r_k(p)e^{ik \cdot q} ,
\]

(77)

with suitable Fourier coefficients \( z_k \) and \( r_k \). Introduce positive numbers \( E, \epsilon, \eta \) by

\[
\| H \|_D \leq E, \quad \| Z + R \|_D \leq \epsilon E, \quad \| R \|_D \leq \eta E
\]

(78)

and assume

\[
\sigma \leq 1, \quad \rho \leq \frac{E}{\alpha} .
\]

(79)

Then, given arbitrary positive numbers \( \delta \) and \( \xi \) satisfying the consistency conditions

\[
\delta < \rho, \quad \xi < \sigma, \quad 2^{3n+3}\xi^{-n-1}\frac{E}{\delta \alpha} \eta < 1 ,
\]

(80)

there exists an analytic canonical transformation \( \psi = (\psi^p, \psi^q): D' \rightarrow D \), where \( D' = F(G', \rho - \delta, \sigma - \xi) \), with \( G' \) compact such that

\[
G \supset G' \supset G - \delta , \quad G \supset \psi^p(G') \supset G - \delta ,
\]

(81)
which is near the identity \( I = (I^p, I^q) \), and in particular satisfies

\[ \| \psi^p - I^p \|_{D'} \leq \frac{\delta}{2} . \]  

In \( D' \) the new Hamiltonian \( H' = H \circ \psi \) takes a form \( H' = h + Z' + R' \) analogous to (75)-(77), and one has the new estimates

\[ \| H' \|_{D'} \leq E, \quad \| Z' + R' \|_{D'} \leq \epsilon' E, \quad \| R' \|_{D'} \leq \eta' E , \]  

with

\[ \epsilon' = \epsilon + \frac{1}{2} \eta' \]  

\[ \eta' = \eta'_1 + \eta'_2 \]  

\[ \eta'_1 = n^{n-1} 2^{n+3} \xi^{-n} e^{-\frac{\xi}{2} \xi} \eta , \]  

\[ \eta'_2 = n(n+1) 2^{6n+10} \xi^{-2(n+1)} \left( \frac{E}{\delta \alpha} \right)^2 \eta . \]

In order to prove the iterative lemma, let us first recall the Lie method for defining canonical transformations. In such a method, a canonical transformation is defined as the “time-one map” of an Hamiltonian flow with a given auxiliary Hamiltonian (generating function) \( \chi \). So, given an analytic function \( \chi \) in \( D = F(G, \rho, \sigma) \), one denotes by \( \psi_t \) the solution at time \( t \) of the auxiliary Hamiltonian system with initial data in a suitable domain \( D' \) such that \( \psi_t(p, q) \in D \) for \( t \in [0,1] \). The canonical transformation \( \psi = \psi_1 : D' \to D \) is then well defined, and any function \( g \) analytic in \( D \) is transformed into a function \( g' = g \circ \psi \) analytic in \( D' \). Thus, recalling the well known property \( \frac{d}{dt}(g \circ \psi_t) = \{ \chi, g \} \circ \psi_t \), one immediately finds

\[ g' = g + \{ \chi, g \} + \frac{1}{2} \{ \chi, \{ \chi, g \} \} \circ \psi , \]  

with a suitable \( t' \in (0,1) \). The relevant estimates are contained in the following lemma

**Lemma on canonical transformations.** Let \( \chi \) be analytic in \( D = F(G, \rho, \sigma) \), where \( G \subset \mathbb{R}^n \) is compact and \( \rho \) and \( \sigma \) are positive parameters; for given positive numbers \( \delta < \rho \) and \( \xi < \sigma \), assume

\[ \left\| \frac{\partial \chi}{\partial p} \right\|_{D - \left( \frac{\delta}{2}, \frac{\xi}{2} \right)} \leq \frac{\xi}{2} , \quad \left\| \frac{\partial \chi}{\partial q} \right\|_{D - \left( \frac{\delta}{2}, \frac{\xi}{2} \right)} \leq \frac{\delta}{2} . \]  

Then the time-one map \( \psi = (\psi^p, \psi^q) \) of the Hamiltonian flow with Hamiltonian \( \chi \) is defined in a suitable domain \( D' = F(G', \rho - \delta, \sigma - \xi) \subset D \), with

\[ G - \delta \subset G' \subset G - \frac{\delta}{2} , \quad G - \delta \subset \psi^p(G') \subset G - \frac{\delta}{2} . \]
A PROOF OF NEKHOROSEV'S THEOREM

The mapping $\psi$ is such that

$$
\|\psi^p - I^p\|_{D'} \leq \frac{\delta}{2}, \quad \|\psi^q - I^p\|_{D'} \leq \frac{\xi}{2},
$$

(91)

and moreover for any function $g$ analytic in $D$ the transformed function $g' = g \circ \psi$ is analytic in $D'$, and satisfies

$$
\|g'\|_{D'} \leq \|g\|_D
$$

(92)

$$
\|g' - g - \{x, g\}\|_{D'} \leq \frac{1}{2} \|\{x, \{x, g\}\}\|_D.
$$

(93)

Proof. Define

$$
\tilde{D} = \{(p, q) \in D''; \psi_t(p, q) \in D'' \text{ for } t \in [0, 1]\},
$$

(94)

where $\psi_t$ is the auxiliary Hamiltonian flow and $D'' = D - \left(\frac{\delta}{2}, \frac{\xi}{2}\right)$. Then, in virtue of (89), if the initial datum $(p, q)$ for $\psi_t$ is chosen in $D'' - \left(\frac{\delta}{2}, \frac{\xi}{2}\right) = D - (\delta, \xi)$, then $\psi_t(p, q)$ certainly remains in $D''$ for $t \in [0, 1]$, and one obtains (91), as well as $D - (\delta, \xi) \subset \tilde{D} \subset D''$; from this, the first of the inclusion relations (90) follows. Moreover, we notice that, from the very definition (94) of $\tilde{D}$, one has

$$
\psi \left(\tilde{D}\right) = \{(p, q) \in D''; \psi_t^{-1}(p, q) \in D'' \text{ for } t \in [0, 1]\}
$$

$$
= \{(p, q) \in D''; \psi_t(p, q) \in D'' \text{ for } t \in [-1, 0]\}
$$

from which in an analogous way the second inclusion relation also follows. Finally, (92) is trivial, while (93) is a consequence of relation (88) recalled above.

To prove the iterative lemma, we also need a few elementary estimates on Fourier series, which are contained in the following

Technical Lemma 2. Let $g(p, q)$ be analytic in $D = F(G, \rho, \sigma)$, where $G \subset \mathbb{R}^n$ is compact and $\rho$ and $\sigma$ are positive parameters. Define Fourier coefficients $g_k(p)$ by $g(p, q) = \sum_{k \in \mathbb{Z}^n} g_k(p) e^{ik \cdot q}$ and define the "ultraviolet part" $g^{>K}$ of $g$ by

$$
g^{>K}(p, q) = \sum_{k \in \mathbb{Z}^n, |k| \geq K} g_k(p) e^{ik \cdot q}.
$$

(95)

Then one has

$$
\|g_k\|_D \leq e^{-|k| \sigma} \|g\|_D,
$$

(96)

and, for any positive $\xi < \sigma$,

$$
\|g^{>K}\|_{D-(0, \xi)} \leq n^{n-1} 2^{n+2} \xi^{-n} e^{-\frac{1}{2} \xi K} \|g\|_D.
$$

(97)
Furthermore, if one defines \( \chi(p,q) \) in \( D = F(G,\rho,\sigma) \) by
\[
\chi(p,q) = \sum_{|k| \leq K} x_k(p) e^{ik\cdot q},
\]
where the coefficients \( x_k(p) \) satisfy
\[
\| x_k \|_D \leq C e^{-|k|\sigma}, \tag{98}
\]
with a given positive \( C \), then one has
\[
\| \chi \|_{D-(0,\frac{1}{2})} \leq C 2^{3n} \xi^{-n}, \tag{99}
\]
\[
\left\| \frac{\partial \chi}{\partial p} \right\|_{D-(\frac{1}{2},\frac{1}{2})} \leq 2\delta^{-1} \| \chi \|_{(0,\frac{1}{2})}, \tag{100}
\]
\[
\left\| \frac{\partial \chi}{\partial q} \right\|_{D-(0,\frac{1}{2})} \leq 2\xi^{-1} \| \chi \|_{D-(0,\frac{1}{2})}, \tag{101}
\]
and also, for any function \( g \) analytic in \( D \),
\[
\| \{ \chi, g \} \|_{D-(\frac{1}{2},\frac{1}{2})} \leq n 2^{3} \delta^{-1} \xi^{-1} \| g \|_D \| \chi \|_{D-(0,\frac{1}{2})}, \tag{102}
\]
\[
\frac{1}{2} \| \{ \chi, g \} \|_{D-(\frac{1}{2},\frac{1}{2})} \leq 2^{n} (4n+3) \delta^{-2} \xi^{-2} \| g \|_D \| \chi \|_{D-(0,\frac{1}{2})}^2. \tag{103}
\]

The proof of the lemma is deferred to Appendix C. We can now come to the

**Proof of the iterative lemma.** First of all, we have to look for a suitable canonical transformation \( \psi \), and so for a suitable generating function \( \chi \). The Hamiltonian \( H = h + Z + R \) will then be transformed to \( H' = H \circ \psi \), and a suitable decomposition for \( H' \), suggested by (88), is
\[
H' = h + Z + R^{\leq K} + \{ \chi, h \} + \tilde{R}, \tag{104}
\]
where
\[
\tilde{R} = R^{> K} + \{ \chi, Z + R \} + H' - H - \{ \chi, H \}, \tag{105}
\]
and the ultraviolet part \( R^{> K} \) is defined as in (95), while \( R^{\leq K} = R - R^{> K} \). One can then choose \( \chi \) in such a way that \( R^{\leq K} + \{ \chi, h \} = 0 \), or
\[
\omega(p) \cdot \frac{\partial \chi}{\partial q} = R^{\leq K}, \tag{106}
\]
while the remainder \( \tilde{R} \) will be estimated using (93) and Technical lemma 2. If \( x_k \) and \( r_k \) denote the Fourier coefficients of \( \chi \) and \( R^{\leq K} \) respectively, then equation (106) is solved by
\[
x_k = -i \frac{r_k}{\omega \cdot k} \quad \text{for} \quad |k| \leq K, \quad x_k = 0 \quad \text{otherwise.}
\]
Now, using (96) and (74) to estimate \( r_k \) and \( \omega \cdot k \) respectively, one gets the estimate (98), with \( C = 2\alpha^{-1} \| R \|_D \leq 2\alpha^{-1} \eta E \), and so, by (99),
\[
\| \chi \|_{D-(0,\frac{1}{2})} \leq 2^{3n+1} \alpha^{-1} \xi^{-n} \eta E. \tag{107}
\]
By inserting this expression into (100) and (101), taking into account hypothesis (80) of the iterative lemma, we are then sure that the hypothesis (89) of the lemma on canonical
transformations is satisfied, so that, as claimed, a canonical transformation \( \psi : D' \to D \) is properly defined, and the inclusion property (81), together with the estimate (82), immediately follow.

Concerning estimates (83), they are quite easily obtained from the decomposition (104), by means of Technical lemma 2. Indeed, the first of them trivially coincides with (92), with \( g = H \). Then, from (97), with \( g = R \), one finds

\[
\| R^{>K} \|_{D'} \leq n^{n-1}2^{n+2}\xi^{-n}\eta^{2\xi K}E = \frac{1}{2}\eta'^{1}E ,
\]

while from (102) and (103), with \( g = Z + R \) and \( g = H \) respectively, taking into account (93) one finds

\[
\|\{x,Z+R\}\|_{D'} + \|H'-H-\{x,H\}\|_{D'} \\
\leq n2^{3n+4}\alpha^{-1}\delta^{-1}\xi^{-n-1}\eta E^{2} + n(4n+3)2^{6n+6}\alpha^{-2}\delta^{-2}\xi^{-2n-2}\eta^{2}E^{3} \\
< n(n+1)2^{6n+9}\left(\frac{E}{\delta}\right)^{2}\xi^{-2n-2}\eta E \\
= \frac{1}{2}\eta'^{1}E ,
\]

where \( \frac{E}{\delta}\alpha > 1 \), \( \xi < 1 \) and \( \eta \leq 2\epsilon \) were used in the intermediate step.

To accomplish the proof of the lemma, we must now properly identify \( Z' \) and \( R' \). Clearly, \( Z' \) is to be taken as the sum of \( Z \) and of those Fourier harmonics of the remainder \( \tilde{R} \) which belong to the modulus \( M \), that we can denote by \( \tilde{R}_{M} \), while \( R' \) is the sum of the remaining harmonics, that is \( R' = \tilde{R} - \tilde{R}_{M} \). As \( \|\tilde{R}\|_{D'} \leq \frac{1}{2}(\eta'_1 + \eta'_2)E = \frac{1}{2}\eta' E \), recalling that \( \tilde{R}_{M} \) and \( Z \) can be expressed as averages of \( \tilde{R} \) and \( Z + R \) respectively over some combinations of angles, one easily finds

\[
\| Z' + R' \|_{D'} = \| Z + \tilde{R} \|_{D} \leq \| Z \|_{D} + \| \tilde{R} \|_{D} \leq (\epsilon + \frac{1}{2}\eta')E \\
\| R' \|_{D'} = \| \tilde{R} - \tilde{R}_{M} \|_{D} \leq \eta' E ,
\]

as claimed. The iterative lemma is thus proven.

\textbf{B. Proof of the analytic lemma}

In the hypotheses of the analytic lemma we are given an Hamiltonian \( H = h + f \), defined in the domain \( D = F(G,\rho,\sigma) \), and characterized by the parameters \( E,\epsilon \) and \( \alpha \). Having in mind to apply recursively the iterative lemma, we denote \( H_{0} = H, f_{0} = f, D_{0} = D, G_{0} = G, \rho_{0} = \rho, \sigma_{0} = \sigma, \epsilon_{0} = \epsilon \) and introduce the decomposition \( f_{0} = Z_{0} + R_{0} \), according to (75) -(77). Notice that \( E \) is considered as a fixed parameter; \( h \) too will remain unchanged, as well as the quantities entering condition (74), i.e. \( \alpha, K \) and \( M \). Denoting also, as is natural, \( \eta_{0} = \|R_{0}\| \), we immediately obtain \( \eta_{0} \leq 2\epsilon_{0} \). Indeed, due to the fact that \( M \) is a modulus, \( Z_{0} \) can be expressed as an average of \( f_{0} \) over a suitable
combinations of angles, so that one obtains \( \| Z_0 \|_{D_0} \leq \| f_0 \|_{D_0} \leq \varepsilon_0 E \), and consequently \( \| R_0 \|_{D_0} = \| f_0 - Z_0 \|_{D_0} \leq 2\varepsilon_0 E \).

We want to show now that the iterative lemma can be applied \( L \) times, with

\[
L = \text{integer part of } (K^\frac{1}{4} - 1) ,
\]

thus giving, for any \( 0 \leq j \leq L \), a Hamiltonian \( H_j = h + Z_j + R_j \), defined in a suitable domain \( D_j = F(G_j, \rho_j, \sigma_j) \), with \( \| Z_j + R_j \|_{D_j} \leq \varepsilon_j E \) and \( \| R_j \|_{D_j} \leq \eta_j E \); moreover, one will obtain the inequalities

\[
\eta_j \leq 2\varepsilon_0 e^{-j} , \quad (j = 0, 1, \ldots, L) .
\]

To prove this statement, we preliminarily choose the arbitrary sequences \( \delta_j, \xi_j \) \( (j = 0, 1, \ldots, l) \) by

\[
\delta_j = \frac{\delta}{2} \frac{1}{(j+1)^2} , \quad \xi_j = \frac{\sigma}{4} \frac{1}{(j+1)^2} , \quad (j = 1, \ldots, L) ,
\]

where \( \delta < \rho \) is the quantity appearing in the statement of the analytic lemma; due to \( \sum_{0}^{\infty} \frac{1}{(j+1)^2} = \frac{\pi^2}{6} < 2 \), one immediately finds \( \sum_{0}^{L} \delta_j < \delta < \rho \) and \( \sum_{0}^{L} \xi_j < \frac{\sigma}{2} \). The proof is made now recursively. Precisely, let us assume we already could use \( l \) times the iterative lemma, \( 0 \leq l \leq L - 1 \), property (109) being satisfied up to \( j = l \). For \( l = 0 \) this assumption is true, as we do not apply at all the lemma, while condition (109) was already shown to be satisfied. We have to show that the \( (l+1) \)-th iteration of the lemma is possible, that is, that the consistency conditions (80) are satisfied by \( \delta_l, \rho_l, \xi_l, \sigma_l, \eta_l \), and that the new quantity \( \eta_{l+1} \) satisfies (109). The first two consistency conditions are trivially satisfied by the choice (110) of the sequences \( \delta_j, \xi_j \). The last one, using (109) and (110), takes the form

\[
2^{5n+7} \sigma^{-(n+1)} \frac{E}{\delta \alpha} (l+1)^{2n+4} e^{-l} \varepsilon_0 < 1 .
\]

Now, the maximum value of the left hand side is taken for \( l + 1 = 2n + 4 \), as can be easily checked; thus, it is enough to require

\[
B' \sigma^{-(n+1)} \frac{E}{\delta \alpha} \varepsilon_0 < 1
\]

with

\[
B' = 2^{5n+7} (2n + 4)^{2n+4} e^{-(2n+3)} .
\]

As \( B' < B \), with \( B \) given by (25), we see that, by (24), the last of the consistency conditions is also satisfied, so that the \( (l+1) \)-th iteration of the iterative lemma can be performed.

There remains to prove that \( \eta_{l+1} \) satisfies

\[
\eta_{l+1} \leq 2\varepsilon_0 e^{-(l+1)} .
\]

In fact we show that each of the two terms defining \( \eta_{l+1} \) as in (85)-(87) is less than \( \varepsilon_0 e^{-(l+1)} \). Considering the first term, one has to guarantee

\[
n^{-1} 2^{n+3} \xi_l^{n} e^{-\frac{c}{2} \xi_l^K} 2\varepsilon_0 e^{-l} < \varepsilon_0 e^{-(l+1)} ,
\]

i.e., by (110),

\[
2^{3n+4} \varepsilon n^{-1} \alpha^{-(n)} (l+1)^{2n} e^{-\frac{c}{2} \xi_l^K} < 1 .
\]

© Kluwer Academic Publishers • Provided by the NASA Astrophysics Data System
A PROOF OF NEKHOROSHEV'S THEOREM

Restricting $l$ by $(l + 1) \leq K^{\frac{1}{4}}$, one has

$$2^{3n+4}e^{n-1}\sigma^{-n}K^{\frac{3}{2}}e^{-\frac{n}{2}K^{\frac{1}{2}}} < 1,$$

or equivalently

$$\frac{\sigma}{8}K^{\frac{1}{2}} - \frac{1}{2}n \log K > \log (2^{3n+4}e^{n-1}\sigma^{-n}) . \tag{112}$$

If one assumes

$$\frac{\sigma}{8}K^{\frac{1}{2}} > n \log K , \tag{113}$$

one easily checks that the inequality $\log (2^{3n+4}e^{n-1}\sigma^{-n}) < \frac{1}{2}n \log K$ also holds (for $n > 1$), so that (112) is satisfied. On the other hand, one immediately sees that (113) is weaker than (23), which was assumed in place of (113) only because it is more manageable.

Considering now the second term defining $\eta_{l+1}$, one has to show

$$n(n + 1)2^{6n+10}\xi^{-2(n+1)}\left(\frac{E}{\delta \alpha}\right)^2 \epsilon_1 \epsilon_0 e^{-l} < \epsilon_0 e^{-(l+1)} ,$$

i.e.

$$n(n + 1)e^{2^{10n+17}\sigma^{-2(n+1)}\left(\frac{E}{\delta \alpha}\right)^2 (l + 1)^4 e^{8\epsilon_1} < 1 ,$$

which requires to previously have a control on $\epsilon_l$. In this connection, from (84), (85) and (86) one sees that, having assigned $\epsilon_0$ and the sequence $\eta_1, \ldots, \eta_l$, the sequence $\epsilon_1, \ldots, \epsilon_l$ is also given, with $\epsilon_i = \epsilon_0 + \frac{1}{2} \sum_{j=1}^{i} \eta_j$, $i = 1, \ldots, l$, so that in particular from (109) one gets $\epsilon_l < \epsilon_0 + \sum_{j=1}^{\infty} \epsilon_0 e^{-j} < 2\epsilon_0$. By inserting $(l + 1) \leq K^{\frac{1}{4}}$ one thus gets the condition

$$B''\sigma^{-2n-2}\left(\frac{E}{\delta \alpha}\right)^2 K^{n+2}\epsilon_0 < 1$$

with

$$B'' = 2^{10n+18}n(n + 1)e < B ,$$

where $B$ given by (25). So this condition is satisfied by (24) of the analytic lemma.

The proof of the analytic lemma is now immediately accomplished. Indeed, identifying as it is obvious $H', Z', R', D', G'$ with the corresponding quantities $H_L, Z_L, R_L, D_L, G_L$, and putting $\psi = \psi_{L-1} \circ \cdots \circ \psi_0$, we see that (26) is an immediate consequence of the inclusion properties (81), while (27) directly follows from (82), because one has $\|\psi^p - I^p\|_{D'} \leq \frac{1}{2} \sum_{j=0}^{L} \delta_j < \frac{\epsilon}{2}$. On the other hand, the Hamiltonian $H'$ has the correct decomposition, and the remainder $R'$ satisfies

$$\|R'\|_{D'} \leq \eta_L E < 2\epsilon_0 E e^{K^{\frac{1}{2}}-2} < 16\epsilon E e^{-K^{\frac{1}{4}}} ,$$

as claimed. The analytic lemma is proven.
Appendices

A. Cauchy estimates

We recall here the well known Cauchy estimates for the derivatives of analytic functions, in the way they are used in the present paper. For a function $g$ analytic in $D = F(G, \rho, \sigma)$ with $G$ compact and $\rho, \sigma$ positive and $F$ defined as in (6), one has

$$
\frac{\|\frac{\partial g}{\partial p}\|_{D-(\delta,0)}}{\delta} \leq \frac{\|g\|_D}{\delta}, \quad \frac{\|\frac{\partial g}{\partial q}\|_{D-(0,\xi)}}{\xi} \leq \frac{\|g\|_D}{\xi}, \quad (A1)
$$

$$
\frac{\|\frac{\partial^2 g}{\partial p_i \partial p_j}\|_{D-(\delta,0)}}{\delta^2} \leq \begin{cases} 
\frac{\|g\|_D}{\delta^2} & \text{if } i \neq j \\
2\frac{\|g\|_D}{\delta} & \text{if } i = j 
\end{cases}, \quad \frac{\|\frac{\partial^2 g}{\partial p_i \partial q_j}\|_{D-(0,\xi)}}{\xi \delta} \leq \frac{\|g\|_D}{\xi \delta}, \quad (A2)
$$

where $\frac{\partial g}{\partial p} = \left(\frac{\partial g}{\partial p_1}, \ldots, \frac{\partial g}{\partial p_n}\right)$, and $\frac{\partial g}{\partial q} = \left(\frac{\partial g}{\partial q_1}, \ldots, \frac{\partial g}{\partial q_n}\right)$.

B. Proof of technical lemma 1

We will here in fact prove the following more general proposition, which includes technical lemma 1 as a particular case.

**Proposition.** Let $u^{(1)}, \ldots, u^{(r)}$ be $r$ linearly independent vectors of $\mathbb{R}^n$, with $|u^{(j)}| = |u_1^{(j)}| + \cdots + |u_n^{(j)}| \leq K$, ($j = 1, \ldots, r$), and let $w \in \mathbb{R}^n$ be any linear combination of $u^{(1)}, \ldots, u^{(r)}$ satisfying

$$
|w \cdot u^{(j)}| < \alpha, \quad (j = 1, \ldots, r). \quad (B1)
$$

Then one has

$$
\|w\|_e \mathcal{V} (u^{(1)}, \ldots, u^{(r)}) < rK^{r-1}\alpha, \quad (B2)
$$

where $\mathcal{V} (u^{(1)}, \ldots, u^{(r)})$ denotes the $r$-dimensional Euclidean volume of the parallelepiped generated by the $r$ vectors $u^{(1)}, \ldots, u^{(r)}$, and $\|w\|_e$ is the Euclidean norm of $w$.

This statement clearly implies technical lemma 1: indeed, if $u^{(1)}, \ldots, u^{(r)}$ have integer components, then $\mathcal{V} (u^{(1)}, \ldots, u^{(r)})$ is integer too, so that, the vectors $u^{(1)}, \ldots, u^{(r)}$ being independent, one has $\mathcal{V} (u^{(1)}, \ldots, u^{(r)}) \geq 1$, and the lemma is proven.

**Proof of the proposition.** For $r = 1$ (B2) is trivially verified. We then prove (B2) for any $r$, $2 \leq r \leq n$, assuming it is satisfied for $r - 1$. Let $\mathcal{E}$ be the linear space spanned by $u^{(1)}, \ldots, u^{(r)}$, $\mathcal{E}'$ the subspace of $\mathcal{E}$ spanned by $u^{(1)}, \ldots, u^{(r-1)}$ and $\mathcal{E}''$ the (one-dimensional) subspace of $\mathcal{E}$ orthogonal to $\mathcal{E}'$. Introduce the decomposition

$$
w = w' + w'' \quad u^{(r)} = u' + u'', \quad (B3)
$$
A PROOF OF NEKHOROSHEV'S THEOREM

with \( w', u' \in \mathcal{E}' \) and \( w'', u'' \in \mathcal{E}'' \). Clearly, as \( u^{(j)} \in \mathcal{E}' \) for \( j = 1, \ldots, r \), one has \( |w' \cdot u^{(j)}| = |w \cdot u^{(j)}| < \alpha \) for \( j = 1, \ldots, r \), and thus, by the inductive hypothesis, also

\[
\|w'\|_e \mathcal{V}_{r-1} \left( u^{(1)}, \ldots, u^{(r-1)} \right) < (r-1)K^{r-2}\alpha ,
\]

so that one immediately obtains

\[
\|w'\|_e \mathcal{V}_r \left( u^{(1)}, \ldots, u^{(r)} \right) = \|w'\|_e \mathcal{V}_{r-1} \left( u^{(1)}, \ldots, u^{(r-1)} \right) \|u''\|_e \\
< (r-1)K^{r-2}\alpha \|u''\|_e .
\]

On the other hand, from \( \|w''\|_e \|u''\|_e = |w'' \cdot u''| = |w \cdot u^{(r)} - w' \cdot u'| < \alpha + \|w'\|_e \|u''\|_e \), one also obtains

\[
\|w''\|_e \mathcal{V}_r \left( u^{(1)}, \ldots, u^{(r)} \right) = \|w''\|_e \|u''\|_e \mathcal{V}_{r-1} \left( u^{(1)}, \ldots, u^{(r-1)} \right) \\
< (K + (r-1)\|u''\|_e)K^{r-2}\alpha ,
\]

where (B4) and the inequality

\[ \mathcal{V}_{r-1} \left( u^{(1)}, \ldots, u^{(r-1)} \right) \leq \|u^{(1)}\|_e \ldots \|u^{(r-1)}\|_e \leq \|u^{(1)}\| \ldots \|u^{(r-1)}\| \leq K^{r-1} \]

were used. From (B5) and (B6) one can then conclude

\[
\|w\|_e \mathcal{V}_r \left( u^{(1)}, \ldots, u^{(r)} \right) = \left( \|w'\|_e^2 + \|w''\|_e^2 \right)^{\frac{1}{2}} \mathcal{V}_r \left( u^{(1)}, \ldots, u^{(r)} \right) \\
< \left[ (r-1)^2 \|u''\|_e^2 + (K + (r-1)\|u''\|_e)^2 \right]^{\frac{1}{2}} K^{r-2}\alpha \\
= \left[ (r-1)^2 \|u\|_e^2 + K^2 + 2K(r-1)\|u''\|_e \right]^{\frac{1}{2}} K^{r-2}\alpha \\
\leq \left[ (r-1)^2 K^2 + K^2 + 2K^2(r-1) \right]^{\frac{1}{2}} K^{r-2}\alpha = rK^{r-1}\alpha ,
\]

as claimed.

C. Proof of technical lemma 2

Concerning inequality (96), it expresses a well known property of analytic functions (actually, it is an elementary application of Cauchy's inequalities); for the proof, see for example ref.2 or the appendix of ref.4. On the basis of this inequality, (97) is then easily deduced. Indeed, from the definition (95) one gets \( \|g^{>K}\|_{D-(0,\xi)} \leq \|g\|_D \sum |k| > K \delta^{-|k|}\xi \). On the other hand, the number of vectors \( k \in \mathbb{Z}^n \) with \( |k| = l \) can be overestimated by \( 2^n(l+1)^{n-1} \), while, for any \( r, s, \lambda > 0 \) one has the inequality

\[
r^s \leq \left( \frac{s}{e\lambda} \right)^{\frac{s}{\lambda}} e^{r\lambda} ;
\]

indeed, such inequality is equivalent to \( r^s \leq e^{r^s - 1} \), which is evidently true. So one has
(with \(l + 1, n - 1\) and \(\frac{\xi}{4}\) in place of \(r, s\) and \(\lambda\))

\[
\sum_{|k| > K} e^{-|k| \xi} \leq 2^n \sum_{l=K+1}^{\infty} (l + 1)^{n-1} e^{-l \xi}
\leq 2^n \left( \frac{2(n-1)}{\xi} \right)^{\frac{n-1}{2}} \sum_{l=K+1}^{\infty} e^{-\frac{1}{2} l \xi}
\leq 2^n n^{n-1} \xi^{-(n-1)} \frac{e^{-\frac{1}{2} K \xi}}{1 - e^{-\frac{1}{2} \xi}}
\leq 2^{n+2} n^{n-1} \xi^{-n} e^{-\frac{1}{2} K \xi},
\]

where use was made of the trivial inequality \(1 - e^{-\frac{1}{2} \xi} \geq \frac{\xi}{4}\) for \(0 \leq \xi \leq 1\). Inequality (97) then immediately follows. In an analogous way (with \(l\) and \(\frac{\xi}{4}\) in place of \(r\) and \(\lambda\) in (C1)), one also proves the inequality

\[
\sum_{l=1}^{\infty} l^s e^{-\frac{1}{2} l \xi} < 2^{s+2} s^{s} \xi^{-s-1},
\]

which will be used below.

We come now to the proof of (99)-(101). For (99), recalling the definition \(|k| = |k_1| + \cdots + |k_n|\), one has

\[
\|x\|_{D-((0, \frac{\xi}{2}))} < C \sum_{k \in \mathbb{Z}^n} e^{-|k| \frac{\xi}{2}} = C \left( \sum_{l=|l|}^{\infty} e^{-\frac{1}{2} |l| \xi} \right)^n
\leq C \left( \frac{2}{1 - e^{-\frac{1}{2} \xi}} \right)^n \leq C 2^{3n} \xi^{-n},
\]

while (100) is nothing but Cauchy's inequality applied to (99). Instead, concerning (101), for any \(j = 1, \ldots, n\) one can write

\[
\left\| \frac{\partial x}{\partial q_j} \right\|_{D-((0, \frac{\xi}{2}))} < C \sum_{k \in \mathbb{Z}^n} |k_j| e^{-\frac{1}{2} |k| \xi}
\leq C \left( \sum_{l=|l|}^{\infty} e^{-\frac{1}{2} |l| \xi} \right)^{n-1} \left( \sum_{l=|l|}^{\infty} |l| e^{-\frac{1}{2} |l| \xi} \right)
\leq C \left( 2^{3(n-1)} \xi^{-(n-1)} \right) \left( 2^{4} \xi^{-2} \right)
\leq C 2^{3n+1} \xi^{-n-1}
\]

where (C2) was used.

The estimates (100)-(101) for the first derivatives of \(x\) already allow one to deduce inequality (102), by a trivial counting of terms, using Cauchy's inequality to estimate the
derivatives of $g$. Finally, to obtain (103) one needs estimates for the second derivatives of $\chi$. From Cauchy's inequality applied to (99) and to (101) one gets

$$\left\| \frac{\partial^2 \chi}{\partial p_i \partial p_j} \right\|_{D^-\left(\frac{\xi}{2},\frac{\xi}{2}\right)} \leq \begin{cases} 4 \delta^{-2} \| \chi \|_{D^-\left(0,\frac{\xi}{2}\right)} & \text{if } i \neq j \\ 8 \delta^{-2} \| \chi \|_{D^-\left(0,\frac{\xi}{2}\right)} & \text{if } i = j \end{cases}$$

(C5)

$$\left\| \frac{\partial^2 \chi}{\partial p_i \partial q_j} \right\|_{D^-\left(\frac{\xi}{2},\frac{\xi}{2}\right)} \leq 4 \delta^{-1} \xi^{-1} \| \chi \|_{D^-\left(0,\frac{\xi}{2}\right)}$$

(C6)

while, proceeding as for (C4), one obtains

$$\left\| \frac{\partial^2 \chi}{\partial q_i \partial q_j} \right\|_{D^-\left(0,\frac{\xi}{2}\right)} \leq \begin{cases} 4 \xi^{-2} \| \chi \|_{D^-\left(0,\frac{\xi}{2}\right)} & \text{if } i \neq j \\ 16 \xi^{-2} \| \chi \|_{D^-\left(0,\frac{\xi}{2}\right)} & \text{if } i = j \end{cases}$$

(C7)

Then, estimating the second derivatives of $g$ by Cauchy's inequality, by a careful counting of the number of terms in the double Poisson bracket, inequality (103) follows.

References


University of Tokyo Press.

359-426, Birkhauser, Boston.

201-223.


