A Proof of Kolmogorov's Theorem on Invariant Tori Using Canonical Transformations Defined by the Lie Method.

G. BENETTIN

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

L. GALGANI

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

A. GIORGILLI

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

J.-M. STRELCYN

Department de Mathématiques, Centre Scientifique et Polytéchnique
Université Paris-Nord - 93 Villetaneuse, France

(ricevuto il 7 Luglio 1983)

Summary. — In this paper a proof is given of Kolmogorov's theorem on the existence of invariant tori in nearly integrable Hamiltonian systems. The scheme of proof is that of Kolmogorov, the only difference being in the way canonical transformations near the identity are defined. Precisely, use is made of the Lie method, which avoids any inversion and thus any use of the implicit-function theorem. This technical fact eliminates a spurious ingredient and simplifies the establishment of a central estimate.

PACS. 03.20. – Classical mechanics of discrete systems: general mathematical aspects.

1. – Introduction and formulation of Kolmogorov's theorem.

1`1. – The aim of this paper is to give a proof of Kolmogorov's theorem on the existence of invariant tori in nearly integrable Hamiltonian systems [1.12] with

---

[1] A. N. KOLMOGOROV: *Dokl. Akad. Nauk SSSR*, **98**, 527 (1954) (*Math. Rev.*, **16**, No. 924); English translation in G. CASATI and J. FORD (Editors): *Lecture Notes in*

a mostly pedagogical intent; thus we consider the simplest case, namely that of analytic and nondegenerate Hamiltonians, but give, however, explicitly all relevant estimates. In particular, in dealing whith canonical transformations near the identity we make use of the Lie method which avoids any inversion and thus any reference to the implicit-function theorem. This fact, which is a purely technical one, eliminates a spurious ingredient and simplifies the establishment of a central estimate. Apart from this technical fact, we follow here the original scheme of Kolmogorov, which can be considered to be, especially from a pedagogical point of view, somehow simpler (although possibly less powerful) than the scheme of Arnold. Moreover, the proof of convergence of the iteration scheme is given in a rather simple way which, in particular, makes use only of trivial geometric series. Thus it is hoped that the present version will be useful at least in expanding the familiarity with this important theorem, which still seems to be considered, especially among physicists, as an extremely difficult one.

1'2. – In the theory of perturbations one is concerned with an Hamiltonian

$$H(p, q) = H^0(p) + H^1(p, q) ,$$

Physics, No. 93 (Berlin, 1979), p. 51.

(²) V. I. ARNOLD: *Usp. Mat. Nauk*, **18**, 13 (1963); *Russ. Math. Surv.*, **18**, 9 (1963); *Usp. Mat. Nauk.*, **18**, No. 6, 91; *Russ. Math. Surv.*, **18**, No. 6, 85 (1963).

(³) V. I. ARNOLD and A. AVEZ: *Problèmes ergodiques de la mécanique classique*, (Paris, 1967).

(⁴) J. MOSER: *Proceedings of the International Conference on Functional Analysis and Related Topics* (Tokyo, 1969), p. 60; *Stable and random motions in dynamical systems* (Princeton, 1973).

(⁵) S. STERNBERG: *Celestial Mechanics* (New York, N. Y., 1968), Part. 2.

(⁶) E. A. GREBENNIKOV and JU. A. RIABOV: *New Qualitative Methods in Celestial Mechanics* (Moskow, 1971), in Russian.

(⁷) H. RÜSSMANN: *Nachr. Akad. Wiss. Gott. Math. Phys. Kl.*, 67 (1970); 1, (1972); *Lecture Notes in Physics*, No. 38 (Berlin, 1975), p. 598; in *Selecta Mathematica*, Vol. **5** (Heidelberg, Berlin, 1979); in G. HELLEMAN (Editors): *Ann. N. Y. Acad. Sci.*, **357**, 90 (1980); see also *On the construction of invariant tori of nearly integrable Hamiltonian systems with applications*, notes of lectures given at the Institut de Statistique, Laboratoire de Dynamique Stellaire, Université Pierre et Marie Curie, Paris.

(⁸) R. BARRAR: *Celes. Mech.*, **2**, 494 (1970).

(⁹) E. ZEHNDER: *Commun. Pure Appl. Math.*, **28**, 91 (1975); **29**, 49 (1976); see also *Stability and instability in celestial mechanics*, lectures given at Départment de Physique, Laboratoire de Physique Théorique, Ecole Polytechnique Fédérale, Lausanne.

(¹⁰) A. I. NEICHSTADT: *J. Appl. Math. Mech.*, **45**, 766 (1981).

(¹¹) G. GALLAVOTTI: *Meccanica elementare* (Torino, 1980); *The Elements of Mechanics* (Berlin, 1983); in J. FRÖLICH (Editor): *Progress in Physics* (Boston, 1982); L. CHIERCHIA and G. GALLAVOTTI: *Nuovo Cimento B*, **67**, 277 (1982).

(¹²) J. POESCHEL: *Commun. Pure Appl. Math.*, **35**, 653 (1982); *Celest. Mech.*, **28**, 133 (1982).

where

$$p = (p_1, ..., p_n) \in B \subset \mathbf{R}^n \quad \text{and} \quad q = (q_1, ..., q_n) \in \mathbf{T}^n,$$

B being an open ball of $\mathbf{R}^n$ and $\mathbf{T}^n$ the n-dimensional torus. The variables p and q are called the actions and the angles, respectively. The functions on the torus $\mathbf{T}^n$ are naturally identified with the functions defined on $\mathbf{R}^n$ and of period 2π in $q_1, ..., q_n$. For any $p \in B$ the unperturbed angular frequencies $\omega = (\omega_1, ..., \omega_n) = (\partial H^0/\partial p_1, ..., \partial H^0/\partial p_n) = \partial H^0/\partial p$ are defined. In Kolmogorov's theorem the nondegeneracy condition $\det(\partial \omega_i/\partial p_j) \neq 0$ is assumed; as is well known, the extension to degenerate systems was accomplished by ARNOLD.

In the unperturbed case ($H^1 = 0$), the equations of motion reduce to $\dot{p} = -\partial H^0/\partial q = 0$, $\dot{q} = \partial H^0/\partial p = \omega(p)$ with solutions $p(t) = p^0$, $q(t) = q^0 + \omega(p^0)t \pmod{2\pi}$. Thus the phase space $B \times \mathbf{T}^n$ is foliated into tori $\{p\} \times \mathbf{T}^n$, $p \in B$, each of which is invariant for the corresponding Hamiltonian flow and supports quasi-periodic motions characterized by a frequency $\lambda = \omega(p)$. We recall that λ is said to be nonresonant if there does not exist $k \in \mathbf{Z}^n$, $k \neq 0$, such that $\lambda \cdot k = 0$, where we denote $\lambda \cdot k = \sum_{i=1}^{n} \lambda_i k_i$.

In Kolmogorov's theorem the attention is restricted to those tori which support quasi-periodic motions with appropriate nonresonant frequencies, namely frequencies belonging to the set Ω_γ defined by

(1.1) $$\Omega_\gamma = \{\lambda \in \Omega \subset \mathbf{R}^n, \ |\lambda \cdot k| \geq \gamma |k|^{-n} \ \forall k \in \mathbf{Z}^n, \ k \neq 0\}$$

for a given positive γ, where $\Omega = \omega(B)$ is the image of B by the map ω (Ω is assumed to be bounded, which can always be obtained by possibly reducing the radius of the ball B). Here, as above, $\lambda \cdot k = \sum_{i=1}^{n} \lambda_i k_i$ and, moreover, $|k| = \max_i |k_i|$; more generally, for $v, w \in \mathbf{C}^n$, we will denote $v \cdot w = \sum_{i=1}^{n} v_i w_i$ and $|v| = \max_i |v_i|$; analogously, the norm $|C|$ of a matrix $\{C_{ij}\}$ will be defined as for a n^2-vector $C \in \mathbf{C}^{n^2}$, namely by $|C| = \max_{ij} |C_{ij}|$. It is a simple classical result (see, for example, ref. (^{2})) that, for any fixed γ, the complement of Ω_γ in Ω is dense and open and that its relative Lebesgue measure tends to zero with γ. Thus for almost all frequencies $\lambda \in \Omega$ one can find a positive γ such that $\lambda \in \Omega_\gamma$. Because of the relation $\Omega_{\gamma'} \subset \Omega_\gamma$, if $\gamma < \gamma'$, it will be no restriction to take $\gamma < 1$. One is led to the diophantine condition defining Ω_γ, by the problem of solving a partial differential equation of the form

$$\sum_i \lambda_i \frac{\partial F}{\partial q_i}(q) = G(q),$$

where the unknown function F and the given function G are defined on the torus $\mathbf{T}^n$ and G has vanishing average. An analogous arithmetical condition in a different but strongly related problem was first introduced by SIEGEL ([13]).

Let us fix some more notations. Being interested in the analytic case, we will have to consider complex extensions of subsets of $\mathbf{R}^{2n}$ and of real analytic functions defined there. Since we have fixed $p^* \in B$ and a positive $\varrho < 1$ so small that the real closed ball of radius ϱ centred at p^* is contained in B, a central role will be played by the subsets D_{ϱ,p^*} of $\mathbf{C}^{2n}$ defined by

$$(1.2) \qquad D_{\varrho,p^*} := \{(p, q) \in \mathbf{C}^{2n};\ \|p - p^*\| \leqslant \varrho,\ \|\operatorname{Im} q\| \leqslant \varrho\},$$

where $\operatorname{Im} q = (\operatorname{Im} q_1, ..., \operatorname{Im} q_n)$. For what concerns functions, let $\mathscr{A}_{\varrho,p^*}$ be the set of all complex continuous functions defined on D_{ϱ,p^*}, analytic in the interior of D_{ϱ,p^*}, which have period 2π in $q_1, ..., q_n$ and are real for real values of the variables. If $f \in \mathscr{A}_{\varrho,p^*}$, its norm is then taken to be $|f|_{\varrho,p^*} = \sup\limits_{x \in D_{\varrho,p^*}} |f(x)|$; in the case of functions $f = (f_1, ..., f_n)$ with values in $\mathbf{C}^n$, we also write $f \in \mathscr{A}_{\varrho,p^*}$ if $f_i \in \mathscr{A}_{\varrho,p^*}$ $(i = 1, ..., n)$, and we set $\|f\|_{\varrho,p^*} = \max\limits_i \|f_i\|_{\varrho,p^*}$. In agreement with the convention made above, if C is a $n \times n$ matrix whose elements C_{ij} belong to $\mathscr{A}_{\varrho,p^*}$, we set $\|C\|_{\varrho,p^*} = \max\limits_{i,j} \|C_{ij}\|_{\varrho,p^*}$; thus, in particular, for any $v \in \mathbf{C}^n$ one has the inequality $\|Cv\|_{\varrho,p^*} \leqslant n \|C\|_{\varrho,p^*} \|v\|$. For notational simplicity we also set $D_\varrho = D_{\varrho,0}$, $\mathscr{A}_\varrho = \mathscr{A}_{\varrho,0}$, $|f|_\varrho = |f|_{\varrho,0}$. Averaging over angles will be denoted by a bar: namely, if $f \in \mathscr{A}_{\varrho,p^*}$, then we denote

$$(1.3) \qquad \bar{f}(p^*) = \frac{1}{(2\pi)^n} \int\limits_0^{2\pi} \cdots \int\limits_0^{2\pi} f(p^*, q)\, \mathrm{d}q_1 ... \mathrm{d}q_n.$$

Functions $f(p)$, $f(q)$ of the actions only or of the angles only will be thought of as defined in D_{ϱ,p^*} by trivial extension.

Let us finally remark that, for the given Hamiltonian $H \in \mathscr{A}_{\varrho,p^*}$, one can assume $\|H\|_{\varrho,p^*} < 1$; indeed this can always be obtained by the change of variables $(p, q) \mapsto (xp, q)$ with a suitable positive α, which leaves the equations of motion in Hamiltonian form with a new Hamiltonian $H' = \alpha H$.

1.3. – Kolmogorov's theorem can then be stated in the following way:

Theorem 1. Consider the Hamiltonian $H(p, q) = H^0(p) + H^1(p, q)$ defined in $B \times \mathbf{T}^n$ and, after having fixed $p^* \in B$, denote

$$(1.4) \qquad \lambda = \omega(p^*) = \frac{\partial H^0}{\partial p}(p^*), \qquad C_{ij}^* = \frac{\partial \omega_i}{\partial p_j}(p^*) = \frac{\partial^2 H^0}{\partial p_i \partial p_j}(p^*).$$

([13]) C. L. SIEGEL: *Ann. Math.*, **43**, 607 (1942).

Assume there exist positive numbers γ, ϱ, $d < 1$ such that

(1.5)   i) $\lambda \in \Omega_\gamma$,

(1.6)   ii) $H^0, H^1 \in \mathscr{A}_{\varrho, \nu \bullet}$,

(1.7)   iii) $d\|v\| \leqslant \|C^* v\| \leqslant d^{-1}\|v\|$ for any $v \in \mathbf{C}^n$

and, moreover, $\|H\|_{\varrho, \nu \bullet} \leqslant 1$. Then there exist positive numbers E and ϱ' with $\varrho' < \varrho$ such that, if the norm $\|H^1\|_{\varrho, \nu \bullet}$ of the perturbation H^1 satisfies

(1.8) $\|H^1\|_{\varrho, \nu \bullet} \leqslant E$,

one can construct a canonical analytic change of variables $(p, q) = \psi(P, Q)$, $\psi: D_{\varrho'} \to D_{\varrho, \nu \bullet}$, $\psi \in \mathscr{A}_{\varrho'}$, which brings the Hamiltonian H into the form $H' = H \circ \psi$ given by

(1.9) $H'(P, Q) = (H \circ \psi)(P, Q) = a + \lambda \cdot P + R(P, Q);$

here $a \in \mathbf{R}$ and the remainder $R \in \mathscr{A}_{\varrho'}$ is, as a function of P, of the order $\|P\|^2$. In particular, one can take

(1.10) $E = c_n \gamma^4 d^6 \varrho^{3(n+1)}$,

where

(1.11) $c_n = \left(\dfrac{1}{52}\right)^{8n+7} \left(\dfrac{1}{4n+1}\right)^4 \left(\dfrac{e}{n+1}\right)^{4(n+1)}$.

The change of variables is near the identity, in the sense that

$$\|\psi - \text{identity}\|_{\varrho'} \to 0 , \qquad\qquad \text{as } \|H^1\|_{\varrho, \nu \bullet} \to 0 .$$

1`4. – Let us now add some comments on the interpretation of Kolmogorov's theorem. In classical perturbation theory one attempted at constructing a canonical change of variables with the aim of eliminating the angles Q in the new Hamiltonian H'; in such a way the phase space would turn out to be foliated into invariant tori $P = $ const, which is possible only for the exceptional class of integrable systems. As one sees, in the method of Kolmogorov one looks instead, in correspondence with any good frequency λ (*i.e.* any $\lambda \in \Omega_\gamma$ for some γ), for a related change of variables which eliminates the angles Q only at first order in P. In such a way, one renounces to have the general solution $P(t) = P^0$, $Q(t) = Q^0 + \omega(P^0)t$ (mod 2π) corresponding to all

initial data $(P^0, Q^0) \in D_{\varrho'}$, but one gets, however, for any good frequency λ, the particular solutions $P(t) = P^0$, $Q(t) = Q + \lambda t \pmod{2\pi}$ for all initial data (P^0, Q^0), with $P^0 = 0$, $Q^0 \in T^n$; this is immediately seen by writing down the equations of motion for H', namely $\dot{P} = -\partial R/\partial Q$, $\dot{Q} = \lambda + \partial R/\partial p$ and using the fact that $R(P, Q)$ is of order $\|P\|^2$. Thus with the Hamitonian in the new form H', adapted to the chosen frequency λ, one is not guaranteed to have a foliation into invariant tori, but one just sees by inspection the invariance of one torus supporting quasi-periodic motions with angular frequency λ. In terms of the original variables (p, q) canonically conjugated to (P, Q) by the mapping ψ, this torus is described by the parametric equations $(p, q) = \psi(0, Q)$, $Q \in T^n$, and is invariant for the Hamiltonian flow induced by H. This torus is a small perturbation of the torus $p = p^*$, $q \in T^n$, which is invariant for the Hamiltonian H^0, supporting quasi-periodic motions with the same angular frequency.

In such a way, having fixed a positive $\gamma \leqslant 1$, to any $p^* \in B$ with a good frequency $\lambda = \omega(p^*) \in \Omega_\gamma$ (and also satisfying the two further conditions of theorem 1), one can associate a torus which is invariant for the original Hamiltonian H supporting quasi-periodic motions with frequency λ. Now, as in virtue of the nondegeneracy condition the mapping $\omega : B \to \Omega = \omega(B)$ is a diffeomorphism, one has that the set of points $\{p^* \in B; \omega(p^*) \notin \Omega_\gamma\}$ has a Lebesgue measure which tends to zero as $\gamma \to 0$. This remark suggests that the $2n$-dimensional Lebesgue measure of the set of tori whose existence is guaranteed by Kolmogorov's theorem is positive and that the measure of its complement in $B \times T^n$ tends to zero as the size of the perturbation tends to zero. Actually this fact, which was already stated in the original paper of Kolmogorov, was proved by ARNOLD (see also [10] for a recent improvement and [12]) using a variant of Kolmogorov's method, but will not be proved in the present paper.

2. – Reformulation of the theorem.

2`1. – The proof of the theorem starts with a trivial rearrangement of the Hamiltonian which reduces it to the form actually considered by KOLMOGOROV. Indeed, after a translation of p^* to the origin, by a Taylor expansion in p one can write the Hamiltonian H in the form

$$(2.1) \quad H(p, q) = a + A(q) + [\lambda + B(q)] \cdot p + \tfrac{1}{2} \sum_{i,j} C_{ij}(q) p_i p_j + R(p, q),$$

where $a \in R$ is a constant which is uniquely defined by the condition $\bar{A} = 0$, while A, B_i, C_{ij}, $R \in \mathscr{A}_\varrho$ and R is, as a function of p, of the order $\|p\|^3$.

One has clearly

$$(2.2) \quad \begin{cases} a = H^0(0) + \bar{H}^1(0) = \bar{H}(0) , \\[2mm] A(q) = H^1(0, q) - \bar{H}^1(0) = H(0, q) - a , \\[2mm] B_i(q) = \dfrac{\partial H^1}{\partial p_i}(0, q) = \dfrac{\partial H}{\partial P_i}(0, q) - \lambda_i , \\[2mm] C_{ij}(q) = \dfrac{\partial^2 H}{\partial p_i \, \partial p_j}(0, q) . \end{cases}$$

This Hamiltonian H has then already the wanted form (1.9) of theorem 1, apart from the disturbing term $A(q) + B(q) \cdot p$, which thus constitutes the actual perturbation in the problem at hand and will be eliminated in the following by a sequence of canonical transformations.

2'2. – One is thus interested in giving estimates on A, $B = \{B_i\}$ and $C = \{C_{ij}\}$. To this end we will make use of the familiar Cauchy's inequality, which will also be used repeatedly in the future sections. This inequality will be used in our framework in two forms: given $f \in \mathscr{A}_\varrho$, a positive $\delta < \varrho$ and nonnegative integers k_i, l_i $(i = 1, ..., n)$, then one has

$$(2.3) \quad \left| \frac{\partial^{k_1 + \dots + k_n + l_1 + \dots + l_n}}{\partial p_1^{k_1} \dots \partial p_n^{k_n} \partial q_1^{l_1} \dots \partial q_n^{l_n}} f(p, q) \right| \leq \frac{k_1! \dots k_n! \, l_1! \dots l_n!}{\delta^{k_1 + \dots + k_n + l_1 + \dots + l_n}} \, \|f\|_\varrho$$

for all points $(p, q) \in D_{\varrho - \delta}$, and

$$(2.4) \quad \left| \frac{\partial^{k_1 + \dots + k_n}}{\partial p_1^{k_1} \dots \partial p_n^{k_n}} f(0, q) \right| \leq \frac{k_1! \dots k_n!}{\varrho^{k_1 + \dots + k_n}} \, \|f\|_\varrho$$

for all points $(0, q)$ with $\|\mathrm{Im}(q)\| < \varrho$. The proof can be given exactly along the same lines as in the case of a polydisk (see, for example, ref. [14]).

It is easily seen that one has

$$(2.5) \quad \max(\|A\|_\varrho, \|B\|_\varrho) \leq 2E/\varrho$$

and that there exists a positive number $m < 1$ such that one has

$$(2.6) \quad m\|v\| \leq \|\bar{C}v\| , \qquad \|Cv\|_\varrho \leq m^{-1}\|v\| \qquad \text{for all } v \in \mathbf{C}^n;$$

precisely one can take

$$(2.7) \quad m = d/2 .$$

---

[14] B. A. FUCHS: *Introduction to the Theory of Analytic Functions of Several Complex Variables, Translations of mathematical monographs*, Vol. **8**. (Providence, R. I., 1963).

For what concerns (2.5), let us remark that one has

$$(2.8) \qquad \|A\|_\varrho \leqslant 2 \|H^1\|_\varrho, \qquad \|B\|_\varrho \leqslant \frac{1}{\varrho} \|H^1\|_\varrho,$$

the first one following immediately from the very definition (2.2) of A, while the second one follows from Cauchy's inequality (2.4) applied to

$$\|B\|_\varrho = \max_i \sup_{\|\mathrm{Im}\, q_i\| \leqslant \varrho} \left| \frac{\partial H^1}{\partial p_i}(0, q) \right|.$$

Inequality (2.5) then follows from $\varrho \leqslant 1$ and from (1.8), which reads now $\|H^1\|_\varrho \leqslant E$.

Coming now to (2.6) and (2.7), let us remark that, by definitions (2.2) of C and (1.4) of C^*, one has

$$(2.9) \qquad C(q) - C^* = \left\{ \frac{\partial^2 H^1}{\partial p_i \, \partial p_j}(0, q) \right\},$$

so that from Cauchy's inequality (2.4) with our definition of the norms one deduces

$$(2.10) \qquad \|(C - C^*) v\|_\varrho \leqslant \frac{2n}{\varrho^2} \|H^1\|_\varrho \|v\|.$$

By $\|H^1\|_\varrho \leqslant E$ and the estimates (1.10) and (1.11) for E one has then surely also

$$(2.11) \qquad \|(C - C^*) v\|_\varrho \leqslant (d/2) \|v\|,$$

which guarantees

$$(2.12) \qquad \|C^* v\| - \|\overline{(C - C^*)} v\| \geqslant (d/2) \|v\| \geqslant 0,$$

as is seen by using $\|\overline{\ldots}\| \leqslant \|\ldots\|_\varrho$ and (1.7). Using now the identity

$$(2.13) \qquad C = C^* + (C - C^*),$$

or also $\bar{C} = C^* + \overline{(C - C^*)}$, by (2.12) one gets

$$(2.14) \qquad \|\bar{C} v\| \geqslant \left| \|C^* v\| - \|\overline{(C - C^*)} v\| \right| \geqslant (d/2) \|v\|,$$

namely the first of (2.6) with $m = d/2$. Analogously, from (2.13) one gets

$$(2.15) \qquad \|C v\|_\varrho \leqslant \|C^* v\| + \|(C - C^*) v\|_\varrho,$$

or with (1.7) and (2.11)

$$(2.16) \qquad \| Cv \|_\varrho \leqslant (d^{-1} + d/2) \| v \| < 2d^{-1} \| v \| \ .$$

2·3. – It is then immediate to deduce theorem 1 from the following

Theorem 2. For given positive numbers γ, ϱ, $m < 1$, consider the Hamiltonian $H(p, q) = H^0(p, q) + H^1(p, q)$ defined in the domain D_ϱ by

$$(2.17) \qquad H^0(p, q) = a + \lambda \cdot p + \tfrac{1}{2} \sum_{i,j} C_{ij}(q) p_i p_j + R(p, q) ,$$

$$(2.18) \qquad H^1(p, q) = A(q) + \sum_i B_i(q) p_i$$

with $\| H \|_\varrho < 1$; here $a \in \mathbf{R}$, $\lambda \in \Omega_\gamma$, A, B_i, C_{ij}, $R \in \mathscr{A}_\varrho$ and $\bar{A} = 0$, while R is, as a function of p, of the order $|p|^3$. For the matrix $C(q) = \{C_{ij}(q)\}$, assume

$$(2.19) \qquad m \| v \| \leqslant \| \bar{C} v \| , \qquad |Cv|_\varrho \leqslant m^{-1} \| v \| \qquad \text{for any } v \in \mathbf{C}^n .$$

Then there exist positive numbers ε and ϱ' with $\varrho' < \varrho$ such that, if

$$(2.20) \qquad \max(\| A \|_\varrho, \| B \|_\varrho) \leqslant \varepsilon ,$$

one can construct a canonical analytical change of variables $\varphi : D_{\varrho'} \to D_\varrho$, $\varphi \in \mathscr{A}_{\varrho'}$, which brings the Hamiltonian H into the form

$$(2.21) \qquad H'(P, Q) = (H \circ \varphi)(P, Q) = a' + \lambda \cdot P + R'(P, Q) ,$$

where $a' \in \mathbf{R}$, while the function $R' \in \mathscr{A}_{\varrho'}$ is, as a function of P, of the order $|P|^2$. The change of variables is near the identity, in the sense that $\| \varphi - \text{identity} \|_{\varrho'} \to 0$ as $\| H^1 \|_\varrho \to 0$.

In particular, one can take

$$(2.22) \qquad \varepsilon = \frac{m^6 \varrho}{26 \varLambda^2} \left(\frac{\varrho}{13} \right)^{8n+6} ,$$

where

$$(2.23) \qquad \varLambda = 2(4n + 1)^2 \frac{\sigma^2}{\gamma^2} ,$$

$$(2.24) \qquad \sigma = 2^{4n+1} \left(\frac{n + 1}{e} \right)^{n+1} .$$

The scheme of proof of theorem 2 is as follows. One performs a sequence of canonical analytical changes of variables such that the disturbing term H^1

of the Hamiltonian H at step k decreases with k, its norm $\lfloor H_k^1 \rfloor_{\varrho_k}$ being essentially of the order of $\|H_{k-1}^1\|_{\varrho_{k-1}}^2$, while the other parameters ϱ_k and m_k are kept controlled. The convergence of the scheme with $\|H_k^1\|_{\varrho_k} \to 0$, $\varrho_k \to \varrho_\infty > 0$ and $m_k \to m_\infty > 0$, as $k \to \infty$, is then established. The iterative lemma and the proof of convergence are given in sect. **4** and **5**, respectively, while two auxiliary lemmas (giving suitable estimates in connection with canonical changes of variables defined by the Lie method and in connection with the use of the diophantine condition $\lambda \in \Omega_\gamma$ in solving a differential equation of the form $\sum_i \lambda_i (\partial F/\partial q_i) = G$ with $\bar{G} = 0$) will be recalled in the next section.

3. – Canonical transformations and small denominators, lemma.

3'1. For a given function $\chi(p, q)$, $\chi \in \mathscr{A}_\varrho$, let χ_ϱ^* be defined by

$$(3.1) \qquad \chi_\varrho^* = \max\left(\left\| \frac{\partial \chi}{\partial p} \right\|_\varrho, \left\| \frac{\partial \chi}{\partial q} \right\|_\varrho \right)$$

and let $\{\cdot, \cdot\}$ denote the Poisson bracket

$$(3.2) \qquad \{f, g\} = \sum_i \left(\frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} \right).$$

Then, for χ, $f \in \mathscr{A}_\varrho$ and any positive $\delta < \varrho$, the inequalities

$$(3.3) \qquad \|\{\chi, f\}\|_{\varrho-\delta} \leqslant 2n(\chi_\varrho^*/\delta)\|f\|_\varrho,$$

$$(3.4) \qquad \|\{\chi, \{\chi, f\}\}\|_{\varrho-\delta} \leqslant 4n(2n+1)(\chi_\varrho^*/\delta)^2\|f\|_\varrho$$

are deduced from Cauchy's inequality and from an enumeration of terms in Poisson brackets.

We come now to the definition of canonical transformations, or changes of variables, near the identity by means of the Lie method, which has, with respect to the standard method, the advantages of avoiding any use of the implicit-function theorem and of providing the relevant estimates in a simple way. In such a method a canonical transformation $(p, q) = \varphi(P, Q)$ is defined through the solution at time one of a system of canonical differential equations with Hamiltonian (here called « generating function ») χ; then the functions f defined in phase space change correspondingly by $f \mapsto \mathscr{U}f = f \circ \varphi$ and it is well known that one has formally $\mathscr{U}f = f + \{\chi, f\} + \frac{1}{2}\{\chi, \{\chi, f\}\} + \ldots$; indeed this is nothing but Taylor's expansion of f, using $\mathrm{d}f/\mathrm{d}t = \{\chi, f\}$, *i.e.*

$$\frac{\mathrm{d}f}{\mathrm{d}t}(p(t), q(t)) = \{\chi, f\}(p(t), q(t)).$$

The existence of such a canonical transformation φ and operator $\mathscr{U}$ and the relevant estimates are then afforded by the following

Lemma 1. Take an analytic function $\chi(p, q)$ defined in D_ϱ, whose derivatives belong to $\mathscr{A}_\varrho$ for a given positive ϱ, and consider the corresponding system of canonical differential equations

$$(3.5) \qquad\qquad \dot p = -\frac{\partial \chi}{\partial q}, \qquad \dot q = \frac{\partial \chi}{\partial p} .$$

With χ_ϱ^* defined by (3.1), for a positive $\delta < \varrho$ assume

$$(3.6) \qquad\qquad \chi_\varrho^* < \delta/2 .$$

Then for all initial data (P, Q) $D_{\varrho-\delta}$ the solution (p, q) of (3.5) at $t = 1$ exists in D_ϱ, thus defining a canonical transformation $\varphi : D_{\varrho-\delta} \to D_\varrho$, $\varphi \in \mathscr{A}_{\varrho-\delta}$. The operator $\mathscr{U} : \mathscr{A}_\varrho \to \mathscr{A}_{\varrho-\delta}$ with

$$(3.7) \qquad\qquad \mathscr{U} f = f \circ \varphi$$

is then well defined and one has the estimates

$$(3.8) \qquad\qquad \| \varphi\text{-identity} \|_{\varrho-\delta} \leqslant \chi_\varrho^*$$

and

$$(3.9) \qquad \begin{cases} \|\mathscr{U} f\|_{\varrho-\delta} \leqslant \|f\|_\varrho , \\[2mm] \|\mathscr{U} f - f\|_{\varrho-\delta} \leqslant 4n \dfrac{\chi_\varrho^*}{\delta} \|f\|_\varrho , \\[2mm] \|\mathscr{U} f - f - \{\chi, f\}\|_{\varrho-\delta} \leqslant 16n(2n+1) \left(\dfrac{\chi_\varrho^*}{\delta}\right)^2 \|f\|_\varrho . \end{cases}$$

Proof. First of all, using Cauchy's inequality (2.3), one guarantees that in the subdomain $D_{\varrho-\delta/2}$ the second members of system (3.5) have finite derivatives, so that the Lipschitz constant K of the system is finite in $D_{\varrho-\delta/2}$ (precisely one has $K < 4n\chi_\varrho^*/\delta$). Moreover, as $\max(\|\dot p\|, \|\dot q\|) \leqslant \chi_\varrho^* < \delta/2$, if one takes initial data (P, Q) in $D_{\varrho-\delta}$, the standard existence and uniqueness theorem guarantees that the corresponding solutions $(p(t), q(t))$ with $(p(0), q(0)) = (P, Q)$ exist in $D_{\varrho-\delta/2}$ for any t with $0 \leqslant t \leqslant 1$. Thus, for any such t, one has a mapping $\varphi^t : D_{\varrho-\delta} \to D_{\varrho-\delta/2} \subset D_\varrho$ with $\varphi^t(P, Q) = (p(t), q(t))$ and, in particular, the mapping $\varphi = \varphi^1$ is also defined. By the standard existence theorem such a mapping is analytic; furthermore, the periodicity in $q_1, \ldots, q_n$ is immediately checked. Finally, the mapping φ is well known to be canonical, being the « time-one solution » of a canonical system.

Let us now come to the estimates (3.8) and (3.9). Estimate (3.8) is an immediate consequence of the mean-value theorem, according to which, for any t in $0 < t \leqslant 1$, one has $|\varphi^t - \text{identity}|_{\varrho'} < \overset{*}{\chi_\varrho}$, while the first of (3.9) is trivial. The second and the third ones follow from Taylor's formula for f of first and second order, respectively, namely

$$\mathscr{U}f - f = \frac{df}{dt}\Big|_{t'} = \{\chi, f\}|_{t'} , \qquad \mathscr{U}f - f - \frac{df}{dt} = \frac{1}{2}\frac{d^2 f}{dt^2}\Big|_{t'} = \frac{1}{2}\{\chi, \{\chi, f\}\}|_{t''}$$

with $0 < t', t'' < 1$. Remarking that, as recalled above, one has $\varphi(P, Q) \in D_{\varrho-\delta/2}$ if $0 < t \leqslant 1$ and $(P, Q) \in D_{\varrho-\delta}$, one has then to estimate $\|\{\chi, f\}\|_{\varrho-\delta/2}$ and $\|\{\chi, \{\chi, f\}\}\|_{\varrho-\delta/2}$. Estimates (3.9) then follow by (3.3) and (3.4).

3˙2. In the course of the proof of theorem 2 a fundamental role is played by the possibility of solving a differential equation of the form $\sum_i \lambda_i(\partial F/\partial q_i) = G$ for functions F and G defined on the torus T^n if λ is a « good frequency ».

Lemma 2. Consider the equation

$$(3.10) \qquad\qquad \sum_i \lambda_i \frac{\partial F}{\partial q_i} = G ,$$

where F and G are functions defined on the torus T^n, and assume $\lambda = (\lambda_1, \ldots \ldots, \lambda_n) \in \Omega_\gamma$ for some $\gamma > 0$ and $G \in \mathscr{A}_\varrho$ for some positive ϱ with $\bar{G} = 0$. Then, for any positive $\delta < \varrho$, eq. (3.10) admits a unique solution $F \in \mathscr{A}_{\varrho-\delta}$ with $\bar{F} = 0$, and one has the estimates

$$(3.11) \qquad\qquad \|F\|_{\varrho-\delta} \leqslant \frac{\sigma}{\gamma\delta^{2n}} \|G\|_\varrho ,$$

$$(3.12) \qquad\qquad \left\|\frac{\partial F}{\partial q}\right\|_{\varrho-\delta} \leqslant \frac{\sigma}{\gamma\delta^{2n+1}} \|G\|_\varrho ,$$

where $\sigma = \sigma(n)$ is defined as in (2.24).

This is a well-established result. First of all, a formal solution in terms of Fourier coefficients f_k and g_k ($0 \neq k \in \mathbf{Z}^n$) of F and G, respectively, is immediately obtained, by $f_k = -ig_k(\lambda \cdot k)^{-1}$ if one has $g_0 = 0$, i.e. $\bar{G} = 0$. The convergence is then easily established, as the coefficients g_k decay exponentially with k in virtue of the analyticity of G, while $|\lambda \cdot k|^{-1}$ grows at most as a power by the diophantine condition (1.1) imposed on λ. The details of the proof are deferred to the appendix. Optimal estimates were given more recently by RÜSSMANN [6] with δ^{-2n} and $\delta^{-(2n+1)}$ in (3.11)-(3.12) replaced by δ^{-n} and $\delta^{-(n+1)}$, respectively, and σ replaced by $\sigma_R = 2^{4n+2}n!(2^n-1)^{-1}$.

4. – The iterative lemma.

4$\cdot$1. – As anticipated in the introduction, one applies now a sequence of canonical transformations in order to eliminate the disturbing term H^1 in the original Hamiltonian H considered in theorem 2. To this end we will repeatedly apply the following

Lemma 3. For given positive numbers γ, ϱ, m, $\varepsilon < 1$ and $\varrho_* < \varrho$, $m_* < m$ consider the Hamiltonian $H(p, q) = H^0(p, q) + H^1(p, q)$ defined in D_ϱ by

(4.1)
$$\begin{cases} H^0(p, q) = a + \lambda \cdot p + \tfrac{1}{2} \sum_{i,j} C_{ij}(q) p_i p_j + R(p, q) \, , \\ H^1(p, q) = A(q) + \sum_i B_i(q) p_i \end{cases}$$

with $\|H\|_\varrho < 1$, where $a \in \mathbf{R}$, $\lambda \in \Omega_\gamma$, A, B_i, C_{ij}, $R \in \mathscr{A}_\varrho$, $\bar{A} = 0$ and R is of order $|p|^2$. Assume, furthermore,

(4.2)
$$m|v\| < \|\bar{C}v\| \, , \qquad |Cv|_\varrho < m^{-1}|v\| \qquad \text{for any } v \in \mathbf{C}^n \, ,$$

(4.3)
$$\max(\|A\|_\varrho, |B\|_\varrho) < \varepsilon \, .$$

For any positive δ so small that $\varrho - 3\delta > \varrho_*$, let the quantity η be defined by

(4.4)
$$\eta = \varLambda \frac{\varepsilon}{m^3 \delta^\tau} \, , \qquad\qquad \tau = 4n + 3 \, ,$$

where $\varLambda = \varLambda(n, \gamma)$ is the constant defined by (2.23), namely

(4.5)
$$\varLambda = 2(4n + 1)^2 \frac{\sigma^2}{\gamma^2} \, ,$$

and assume that ε is so small that

(4.6)
$$m - n\eta/\varrho_*^2 > m_* \, .$$

Then one can find an analytical canonical change of variables, $\varphi : D_{\varrho - 3\delta} \to D_\varrho$, $\varphi \in \mathscr{A}_{\varrho - 3\delta}$, such that the transformed Hamiltonian $H' = \mathscr{U}H = H \circ \varphi$ can be decomposed in a way analogous to H with corresponding primed quantities a', A', B', C' and R', but with the same λ, and satisfies analogous conditions with positive parameters ϱ', m', $\varepsilon' < 1$ given by

(4.7)
$$\begin{cases} \varrho' = \varrho - 3\delta > \varrho_* \, , \\ m' = m - n\eta/\varrho_*^2 > m_* \, , \\ \varepsilon' = \eta^2/\varrho_* \end{cases}$$

with, moreover, $\|H'\|_{\varrho'} \leqslant 1$. One has, furthermore, for any $f \in \mathscr{A}_{\varrho}$

(4.8) $$\|\mathscr{U}f - f\|_{\varrho'} \leqslant \frac{\eta}{2}\|f\|_{\varrho}\,.$$

4'2. – Proof. a) If one performs a canonical change of variables with « generating function » χ in the sense of lemma 1, one obtains in place of H the new Hamiltonian $H' = \mathscr{U}H$ and one can make a decomposition $H' = H'^0 + H'^1$ in a way analogous to the decomposition $H = H^0 + H^1$; precisely, using again the symbol (p, q) instead of (P, Q) for a point of the new domain, one has

$$H'^0 = a' + \lambda \cdot p + \tfrac{1}{2}\sum_{i,j} C'_{ij}(q)\,p_i p_j + R'(p, q)\,, \qquad H'^1 = A'(q) + B'(q)\cdot p\,,$$

where

(4.9) $\begin{cases} a' \quad = \bar{H}'(0)\,, \\[2mm] A'(q) = H'(0, q) - a'\,, \\[2mm] B'_i(q) = \dfrac{\partial H'}{\partial p_i}(0, q) - \lambda_i\,, \\[2mm] C'_{ij}(q) = \dfrac{\partial^2 H'}{\partial p_i\,\partial p_j}(0, q) \end{cases}$

and R' is of order $|p|^3$. In the spirit of perturbation theory, one thinks that both H^1 and the generating function χ are of first order and one chooses χ in order to eliminate the undesired terms of the same order in the new Hamiltonian H'. To this end one first writes the identity

(4.10) $$H' = \mathscr{U}H = H^0 + H^1 + \{\chi, H^0\} + [\{\chi, H^1\} + \mathscr{U}H - H - \{\chi, H\}]\,,$$

where in the expression $[\dots]$ are isolated the terms that have to be considered of second order, in agreement with the estimate given in the third of (3.9). Then one tries to choose χ in such a way that the first-order terms in H', namely $H^1 + \{\chi, H^0\}$, do not contribute to H'^1; this is obtained by imposing $H^1 + \{\chi, H^0\} = c + O(|p|^2)$, where c is a constant.

b) Following KOLMOGOROV, we show that this condition is met by a generating function χ of the form

(4.11) $$\chi = \xi \cdot q + X(q) + \sum_i Y_i(q)\,p_i\,,$$

where the constant $\xi \in \mathbf{R}^n$ and the functions $X(q)$, $Y_i(q)$ (of period 2π in $q_1, \dots, q_n$)

have to be suitably determined. A straightforward calculation gives, by recalling the definitions (2.17)-(2.18) of H^0 and H^1,

$$(4.12) \qquad H^1 + \{\chi, H^0\} = -\sum_i \xi_i \lambda_i + A(q) - \sum_i \lambda_i \frac{\partial X}{\partial q_i} +$$

$$+ \sum_j \left[B_j(q) - \sum_i C_{ij}(q) \left(\xi_i + \frac{\partial X}{\partial q_i} \right) - \sum_i \lambda_i \frac{\partial Y_j}{\partial q_i} \right] p_j + O(\|p\|^2) .$$

Thus it is sufficient to impose

$$(4.13) \qquad \sum_i \lambda_i \frac{\partial X}{\partial q_i} = A(q) ,$$

$$(4.14) \qquad \sum_i \lambda_i \frac{\partial Y_j}{\partial q_i} = B_j(q) - \sum_i C_{ij}(q) \left(\xi_i + \frac{\partial X}{\partial q_i} \right), \qquad j = i, ..., n .$$

By the small denominators, lemma 2, eq. (4.13) in the unknown X' can be solved, as $\overline{A} = 0$. Then one has to determine the unknown constant ξ in such a way that the mean value of the r.h.s. of (4.14) vanish. This leads to a linear equation for ξ, which in compact notation can be written as

$$(4.15) \qquad \overline{C}\xi = \overline{B - C \frac{\partial X}{\partial q}} ,$$

and such an equation can be solved by (4.2), which guarantees $\det \overline{C} \neq 0$. Equation (4.14) in the unknown Y can then be solved too.

  c) The existence of the wanted generating function χ is thus ascertained and one remains with the problem of specific estimates. By a twofold application of lemma 2, one obtains by easy calculations the following

  Main estimate. Equations (4.13)-(4.15) in the unknowns ξ, $X(q)$ and $Y_j(q)$, which define by (4.11) the generating function χ, can be solved with X, $Y_j \in \mathscr{A}_\varrho$, $\tilde{\varrho} = -2\delta$ for any positive $\delta < \varrho/2$, and, for the quantity χ_ϱ^* defined by (3.1), one gets

$$(4.16) \qquad \chi_\varrho^* < (4n + 1) \frac{\sigma^2}{\gamma^2} \frac{\varepsilon}{m^3 \delta^{4n+2}}$$

or equivalently

$$(4.17) \qquad \frac{\chi_{\tilde{\varrho}}^*}{\delta} < \frac{1}{2(4n + 1)} \eta .$$

The details are deferred to the end of the present section.

  Remarking now that condition (4.6) evidently implies $\eta < 1$, from (4.17) one gets $\chi_{\tilde{\varrho}}^* < \delta/2$, so that lemma 1 can be applied and χ generates a canonical transformation with domain $D_{\varrho'}$, where $\varrho' = \tilde{\varrho} - \delta = \varrho - 3\delta$.

d) We come now to estimates (4.7)-(4.8) and $\|H'\|_{\varrho'} \leqslant 1$. The inequality $\|H'\|_{\varrho'} \leqslant \|H\|_{\varrho} \leqslant 1$ is just the first of (3.9). Let us then come to ε'. For what concerns A', by definition (4.9), using expression (4.10) for H' and recalling that, in virtue of the choice of χ, only the last term contributes to A', one gets $|A'|_{\varrho'} \leqslant 2\|\{\chi, H^1\} + \mathcal{U}H - H - \{\chi, H\}\|_{\varrho'}$. Using now (3.3) and the third of (3.9), together with $\|H^1\|_{\bar\varrho} \leqslant (n+1)\varepsilon$, $\|H\|_{\bar\varrho} \leqslant 1$, $\varrho' = \bar\varrho - \delta$, and (4.17), one gets

$$\|\{\chi, H^1\} + \mathcal{U}H - H - \{\chi, H\}\|_{\varrho'} \leqslant 2n(n+1)\varepsilon \frac{\chi^*_{\bar\varrho}}{\delta} + 16n(2n+1)\left(\frac{\chi^*_{\bar\varrho}}{\delta}\right)^2 \leqslant$$

$$\leqslant (4n+1)\frac{\sigma^2}{\gamma^2}\frac{\varepsilon}{m^3\delta^\tau}\frac{\chi^*_{\bar\varrho}}{\delta} + 16n(2n+1)\left(\frac{\chi^*_{\bar\varrho}}{\delta}\right)^2 \leqslant$$

$$\leqslant [16n(2n+1)+1]\left[\frac{\eta}{2(4n+1)}\right]^2 \leqslant \eta^2/2 ,$$

where (4.17) was used together with

$$2n(n+1) < (4n+1)\frac{\sigma^2}{\gamma^2 m^3\delta^\tau} \quad \text{and} \quad 16n(2n+1)+1 < 2(4n+1)^2 .$$

Thus one has $|A'|_{\varrho'} \leqslant \eta^2$. Concerning B', for the same reason, from (4.9) and (4.10) and recalling $(\partial H^0/\partial p_i)(0, q) = \lambda_i$, using Cauchy's inequality (2.4) one has

$$\|B'\|_{\varrho'} \leqslant \frac{1}{\varrho'}\|\{\chi, H^1\} + \mathcal{U}H - H - \{\chi, H\}\|_{\varrho'} \leqslant \frac{\eta^2}{2\varrho'} \leqslant \frac{\eta^2}{2\varrho_*} .$$

In conclusion, one has $\max(|A'|_{\varrho'}, |B'|_{\varrho'}) < \varepsilon'$ with $\varrho' = \varrho - 3\delta$ and

(4.18) $$\varepsilon' = \eta^2/\varrho_* .$$

For what concerns inequality (4.8), from the second of (3.9), with $\bar\varrho$ in place of ϱ, one has

(4.19) $$\|\mathcal{U}f - f\|_{\varrho'} < \frac{\chi^*_{\bar\varrho}}{\delta}\|f\|_{\bar\varrho} ,$$

from which (4.8) follows by using (4.17) and $\|f\|_{\bar\varrho} \leqslant \|f\|_{\bar\varrho}$.

Let us finally come to the estimate for m'. To this end, we recall first that, by the definition of C and C', one has

$$C'_{ij}(q) - C_{ij}(q) = \frac{\partial^2}{\partial p_i \partial p_j}[\mathcal{U}H - H](0, q) .$$

Thus, from Cauchy's inequality (2.4) and (4.8), one gets

$$(4.20) \qquad \|C'_{ij} - C_{ij}\|_{\varrho'} < 2 \|\mathscr{U}H - H\|_{\varrho'} < \frac{\eta}{\varrho_*^2},$$

which gives

$$(4.21) \qquad \|(C' - C)v\|_{\varrho'} < (n\eta/\varrho_*^2)\|v\| .$$

Thus from (4.2) and (4.21), using also $\|\overline{\ldots}\| \leqslant \|\ldots\|_{\varrho'}$, one has

$$\|\bar{C}v\| - \|\overline{(C' - C)}v\| \geqslant (m - n\eta/\varrho_*^2)\|v\| > 0$$

in virtue of condition (4.6) and, moreover, from the identity $C' = C + (C' - C)$,

$$(4.22) \qquad \|\bar{C}'v\| > (m - n\eta/\varrho_*^2)\|v\| .$$

Furthermore, one has $\|C'v\|_{\varrho'} \leqslant \|Cv\|_{\varrho'} + \|(C' - C)v\|_{\varrho'}$, or, with (4.21) and $\|Cv\|_{\varrho'} < \|Cv\|_{\varrho} \leqslant m^{-1}\|v\|$,

$$(4.23) \qquad \|C'v\|_{\varrho'} < (m + n\eta/\varrho_*^2)^{-1}\|v\| .$$

Thus with the trivial inequality $a^{-1} + b < (a - b)^{-1}$ for $0 < b < a < 1$ one obtains

$$(4.24) \qquad \|C'v\|_{\varrho'} < (m - n\eta/\varrho_*^2)^{-1}\|v\| ,$$

so that the estimate for m' is proven.

e) To conclude the proof, we give now the main estimate (4.16). First, by lemma 2 applied to eq. (4.13) with $\|A\|_\varrho < \varepsilon$, one finds

$$(4.25) \qquad \|X\|_{\varrho-\delta} \leqslant \frac{\sigma\varepsilon}{\gamma\delta^{2n}} , \qquad \left\|\frac{\partial X}{\partial q}\right\|_{\varrho-\delta} \leqslant \frac{\sigma\varepsilon}{\gamma\delta^{2n+1}} .$$

Coming to eq. (4.15) we first establish

$$(4.26) \qquad \left\|B - C\frac{\partial X}{\partial q}\right\|_{\varrho-\delta} \leqslant 2\frac{\sigma\varepsilon}{\gamma m\delta^{2n+1}} ,$$

making use of

$$\|B\|_{\varrho-\delta} \leqslant \varepsilon , \qquad \left\|C\frac{\partial X}{\partial q}\right\|_{\varrho-\delta} \leqslant m^{-1}\left\|\frac{\partial X}{\partial q}\right\|_{\varrho-\delta}$$

and of $\sigma/(\gamma m\delta^{2n+1}) > 1$. Thus, by $\|\overline{\ldots}\| \leqslant \|\ldots\|_{\varrho-\delta}$, from (4.15) and $\|\xi\| <$

$\leqslant m^{-1}\|\bar{C}\xi\|$, *i.e.* the definition of m, one gets

(4.27)
$$\|\xi\| < 2 \frac{\sigma \varepsilon}{\gamma m^2 \delta^{2n+1}}.$$

Coming then to eq. (4.14), in order to estimate the r.h.s. we remark that

$$\left\| B - C\frac{\partial X}{\partial q} - C\xi \right\|_{\varrho-\delta} < \left\| B - C\frac{\partial X}{\partial q} \right\|_{\varrho-\delta} + m^{-1}\|\xi\|,$$

so that, by (4.26), (4.27) with $m < 1$, one has

(4.28)
$$\left\| B - C\frac{\partial X}{\partial q} - C\xi \right\|_{\varrho-\delta} < 4 \frac{\sigma \varepsilon}{\gamma m^3 \delta^{2n+1}}.$$

A direct application of lemma 2 then gives

$$\|Y\|_{\varrho-2\delta} \leqslant 4 \frac{\sigma^2 \varepsilon}{\gamma^2 m^3 \delta^{4n+1}}, \qquad \left\| \frac{\partial Y}{\partial q} \right\|_{\varrho-2\delta} \leqslant 4 \frac{\sigma^2 \varepsilon}{\gamma^2 m^3 \delta^{4n+2}}.$$

As a consequence, recalling definition (3.1) of $\chi_{\tilde{\varrho}}^*$ (with $\tilde{\varrho} = \varrho - 2\delta$), by definition (4.11) of χ which gives

$$\left\| \frac{\partial \chi}{\partial p} \right\|_{\tilde{\varrho}} = \|Y\|_{\tilde{\varrho}}, \qquad \left\| \frac{\partial \chi}{\partial q} \right\|_{\tilde{\varrho}} < \|\xi\| + \left\| \frac{\partial X}{\partial q} \right\|_{\tilde{\varrho}} + n \left\| \frac{\partial Y}{\partial q} \right\|_{\tilde{\varrho}} \tilde{\varrho},$$

and using δ, γ, $\tilde{\varrho}$, $m < 1$ with $3\sigma^{-1} < 1$, one gets

(4.29)
$$\chi_{\tilde{\varrho}}^* < (4n+1) \frac{\sigma^2 \varepsilon}{\gamma^2 m^3 \delta^{4n+2}}.$$

The main estimate has thus been checked and the proof of lemma 3 completed.

5. – Conclusion of the proof.

5'1. – One has now to apply repeatedly the iterative lemma 3 in order to eliminate the perturbation $H^1(p, q) = A(q) + B(q) \cdot p$. Thus, starting from given positive numbers ϱ, m, $\varepsilon < 1$, which we denote now by ϱ_0, m_0, ε_0, by assigning a sequence $\{\delta_k\}_{k=0}^{\infty}$ one can define recursively ϱ_{k+1}, m_{k+1}, ε_{k+1} by relations analogous to (4.7), which we write now in the form

(5.1)
$$\begin{cases} \varrho_{k+1} = \varrho_k - 3\delta_k, \\[2mm] \varepsilon_{k+1} = \frac{\Lambda^2}{\varrho_*} \left(\frac{\varepsilon_k}{m_k^3 \delta_k^\tau} \right)^2, \\[2mm] m_{k+1} = m_k - (n/\varrho_*^{\frac{1}{2}}) \varepsilon_{k+1}^{\frac{1}{2}}, \end{cases}$$

where definition (4.4) of η was used and, moreover, in the relation between m' and m, η was expressed through ε'. One has, however, to satisfy the two consistency conditions $\varrho_{k+1} > \varrho_*$, $m_{k+1} > m_*$ $(k = 0, 1, \ldots)$, where $\varrho_* < \varrho_0$ and $m_* < m_0$ are arbitrary positive numbers. In virtue of

$$\varrho_\infty = \varrho_0 - \sum_{k=0}^{\infty} 3\delta_k \quad \text{and} \quad m_\infty = m_0 - (n/\varrho_*^{3}) \sum_{k=0}^{\infty} \varepsilon_{k+1}^{1/2},$$

these consistency conditions are satisfied if one guarantees

(5.2)
$$3 \sum_{k=0}^{\infty} \delta_k < \varrho_0 - \varrho_*,$$

(5.3)
$$(n/\varrho_*^{3}) \sum_{k=0}^{\infty} \varepsilon_{k+1}^{\frac{1}{2}} < m_0 - m_*.$$

We notice, however, that the relation connecting ε_{k+1} with m_k, δ_k and ε_k can also be read as defining δ_k if ε_{k+1} is given; precisely one has

(5.4)
$$\delta_k^{2\tau} = \frac{\varLambda^2}{\varrho_* m_k^6} \frac{\varepsilon_k^2}{\varepsilon_{k+1}}.$$

Notice that m_k is defined in terms of ε_k and m_0 by the last of (5.1). Thus one can think of defining arbitrarily the sequence $\{\varepsilon_k\}$, instead of the sequence $\{\delta_k\}$. Making also the position

(5.5)
$$\varepsilon_k = c_k \varepsilon_0, \qquad k = 0, 1, 2, \ldots (c_0 = 1),$$

one has thus to define a sequence of positive numbers c_k with $c_0 = 1$ so that the two series

(5.6)
$$\sum_{k=0}^{\infty} s_k = \sum_{k=0}^{\infty} (c_k^2/c_{k+1})^{1/2\tau} = s,$$

(5.7)
$$\sum_{k=0}^{\infty} t_k = \sum_{k=0}^{\infty} c_{k+1}^{\frac{1}{2}} = t$$

converge. Then the two consistency conditions (5.2) and (5.3) determine ε_0; indeed (5.3) gives

(5.8)
$$\varepsilon_0 < \frac{\varrho_*^3 (m_0 - m_*)^2}{t^2 n^2},$$

while (5.2), by using $m_k > m_*$, gives

(5.9)
$$\varepsilon_0 < \frac{\varrho_* m_*^6}{\varLambda^2} \left(\frac{\varrho_0 - \varrho_*}{3s}\right)^{2\tau}.$$

A rough estimate for ε_0 is obtained as follows. Choose $c_k = 2^{-2\tau k}$; this gives $s_k = 2 \cdot 2^{-k}$ and $t_k = 2^{-\tau(k+1)}$, so that $s = 4$ and $t = 1/(2^\tau - 1) < 1$. For the two up to now arbitrary numbers ϱ_* and m_* make the simple choice $\varrho_* = \varrho_0/13$ and $m_*^6 = m_0^6/2$. This ensures that the r.h.s. of (5.8) is larger than that of (5.9), so that we conclude that the two consistency conditions (5.2) and (5.3) are satisfied if one assumes, for example,

$$(5.10) \qquad \varepsilon_0 < \frac{\varrho_0 m_0^6}{26 A^2} \left(\frac{\varrho_0}{13} \right)^{2\tau} .$$

5'2. – We finally come to the convergence of the sequence of canonical transformations, thus giving the

Proof of theorem 2. Starting from the Hamiltonian $H_0 = H$ defined in D_{ϱ_0} with $\|H_0\|_{\varrho_0} < 1$, characterized by positive parameters γ, ϱ_0, $m_0 < 1$, consider the quantity ε_0 satisfying (5.10) and assume $\max(\|A\|_{\varrho_0}, \|B\|_{\varrho_0}) < \varepsilon_0$. Then one can apply recursively the iterative lemma 3, defining at each step $k \geqslant 1$ a canonical transformation $\varphi_k : D_{\varrho_k} \to D_{\varrho_{k-1}}$, with the corresponding operator $\mathcal{U}_k : \mathcal{A}_{\varrho_k} \to \mathcal{A}_{\varrho_{k-1}}$; furthermore, from (4.7) and (4.8) one has the estimate

$$(5.11) \qquad \|\mathcal{U}_k f - f\|_{\varrho_k} \leqslant \eta_{k-1} \|f\|_{\varrho_{k-1}}$$

with $\eta_{k-1} = \varrho_*^{\frac{1}{2}} \varepsilon_k^{\frac{1}{2}}$, so that, in particular, the series $\sum\limits_{k=1}^{\infty} \eta_k$ is known to be convergent. We can now define the composite canonical transformation $\hat{\phi}_k : D_{\varrho_k} \to D_{\varrho_0}$ by $\hat{\phi}_k = \varphi_1 \circ \ldots \circ \varphi_k$ and the corresponding composite operator $\hat{\mathcal{U}}_k : \mathcal{A}_{\varrho_0} \to \mathcal{A}_{\varrho_k}$, defined by $\hat{\mathcal{U}}_k f = f \circ \hat{\phi}_k$, or equivalently by $\hat{\mathcal{U}}_k = \mathcal{U}_k \circ \hat{\mathcal{U}}_{k-1}$, $\hat{\mathcal{U}}_0 = $ identity. Clearly, in order to prove the convergence of the sequence $\{\hat{\phi}_k\}$ of canonical transformations restricted to D_{ϱ_∞}, it is sufficient to prove the convergence of the corresponding sequence $\{\hat{\mathcal{U}}_k\}$ of operators for every $f \in \mathcal{A}_{\varrho_0}$. This in turn is seen by remarking that, from (5.11) and the first of (3.9), one has

$$\|(\hat{\mathcal{U}}_{k+1} - \hat{\mathcal{U}}_k) f\|_{\varrho_{k+1}} = \|\mathcal{U}_{k+1}(\hat{\mathcal{U}}_k f) - \hat{\mathcal{U}}_k f\|_{\varrho_{k+1}} \leqslant \eta_k \|\hat{\mathcal{U}}_k f\|_{\varrho_k} \leqslant \eta_k \|f\|_{\varrho_0}$$

and so also, for any $l \geqslant 1$,

$$(5.12) \qquad \|(\hat{\mathcal{U}}_{k+l} - \hat{\mathcal{U}}_k) f\|_{\varrho_{k+l}} \leqslant \|f\|_{\varrho_0} \sum\limits_{p=k}^{k+l-1} \eta_p .$$

Thus, as the series $\sum\limits_{k=1}^{\infty} \eta_k$ converges, one deduces that the sequence $\hat{\mathcal{U}}_k f$ converges uniformly, for any $f \in \mathcal{A}_{\varrho_0}$; by Weierstrass' theorem one has then

$$\lim_{k \to \infty} \hat{\mathcal{U}}_k f = \hat{\mathcal{U}}_\infty f \in \mathcal{A}_{\varrho_\infty};$$

furthermore one immediately gets the estimate

(5.13)
$$\|\hat{\mathscr{U}}_\infty f - f\|_{\varrho_\infty} \leqslant (\varrho_*^{\frac{1}{2}} \varepsilon_0^{\frac{1}{2}} t) \|f\|_{\varrho_*},$$

where t is the constant defined by (5.7). In particular, for the Hamiltonian $H_\infty = \lim_{k\to\infty} \hat{\mathscr{U}}_k H$, one has $H_\infty = H_\infty^0 + H_\infty^1$, where by construction $H_\infty^1(p, q) = = A_\infty(q) + B_\infty(q) \cdot p = 0$. Finally the mapping $\hat{\phi}_\infty = \lim_{k\to\infty} \hat{\phi}_k$ turns out to be canonical again in virtue of Weierstrass' theorem, as a uniform limit of canonical mappings.

APPENDIX

Proof of lemma 2.

As we are here interested in functions of the angles q only, in the definitions of D_ϱ and $\mathscr{A}_\varrho$ the actions p will be now considered as parameters and completely disregarded. Moreover, in addition to our standard norm $\|k\| = = \max_i |k_i|$, we will also introduce the norm $|k| = \sum_i |k_i|$. We recall first two elementary properties concerning analytic functions on the torus T^n, and an elementary inequality, precisely we prove what follows:

  i) if $F \in \mathscr{A}_\varrho$, $F(q) = \sum_{k\in Z^n} f_k \exp[ik \cdot q]$, then for every $k \in Z^n$ one has

(A.1)
$$|f_k| \leqslant \|F\|_\varrho \exp[-|k|\varrho];$$

  ii) suppose that for some positive constants C and ϱ with $\varrho \leqslant 1$ and every $k \in Z^n$ one has $|f_k| \leqslant C \exp[-|k|\varrho]$, and consider the function $F(q) = = \sum_{k\in Z^n} f_k \exp[ik \cdot q]$; then for any positive $\delta < \varrho$ one has $F \in \mathscr{A}_{\varrho-\delta}$ and

(A.2)
$$\|F\|_{\varrho-\delta} \leqslant C \left(\frac{4}{\delta}\right)^n;$$

  iii) for any K, s, $\delta > 0$ one has the inequality

(A.3)
$$K^s \leqslant \left(\frac{s}{e\delta}\right)^s \exp[K\delta].$$

Proof. i) By definition one has

$$f_k = \frac{1}{(2\pi)^n} \int_{T^n} F(q_1, \ldots, q_n) \exp\left[-i\sum_j k_j q_j\right] dq_1 \ldots dq_n.$$

As $F \in \mathcal{A}_\varrho$, one can suitably shift the integration paths and write

$$f_k = \frac{1}{(2\pi)^n} \int\limits_{T^n} F\left(q_1 - i\frac{k_1}{|k_1|}\varrho, \ldots, q_n - i\frac{k_n}{|k_n|}\varrho\right) \prod_{j=1}^{n} \exp\left[-ik_j\left(q_j - i\frac{k_j}{|k_j|}\varrho\right)\right] dq_1 \ldots dq_n ,$$

where, by convention, we may assume $\dfrac{0}{|0|} = 0$. One has then

$$|f_k| \leqslant \|F\|_\varrho \exp\left[-\left(\sum_j |k_j|\right)\varrho\right] \frac{1}{(2\pi)^n} \int\limits_{T^n} |\exp\left[-ik\cdot q\right]| \, dq_1 \ldots dq_n = \|F\|_\varrho \exp\left[-|k|\varrho\right] .$$

ii) Let $\|\operatorname{Im} q\| \leqslant \varrho - \delta$, where $\operatorname{Im} q = (\operatorname{Im} q_1, \ldots, \operatorname{Im} q_n)$. Then one has

$$\|F\|_{\varrho-\delta} = \sup\left|\sum_{k\in Z^n} f_k \exp\left[ik\cdot q\right]\right| \leqslant$$

$$\leqslant C \sum_{k\in Z^n} \exp\left[-|k|\varrho\right] \exp\left[|k|(\varrho-\delta)\right] = C \sum_{k\in Z^n} \exp\left[-|k|\delta\right] <$$

$$< C2^n \sum_{\substack{(k_1,\ldots,k_n)\in Z^n \\ k_1\geqslant 0,\ldots,k_n\geqslant 0}} \exp\left[-\delta \sum_{j=1}^{n} k_j\right] = C2^n \left(\sum_{k=0}^{\infty} \exp\left[-\delta k\right]\right)^n =$$

$$= C2^n \frac{1}{(1-\exp\left[-\delta\right])^n} < C\left(\frac{4}{\delta}\right)^n ,$$

because $1/(1 - \exp\left[-\delta\right]) < 2/\delta$, for any positive $\delta < 1$, as is immediately seen by comparing $1 - \delta/2$ and $\exp\left[-\delta\right]$.

iii) The considered inequality is equivalent to $K\delta/s \leqslant \exp\left[K\delta/s - 1\right]$, i.e. to $x \leqslant \exp\left[x - 1\right]$, which is evidently true even for any real x.

We come then to the

Proof of the lemma. In terms of the Fourier coefficients f_k and g_k of F and G, respectively, one formally satisfies eq. (3.10) by

$$(A.4) \qquad\qquad\qquad f_k = \frac{-ig_k}{\lambda \cdot k} , \qquad\qquad\qquad 0 \neq k \in \mathbf{Z}^n ,$$

where, by the condition $\lambda \in \Omega_\gamma$, all denominators are, in particular, nonvanishing. By $G \in \mathcal{A}_\varrho$, property i) and $\lambda \in \Omega_\gamma$, one has then

$$(A.5) \qquad\qquad\qquad |f_k| \leqslant \gamma^{-1} \|G\|_\varrho |k|^n \exp\left[-|k|\varrho\right] ,$$

where the inequality $\|k\| \leqslant |k|$ has been used. From inequality (A.3) with $K = |k|$ and $s = n$, for any positive δ this gives

$$|f_k| \leqslant \gamma^{-1} \|G\|_\varrho \left(\frac{n}{e\delta}\right)^n \exp\left[-|k|(\varrho-\delta)\right] = C \exp\left[-|k|(\varrho-\delta)\right] ,$$

where

$$C = \gamma^{-1} \|G\|_\varrho \left(\frac{n}{e\delta}\right)^n .$$

Now, for any $\delta < \varrho$, one can make use of property ii) (with $\varrho - \delta$ in place of ϱ) and one obtains for $F(q) = \sum\limits_{0 \neq k \in Z^n} f_k \exp [ik \cdot q]$ that $F \in \mathscr{A}_{\varrho - 2\delta}$ and that

$$\|F\|_{\varrho - 2\delta} \leqslant \gamma^{-1} \left(\frac{4n}{e\delta^2} \right)^n \|G\|_\varrho .$$

Thus (3.11) is proved by taking $\delta/2$ in place of δ, because $\sigma > (16n/e)^n$.

Let us finally prove (3.12). The Fourier coefficients of $\partial F/\partial q_j$ are of the form $h_{k_j} = k_j g_k/(\lambda \cdot k)^{-1}$, $j = 1, ..., n$, so that as for (A.5) one gets

$$|h_{k_j}| \leqslant \gamma^{-1} \|G\|_\varrho |k|^{n+1} \exp [- |k|\varrho] .$$

By iii) with $K = |k_j|$ and $s = n + 1$, this gives

$$|h_{k_j}| \leqslant \gamma^{-1} \left(\frac{n+1}{e\delta} \right)^{n+1} \|G\|_\varrho \exp [- |k|(\varrho \cdot \delta)] ,$$

so that one obtains

$$\left\| \frac{\partial F}{\partial q} \right\|_{\varrho - 2\delta} \leqslant \frac{4^n}{\gamma \delta^{2n+1}} \left(\frac{n+1}{e} \right)^{n+1} \|G\|_\varrho .$$

By taking again $\delta/2$ in place of δ, one thus gets (3.12) with

$$\sigma = 2^{4n+1} \left(\frac{n+1}{e} \right)^{n+1} .$$

● RIASSUNTO

Nel presente lavoro si dimostra il teorema di Kolmogorov sull'esistenza di tori invarianti in sistemi Hamiltoniani quasi integrabili. Si usa lo schema di dimostrazione di Kolmogorov, con la sola variante del modo in cui si definiscono le trasformazioni canoniche prossime all'identità. Si usa infatti il metodo di Lie, che elimina la necessità d'inversioni e quindi dell'impiego del teorema delle funzioni implicite. Questo fatto tecnico evita un ingrediente spurio e semplifica il modo in cui si ottiene una delle stime principali.

Доказательство теоремы Колмогорова на инвариантных торах, используя канонические преобразования, определенные с помощью метода Ли.

Резюме (*). — В этой работе предлагается доказательство теоремы Колмогорова о существовании инвариантных торов в квази-интегрируемых Гамильтоновых системах. Используется схема доказательства, предложенная Колмогоровым, единственное отличие состоит в способе, которым определяются канонические преобразования. В этой работе используется метод Ли, который исключает необходимость инверсии и, следовательно, использование теоремы для неявной функции. Этот технический прием исключает ложный ингрдеиент и упрощает получение главной оценки.

(*) *Переведено редакцией.*