QUANTITATIVE METHODS IN CLASSICAL PERTURBATION THEORY

ANTONIO GIORGILLI
Dipartimento di Matematica dell' Università di Milano
Via Saldini 50, 20133 Milano, Italy.

1. Overview

At the beginning of the second volume of his *Méthodes nouvelles de la Mécanique Céleste* Poincaré devoted the chapter VIII to the problem of the reliability of the formal expansions of perturbation theory. He proved that the series commonly used in Celestial mechanics are typically non convergent, although their usefulness is generally evident. In particular, he pointed out that these series could have the same character of the Stirling's series. Recent work in perturbation theory has enlighten this conjecture of Poincaré, bringing into evidence that the series of perturbation theory, although non convergent in general, furnish nevertheless valuable approximations to the true orbits for a very large time, which in some practical cases could be comparable with the age of the universe.

The aim of my lectures is to introduce the quantitative methods of perturbation theory which allow to obtain such powerful results.

It is well known that the hardest problem in dealing with the possible convergence of the series is the appearance of the so called *small denominators*. It is also known that Kolmogorov (1954) proved that in some special cases these series do converge, but only for a strange set of initial conditions which has large measure, but is nowhere dense in phase space. I refer here to the celebrated theorem on the persistence under small perturbation of conditionally periodic motions on invariant tori, usually referred to as KAM theorem.

On the other hand, the problem of establishing the asymptotic character of the formal expansions seems to have been less considered by mathematicians: as far as I know, it has been discussed by Moser (1955) and Littlewood (1959), but these works have essentially been forgotten. More recently, a general theorem in this direction was proven by Nekhoroshev (1977). This will be the subject of part of my lectures.

The theorem of Nekhoroshev can be considered as a quantitative approach to the normal form methods in perturbation theory. The normal form theory is a rather well known topic, at least as far the formal aspect is concerned (an application to the problem of the solar system has been given in the lectures of Message). The problem can be stated as follows: given an Hamiltonian of the form $H(p,q) = h(p) + \varepsilon f(p,q)$, where p, q are action angle variables and ε a small parameter, to find a canonical transformation which gives the new Hamiltonian, Z say, a "simple" form (the so called *normal form*). A rather detailed discussion is required to explain what is meant by "simple". In the most naive approach one tries to remove the dependence on the angle variables, so that Z depends only on the action variables, and the system turns out to be integrable. But such an approach turns out to be too naive: the existence of resonances, which manifests itself through the appearance of small denominators, makes the procedure inconsistent, in general. This was first stated by Poincaré in his celebrated theorem on the nonexistence of uniform integrals (see Poincaré, 1892, chapt. V).

However, it is a common procedure in Celestial mechanics (and more generally in perturbation theory) to perform a partial normalization of the system. The naive argument is more or less the following. Let us consider the Hamiltonian above, and truncate the Fourier expansion of $f(p,q)$ in the angles to some (not too high) order. Then we can find a canonical transformation which gives the Hamiltonian a normal form up to terms of order ε^2, i.e., (denoting by p', q' the new variables) the form $H'(p',q') = h(p') + \varepsilon\overline{f}(p') + \varepsilon^2 f'(p,q,\varepsilon)$. Here, $\overline{f}$ is in fact the average of f with respect to the angles. Forget now the terms of order ε^2; then we have an integrable system, the solutions of which can be considered as reliable up to a time of the order of ε^{-2}. If we iterate the same procedure a finite number of times, r say, then we obtain an Hamiltonian which is integrable up to remainder terms of order ε^r. This gives results which are reliable up to a time of the order of ε^{-r}.

As is, this argument is wrong, for two reasons. The first reason is that the integrable normal form can be obtained only in regions of the phase space which are far from resonances among the frequencies. This fact is well known, and leads to theories adapted to resonant situations. The second, and more hidden, reason is that when one says "a remainder of order ε^r" one actually means "a term of the form $\varepsilon^r f^{(r)}(p,q,\varepsilon)$, with some function $f^{(r)}$". But nothing is said about the *size* of the function $f^{(r)}$. In fact, according to quantitative estimates, the size of $f^{(r)}$ grows typically as fast as some power of $r!$, which makes the series to diverge.

Now, the underlying idea of Nekhoroshev's theorem is to use a standard argument in the theory of asymptotic expansions. In its simplest formulation the argument is the following. Forget for a moment the problems due to the resonances; let the size of the remainder after r normalization steps be $r!\,\varepsilon^r$, and compare it with the size $(r-1)!\varepsilon^{r-1}$ after $r-1$ steps; remark that $r!\varepsilon^r = r\varepsilon \cdot (r-1)!\varepsilon^{r-1}$. Thus, performing the step r is useful (i.e., the size of the remainder actually decreases) only if $r\varepsilon < 1$. This means on the one hand that the normalization procedure is useful only if ε is small enough, and on the other hand that there is an optimal choice of the number r of normalization steps as a function of ε, namely $r \simeq 1/\varepsilon$. Replace this value in $r!\varepsilon^r$, also using the Stirling's formula $r! \simeq r^r e^{-r}$; this gives $r!\varepsilon^r \simeq e^{-r} \simeq \exp(-1/\varepsilon)$. The final conclusion is that one has results which are reliable up to a time of the order of $\exp(1/\varepsilon)$. In the words of Littlewood, "if not eternity, this is a considerable slice of it".

Taking into account the resonances is in fact, in my opinion, the main contribution of Nekhoroshev. This requires the construction of a kind of geography of the phase space, with some suitable properties. This part is left for the detailed discussion in sect. 3. The final result can be roughly formulated as follows: the action variables satisfy the bound

$$|p(t) - p(0)| < C\varepsilon^a \quad \text{for} \quad |t| < T_* \exp\left[\left(\frac{\varepsilon_*}{\varepsilon}\right)^a\right] ,$$

with constants C, T_* and ε_*, and with typically $a \sim 1/n$, where n is the number of degrees of freedom.

The lectures are organized in two parts. The first part has a tutorial character. Its goal is to introduce the technical tools of quantitative perturbation theory. I will first recall, at a formal level, the definition of Lie series and of Lie transform. Next, I will introduce some elements of the theory of analytic functions, with particular attention to the Cauchy estimates. Finally, I will show how all these things work together by giving the conditions for the convergence of the Lie series. The second part will be devoted to a general discussion of the theorem of Nekhoroshev. In order to avoid unnecessary complications of a purely technical character, I will give the proof for a particular, but relevant, system, which contains almost all the essential elements of the theorem. This part is based on a recent paper by Zehnder and myself (see Giorgilli and Zehnder, 1992); thus, I will stress the relevant points of the proof, without insisting on technical details that can be found in that paper.

2. Lie series and Lie transforms

The use of Lie series and Lie transforms is quite recent in Celestial Mechanics, and is generally known at a formal level. I give here a quantitative formulation of the theory.

2.1 Formal definitions

On a $2n$–dimensional phase space endowed with canonical coordinates p, q, consider an analytic function $\chi(p,q)$, which will be called a *generating function.* The *Lie derivative* L_χ is a linear differential operator acting on functions on the phase space, defined as

$$L_\chi \cdot = \{\chi, \cdot\} \ . \tag{1}$$

This is nothing but the derivative along the Hamiltonian field generated by χ. The *Lie series* operator is then defined as the exponential of L_χ, namely

$$\exp(L_\chi) = \sum_{s\geq 0} \frac{1}{s!} L_\chi^s \tag{2}$$

(see for example Gröbner, 1960). It is known that this operator represents the time one evolution of the canonical flow generated by the autonomous Hamiltonian χ.

As an example, consider the case of action–angle variables $p \in \mathbf{R}^n$ and $q \in \mathbf{T}^n$. The function $\chi = \sum_l \xi_l q_l$, with $\xi \in \mathbf{R}^n$ generates through the Lie series operator the canonical transformation $p' = \exp(L_\chi)p = p + \xi$, $q' = \exp(L_\chi)q = q$, namely a translation in the action space. Similarly, a generating function $\chi = \chi(q)$ (independent of p) generates the transformation $p' = p + \frac{\partial \chi}{\partial q}$, $q' = q$, namely a deformation of the action variables.

The *Lie transform* differs from the Lie series in that it is connected with the flow of a *nonautonomous* (i.e., time dependent) Hamiltonian. In such a form, it has been introduced by Hori (1966) and Deprit (1969). In fact, several algorithms have been devised in order to give the Lie transform an algorithmic recursive form: everybody, of course, has his favourite one. To make a definite choice, I will make use here of my favourite one, which is related to the "algorithm of the inverse", introduced by Henrard. Consider a *generating sequence* $\chi = \{\chi_s\}_{s\geq 1}$ of analytic functions on the phase space. The Lie transform operator T_χ is defined as

$$T_\chi = \sum_{s\geq 0} E_s \ , \tag{3}$$

where the sequence $\{E_s\}_{s\geq 0}$ of operators is recursively defined as

$$E_0 = \mathrm{Id} \ , \quad E_s = \sum_{j=1}^{s} \frac{j}{s} L_{\chi_j} E_{s-j} \ . \tag{4}$$

As a simple example, consider the case $\chi = \{\chi_1, 0, 0, \ldots\}$, namely a generating sequence containing only the first term. Then the Lie transform generated by χ coincides with the Lie series generated by χ_1, i.e.,

$$T_\chi = \exp(L_{\chi_1}) \ .$$

From now on, it will be useful to look at the algebraic aspect of the definitions of Lie series and Lie transform, rather than to their property of being related to the canonical flow of some Hamiltonian. In particular, all the properties stated here can be proved on a purely algebraic basis, without any reference to the canonical flow (see for example Giorgilli and Galgani, 1978). The relevant properties are the following: both the Lie series and the Lie transform operators are linear, and preserve products and Poisson brackets, namely

$$T_\chi(fg) = (T_\chi f)(T_\chi g) \ , \quad T_\chi \{f,g\} = \{T_\chi f, T_\chi g\} \ ,$$

with analogous formulæ for the Lie series. Moreover, both the operators are invertible. Finding the inverse of the Lie series is an easy matter: recalling that it represents the flow of an autonomous canonical system, one immediately concludes that the inverse of $\exp(L_\chi)$ is $\exp(L_{-\chi}) = \exp(-L_\chi)$, i.e., the generating function of the inverse is $-\chi$. Concerning the Lie transform, a careful analysis of the algorithm(3)–(4) shows that finding the inverse transformation is an easy matter, although finding the generating sequence of the inverse is not as easy. Indeed, assume we are given a function $f = f_0 + f_1 + \ldots$, and denote by $g = g_0 + g_1 + \ldots$ its transformed function $g = T_\chi f$; using the linearity of the Lie transform, we can apply T_χ separately to every term of f. It is useful to rearrange the terms according to the triangular diagram

$$\begin{array}{cccccc}
g_0 & f_0 & & & & \\
 & \downarrow & & & & \\
g_1 & E_1 f_0 & f_1 & & & \\
 & \downarrow & \downarrow & & & \\
g_2 & E_2 f_0 & E_1 f_1 & f_2 & & \\
 & \downarrow & \downarrow & \downarrow & & \\
g_3 & E_3 f_0 & E_2 f_1 & E_1 f_2 & f_3 & \\
 & \downarrow & \downarrow & \downarrow & \downarrow & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array} \tag{5}$$

where terms of the same order appear on the same line. Remark that the operator T_χ acts *by columns*, as indicated by the arrows: the knowledge of f_j and of the generating sequence allows one to construct the whole column below f_j. Thus, the first line gives g_0, the second line gives $g_1 = E_1 f_0 + f_1$, and so on. This shows how to practically perform the transformation. Concerning the inverse, assume that g is given, and f is unknown. Then, the first line gives $f_0 = g_0$; having determined f_0, all the column below f_0 can be constructed, and the second line gives immediately $f_1 = g_1 - E_1 f_0$; having determined f_1, all the corresponding column can be constructed, so that f_2 can be determined from the third line as $f_2 = g_2 - E_2 f_0 - E_1 f_1$, and so on. In fact, there is also an explicit formula for the inverse, namely

$$T_\chi^{-1} = \sum_{s\geq 0} D_s \ ,$$

where

$$D_0 = \mathrm{Id} \ , \quad D_s = -\sum_{j=1}^{s} \frac{j}{s} D_{s-j} L_{\chi_j} \ .$$

This expression is actually useless for a practical computation: the algorithm described above is much more efficient. Nevertheless, the explicit recursive formula is useful for quantitative estimates.

The property of preserving the Poisson brackets is relevant, as it means that the coordinate transformation

$$p = T_\chi p' \ , \quad q = T_\chi q' \ ,$$

where the generating sequence χ depends on the "new" variables p', q', is canonical (I omit here the analogous statement concerning the Lie series). Now, if we consider a function $f(p, q)$ on the phase space, we should in principle compute the transformed function, $f'(p', q')$ say, as

$$f'(p', q') = f(p, q)\big|_{p=T_\chi p', q=T_\chi q'} \ ,$$

or, equivalently,

$$f'(p', q') = f(T_\chi p', T_\chi q') \ ,$$

namely by substitution of the coordinate transformation. Now, the properties of being linear and of preserving products can be used to prove that in fact one has

$$f(T_\chi p', T_\chi q') = (T_\chi f)(p', q') \ . \tag{6}$$

Briefly, "don't try to make substitutions: just transform the function, and change the name of the variables".

2.2 Analytic framework

The purpose of this section is to introduce the basic elements of the quantitative perturbation theory. Most of the technical elements are concerned with the theory of analytic functions, and so with functions defined on complex domains.

First of all, to avoid unnecessary complications of mathematical character, it is useful to have some precise characterization of the phase space. Thus, I shall consider the common case of action–angle variables $q \in \mathbf{T}^n$ and $p \in \mathcal{G} \in \mathbf{R}^n$, where $\mathcal{G}$ is a subset of $\mathbf{R}^n$.

Next, I will introduce a *complexification* of the domains. For $p \in \mathcal{G}$ consider the complex polydisk $\Delta_\varrho(p)$ of radius $\varrho > 0$ with center p defined as

$$\Delta_\varrho(p) = \{p' \in \mathbf{C}^n \ : \ |p' - p| < \varrho\} \ , \tag{7}$$

where

$$|p' - p| = \max_j \left|p'_j - p_j\right| \ ;$$

this is nothing but the cartesian product of complex disks of radius ϱ in the complex plane. The complexification $\mathcal{G}_\varrho$ of the real domain $\mathcal{G}$ is defined as

$$\mathcal{G}_\varrho = \bigcup_{p \in \mathcal{G}} \Delta_\varrho(p) \ . \tag{8}$$

Similarly, the complexification $\mathbf{T}^n_\sigma$ for $\sigma > 0$ of the n–torus is defined as

$$\mathbf{T}^n_\sigma = \{q \in \mathbf{C}^n \ : \ |\operatorname{Im} q| < \sigma\} \ , \tag{9}$$

namely the cartesian product of strips of width σ around the real axis on the complex plane. From now on, the phase space will be the domain

$$\mathcal{D}_{(\varrho,\sigma)} = \mathcal{G}_\varrho \times \mathbf{T}^n_\sigma \ .$$

The next step is concerned with the introduction of *norms* on the space of analytic functions on the domain $\mathcal{D}_{(\varrho,\sigma)}$. The most natural norm is the *supremum norm*. For a function f analytic on $\mathcal{G}_\varrho$ the supremum norm $|f|_\varrho$ of f is defined as

$$|f|_\varrho = \sup_{p \in \mathcal{G}_\varrho} |f(p)| \ . \tag{10}$$

Similarly, for a function g analytic on $\mathcal{D}_{(\varrho,\sigma)}$ the supremum norm $|g|_{(\varrho,\sigma)}$ is

$$|g|_{(\varrho,\sigma)} = \sup_{(p,q)\in\mathcal{D}_{(\varrho,\sigma)}} |g(p,q)| \ . \tag{11}$$

As a matter of fact, the supremum norm is not the best one, to my knowledge, when dealing with small denominators. A more useful norm is the so called *weighted Fourier norm*, which can be introduced on the basis of the following considerations. It is known that an analytic function f on the domain $\mathcal{D}_{(\varrho,\sigma')}$ admits the Fourier expansion

$$f(p,q) = \sum_{k\in\mathbf{Z}^n} f_k(p) e^{ik\cdot q} \ .$$

It is also known that the coefficients $f_k(p)$ of the Fourier expansion decay exponentially with k. More precisely, one has

$$|f_k|_\varrho \le |f|_{(\varrho,\sigma')} e^{-|k|\sigma'} \ ,$$

where $|k| = |k_1| + \cdots + |k_n|$. Thus, it is natural to define the weighted Fourier norm $\|f\|_{(\varrho,\sigma)}$ for $\sigma < \sigma'$ as

$$\|f\|_{(\varrho,\sigma)} = \sum_{k\in\mathbf{Z}^n} |f_k|_\varrho e^{|k|\sigma} \ . \tag{12}$$

The exponential decay of the coefficients and the condition $\sigma < \sigma'$ ensure that the series defining the norm converges. It is also an easy matter to check that the supremum norm is bounded by the weighted Fourier norm, i.e.,

$$|f|_{(\varrho,\sigma)} \le \|f\|_{(\varrho,\sigma)} \ .$$

2.3 Cauchy estimates

The need for analytic functions on complex domains is due to the availability of the Cauchy estimates for the derivatives of analytic functions. Let me briefly recall the essentials. For simplicity, consider a disk $\Delta_\varrho(0)$ centered at the origin of the complex plane $\mathbf{C}$. Consider a function f analytic and bounded on the disk $\Delta_\varrho(0)$. The Cauchy estimate for the derivative f' of f at the origin states that

$$|f'(0)| \le \frac{1}{\varrho}|f|_\varrho \ .$$

More generally, for the s–th derivative $f^{(s)}$ one has the estimate

$$|f^{(s)}(0)| \le \frac{s!}{\varrho^s}|f|_\varrho \ .$$

For instance, let $\varrho = 1$, and consider the function $f(z) = z^s$; it is an easy matter to check that $|f|_1 = 1$, so that the Cauchy estimate gives $|f^{(s)}(0)| \le s!$. This shows that the estimate cannot be improved in general. Similarly, consider a polydisk $\Delta_\varrho(0)$ centered at the origin of $\mathbf{C}^n$. Then one has the estimate

$$\left|\frac{\partial f}{\partial z_j}(0)\right| \le \frac{1}{\varrho}|f|_\varrho \ , \quad 1 \le j \le n \ .$$

I omit here the estimates for higher order derivatives, since we shall not need them.

Let us now come back to our analytic setting. Consider a function f analytic and bounded on the complex domain $\mathcal{G}_\varrho$. Consider now a point $p \in \mathcal{G}_{\varrho-\delta}$ for $0 < \delta \le \varrho$ (that is, the union of polydisks of radius $\varrho - \delta$ centered at every point of $\mathcal{G}$). Remark that the polydisk $\Delta_\delta(p)$ is a subset of $\mathcal{G}_\varrho$, so that f is analytic and bounded on $\Delta_\delta(p)$, and moreover we have the estimate $|f(p')| \le |f|_{\varrho}$ for all $p' \in \Delta_{\delta}(p)$. By the Cauchy estimate, we immediately get

$$\left|\frac{\partial f}{\partial p_j}(p)\right| \le \frac{1}{\delta}|f|_{\varrho}\ , \quad 1 \le j \le n\ .$$

Since this is true for every point $p \in \mathcal{G}_{\varrho-\delta}$, we immediately get

$$\left|\frac{\partial f}{\partial p_j}\right|_{\varrho-\delta} \le \frac{1}{\delta}|f|_{\varrho}\ , \quad 1 \le j \le n\ .$$

The relevant fact here is that we have an explicit bound on the derivative of a function, but we must pay this information with a restriction of the complex domain.

2.4 Generalized Cauchy estimates

The next step is concerned with the generalization of the Cauchy estimates, taking into account that we are interested in Lie derivatives (or Poisson brackets), and that we are going to use the weighted Fourier norm instead of the supremum one.

A first result concerns the Poisson bracket between two functions, and the derivative of a function with respect to one of the canonical coordinates (i.e., the Poisson bracket with the conjugate one).

Lemma 1: *Let f be analytic on the domain $\mathcal{D}_{(\varrho,\sigma)}$, and g be analytic in $\mathcal{D}_{(1-d')(\varrho,\sigma)}$ for some $0 \le d' < 1$; assume moreover that $\|f\|_{(\varrho,\sigma)}$ and $\|g\|_{(1-d')(\varrho,\sigma)}$ are finite. Then:*

i. for $0 < d < 1$ and for $1 \le j \le n$ one has

$$\left\|\frac{\partial f}{\partial p_j}\right\|_{(1-d)(\varrho,\sigma)} \le \frac{1}{d\varrho}\|f\|_{(\varrho,\sigma)}\ , \quad \left\|\frac{\partial f}{\partial q_j}\right\|_{(1-d)(\varrho,\sigma)} \le \frac{1}{ed\sigma}\|f\|_{(\varrho,\sigma)}\ ; \tag{13}$$

ii. for $0 < d < 1 - d'$ one has

$$\|\{f,g\}\|_{(1-d'-d)(\varrho,\sigma)} \le \frac{2}{ed(d+d')\varrho\sigma}\|f\|_{(\varrho,\sigma)}\|g\|_{(1-d')(\varrho,\sigma)}\ . \tag{14}$$

Proof. Using the Fourier expansion, compute

$$\frac{\partial f}{\partial p_j} = \sum_k \frac{\partial f_k}{\partial p_j}(p)e^{ik\cdot q}\ .$$

According to the definition of the norm, we have

$$\left\|\frac{\partial f}{\partial p_j}\right\|_{(1-d)(\varrho,\sigma)} = \sum_k \left|\frac{\partial f_k}{\partial p_j}\right|_{(1-d)\varrho} e^{(1-d)|k|\sigma}\ ;$$

by the Cauchy estimate, and taking out d from the exponent, the right hand side of the latter expression is bounded by

$$\frac{1}{d\varrho}\sum_k |f_k|_{\varrho}e^{|k|\sigma}\ ,$$

and the first of (13) follows in view of the definition of the norm. Coming to the second one, compute

$$\frac{\partial f}{\partial q_j} = i\sum_k k_j f_k(p)e^{ik\cdot q}\ ,$$

and use again the definition of the norm to estimate

$$\left\| \frac{\partial f}{\partial q_j} \right\|_{(1-d)(\varrho,\sigma)} = \sum_k |k_j|\, |f_k|_{(1-d)\varrho} e^{(1-d)|k|\sigma} \le \sum_k |f_k|_\varrho e^{|k|\sigma} |k| e^{-d|k|\sigma} \ .$$

By the general inequality

$$x^\alpha e^{-\delta x} \le \left(\frac{\alpha}{e\delta}\right)^\alpha \qquad \text{for positive } \alpha,\ x,\ \delta \tag{15}$$

(just compute the supremum over x of $x^\alpha e^{-\delta x}$), with $\alpha = 1$ and $|k|$ and $d\sigma$ in place of x and δ respectively, one gets $|k|e^{-d|k|\sigma} \le 1/(ed\sigma)$; the second of (13) then follows in view of the definition of the norm. The proof of (14) requires some more work. Compute

$$\{f,g\} = i \sum_{k,k'} \left[\sum_{l=1}^{n} \left(k_l \frac{\partial g_{k'}}{\partial p_l} f_k + k'_l \frac{\partial f_k}{\partial p_l} g_{k'} \right) \right] e^{ik\cdot q} e^{ik'\cdot q} \ ,$$

and use the definition of the norm to estimate

$$\| \{f,g\} \|_{(1-d'-d)(\varrho,\sigma)} < \sum_{k,k'} \left[\sum_{l=1}^{n} \left(|k_l| \left| \frac{\partial g_{k'}}{\partial p_l} \right|_{(1-d'-d)\varrho} |f_k|_\varrho + |k'_l| \left| \frac{\partial f_{k'}}{\partial p_l} \right|_{(1-d'-d)\varrho} |g_k|_{(1-d')\varrho} \right) \right] e^{(1-d'-d)|k+k'|\sigma} \ .$$

Using now the Cauchy estimates, the right hand side of the latter expression is found to be smaller than

$$\begin{aligned} &\frac{1}{d\varrho} \sum_{k,k'} |g'_k|_{(1-d')\varrho} e^{(1-d')|k'|\sigma} |f_k|_\varrho e^{|k|\sigma} \sum_{l=1}^{n} |k_l| e^{-(d'+d)|k|\sigma} \\ &\quad + \frac{1}{(d'+d)\varrho} \sum_{k,k'} |g_{k'}|_{(1-d')\varrho} e^{(1-d')|k'|\sigma} |f_k|_\varrho e^{|k|\sigma} \sum_{l=1}^{n} |k'_l| e^{-d|k'|\sigma} \ . \end{aligned} \tag{16}$$

Remark now that the only terms depending on the index l are $\sum_l |k_l| = |k|$ and $\sum_l |k'_l| = |k'|$; moreover, using the inequality (15) one concludes

$$|k| e^{-(d'+d)|k|\sigma} \le \frac{1}{(d'+d)e\sigma} \ , \quad |k'| e^{-d|k'|\sigma} \le \frac{1}{de\sigma} \ .$$

Reordering the terms, the expression in (16) is found to be smaller than

$$\frac{2}{ed(d'+d)\varrho\sigma} \sum_{k'} |g'_k|_{(1-d')\varrho} e^{(1-d')|k'|\sigma} \cdot \sum_k |f_k|_\varrho e^{|k|\sigma} \ ,$$

and the conclusion follows in view of the definition of the norms $\|f\|_{(\varrho,\sigma)}$ and $\|g\|_{(1-d')(\varrho,\sigma)}$. *Q.E.D.*

The next step concerns the estimate of multiple Poisson brackets, namely of an expression of the form $L^s_\chi f$. Since $L^s_\chi = L_\chi \circ L^{s-1}_\chi$, one can iterate s times the estimate of lemma 2.1. However, one must take into account the restriction of the domains. To this end, having fixed the final restriction d, we choose s positive quantities $\delta_1, \ldots, \delta_s$ the sum of which is d, and estimate $L^j_\chi f$ on the domain restricted by the factor $1 - \delta_1 - \ldots - \delta_j$. The simplest choice is $\delta_1 = \ldots = \delta_s = d/s$. The final result is given by

Lemma 2: *Let f and χ be analytic on the domain $\mathcal{D}_{(\varrho,\sigma)}$, and assume that $\|f\|_{(\varrho,\sigma)}$ and $\|\chi\|_{(\varrho,\sigma)}$ are finite. Then for every positive $d<1$ one has*

$$\left\|L_\chi^s f\right\|_{(1-d)(\varrho,\sigma)} \le \frac{s!}{e^2}\left(\frac{2e}{d^2\varrho\sigma}\right)^s \|\chi\|_{(\varrho,\sigma)}^s \|f\|_{(\varrho,\sigma)} \tag{17}$$

Proof. For $s=1$ this is nothing but lemma 1. For $s>1$ let $\delta=d/s$. A straightforward application of lemma 1 gives

$$\|L_\chi f\|_{(1-\delta)(\varrho,\sigma)} \le \frac{2}{e\delta^2\varrho\sigma}\|\chi\|_{(\varrho,\sigma)}\|f\|_{(\varrho,\sigma)} \ .$$

For $j=2,\ldots,s$, recalling that $L_\chi^j f$ can be estimated on the domain restricted by the factor $1-j\delta$, use again lemma 1 with δ in place of d and $(j-1)\delta$ in place of d', and get the recursive estimate

$$\left\|L_\chi^j f\right\|_{(1-j\delta)(\varrho,\sigma)} \le \frac{2}{ej\delta^2\varrho\sigma}\|\chi\|_{(\varrho,\sigma)}\|L_\chi^{j-1} f\|_{(1-(j-1)\delta)(\varrho,\sigma)} \ .$$

By recursive application of this formula up to s, and recalling $\delta=d/s$ we get

$$\left\|L_\chi^s f\right\|_{(1-d)(\varrho,\sigma)} \le \left(\frac{2}{ed^2\varrho\sigma}\right)^s \frac{s^{2s}}{s!}\|\chi\|_{(\varrho,\sigma)}^s\|f\|_{(\varrho,\sigma)} \ .$$

Then, (17) follows in view of the trivial inequality $s^s \le e^{s-1}s!$ for $s\ge 1$ (prove it by induction). *Q.E.D.*

2.5 Analyticity of the Lie series and of the Lie transform

The Lie series and the Lie transform have been defined as series of analytic functions. It is natural to ask whether these series are convergent, or, more precisely, whether the sums of these series are analytic functions on some domain. Sufficient conditions can be found on the basis of the results of the previous section. This is stated by

Proposition 1: *Let f and χ be analytic on the domain $\mathcal{D}_{(\varrho,\sigma)}$, and assume that $\|f\|_{(\varrho,\sigma)}$ and $\|\chi\|_{(\varrho,\sigma)}$ are finite. Then for every positive $d<1/2$ the following statement holds true: if the condition*

$$\frac{2e\|\chi\|_{(\varrho,\sigma)}}{d^2\varrho\sigma} \le \frac{1}{2} \tag{18}$$

is satisfied, then the operator $\exp(L_\chi)$ and its inverse $\exp(-L_\chi)$ define an analytic canonical transformation on the domain $\mathcal{D}_{(1-d)(\varrho,\sigma)}$ with the properties

$$\begin{aligned}\mathcal{D}_{(1-2d)(\varrho,\sigma)} &\subset \exp(L_\chi)\mathcal{D}_{(1-d)(\varrho,\sigma)} \subset \mathcal{D}_{(\varrho,\sigma)} \ ,\\ \mathcal{D}_{(1-2d)(\varrho,\sigma)} &\subset \exp(-L_\chi)\mathcal{D}_{(1-d)(\varrho,\sigma)} \subset \mathcal{D}_{(\varrho,\sigma)} \ .\end{aligned}$$

The proof of the proposition is based on the following

Lemma 3: *With the hypotheses of proposition 1 the series $\exp(\pm L_\chi)f$, $\exp(\pm L_\chi)p$ and $\exp(\pm L_\chi)q$ are absolutely convergent on $\mathcal{D}_{(1-d)(\varrho,\sigma)}$, and for integer $r>0$ one has the estimates*

$$\begin{aligned}&\|\exp(\pm L_\chi)p - p\|_{(1-d)(\varrho,\sigma)} \le d\varrho \ ,\\ &\|\exp(\pm L_\chi)q - q\|_{(1-d)(\varrho,\sigma)} \le d\sigma \ ,\\ &\left\|\exp(\pm L_\chi)f - \sum_{s=0}^{r}\frac{1}{s!}L_\chi^s f\right\|_{(1-d)(\varrho,\sigma)} \le \frac{2}{e^2}\left(\frac{2e\|\chi\|_{(\varrho,\sigma)}}{d^2\varrho\sigma}\right)^{r+1}\|f\|_{(\varrho,\sigma)} \ .\end{aligned} \tag{19}$$

Proof. By lemma 2 one has

$$\sum_{s\geq 0}\frac{1}{s!}\left\|L_\chi^s f\right\|_{(1-d)(\varrho,\sigma)} \leq \|f\|_{(\varrho,\sigma)} + \frac{1}{e^2}\sum_{s>0}\left(\frac{2e\|\chi\|_{(\varrho,\sigma)}}{d^2\varrho\sigma}\right)^s \|f\|_{(\varrho,\sigma)}\ ;$$

in view of condition (18) the series in the left hand side of the latter inequality is a convergent geometric series. Since the supremum norm is bounded by the weighted Fourier norm, this implies that the series $\exp(L_\chi)f$ is absolutely convergent on the domain $\mathcal{D}_{(1-d)(\varrho,\sigma)}$, as claimed. The third of (19) follows by simply computing the sum of the geometric series above starting from $s = r+1$ and taking into account the condition (18). Concerning the estimate on the transformation of the coordinates p, q, a similar argument applies, with a minor difference: the first terms $L_\chi p$ and $L_\chi q$ must be estimated using (13) of lemma 1; then lemma 2 is used to estimate the remaining terms. Then the same convergence argument is used. *Q.E.D.*

Proof of proposition 1. By lemma 3, the series of analytic functions defining the canonical transformation are absolutely convergent in $\mathcal{D}_{(1-d)(\varrho,\sigma)}$. Thus, the same series are uniformly convergent on every compact subset of $\mathcal{D}_{(1-d)(\varrho,\sigma)}$. By Weierstrass theorem, this implies that the sums of the series are analytic functions on $\mathcal{D}_{(1-d)(\varrho,\sigma)}$, as claimed. The statement concerning the inclusions of the domains follows from the estimates (19). *Q.E.D.*

A straightforward consequence of proposition 1 is the following: if $f(p,q)$ is analytic in $\mathcal{D}_{(\varrho,\sigma)}$, then the transformed function $f\big(\exp(L_\chi)p, \exp(L_\chi)q)\big)$ is analytic in $\mathcal{D}_{(1-d)(\varrho,\sigma)}$, being a composition of analytic functions. However, as has been remarked at the end of sect. 2.1, the transformed function is nothing but $\exp(L_\chi)f$. The last estimate (19) gives in fact a direct proof of the analyticity of $\exp(L_\chi)f$. Moreover, taking into account that an explicit computation must be truncated at some order r, it gives an estimate of the error.

A similar result holds for the Lie transform. I will give here only the statement, without entering the details of the proof. The technical tools of the proof are again the results of sect. 2.4; however, some additional complication arises, due to the fact that the definition of the Lie transform is given in recursive form. The scheme of the proof can be found in (Giorgilli and Galgani, 1985), and requires only some trivial adaptation, due to the use of a different norm.

Proposition 2: *Let f and the generating sequence $\chi = \{\chi_s\}_{s\geq 1}$ be analytic on the domain $\mathcal{D}_{(\varrho,\sigma)}$, and assume that $\|f\|_{(\varrho,\sigma)}$ is finite, and that*

$$\|\chi_s\|_{(\varrho,\sigma)} \leq \frac{b^{s-1}}{s}G$$

for some $b\geq 0$ and $G>0$. Then for every positive $d<1/2$ the following statement holds true: if the condition

$$\frac{2eG}{d^2\varrho\sigma} + b \leq \frac{1}{2} \tag{20}$$

is satisfied, then the operator T_χ and its inverse T_χ^{-1} define an analytic canonical transformation on the domain $\mathcal{D}_{(1-d)(\varrho,\sigma)}$ with the properties

$$\mathcal{D}_{(1-2d)(\varrho,\sigma)} \subset T_\chi\mathcal{D}_{(1-d)(\varrho,\sigma)} \subset \mathcal{D}_{(\varrho,\sigma)}\ ,$$
$$\mathcal{D}_{(1-2d)(\varrho,\sigma)} \subset T_\chi^{-1}\mathcal{D}_{(1-d)(\varrho,\sigma)} \subset \mathcal{D}_{(\varrho,\sigma)}\ .$$

Remark that the statement is essentially the same as for the Lie series. The difference is due to the use of a generating sequence instead of a function, which requires a condition on the behaviour of the sequence, and a corresponding minor modification in the convergence condition (20).

3. The Nekhoroshev's theorem

In this section I will give the lines of the proof of the theorem of Nekhoroshev. In order to have a well definite framework, consider the special nonautonomous canonical system with Hamiltonian

$$H(p,q,t) = \sum_{j=1}^{n} \frac{p_j^2}{2} + V(q,t) \ , \tag{21}$$

where $p \in \mathbf{R}^n$ and $q \in \mathbf{T}^n$ are action angle variables. The potential energy $V(q,t)$ is assumed to be real analytic and bounded in the complex strip $|\,\mathrm{Im}\, x| < 2\sigma$ and $|\,\mathrm{Im}\, t| < 2\sigma$ for some positive σ. The potential energy must be periodic in the angles, while no condition, apart from analyticity, is assumed for the dependence on time.

Such a formulation of the problem might look obscure in view of the absence of a small perturbation parameter. To clarify this point, one should observe that the potential energy in the Hamiltonian (21) is small in case the kinetic energy is large. The usual approach is recovered by rescaling the action variables and the time via the transformation

$$q = q' \ , \quad p = \frac{p'}{\varepsilon} \ , \quad t' = \varepsilon t \ ,$$

and considering the Hamiltonian

$$H'(q',p',t') = \sum_j \frac{{p'_j}^2}{2} + \varepsilon^2 V(q', t'/\varrho) \ .$$

This model has been studied in (Giorgilli and Zehnder, 1992), so that it is not necessary to give here a detailed exposition. Rather, I will try to bring into evidence some central ideas of the proof, without insisting on technical details.

3.1 Local normal form

Let me first recall in a few words the usual formulation, using the formalism of the Lie series. Starting with an Hamiltonian $H(p,q) = h(p) + \varepsilon f(p,q)$, one looks for a generating function $\chi(p,q)$ such that the transformed Hamiltonian

$$\exp(L_\chi) H = h_0 + \varepsilon\,(f + \{\chi, h_0\}) + \varepsilon^2 \ldots$$

has some desired property (e.g., it is integrable) up to terms of order ε^2. This means that one must solve for χ and an auxiliary function Z the equation

$$L_{h_0}\chi + Z = f \ ,$$

with some condition on Z (e.g., one requires $Z = Z(p)$). The latter equation is solved as follows. Since the q variables are angles, expand f in Fourier series as

$$f(p,q) = \sum_{k \in \mathbf{Z}^n} f_k(p) \exp(ik \cdot q) \ ,$$

with known coefficients $f_k(p)$, and consider the same expansion for χ and Z, with unknown coefficients $c_k(p)$ and $z_k(p)$ respectively. Then the equation above splits into the infinite system of equations for the coefficients of the Fourier expansion

$$ik \cdot \omega(p) c_k(p) + z_k(p) = f_k(p) \ , \quad \omega(p) = \frac{\partial h_0}{\partial p} \ .$$

where $\omega(p)$ are the frequencies of the unperturbed system. The key point here is that the coefficient $c_k(p)$ should be determined as

$$c_k(p) = -i\frac{f_k(p)}{k\cdot\omega(p)} \ ;$$

this is formally consistent only if the denominator $k\cdot\omega(p)$ does not vanish. Thus, the simplest rule seems to be the following: if $k\cdot\omega = 0$, then put $z_k = f_k$ and $c_k = 0$; else put $z_k = 0$ and determine c_k as above.

However, a careful consideration shows that such an approach is too naive. Indeed, the set $\omega \in \mathbf{R}^n$ such that $k\cdot\omega = 0$ for some nonzero $k \in \mathbf{Z}^n$ is clearly dense in $\mathbf{R}^n$; thus, unless some strong restriction is imposed on h_0 and/or on f, one has to expect that arbitrarily close to every value p of the actions there is some value, p', say, such that at least one denominator $k\cdot\omega(p')$ does vanish. This is indeed, in very rough terms, the basis of the Poincaré's theorem on the nonexistence of uniform integrals. The way out of these troubles consists in making some truncation of the Fourier expansion of the Hamiltonian (in fact, it is a common practice to consider only a *finite* set of Fourier components), and constructing a *local* theory (i.e., one restricts the action variables to some suitable domain).

A more precise formulation of these remarks is based on the following technical definitions.

(i) A resonance module is defined as a subgroup $\mathcal{M} \in \mathbf{Z}^n$ satisfying $\mathrm{span}(\mathcal{M}) \cap \mathbf{Z}^n = \mathcal{M}$; here, both $\mathcal{M}$ and $\mathbf{Z}^n$ are considered as subset of $\mathbf{R}^n$, and $\mathrm{span}(\mathcal{M})$ is the linear subspace in $\mathbf{R}^n$ spanned by $\mathcal{M}$.

(ii) A function $Z(p,q)$ is said to be in *normal form with respect to the resonance module* $\mathcal{M}$ in case its Fourier expansion has the form

$$Z(p,q) = \sum_{k\in\mathcal{M}} z_k(p)\exp(ik\cdot q) \ ,$$

namely, if it contains only harmonics belonging to $\mathcal{M}$ (resonant harmonics).

(iii) A set $\mathcal{G} \subset \mathbf{R}^n$ is said to be a nonresonance domain of type $(\mathcal{M},\alpha,\varrho,N)$ in case one has

$$|k\cdot\omega| > \alpha \text{ for all } p \in \mathcal{G}_\varrho \ , \quad k \in \mathbf{Z}^n \setminus \mathcal{M} \text{ and } |k| \le N \ .$$

Here, $|k| = |k_1| + \ldots + |k_n|$, α and ϱ are positive parameters, N is a positive integer, $\mathcal{M}$ is a resonance module, and $\mathcal{G}_\varrho$ is the complex extension of the real domain $\mathcal{G}$, as described in sect. 2.2.

The nonresonance domain appears as the subset of the action space where the Hamiltonian can be given a normal form with respect to the given resonance module. The only point that needs perhaps to be clarified is related to the use of a *finite* number harmonics in the Fourier expansion of the Hamiltonian. This can be justified by recalling that the coefficients of the Fourier expansion of an analytic function decay exponentially (this has been discussed in sect. 2.2). Thus, one can first consider only the harmonics k satisfying the condition $|k| \le K$ for some K, and perform the first normalization step; then one can perform a second step considering harmonics up to order $2K$, and so on up to a *finite* number r of normalization steps, which takes into consideration harmonics up to order rK. On a nonresonance domain with $N = rK$ the procedure can be consistently performed. A formal statement concerning the Hamiltonian system (21) is given by the following

Proposition 3: *Let r and K be positive integers, and $N = rK$. Moreover, let $\mathcal{G} \subset \mathbf{R}^n$ be a nonresonance domain of type $(\mathcal{M},\alpha,\varrho,N)$. Then, there are positive constants A*

and B depending on $|V|_\sigma$, σ, ϱ and n such that if

$$\varepsilon := \left(\frac{rA}{\alpha} + 2e^{-K\sigma/2}\right) \le \frac{1}{2} \tag{22}$$

then there exists a real analytic symplectic diffeomorphism φ such that φ and φ^{-1} are defined on the complex domain $\mathcal{G}_{(\varrho,\sigma)\frac{1}{2}}$ and satisfy

$$\mathrm{dist}(p,\varphi(p)) \le \frac{1}{4}\varrho\ , \quad \mathrm{dist}(p,\varphi^{-1}(p)) \le \frac{1}{4}\varrho\ . \tag{23}$$

The map φ transforms the Hamiltonian H into the following normal form on $\mathcal{G}_{\frac{1}{2}(\varrho,\sigma)}$:

$$H\circ\varphi = \frac{1}{2}\sum_j p_j^2 + Z_{\mathcal{M}} + \mathcal{R}\ , \tag{24}$$

where $Z_{\mathcal{M}}$ is in normal form with respect to the resonance module $\mathcal{M}$. The remainder $\mathcal{R}$ satisfies, on $\mathcal{G}_{\frac{1}{2}(\varrho,\sigma)}$, the estimate

$$|\mathcal{R}| \le B\varepsilon^r\ . \tag{25}$$

Moreover:

$$\mathcal{G}_{(\varrho,\sigma)\frac{1}{4}} \subset \varphi\big(\mathcal{G}_{(\varrho,\sigma)\frac{1}{2}}\big) \subset \mathcal{G}_{(\varrho,\sigma)\frac{3}{4}}\ . \tag{26}$$

In addition, the same statement holds true for φ^{-1} instead of φ. The constants are given by

$$A = \frac{2^5 en}{\varrho\sigma}\left[\left(\frac{1+e^{-\sigma/2}}{1-e^{-\sigma/2}}\right)^n |V|_{2\sigma} + \frac{e\varrho}{2}\right]$$
$$B = 2\left(\frac{1+e^{-\sigma/2}}{1-e^{-\sigma/2}}\right)^n |V|_{2\sigma}\ .$$

The local normal form allows us to have some insight on the dynamics in the nonresonance domain. Indeed, forget for a moment the remainder $\mathcal{R}$, and write the Hamilton's equations for the normalized part of the Hamiltonian: it is immediately seen that $\dot p$ is expressed as a sum of vectors belonging to span($\mathcal{M}$). The resulting picture is the following. Consider the nonresonance domain endowed with the new coordinates (i.e., the coordinates giving the Hamiltonian the normal form). Through any point p_0 belonging to the resonance domain draw the plane generated by span($\mathcal{M}$): this will be called the *plane of fast drift*. Then the actions p evolve essentially along this plane, in the sense that the motion transversal to the plane is slow, being due only to the remainder $\mathcal{R}$. Try now to describe the motion in the original coordinates. The plane of fast drift is replaced by a surface which can be seen as generated by a *deformation* of a plane, due to the coordinate transformation; the size of such a deformation is estimated by (23). Thus, the orbit is contained in a kind of *cylinder* having the plane of fast drift as axis, and with radius estimated by (23); a small increase of the radius of the cylinder (e.g., multiply it by 2) takes into account the effect of the drift for a long time.

Thus, we know the behaviour of the orbit as far as it remains in the nonresonance domain. The natural question is: what happens if the motion along the plane of fast drift moves the orbit out of the nonresonance domain? The answer to this question is the achievement of the geometric part of the Nekhoroshev's theorem.

3.2 The geometric part

Let the domain $\mathcal{G}$ of the action variables be an open subset of $\mathbf{R}^n$, and assume that for some positive ϱ the set $\mathcal{G} - \varrho$ (i.e., the points $\in \mathcal{G}$ such that the ball of radius ϱ and center p is still contained in $\mathcal{G}$) is not empty. In the case of the Hamiltonian (21) the choice of $\mathcal{G}$ is quite arbitrary; it can be taken for instance to be a ball of radius larger than ϱ, or even the whole space $\mathbf{R}^n$.

Fix now a positive integer N, to be identified with the final truncation order rK of the analytic part. Let $\mathcal{M}$ be a resonance module with dimension s, and consider the set $\mathcal{M}_N = \{k \in \mathcal{M} \,:\, |k| \le N\}$; we call $\mathcal{M}$ an N–module in case $\mathcal{M}_N$ contains s independent vectors k. In view of the results of the analytic part, we shall not need to consider resonance moduli which are not N–moduli. To a resonant N–module of dimension s we associate a real parameter β_s, with $\beta_0 < \ldots < \beta_n$, to be determined.

The aim of the geometric construction is to cover $\mathcal{G} - \varrho$ with nonresonance domains to which the results of the analytic part apply. This is the most difficult part. I give here the definitions for a generic unperturbed Hamiltonian $H_0(p)$; it is an useful exercise to consider the simple case of the Hamiltonian (21) for $n = 2$.

1. The *resonant zone* $\mathcal{Z}_{\mathcal{M}}$ associated to a N–module $\mathcal{M}$ of dimension s is defined as the set of points $p \in \mathcal{G} - \varrho$ satisfying $|k \cdot \omega(p)| < \beta_s$ for s independent vectors of $\mathcal{M}_N$, namely, the points close "within β_s" to resonance. The dimension s of $\mathcal{M}$ will be called the multiplicity of the resonance, or also, by abuse of language, the multiplicity of the zone. The union of all the resonant zones with the same multiplicity s will be denoted by $\mathcal{Z}^*_s$, and will be called the *resonant region of order* s. In rough terms, $\mathcal{Z}^*_s$ is the set of points which admit at least s resonances within β_s; further resonances are not excluded till now. In particular, resonant zones $\mathcal{Z}_{\mathcal{M}}$ corresponding to different $\mathcal{M}$'s can intersect, giving rise to regions characterized by resonances of higher multiplicity.
2. A resonant block $\mathcal{B}_{\mathcal{M}}$ associated to $\mathcal{M}$ of dimension s is defined as the set $\mathcal{B}_{\mathcal{M}} = \mathcal{Z}_{\mathcal{M}} \setminus \mathcal{Z}^*_{s+1}$. In view of this definition the block appears as the set of points admitting s resonances within β_s, but no further resonances; it is a good basis for the construction of nonresonant domains.
3. To every point p of a given resonant block $\mathcal{B}_{\mathcal{M}}$ we associate the plane in $\mathbf{R}^n$ passing through p and parallel to span$(\mathcal{M})$ (i.e., the plane of fast drift), and define the *extended plane* $P_{\mathcal{M},\delta_s}(y)$ as the set of points of $\mathbf{R}^n$ the distance of which from the plane above is less than $\delta_s = \beta_s/(2N)$. The *cylinder* $\mathcal{C}_{\mathcal{M},\delta_s}$ is defined as the intersection $\mathcal{P}_{\mathcal{M},\delta_s} \cap \mathcal{Z}_{\mathcal{M}}$ of the extended plane with the corresponding resonant zone, and the bases of the cylinder are defined as the intersection $\mathcal{P}_{\mathcal{M},\delta_s} \cap \partial\mathcal{Z}_{\mathcal{M}}$ of the extended plane with the border of the zone.
4. Finally, the extended resonant block $\mathcal{B}_{\mathcal{M},\delta_s}$ is defined as the union of the cylinders corresponding to all points of the block $\mathcal{B}_{\mathcal{M}}$.

Having completed this construction, one proves that with an hypothesis of convexity on the unperturbed Hamiltonian H_0 and with the constants $\beta_0, \ldots, \beta_n$ defined as

$$\beta_0 = \left[2^{n+1}(n+1)!\, N^{(n^2+n-2)/2}\right]^{-1} \varrho \ ,$$
$$\beta_s = 2^s s!\, N^{[s(s-1)]/2} \beta_0 \quad \text{for } 1 \le s \le n \ ,$$

one has the following properties.

(i) The blocks are a covering of the domain $\mathcal{G} - \varrho$; that is, every point in $\mathcal{G} - \varrho$ belongs to at least one block.

(ii) The extended block $\mathcal{B}_{\mathcal{M},\delta_s}$ is a nonresonance domain of type $(\mathcal{M}, \beta_s/2, \delta_s, N)$ (remark that $\delta_s < \varrho$).

(iii) The complement of a resonant region $\mathcal{Z}_s^*$ of multiplicity s is covered by the union of all block of multiplicity less than s; that is, if a point does not admit a resonance of multiplicity s, then it belongs to at least one block of multiplicity less than s.
(iv) The closure of an extended block $\mathcal{B}_{\mathcal{M},\delta_s}$ and the resonant zone $\mathcal{Z}_{\mathcal{M}'}$ corresponding to a different N–module with the same dimension as $\mathcal{M}$ are disjoint; that is, the basis of a cylinder $\mathcal{C}_{\mathcal{M},\delta_s}(p)$ corresponding to a point $p \in \mathcal{B}_{\mathcal{M}}$ does not belong to any resonant zone of the same multiplicity $\dim(\mathcal{M})$.
(v) The diameter of a cylinder does not exceed ϱ.

Property (ii) means that the extended block are the domains where the analytic theory of the first part can be consistently applied; in particular, setting $N = rK$ with a given K allows to perform r normalization steps in every block. Then one has the following picture of the motion: take an initial value p_0 for the actions; find a resonant block $\mathcal{B}_{\mathcal{M}}$ to which p_0 belongs (may be it is not unique, but this is not relevant), and consider the corresponding cylinder $\mathcal{C}_{\mathcal{M},\delta_s}(p_0)$. Then the orbit either is confined inside the cylinder up to a time of the order of ε^{-r}, or it leaves the cylinder through a basis. In the latter case the intersection of the orbit with the basis of the cylinder is a point not belonging to the resonant region of multiplicity s, by property (iv); moreover, by property (iii), this point belongs to a block of multiplicity less than s. It is now easy to conclude that there is a cylinder containing the orbit for a time interval of the order of ε^{-r}. For, assume that the orbit visits several blocks, and determine the block of minimal multiplicity; pick as initial point any point of the orbit belonging to that block, and consider the corresponding cylinder; this is the wanted cylinder, since the orbit could leave it only by entering a block of lower multiplicity, thus contradicting the assumption that the multiplicity of the block was minimal. Property (v) gives then a bound on the time evolution of the actions p, which can not exceed the diameter of the cylinder. The formal statement for the case of the Hamiltonian (21) is the following.

Proposition 4: *Let $\mathcal{G}$ and ϱ be given, with $\mathcal{G} - 2\varrho$ nonempty; let r, K be positive integers, and $N = rK$; let moreover A and B be defined as in proposition 3, with ϱ replaced by δ_s, and let*

$$\varepsilon_0 = \frac{2rA}{\beta_0} + 2e^{-K\sigma/2} .$$

Then for every solution with $p(0) \in \mathcal{G} - 2\varrho$ one has

$$\mathrm{dist}(p(t), p(0)) < \varrho \quad \text{if} \quad |t| \le T_0^* = \frac{2\sigma}{B}\varepsilon_0^{-r} ,$$

provided $\varepsilon_0 \le 1/2$.

3.3 The exponential estimate

The proposition 4 gives a stability result depending on several parameters. The parameters σ, $|V|_{2\sigma}$ and ϱ depend on the analyticity properties of the Hamiltonian and on the choice of the domain of initial data, while the constants β_s and δ_s are given by the geometric part; these parameters can be considered as the natural ones of the system. The parameters K, namely the truncation of the Fourier expansion of the Hamiltonian, and r, the order of normalization, are instead arbitrary. So, it is natural to look for an optimal choice of these parameters, i.e., the choice giving the best estimate of the stability time. The key point is that the first term in the definition of ε_0, namely $2rA/\beta_0$, behaves roughly as $\varrho^{-1}r^a$, with some constant a (here the value of β_0 given by the geometric part must be used); thus, the estimate of the stability time has the form $r^{ar}\varrho^{-r}$. Following the argument given in the introduction, it is natural to set $N = (\varrho_0/\varrho)^{1/a}$ with a dimensional constant ϱ_0. Then, one determines K in such a way that the two contributions to ε_0 are of the same order. This immediately allows to realize that the stability time is indeed exponential in $(\varrho_0/\varrho)^{1/a}$. A more careful determination of the constants is just matter of æstetics, and leads to the final formulation of the stability theorem for the system (21).

Theorem 1: *Let $\varphi^t\bigl(q(0),p(0)\bigr) = \bigl(q(t),p(t)\bigr)$ be the flow of the Hamiltonian vector field (21) on $\mathbf{T}^n \times \mathbf{R}^n$. Assume the potential $V(q,t)$ is real analytic on $\mathbf{T}^n \times \mathbf{R}$ and has a bounded analytic extension to a complex strip $2\sigma > 0$. Then there are two positive constants ϱ_* and T_* depending on $|V|_{2\sigma}$, σ and the dimension n, such that for $\varrho \geq \varrho_*$ one has*

$$\mathrm{dist}\bigl(p(t),p(0)\bigr) < \varrho$$

for all t in

$$|t| \leq T_* \exp\left(\frac{\varrho}{\varrho_*}\right)^{1/a} \quad \text{with} \quad a = \frac{n^2+n}{2} \ .$$

The constants are given by

$$T_* = \frac{\sigma}{8e^2}$$

$$\varrho_* = K_*^a \, \frac{2^{n+3}(n+1)!}{e} \left(\frac{1+e^{-\sigma/2}}{1-e^{-\sigma/2}}\right)^n |V|_{2\sigma}$$

where K_* is the smallest positive integer satisfying

$$K_* \geq \frac{2^5 e^4 n}{\sigma} \left(\frac{1-e^{-\sigma/2}}{1+e^{-\sigma/2}}\right)^n \frac{1}{|V|_{2\sigma}} \quad \text{and} \quad K_* \geq \frac{5}{\sigma} \ .$$

References

Deprit, A. (1969): "Canonical transformations depending on a small parameter", *Cel. Mech.* **1**, 12–30.

Gröbner, W. (1960): *Die Lie–Reihen und Ihre Anwendungen*, Springer Verlag, Berlin; it. transl.: *Le serie di Lie e le loro applicazioni*, Cremonese, Roma (1973).

Giorgilli, A. and Galgani, L. (1978): "Formal integrals for an autonomous Hamiltonian system near an equilibrium point", *Cel. Mech.* **17**, 267–180.

Giorgilli, A. and Galgani, L. (1985): "Rigorous estimates for the series expansions of Hamiltonian perturbation theory", *Cel. Mech.* **37**, 95–112.

Kolmogorov, A.N. (1954): *On the preservation of conditionally periodic motions*, *Doklady Akademia Nauk SSSR*, **96**, 527–530.

Hori, G. (1966):: "Theory of general perturbations with unspecified canonical variables", *Publ. Astron. Soc. Japan*, **18**, 287–296.

Littlewood, J.E. (1959): "On the equilateral configuration in the restricted problem of three bodies", *Proc. London Math. Soc.* **9**, 343–372;

Littlewood, J. E. (1959): "The Lagrange configuration in celestial mechanics", *Proc. London Math. Soc.* **9**, 525–543.

Moser, J., (1955): "Stabilitätsverhalten kanonisher differentialgleichungssysteme", *Nachr. Akad. Wiss. Göttingen, Math. Phys.* **K1 IIa**, nr.6, 87–120.

Nekhoroshev, N., N., (1977): "Exponential estimates of the stability time of near–integrable Hamiltonian systems.", *Russ. Math. Surveys*, **32**, 1-65.

Nekhoroshev, N., N., (1979): "Exponential estimates of the stability time of near–integrable Hamiltonian systems, 2.", *Trudy Sem. Petrovs.*, **5**, 5-50.

Poincaré, H. (1892): *Le méthodes nouvelles de la méanique célèste*, Gauthier–Villard, Paris.