

CLASSICAL CONSTRUCTIVE METHODS IN KAM THEORY

ANTONIO GIORGILLI

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

Abstract. Two methods for constructing quasiperiodic solutions as expansion in a small parameter are discussed. The first one is the classical Lindstedt's method; the second one an algorithm based on Kolmogorov's paper [1]. Besides a complete formulation of the algorithms, an overview of the main ideas leading to the proof of convergence of the expansions is given. Some comparison is also made, including in particular the analysis of the effectiveness of the algorithms.

1. Overview

I shall discuss here the construction of quasi periodic solutions for the nearly integrable Hamiltonian system

$$(1) \quad H(p, q, \varepsilon) = \frac{1}{2} \langle Cp, p \rangle + \varepsilon f(q) ,$$

where $p \in \mathbf{R}^n$ and $q \in \mathbf{T}^n$ are action-angle variables, C is a $n \times n$ real symmetric matrix and ε is a small parameter. This is a particular case of the class of models considered by Thirring^[2]. In the following I shall also assume that the perturbation $f(q)$ is a trigonometric polynomial of degree K for some positive K . The reason for considering such a simple model is that it contains all the difficulties of the same problem for a general Hamiltonian without extra complications of purely technical nature.

The problem of constructing quasiperiodic solutions has a very long tradition connected with the search for motions of the planetary system that can be represented as purely trigonometric series in the time t (i.e., with no secular terms that are linear or, more generally, polynomial in t). From a theoretical viewpoint this problem has been solved by Kolmogorov^[1], Arnold^[3] and Moser^[4]. However, there is a quite wide gap separating the theoretical results from the actual applications to the dynamics of the solar system, even considering simplified models that include only the major planets. Roughly speaking, the problem is that the applicability of Kolmogorov's theorem

is assured only if the perturbation is small enough, but if one tries to determine how small is “small enough” the conclusions are usually far from being realistic. Up to my knowledge, the best significant attempt to show the applicability of KAM theorem to the simplified model of the Sun–Jupiter–Saturn system has been made by Laskar and Robutel^[5]. However, it is still far from being a realistic result because the applicability of KAM theory is assured only if the masses are by some orders of magnitude smaller than the real ones.

Improving the estimates usually requires a big computational effort in order to explicitly construct the expansions required by perturbation theory. This, of course, stimulates the search for improved explicit methods. On the other hand, achieving a realistic result requires also a careful analysis of the convergence of the perturbation expansions given by the algorithms taken into consideration.

I will concentrate here on two algorithms for constructing quasiperiodic solutions that have been recently reconsidered by some authors. The first method goes back to Lindstedt, and consists in looking directly for the solution as a power expansion in a perturbation parameter the coefficients of which are trigonometric functions of the angles. The second method uses the idea of Kolmogorov of looking for a particular normal form of the Hamiltonian. The actual implementation is slightly different from Kolmogorov’s one, but this is just a technical matter. I will call these methods Lindstedt’s algorithm and Kolmogorov’s algorithm, respectively.

I will pay particular attention to the formal aspect of the algorithms, i.e., the explicit formulæ that allow one to write the expansions in explicit form, possibly with the help of an algebraic manipulator. For both algorithms there exist convergence proofs; I will give here a short sketch, referring for details to other works. My aim is only to outline the main points of the proofs. Finally, I will add a short section concerning the relations between the two methods.

2. Formal aspects

In this section I shall discuss the formal algorithms for constructing the wanted quasiperiodic solutions. The more intriguing question concerning the convergence is deferred to the next section.

2.1 Lindstedt’s series

The starting point of this discussion can be found in Poincaré’s *Méthodes Nouvelles*^[6], Tome II, Chapters IX and XIII. In Chapter IX Poincaré introduces the method of Lindstedt, considering only the *formal* aspect of the construction of series expansions in the small parameter ε that represent quasi periodic solutions.

In the case of the Hamiltonian (1) we can proceed as follows. Considering first

the unperturbed system, $\varepsilon = 0$, let us determine an invariant torus characterized by nonresonant frequencies. That is: determine \bar{p} such that the real vector $\omega = C\bar{p}$ satisfies $\langle k, \omega \rangle \neq 0$ for all non zero $k \in \mathbf{Z}^n$. By the canonical transformation

$$(2) \quad p' = p - \bar{p}, \quad q' = q$$

the Hamiltonian (1) is given the form (omitting primes and an unessential constant)

$$(3) \quad H(p, q, \varepsilon) = \langle \omega, p \rangle + \frac{1}{2} \langle Cp, p \rangle + \varepsilon f(q).$$

This transformation has no special meaning: it just gives the Hamiltonian a form that will be useful in the rest of these notes.

Let us write the canonical equations for the Hamiltonian (3), namely

$$(4) \quad \dot{q} = \omega + Cp, \quad \dot{p} = -\varepsilon \frac{\partial f}{\partial q}.$$

Following Lindstedt and Poincaré, we look for a solution of equations (4) of the form

$$(5) \quad q(t) = \omega t + \varepsilon \varphi^{(1)}(\omega t) + \varepsilon^2 \varphi^{(2)}(\omega t) + \dots, \quad p(t) = \varepsilon \psi^{(1)}(\omega t) + \varepsilon^2 \psi^{(2)}(\omega t) + \dots,$$

where $\varphi^{(s)}$ and $\psi^{(s)}$ for $s > 0$ are vector functions. To this end, let us substitute the latter formulæ in (4). Then we expand $f(\omega t + \varepsilon \varphi^{(1)}(\omega t) + \varepsilon^2 \varphi^{(2)}(\omega t) + \dots)$ in Taylor series in ε and compare the coefficients of the same degree in ε . It is convenient to consider the functions $\varphi^{(s)}$ and $\psi^{(s)}$ as depending on the dummy angle variables $(\vartheta_1, \dots, \vartheta_n)$, and then set $\vartheta = \omega t$. With this notation the time derivative turns out to be $\frac{d}{dt} = \partial_\omega := \langle \omega, \frac{\partial}{\partial \vartheta} \rangle$. Let us denote by

$$(6) \quad \Delta_j(m, s) = \sum_{\substack{l_1 > 0, \dots, l_m > 0 \\ l_1 + \dots + l_m = s}} \varphi_j^{(l_1)} \cdot \dots \cdot \varphi_j^{(l_m)}, \quad 0 < m \leq s,$$

the coefficient of ε^s in $(\varepsilon \varphi_j^{(1)} + \varepsilon^2 \varphi_j^{(2)} + \dots)^m$. For $m = 0$ just set $\Delta_j(0, 0) = 1$ and $\Delta_j(0, s) = 0$ for $s > 0$. With this setting we get the system of equations

$$(7) \quad \begin{aligned} \partial_\omega \psi^{(r)} &= - \sum_{\substack{m \in \{0, \dots, r-1\}^n \\ 0 \leq |m| \leq r-1}} \frac{\partial^{m_1 + \dots + m_n}}{\partial \vartheta_1^{m_1} \dots \partial \vartheta_n^{m_n}} \frac{\partial f}{\partial \vartheta}(\omega t) \sum_{\substack{s_1 \geq m_1, \dots, s_n \geq m_n \\ |s| = r-1}} \prod_{j=1}^n \frac{\Delta_j(m_j, s_j)}{m_j!}, \\ \partial_\omega \varphi^{(r)} &= C\psi^{(r)}, \end{aligned}$$

where $|m| = m_1 + \dots + m_n$. For instance, starting with $r = 1$ we get the equations

$$(8) \quad \partial_\omega \psi^{(1)} = -\frac{\partial f}{\partial \vartheta}, \quad \partial_\omega \varphi^{(1)} = C\psi^{(1)}.$$

Remark that the r.h.s. of the first equation has zero average in ϑ . Considering the Fourier expansion of $f(\vartheta)$, namely

$$(9) \quad f(\vartheta) = \sum_{k \in \mathbf{Z}^n} \alpha_k \exp(i\langle k, \vartheta \rangle)$$

with known coefficients α_k we immediately get a solution

$$(10) \quad \begin{aligned} \psi^{(1)} &= - \sum_k \frac{k \neq 0}{\langle k, \omega \rangle} \alpha_k \exp(i\langle k, \vartheta \rangle) , \\ \varphi^{(1)} &= i \sum_{k \neq 0} \frac{Ck}{\langle k, \omega \rangle^2} \alpha_k \exp(i\langle k, \vartheta \rangle) . \end{aligned}$$

The denominators in the expression above never vanish in view of the nonresonance condition that we have assumed on the frequencies ω . For $r > 1$ one can proceed recursively, remarking that the r.h.s. of equations (7) depend only on quantities that have already been determined. However, a problem of *formal* consistency appears here: the equations (7) can be solved only if the average of their r.h.s. is zero. This problem is pointed out by Poincaré in chapter IX.

The consistency of the construction above is easily established if one assumes, as in [7], that the perturbation $f(q)$ is an even function of the angles q , i.e., $f(-q) = f(q)$. The proof goes as follows. Let $r = 1$. From (8) we see that the r.h.s. of the equation for $\psi^{(1)}$ is an odd function of the angles, and so $\psi^{(1)}$ is even (remark that the operator ∂_ω is a derivative with respect to the angles, and so it changes the parity of the function). Hence the r.h.s. of the equation for $\varphi^{(1)}$ is even, and so $\varphi^{(1)}$ can be determined as an odd function. We prove now by recursion that we can construct a solution of (7) such that $\psi^{(s)}$ is even and $\varphi^{(s)}$ is odd for all $s > 0$. The r.h.s. of the equation for $\psi^{(s)}$ contains $|m| + 1$ derivatives of the even function f , and so it is even for $|m|$ odd and odd for $|m|$ even. On the other hand the product of the functions Δ_j involves exactly $|m|$ odd functions, and so it is even for $|m|$ even and odd for $|m|$ odd. We conclude that the whole expression is written as the product of two functions of different parity, which gives an odd result, and its average must vanish because it is an odd function. This allows us to determine $\psi^{(r)}$ as an even function with zero average. Putting it in the second equation (7) we determine $\varphi^{(r)}$ as an odd function, i.e., with zero average. This proves our claim. Remark that the choice of determining φ with zero average makes the construction unique.

If we remove the parity hypothesis then some extra argument is necessary in order to prove the consistency of the construction. Poincaré's argument in chapter IX makes use of what is usually called Birkhoff's normal form. I omit here this part of the discussion, replacing it with the more recent approach due to Kolmogorov.

2.2 Kolmogorov's normal form

The basis for this section is the short note [1] published by Kolmogorov in 1954. The method exposed here differs from the original one of Kolmogorov only because I shall reformulate his algorithm using the classical expansions in the parameter ε . Moreover, I shall also make use of the formalism of Lie series in order to perform a near to identity canonical transformation (see for instance [8] for a very detailed discussion of the method, or [9] for a short introduction specialized for the canonical case).

Kolmogorov's suggestion is to use the well established formalism of canonical transformations in order to give the Hamiltonian (3) the special form (hereafter called the *Kolmogorov's normal form*)

$$(11) \quad H(p, q, \varepsilon) = \langle \omega, p \rangle + R(p, q, \varepsilon), \quad R(p, q, \varepsilon) = O(p^2),$$

i.e., $R(p, q)$ is a power series in p starting with quadratic terms. In our case this is actually a homogeneous polynomial of degree 2 in the action variables p , with coefficients depending on q and ε .

The canonical equations for the Hamiltonian (11) are

$$(12) \quad \dot{q} = \omega + \frac{\partial R}{\partial p}, \quad \dot{p} = -\frac{\partial R}{\partial q},$$

with $\frac{\partial R}{\partial p} = O(p)$ and $\frac{\partial R}{\partial q} = O(p^2)$. Hence it is evident that the orbit with initial conditions $p = 0$ and arbitrary q lies on the invariant torus $p = 0$, and is characterized by frequencies ω .

The idea of replacing the Birkhoff's normal form (used by Poincaré) with the weaker form above turns out to be crucial in the discussion of the convergence of the series.

Following Kolmogorov, we perform an infinite sequence of near to identity canonical transformations which bring the Hamiltonian to Kolmogorov's normal form order by order in ε . As it will be evident from the algorithm, the property of the Hamiltonian (1) of being at most quadratic in the actions p is preserved by the transformation.

Let us write the Hamiltonian after r normalization steps in the form

$$(13) \quad H^{(r)} = \sum_i \omega_i p_i + \sum_{s=0}^r \varepsilon^s h^{(s)} + \sum_{s>r} \varepsilon^s \left(f_0^{(r,s)} + f_1^{(r,s)} + f_2^{(r,s)} \right),$$

where $h^{(s)}$ is a homogeneous polynomial of degree 2 in the actions p and a trigonometric polynomial of degree sK in the angles q , and the functions $f_l^{(r,s)}$ are homogeneous polynomials of degree l in the actions and trigonometric polynomials of degree sK in the angles. We can always assume that $\langle f_0^{(r,s)} \rangle = 0$, where $\langle \cdot \rangle$ denotes averaging over all angles. The upper index r refers to the current iteration of the normalization algorithm, and is missing in $h^{(s)}$ because this is the part of the Hamiltonian already in normal

form. The original Hamiltonian (3) has clearly this form: just put $r = 0$, and identify $h^{(0)} = \langle Cp, p \rangle / 2$, $f_0^{(0,1)} = f(q)$. All the rest of the expansion is zero.

Assuming that we have performed $r - 1$ normalization steps, thus obtaining $H^{(r-1)}$ as in (13), we perform two canonical transformations with generating functions $\chi_1^{(r)}(q) = X^{(r)}(q) + \langle \xi^{(r)}, q \rangle$ and $\chi_2^{(r)}(p, q)$ respectively, where $X^{(r)}$ does not depend on p , $\xi^{(r)}$ is a real vector, and $\chi_2^{(r)}$ is linear in p . Performing the transformation in two separate steps allows us to easily control the degree in the actions of every term in the expansion.

With the first transformation we compute an intermediate Hamiltonian $\hat{H}^{(r)} = \exp(\varepsilon^r L_{\chi_1^{(r)}}) H^{(r-1)}$ of the form

$$(14) \quad \hat{H}^{(r)} = \sum_i \omega_i p_i + \sum_{s=0}^{r-1} \varepsilon^s h^{(s)} + \sum_{s \geq r} \varepsilon^s \left(\hat{f}_0^{(r,s)} + \hat{f}_1^{(r,s)} + \hat{f}_2^{(r,s)} \right) ,$$

where $\chi_1^{(r)} = X^{(r)} + \langle \xi^{(r)}, q \rangle$ has to be determined so that $\hat{f}_0^{(r,r)} = 0$ and $\hat{f}_1^{(r,r)}$ has zero average over the angles. Then we perform the second transformation, computing $H^{(r)} = \exp(\varepsilon^r L_{\chi_2^{(r)}}) \hat{H}^{(r)}$ of the form (13), where $\chi_2^{(r)}$ has to be determined with the condition that the linear term $\hat{f}_1^{(r,r)}$ is removed.

The generating functions are determined by the equations

$$(15) \quad \partial_\omega X^{(r)} + f_0^{(r-1,r)} = 0 ,$$

$$(16) \quad C \langle \xi^{(r)}, p \rangle + \langle f_1^{(r-1,r)} \rangle = 0 ,$$

$$(17) \quad \partial_\omega \chi_2^{(r)} + \{X^{(r)}, h^{(0)}\} + f_1^{(r-1,r)} - \langle f_1^{(r-1,r)} \rangle = 0 ,$$

where $\partial_\omega \cdot = \{\cdot, \langle \omega, p \rangle\}$ and $\langle \cdot \rangle$ denotes averaging over the angles q . Moreover, the normalized term $h^{(r)}$ in the Hamiltonian is given by

$$h^{(r)} = L_{\chi_2^{(r)}} h^{(0)} + f_2^{(r-1,r)} .$$

Looking at the explicit expressions of the exponential operators $\exp(\varepsilon^r L_{\chi_j^{(r)}})$ one immediately sees that terms of the same degree in ε and in the actions p are easily isolated, so that one can obtain explicit recursive formulæ for all functions $f_l^{(r,s)}$ in the expression (13). Here are all formulæ. The intermediate Hamiltonian $\hat{H}^{(r)}$ in the form (14) is

defined by

$$\begin{aligned}
 \hat{f}_0^{(r,r)} &= 0 ; \\
 \hat{f}_0^{(r,r+m)} &= f_0^{(r-1,r+m)} && \text{for } 0 < m < r ; \\
 \hat{f}_0^{(r,2r+m)} &= \frac{1}{2} L_{\chi_1^{(r)}}^2 h^{(m)} + L_{\chi_1^{(r)}} f_1^{(r-1,r+m)} + f_0^{(r-1,2r+m)} && \text{for } 0 \leq m < r ; \\
 (18) \quad \hat{f}_0^{(r,kr+m)} &= \frac{1}{2} L_{\chi_1^{(r)}}^2 f_2^{(r-1,(k-2)r+m)} + L_{\chi_1^{(r)}} f_1^{(r-1,(k-1)r+m)} + f_0^{(r-1,kr+m)} \\
 &&& \text{for } k > 2 \text{ and } 0 \leq m < r ; \\
 \hat{f}_1^{(r,r+m)} &= L_{\chi_1^{(r)}} h^{(m)} + f_1^{(r-1,r+m)} && \text{for } 0 \leq m < r ; \\
 \hat{f}_1^{(r,kr+m)} &= L_{\chi_1^{(r)}} f_2^{(r-1,(k-1)r+m)} + f_1^{(r-1,kr+m)} && \text{for } k \geq 2 \text{ and } 0 \leq m < r ; \\
 \hat{f}_2^{(r,s)} &= f_2^{(r-1,s)} && \text{for } s \geq r .
 \end{aligned}$$

The final Hamiltonian $H^{(r)}$ is defined by $h^{(r)} = L_{\chi_2^{(r)}} h^{(0)} + f_2^{(r-1,r)}$, and

$$\begin{aligned}
 f_0^{(r,kr+m)} &= \sum_{j=0}^{k-1} \frac{1}{j!} L_{\chi_2^{(r)}}^j \hat{f}_0^{(r,(k-j)r+m)} && \text{for } k \geq 1, 0 \leq m < r, kr+m > r ; \\
 f_1^{(r,kr)} &= \frac{k-1}{k!} L_{\chi_2^{(r)}}^{k-1} \hat{f}_1^{(r,r)} + \sum_{j=0}^{k-2} \frac{1}{j!} L_{\chi_2^{(r)}}^j \hat{f}_1^{(r,(k-j)r)} && \text{for } k \geq 2 ; \\
 (19) \quad f_1^{(r,kr+m)} &= \sum_{j=0}^{k-1} \frac{1}{j!} L_{\chi_2^{(r)}}^j \hat{f}_1^{(r,(k-j)r+m)} && \text{for } k \geq 1, 0 < m < r ; \\
 f_2^{(r,kr+m)} &= \frac{1}{k!} L_{\chi_2^{(r)}}^k h^{(m)} + \sum_{j=0}^{k-1} \frac{1}{j!} L_{\chi_2^{(r)}}^j \hat{f}_2^{(r,(k-j)r+m)} \\
 &&& \text{for } k \geq 1, 0 \leq m < r, kr+m > r .
 \end{aligned}$$

Understanding these formulæ takes some time and a lot of patience. A complete deduction is given in [10]. Let us examine in detail the case $r = 1$ in order to illustrate the procedure. Recall that for $r = 0$ we have

$$H^{(0)} = \langle \omega, p \rangle + h^{(0)} + \varepsilon f_0^{(0,1)}$$

with

$$h^{(0)} = \frac{1}{2} \langle Cp, p \rangle, \quad f_0^{(0,1)} = f(q) .$$

We first determine the generating function $\chi_1^{(1)} = X^{(1)} + \langle \xi^{(1)}, q \rangle$ from (15) and (16),

namely by solving the equations

$$(20) \quad \partial_\omega X^{(1)} + f_0^{(0,1)} = 0, \quad C\langle \xi, p \rangle + \langle f_1^{(0,1)} \rangle = 0.$$

In view of $f_1^{(0,1)} = 0$ the second equation has the trivial solution $\xi = 0$. The first equation is solved in the usual way (see the discussion above concerning the solution of eq. (8), which is similar to this one). The solution is made unique by the condition that $X^{(1)}$ has zero average. Having determined $\chi_1^{(1)}$ we calculate

$$\begin{aligned} \hat{H}^{(1)} &= \exp(\varepsilon L_{\chi_1^{(1)}} H^{(0)}) \\ &= \langle \omega, p \rangle + \frac{1}{2} \langle Cp, p \rangle + \varepsilon \left(f_0^{(0,1)} + \partial_\omega X^{(1)} + \frac{1}{2} L_{\chi^{(1)}} \langle Cp, p \rangle \right) \\ &\quad + \varepsilon^2 \frac{1}{4} L_{\chi^{(1)}}^2 \langle Cp, p \rangle. \end{aligned}$$

Using (20), and by comparison with (14) we get

$$(21) \quad \hat{f}_1^{(1,1)} = \left\langle C \frac{\partial \chi_1^{(1)}}{\partial q}, p \right\rangle, \quad \hat{f}_0^{(1,2)} = \frac{1}{2} \left\langle C \frac{\partial \chi_1^{(1)}}{\partial q}, \frac{\partial \chi_1^{(1)}}{\partial q} \right\rangle,$$

the rest of the expansion being zero. This completely determines $\hat{H}^{(1)}$.

We proceed now by determining $\chi_2^{(1)}$. To this end we solve the equation

$$(22) \quad \partial_\omega \chi_2^{(1)} + \left\langle C \frac{\partial X_1^{(1)}}{\partial q}, p \right\rangle = 0.$$

The r.h.s. of this equation has zero average; hence it can be uniquely solved with the condition that $\chi_2^{(1)}$ has zero average, too. The transformed Hamiltonian is determined as

$$\begin{aligned} H^{(1)} &= \exp(\varepsilon L_{\chi_2^{(1)}} \hat{H}^{(1)}) \\ &= \langle \omega, p \rangle + \frac{1}{2} \langle Cp, p \rangle + \sum_{s \geq 1} \varepsilon^s \left(\frac{1}{s!} L_{\chi_2^{(1)}}^{s-1} \partial_\omega \chi_2^{(1)} + \frac{1}{(s-1)!} L_{\chi_2^{(1)}}^{s-1} \hat{f}_1^{(1,1)} \right) \\ &\quad + \frac{1}{2} \sum_{s \geq 1} \frac{\varepsilon^s}{s!} L_{\chi_2^{(1)}}^s \langle Cp, p \rangle + \sum_{s \geq 0} \frac{\varepsilon^{s+2}}{s!} L_{\chi_2^{(1)}}^s \hat{f}_0^{(1,2)}. \end{aligned}$$

Using again (17) it is an easy matter to write the explicit expressions of $h^{(1)}$, $f_0^{(1,s)}$, $f_1^{(1,s)}$ and $f_2^{(1,s)}$ so that the transformed Hamiltonian $H^{(1)}$ takes the form (13). This completes the first step, giving us a Hamiltonian in Kolmogorov's normal form up to order 1.

By iteration of the algorithm above, we can construct the generating functions $\chi_1^{(2)}, \chi_2^{(2)}, \dots, \chi_1^{(r)}, \chi_2^{(r)}$ up to any wanted order r in ε . At every step s , the transformed Hamiltonians $\hat{H}^{(s)}$ and $H^{(s)}$ must be computed up to terms of order ε^r . At first sight, this seems to make the algorithm of normal form much worse than the algorithm of Lindstedt's series. I will discuss this point in sect 4.

Having determined the generating function, we construct the equations of the torus by transforming the coordinates. Precisely, we construct explicitly the canonical transformation to Kolmogorov's normal form up to order r in ε as

$$(23) \quad \begin{aligned} p &= \exp(\varepsilon^r L_{\chi_2^{(r)}}) \circ \exp(\varepsilon^r L_{\chi_1^{(r)}}) \circ \dots \circ \exp(\varepsilon L_{\chi_2^{(1)}}) \circ \exp(\varepsilon L_{\chi_1^{(1)}}) p' , \\ q &= \exp(\varepsilon^r L_{\chi_2^{(r)}}) \circ \dots \circ \exp(\varepsilon L_{\chi_2^{(1)}}) q' , \end{aligned}$$

where the generating functions are considered as functions of p', q' , and all contributions of order higher than r in ε are neglected. This gives us a transformation of the form

$$q = q' + \varepsilon \varphi^{(1)}(q') + \dots + \varepsilon^r \varphi^{(r)}(q') , \quad p = p' + \varepsilon \psi^{(1)}(p', q') + \dots + \varepsilon^r \psi^{(r)}(p', q') .$$

Remark that the transformations with the generating functions $\chi_1^{(s)}$ leave the angles unchanged; for this reason in the second formula (23) only $\chi_2^{(s)}$ appears.

The solutions (5) are recovered as follows. Forgetting terms of order higher than r in ε , the solutions of the canonical equations for the Hamiltonian in Kolmogorov's normal form are

$$(24) \quad q'(t) = \omega t + q'_0 , \quad p'(t) = 0 ,$$

where q'_0 is the arbitrary vector of the initial phases. Since the functions (5) are solutions of the same canonical equations, they must coincide with those in (24) for some suitable choice of the initial phases q'_0 . The comparison is actually straightforward if one assumes that the Hamiltonian is even in the angles. Indeed, in such a case the functions $\varphi^{(1)}, \dots, \varphi^{(r)}$ above are odd in the angles, so that the identification is trivially made by just putting the initial phases equal to zero. By the way, this proves the consistency of Lindstedt's construction in all cases.

3. Convergence

The question of convergence of the series expansions discussed above is a very old one, being related to the old question about the possibility of expressing the solutions of the equations for the planetary system by series where the time appears only as argument of trigonometric function. That is, as quasiperiodic functions. The difficulty is due to the so called "small divisors" that are introduced by the solution of eq. (7) in Lindstedt's method and of eqs. (15) and (16) in Kolmogorov's method. For, in view of

the nonresonance condition on the frequencies that we have assumed the denominators $\langle k, \omega \rangle$ never vanish for $k \neq 0$; however, they can assume arbitrarily small values.

3.1 Convergence of Lindstedt's series

To start with, we can still refer to the discussion about the ‘‘Divergence des s eries de M. Lindstedt’’, in chap. XIII of Poincar e’s *M ethodes Nouvelles*. We can find there some useful indications that are at the basis of the solution of the problem.

The first remark made by Poincar e is that the small divisors introduced by the solution of the equation

$$(25) \quad \partial_\omega \chi = f(q) := \sum_{0 \neq k \in \mathbf{Z}^n} c_k \exp(i \langle k, q \rangle) ,$$

where $f(q)$ is a known periodic function with zero average, do not prevent the convergence of the formal solution

$$(26) \quad \chi = -i \sum_{0 \neq k \in \mathbf{Z}^n} \frac{c_k}{\langle k, \omega \rangle} \exp(i \langle k, q \rangle)$$

provided some suitable conditions are fulfilled. Precisely, one assumes: (i) that $f(q)$ is analytic in a complex strip $|\operatorname{Im} q| \leq \sigma$ for some positive σ , and (ii) that ω is a diophantine vector. The latter condition means that there are positive numbers γ and τ such that

$$(27) \quad |\langle k, \omega \rangle| \geq \gamma |k|^{-\tau} \quad \text{for all } 0 \neq k \in \mathbf{Z}^n .$$

In particular, condition (ii) is fulfilled if $\omega \in \mathbf{Z}^2$ is the vector $(1, \sqrt{r/s})$, where r, s are mutually prime integers such that r/s is not a perfect square: this is the case considered by Poincar e. The key remark is that the analyticity condition implies that the coefficients c_k decay exponentially with $|k|$, i.e., $|c_k| \leq C e^{-|k|\sigma}$ for some positive C . Hence we get

$$\sum_{0 \neq k \in \mathbf{Z}} \frac{|c_k|}{|\langle k, \omega \rangle|} \leq \frac{C}{\gamma} \sum_{0 \neq k \in \mathbf{Z}} |k|^\tau e^{-|k|\sigma} ,$$

which is a convergent series. This implies the uniform convergence of the series defining the solution χ .

The problem appears to be much simpler in the case of model (3), in view of our assumption that the perturbation $f(q)$ is actually a trigonometric polynomial. As we have already remarked, this implies that the coefficient of ε^s in all expansions generated by both Lindstedt’s and Kolmogorov’s method are trigonometric polynomials, too, so that we do not really need the argument above. On the other hand, our assumption is not really a restriction: if the Fourier expansion of $f(q)$ is a really infinite series we can take advantage of the exponential decay of the coefficients in order to split the series into an infinite sum of trigonometric polynomials of finite degree: see chap. XIII, §147 of *M ethodes nouvelles*.

However, the problem of convergence is not solved by the remarks above. Indeed, the recursive nature of both algorithms causes the small divisors to accumulate, thus raising serious doubts on the possibility that the exponential decay of coefficients and the increase of the power of ε can control the overall effect of the increasing number of small divisors. After excluding several cases, Poincaré comes to the following conclusion:[†] “Supposons, pour simplifier, qu’il y ait deux degrés de liberté; les séries ne pourraient-elles pas, par exemple, converger quand x_0^1 et x_0^2 ont été choisis de telle sorte que le rapport $\frac{n_1}{n_2}$ soit incommensurable, et que son carré soit au contraire commensurable (ou quand le rapport $\frac{n_1}{n_2}$ est assujéti à une autre condition analogue à celle que je viens d’enoncer un peu au hasard)? Les raisonnements de ce Chapitre ne me permettent pas d’affirmer que ce fait ne se présentera pas. Tout ce qu’il m’est permis de dire, c’est qu’il est fort invraisemblable.”

In order to understand the problem of the accumulation of the small divisors let us make a few simple considerations based on formulæ (6) and (7). Considering the Fourier expansion of every function we can identify some paths that could generate very big coefficients. Precisely, there are two kinds of such paths. Just in order to fix the language, let me call these paths the “dangerous path” of first and of second type, respectively. I will consider the case $n = 2$ and $\omega = (1, (\sqrt{5} - 1)/2)$, i.e., the so called “golden torus”. I shall also assume that the perturbation $f(q)$ in (3) contains at least the Fourier modes $(0, -1)$ and $(1, -1)$. The key remark is that the r.h.s. of equation (7) for $\psi^{(r)}$ contains the term

$$\frac{\partial}{\partial q_1} \frac{\partial f}{\partial q} \Delta_1(1, r - 1) .$$

On the other hand, in view of (6), $\Delta_1(1, r - 1)$ contains at least the term $\varphi_1^{(r-1)}$. Hence, the equation for $\psi^{(r)}$ contains at least the expression

$$\frac{\partial}{\partial q_1} \frac{\partial f}{\partial q} \varphi_1^{(r-1)} .$$

Now, if f and $\varphi_1^{(r-1)}$ contain the Fourier modes k and λ , respectively, then $\psi^{(r)}$ will contain both Fourier modes $k + \lambda$ and $k - \lambda$, in view of an elementary property of the product of trigonometric functions. Hence, the solution $\varphi^{(r)}$ of equations (7) contains both these Fourier modes with divisors $\langle (k + \lambda), \omega \rangle$ and $\langle (k - \lambda), \omega \rangle$, respectively. This simple remark allows us to isolate the paths referred to above.

[†] *Méthodes Nouvelles*, ch. XIII, §149. Poincaré denotes by x the action variables that we call p , and by n the frequencies, that we call ω .

The dangerous path of first type is due to the Fourier modes

$$\begin{aligned}
 (28) \quad & k^{(1)} = (1, -1) \quad , \\
 & k^{(2)} = (1, -2) \quad , \\
 & k^{(3)} = (2, -3) \quad , \quad k^{(4)} = (2, -4) \quad , \\
 & k^{(5)} = (3, -5) \quad , \quad k^{(6)} = (4, -6) \quad , \quad k^{(7)} = (4, -7) \quad , \\
 & k^{(8)} = (5, -8) \quad , \quad k^{(9)} = (6, -9) \quad , \quad k^{(10)} = (6, -10) \quad , \quad k^{(11)} = (7, -11) \quad , \quad k^{(12)} = (7, -12) \quad , \\
 & k^{(13)} = (8, -13) \quad , \quad \dots
 \end{aligned}$$

that appear in $\varphi^{(1)}$, $\varphi^{(2)}$, \dots , $\varphi^{(13)}$, \dots , respectively. One will immediately notice that this is the sequence of the best approximants to the direction orthogonal to ω that can be obtained by adding to a given vector k either the vector $(0, -1)$ or the vector $(1, -1)$, as appropriate. In particular, this sequence contains as a subsequence the Fibonacci's approximants, which are the best ones according to the theory of continuous fractions. These approximants appear in the first column. We conclude that in $\varphi^{(r)}$ there is a coefficient with a divisor

$$\left(\langle k^{(1)}, \omega \rangle \langle k^{(2)}, \omega \rangle \cdot \dots \cdot \langle k^{(r)}, \omega \rangle \right)^2 .$$

A naive application of the diophantine estimate (27) leads to the conclusion that the accumulation above of small divisors causes the coefficient to grow as fast as $(r!)^2$, thus preventing the convergence of Lindstedt series. However, such an estimate is very naive and pessimistic. Indeed, it has been shown by Siegel^{[11][12]} that the overall effect of the product above of small divisors is bounded by C^r with some positive constant C . Therefore, the path above is not really dangerous, in that it does not prevent the convergence of Lindstedt's series. The estimate of Siegel has been improved by Eliasson^[13] and Bruno^[14].

The dangerous path of second type is subtler: it is due to a *repetition* of small divisors. Starting with the Fourier mode $(1, -1)$ in f , it is easy to see that the Fourier modes

$$k^{(1)} = (1, -1), \quad k^{(2)} = (1, -2), \quad k^{(3)} = (1, -1), \quad k^{(4)} = (1, -2), \quad k^{(5)} = (1, -1), \quad \dots$$

will successively appear in $\varphi^{(1)}$, $\varphi^{(2)}$, \dots . For, this sequence is simply obtained by adding to the Fourier mode k the vector $(0, -1)$, and by subtracting it at the next step. Hence, the products of divisors

$$\left(\langle k^{(1)}, \omega \rangle \langle k^{(2)}, \omega \rangle \right)^{2r} \quad \text{and} \quad \langle k^{(1)}, \omega \rangle \left(\langle k^{(1)}, \omega \rangle \langle k^{(2)}, \omega \rangle \right)^{2r}$$

will appear in $\varphi^{(2r)}$ and in $\varphi^{(2r+1)}$, respectively. An elementary argument using the ratio criterion for the convergence of series shows that the convergence radius in ε of

Lindstedt's series will be $O(|\langle k^{(2)}, \omega \rangle|)$. This looks not so bad. However, a moment's thought allows us to realize that whenever a Fourier mode corresponding to any of the Fibonacci approximations in (28), $k^{(s)}$ say, appears, e.g. in $\varphi^{(s)}$, then *the same mode* will appear again in $\varphi^{(s+2)}$, $\varphi^{(s+4)}$, \dots , *ad infinitum*. By the argument above on the ratio criterion we find that the convergence radius in ε of Lindstedt's series will be $O(|\langle k^{(s)}, \omega \rangle|)$, i.e., arbitrarily close to zero. This fact appears to be incompatible with the convergence of Lindstedt's expansions.

Now, the existence of quasiperiodic solutions is the main achievement of the theorem of Kolmogorov. Moreover, as an immediate byproduct of Kolmogorov's proof it turns out that the series representing such quasiperiodic motions are analytic in ε . Since these series coincide with Lindstedt's ones (as we have seen above) we must conclude for the convergence of Lindstedt's series. This fact has been pointed out by Moser^[15], who concluded that Lindstedt's series do converge, but a direct proof of their convergence with Cauchy's method of majorant's seems to be impossible.

This complicated state of affairs has been fully clarified with the recent works of Eliasson (see report [16], now published as [17], and [18][13]) and the more recent ones of Gallavotti and coworkers^{[7][19][20][21][22][23]} and of Chierchia and Falcolini^{[24][25]}.

The underlying idea is that there are *compensations* among those terms — like the ones discussed above — that could cause divergence. Indeed, it is easy to realize that a coefficient of a given Fourier mode in $\varphi^{(s)}$ or $\psi^{(s)}$ can be generated via different paths. The contributions of all these paths could compensate each other, leaving only a remainder that can be dominated by the powers of ε . Identifying such paths by just looking at formulæ (6) and (7) is not easy at all. However, there is an elegant method to do it by representing the solutions as trees. I will not give here details about the construction of these trees: it is already well explained, e.g., in [19] or in [24]. I just mention that every term contributing to any Fourier coefficient of the expansion is represented by a tree which makes explicit how that term must be constructed in terms of the Fourier modes that appear in the perturbation f in (3). Thus, the coefficient of a given Fourier mode is represented as a sum over a set of trees.

The discussion above concerning the dangerous paths can be summarized as follows: if we add together the absolute values of all coefficients represented by all trees (thus disregarding the fact that we are working with power series in ε and Fourier series in the angles q) then we find a divergent series. With a non completely standard language we could say that the series produced by Lindstedt's algorithm are *not absolutely convergent*.[†] If instead we add together *algebraically* all trees that generate the same Fourier mode, then compensations take effect, so that the final result is an absolutely convergent series.

[†] We are actually dealing with a sum involving infinite terms, the sum of the absolute values of which diverges.

Unfortunately, it seems that nobody till now has discovered a simple method to add up *all* coefficients of the same Fourier mode. However, the papers above contain explicit algorithms that allow one to isolate the so called *resonant trees*, actually the trees with repetitions of small divisors as illustrated by the dangerous path of second type above. This is enough in order to give a direct proof of the convergence of Lindstedt's series.

3.2 Convergence of Kolmogorov's construction of the normal form

The convergence of Kolmogorov's algorithm was the great result communicated by Kolmogorov in his short note, in 1954.^{[1]‡} After the proof published by Moser^[4] and Arnold^[3] the so called KAM theory has been originated. A detailed account of such a big development is outside the limits of this short note, that is mainly devoted to recent results connected with Lindstedt's series. I will limit my discussion to a sketch of the proof of convergence of the algorithm of sect. 2.2.

The main point concerns precisely the accumulation of small divisors. Let us define the non increasing positive sequence $\{\alpha_s\}_{s>0}$ as

$$(29) \quad \alpha_s = \min_{0 < |k| \leq sK} |\langle k, \omega \rangle| ,$$

i.e., the smallest divisor (in absolute value) that can appear in the solution of the equations (15) and (17) for the generating functions $X^{(s)}$ and $\chi^{(s)}$; here, K is the trigonometric degree of the initial perturbation $f(q)$ in the Hamiltonian (3). Using the

‡ In his note Kolmogorov suggested also to use a fast convergence method (in his own words "reminiscent of Newton's method") in order to dominate the accumulation of small divisors. That method, sometimes called "quadratic" has become the standard basis of the so called "KAM theory". As noted by Rüssman:^[26] "It has often been said that the rapid convergence of the Newton iteration is necessary for compensating the influence of small divisors. But a deeper analysis shows that this is not true. (...) Historically, the Newton method was surely necessary to establish the main theorems of the KAM-theory. But for clarifying the structure of the small divisor problems the Newton method is not useful because it compensates not only the influence of small divisors, but also many bad estimates veiling the structure of the problems." Up to my knowledge, the first proof of existence of invariant tori that does not make use of the fast convergence assured by the quadratic method has been published by Rüssman^[27]. The present discussion is based on a recent scheme introduced by U. Locatelli and myself. The differences with respect to Rüssman's approach are: (i) we use the classical expansion in a parameter ε , as illustrated in the formal section 2.2, while Rüssman uses a different reordering of terms without making use of an expansion in a small parameter, and, (ii) our attention is concentrated on a direct analysis of the accumulation of small divisors, while Rüssman dominates the problem by using an approximating function. Concerning the nonresonance condition on the frequencies, in both cases the proof is worked out using Bruno's condition, which is weaker than the diophantine one used in all works quoted above on Lindstedt's series.

explicit formulæ (18) and (19) we can follow the accumulation of divisors disregarding all other quantities. The rules are quite simple: (a) when solving the equations for the generating functions $X^{(s)}$ and $\chi_2^{(s)}$ we just add a new divisor α_s ; (b) when we perform a Poisson bracket between two functions, the divisors are multiplied together; (c) when we perform a sum, we keep trace only of the worst accumulation, taking into account that the sequence $\{\alpha_s\}_{s>0}$ is not increasing. For instance, eq. (20) tells us that $X^{(1)}$ has a divisor α_1 . The explicit expressions (21) tell us that $\hat{f}_1^{(1,1)}$ and $\hat{f}_0^{(1,2)}$ will contain the divisors α_1 and α_1^2 , respectively. By (22) we see that the worst product of divisors in $\chi_2^{(1)}$ will be α_1^2 , and by the explicit expression of $H^{(1)}$ we get that the worst divisor in the coefficient of ε^s is $\alpha_1^{2^s}$. If, with some patience, we follow the first few normalization steps we realize that the accumulation of small divisors seems not so bad. Identifying a rule for the accumulation of small divisors requires some effort. We found the following rule: *the denominator of the coefficient of ε^s of any function produced by Kolmogorov's algorithm contains at most $2s$ divisors $\langle k, \omega \rangle$; among them there are at most $2 \cdot \lfloor s/(2^{m-1} + 1) \rfloor$ divisors with $2^{m-1} < |k|/K \leq 2^m$, for every $m \geq 1$.*

Let us show that this property is enough in order to state that the overall effect of the small divisors is bounded by C^s , with some positive constant C . Considering the worst case, and recalling that the sequence $\{\alpha_s\}_{s \geq 1}$ is monotonically not increasing, we have for every coefficient of ε^s the bound

$$\prod_{m=1}^{\lceil \log_2 s \rceil} \left(\frac{1}{\alpha_{2^m}} \right)^{s/2^{m-1}} .$$

With an elementary calculation we get

$$\log_2 \prod_{m=1}^{\lceil \log_2 s \rceil} \left(\frac{1}{\alpha_{2^m}} \right)^{s/2^{m-1}} < -s \sum_{m \geq 1} \frac{\log_2 \alpha_{2^m}}{2^{m-1}} .$$

If we assume the condition

$$(30) \quad - \sum_{j \geq 0} \frac{1}{2^{j-1}} \log_2 \alpha_{2^j} = B < \infty ,$$

then we conclude that the accumulation of small divisors is bounded by 2^{sB} , as claimed. This is a condition on the frequencies ω , that is a generalization to a generic dimension n of a condition introduced by Bruno for the case of the Siegel center's problem, the latter one corresponding to $n = 2$. For comparison with the diophantine condition (27), remark that if ω satisfies, e.g.,

$$(31) \quad |\langle k, \omega \rangle| \geq \beta \exp \left(- \frac{|k|}{\ln^{1+\delta} |k|} \right)$$

with some positive constants β and δ then it fulfills the Bruno's condition above.

A complete proof of the convergence of the transformation to Kolmogorov's normal form requires also an estimate of the number of terms generated by the algorithm, and of the accumulation of other constants due to Poisson brackets, and so on. However, this is just a standard technical matter that constitutes the technological background of this type of problems. The interested reader will find a detailed proof in [28].

4. Comparing the algorithms

An immediate remark that comes out from the discussion of the formal algorithms in sect. 2 is that the algorithm of Lindstedt seems to be definitely simpler and faster than Kolmogorov's one. For, a straightforward comparison of the calculation of the coefficient of order 1 in ε shows that the algorithm of Lindstedt can be easily applied by hand in a few minutes, while Kolmogorov's algorithm requires the calculation of two generating functions, namely $X^{(1)}$ and $\chi_2^{(1)}$ and the construction of the first order term of the canonical transformation putting the Hamiltonian in Kolmogorov's normal form up to order 1. A few minutes are not enough, of course, but a few hours are. . .

One is tempted to explain such a big difference by remarking that with the algorithm of Lindstedt we just find a solution of the canonical equations. Nothing related to a canonical transformation seems to come out. Hence, the algorithm of Kolmogorov seems to be less efficient because it gives more information. Furthermore, one is also tempted to stress that Lindstedt's algorithms does not take into account the canonical character of the differential equations.

On the other hand, the discussion of the convergence seems to be easier with the Kolmogorov's approach: no dangerous paths for the accumulation of small divisors occur, so that the control is based on a simple algebraic remark; there is no need of compensations, nor do we need to use the Siegel–Eliasson–Bruno estimate. This looks quite puzzling, because we *know* that both algorithms must give the *same* solution. Why does this difference occur?

My aim in this section is to discuss in some detail these points. I do not pretend to explain all relations between Lindstedt's and Kolmogorov's algorithms. I will just point out a few considerations that seem to me relevant.

4.1 Equivalence of the algorithms

Let me first state that the series produced by Lindstedt's algorithms contain the same information as the series of Kolmogorov. More precisely, the canonical transformation putting the Hamiltonian in Kolmogorov's normal form is implicitly contained in Lindstedt's expansions.

The argument makes use of a well known fact concerning the general theory of canonical transformations. Assume that we are given a transformation $q = \varphi(q')$ involving only the coordinates (as opposed to the canonical momenta). Then we can construct a complete canonical transformation in a very standard manner by just using a generating function $S(p', q)$ of the form

$$(32) \quad S(p, q') = \langle p, \varphi(q') \rangle + w(q') ,$$

where $w(q')$ is an arbitrary function. This is often called the extended point transformation. Now, the series produced by Lindstedt's algorithm are of the form

$$(33) \quad q = \varphi(q') , \quad p = \psi(q') .$$

On the other hand, it is an easy matter to prove that the invariant torus is a Lagrangian manifold, so that the differential form $p dq$ is exact on it. Therefore we can construct a generating function of the form (32), with a function $w(q')$ that is determined by the functions $\psi(q')$. To this end, just write the canonical transformation generated by (32) in the implicit form

$$(34) \quad q_j = \varphi_j(q') , \quad p'_j = \sum_k p_k \frac{\partial \varphi_k}{\partial q'_j} + \frac{\partial w}{\partial q'_j} .$$

Invert the second relation so as to obtain $p = \tilde{\psi}(p', q')$, still leaving the function $w(q')$ undetermined; then determine $w(q)$ so as to satisfy the identity $\tilde{\psi}(0, q') = \psi(q')$ as given by Lindstedt's expansion. This completely determines the generating function. That the transformed Hamiltonian is in normal form follows by considering that $p' = \text{const}$ is a consequence of the form of Lindstedt's solution.

We conclude that Kolmogorov's method requires the construction of a canonical transformation which is only implicit in Lindstedt's method, but does not actually produce more information than Lindstedt's algorithm.

4.2 On the accumulation of small divisors

The different behaviour in the accumulation of the small divisors has a quite simple explanation: the algorithms of Lindstedt and of Kolmogorov produce different expansions of one and the same function. In more precise terms, the situation is the following. We are dealing with power series in ε and with trigonometric series in the angles q (trigonometric polynomials in the case of the Hamiltonian (3)). Each algorithm produces the coefficient of ε^s as a sum over the Fourier modes, and the coefficient of each Fourier mode as a sum of several terms; every such term contains a product of small divisors. It is not difficult to devise an algorithm which produces separately every term, without making any resummation: see for instance [28]. Consider now the coefficient of a given Fourier mode: it will turn out that the expressions generated by the two algorithms are formally different, but are algebraically equivalent.

This state of affairs is better illustrated by an explicit example. Consider the Hamiltonian (even in the angles)

$$H(p, q) = \omega_1 p_1 + \omega_2 p_2 + \frac{1}{2}(p_1^2 + p_2^2) + \varepsilon(\cos q_1 + \cos(q_1 - q_2)) ,$$

so that $K = 2$. The calculation of quasi periodic solutions has been performed up to order ε^5 using the commercial algebraic manipulator Mathematica. It turns out that the expressions at order ε are identical, but the expressions at order ε^2 are not. For instance, for the function $\varphi_1^{(2)}$ the construction of Lindstedt gives

$$\begin{aligned} \varphi_1^{(2)} = & \frac{\sin 2q_1}{8\omega_1^4} + \frac{\sin(2q_1 - 2q_2)}{4(\omega_1 - \omega_2)^4} + \frac{\sin(2q_1 - q_2)}{2\omega_1^2(2\omega_1 - \omega_2)^2} \\ & + \frac{\sin(2q_1 - q_2)}{2(2\omega_1 - \omega_2)^2(\omega_1 - \omega_2)^2} + \frac{\sin q_2}{2\omega_1^2\omega_2^2} - \frac{\sin q_2}{2\omega_2^2(\omega_1 - \omega_2)^2} ; \end{aligned}$$

The expression given by Kolmogorov's algorithm is

$$\begin{aligned} \varphi_1^{(2)} = & \frac{\sin 2q_1}{8\omega_1^4} + \frac{\sin(2q_1 - 2q_2)}{4(\omega_1 - \omega_2)^4} + \frac{\sin(2q_1 - q_2)}{2\omega_1^2(\omega_1 - \omega_2)^2} - \frac{\sin(2q_1 - q_2)}{\omega_1(\omega_1 - \omega_2)(2\omega_1 - \omega_2)^2} \\ & - \frac{\sin q_2}{2\omega_1\omega_2(\omega_1 - \omega_2)^2} - \frac{\sin q_2}{2\omega_1^2(\omega_1 - \omega_2)\omega_2} . \end{aligned}$$

At first sight, the expressions look different. However, it is not difficult to check that they are algebraically equivalent: just use the elementary algebraic identity

$$\frac{1}{ab} = \frac{1}{b(a-b)} - \frac{1}{a(a-b)} .$$

These expressions do not allow us to show that the Lindstedt's expansion actually contains accumulations of small divisors that violate the rule stated in sect. 3.2. Such terms start to appear at order ε^4 . We can not insert here the complete expression, since it contains more than hundred terms; therefore we report here only the expansion of the coefficient of the Fourier mode $(4, -3)$. The algorithm of Lindstedt gives

$$\begin{aligned} \frac{1}{2\pi^2} \left\langle \varphi_1^{(4)}, \sin(4q_1 - 3q_2) \right\rangle = & \frac{3}{8(\omega_1)^2(\omega_1 - \omega_2)^4(4\omega_1 - 3\omega_2)^2} \\ & + \frac{1}{2(\omega_1)^2(\omega_1 - \omega_2)^2(2\omega_1 - \omega_2)^2(4\omega_1 - 3\omega_2)^2} + \frac{1}{2(\omega_1)^2(\omega_1 - \omega_2)^2(3\omega_1 - 2\omega_2)^2(4\omega_1 - 3\omega_2)^2} \\ & + \frac{1}{2(\omega_1)^2(2\omega_1 - \omega_2)^2(3\omega_1 - 2\omega_2)^2(4\omega_1 - 3\omega_2)^2} + \frac{1}{8(\omega_1 - \omega_2)^6(4\omega_1 - 3\omega_2)^2} \\ & + \frac{1}{4(\omega_1 - \omega_2)^4(2\omega_1 - \omega_2)^2(4\omega_1 - 3\omega_2)^2} + \frac{1}{8(\omega_1 - \omega_2)^4(3\omega_1 - 2\omega_2)^2(4\omega_1 - 3\omega_2)^2} \\ & + \frac{1}{4(\omega_1 - \omega_2)^2(2\omega_1 - \omega_2)^2(3\omega_1 - 2\omega_2)^2(4\omega_1 - 3\omega_2)^2} \end{aligned}$$

On the other hand, Kolmogorov's algorithm gives

$$\begin{aligned}
 \frac{1}{2\pi^2} \left\langle \varphi_1^{(4)}, \sin(4q_1 - 3q_2) \right\rangle &= \frac{19}{96(\omega_1)^2(\omega_1 - \omega_2)^6} + \frac{3}{32(\omega_1)^2(\omega_1 - \omega_2)^5(2\omega_1 - \omega_2)} \\
 &+ \frac{7}{24(\omega_1)^2(\omega_1 - \omega_2)^5(3\omega_1 - 2\omega_2)} + \frac{9}{16(\omega_1)^2(\omega_1 - \omega_2)^5(4\omega_1 - 3\omega_2)} \\
 &- \frac{1}{16(\omega_1)^2(\omega_1 - \omega_2)^4(2\omega_1 - \omega_2)(4\omega_1 - 3\omega_2)} - \frac{5}{4(\omega_1)^2(\omega_1 - \omega_2)^4(3\omega_1 - 2\omega_2)^2} \\
 &+ \frac{2}{2(\omega_1)^2(\omega_1 - \omega_2)^4(3\omega_1 - 2\omega_2)(4\omega_1 - 3\omega_2)} - \frac{21}{4(\omega_1)^2(\omega_1 - \omega_2)^4(4\omega_1 - 3\omega_2)^2} \\
 &- \frac{3}{32\omega_1(\omega_1 - \omega_2)^6(2\omega_1 - \omega_2)} - \frac{1}{24\omega_1(\omega_1 - \omega_2)^6(3\omega_1 - 2\omega_2)} + \frac{1}{16\omega_1(\omega_1 - \omega_2)^6(4\omega_1 - 3\omega_2)} \\
 &- \frac{3}{8\omega_1(\omega_1 - \omega_2)^5(2\omega_1 - \omega_2)^2} + \frac{3}{2\omega_1(\omega_1 - \omega_2)^5(2\omega_1 - \omega_2)(3\omega_1 - 2\omega_2)} \\
 &- \frac{11}{16\omega_1(\omega_1 - \omega_2)^5(2\omega_1 - \omega_2)(4\omega_1 - 3\omega_2)} - \frac{15}{8\omega_1(\omega_1 - \omega_2)^5(3\omega_1 - 2\omega_2)^2} \\
 &+ \frac{15}{4\omega_1(\omega_1 - \omega_2)^5(3\omega_1 - 2\omega_2)(4\omega_1 - 3\omega_2)} - \frac{4}{\omega_1(\omega_1 - \omega_2)^5(4\omega_1 - 3\omega_2)^2} \\
 &- \frac{3}{2\omega_1(\omega_1 - \omega_2)^4(2\omega_1 - \omega_2)(4\omega_1 - 3\omega_2)^2}
 \end{aligned}$$

Again, these expressions are algebraically equivalent. However, the reader will easily check that in the expression given by Lindstedt's algorithm there are four terms that violate our rule on the accumulation of small divisors, while all terms in the expression given by Kolmogorov's algorithm respect that rule. This shows that the dangerous paths do actually appear in Lindstedt's construction, so that the compensations discovered by the authors of the quoted papers on Lindstedt's series are really needed.

Discovering the complete algebraic relations between the algorithms seems to be a very hard matter, and we did not succeed in finding it.

4.3 On the effectiveness of the algorithms

The discussion of the previous section shows that in the algorithm of Kolmogorov compensations between big coefficients do not occur. In a practical construction, e.g., by computer, this could help a lot, because it avoids the numerical errors generated by such cancellations.

Let me now come to the problem of the complexity of the algorithm. My point is that the remark made at the beginning of this section about the calculation of the first order terms is correct, but misleading. It only means that trying to apply the algorithm of Kolmogorov by hand is not recommended. The question is: what happens if one implements both algorithms on a computer, using an algebraic manipulator, and tries to construct the series up an arbitrary order r ? I think that the considerations in this section will show that the effectiveness of Lindstedt's algorithm is at least questionable.

The algorithm of Lindstedt is based essentially on an operation of substitution of functions. Indeed, the equations (7) are obtained by substituting the form (5) of the solution in the differential equations (4). On the other hand, our formulation of Kolmogorov's algorithm is based essentially on performing a transformation by the method of Lie series, and using the same method in order to construct the explicit form of the canonical transformation. In both cases, at every step we must also solve twice an equation of the form $\partial_\omega g = f$, where f is a known function and g is unknown. Therefore the solution of the latter equation has no real influence on the effectiveness of the algorithm.[†] It is legitimate to conclude that it is enough to compare a substitution of variables with a transformation with the algorithm of Lie series.

In order to simplify the problem as much as possible, let us consider the following one-dimensional situation: we are given a function $f(q)$, $q \in \mathbf{T}$, and want to perform a substitution of variables of the form

$$(35) \quad q = q' + \varepsilon\varphi_1(q') + \varepsilon^2\varphi_2(q') + \dots + \varepsilon^r\varphi_r(q') .$$

We know also that the transformation (35) can be generated as

$$(36) \quad q = \exp(\varepsilon^r L_{\chi_r}) \circ \dots \circ \exp(\varepsilon L_{\chi_1}) x ,$$

where q must be replaced by q' after performing all operation indicated on the r.h.s., and the generating functions $\chi_1(p, q), \dots, \chi_r(p, q)$ are known, and linear in p . All expressions are to be expanded in powers of ε , and truncated at order ε^r . My aim is to show that the Lie series method is more effective than a direct substitution.

Let us first examine the substitution. Writing $x = x' + \delta(x')$, and expanding in Taylor series we get

$$(37) \quad f(x' + \delta(x')) = f(x') + \delta(x') \left. \frac{df}{dx} \right|_{x'} + \dots + \delta^r(x') \left. \frac{d^r f}{dx^r} \right|_{x'} .$$

This expression must be further expanded and reordered by substituting

$$\delta(x') = \varepsilon\varphi_1(x') = \dots + \varepsilon^r\varphi_r(x') ;$$

Let us now evaluate the complexity of the algorithm, taking as reference the number of

[†] If any role is played by this equation in determining the effectiveness of the algorithm, then it is in favor of Kolmogorov's one. Indeed, Lindstedt's method requires to solve the equation twice for every pair of canonical coordinates, with a total of $2n$ equations, while Kolmogorov's algorithm requires only the determination of the generating functions. However, the computational cost of this operation is negligible with respect to that of the rest of the calculation.

products of functions that we must perform.[‡] The evaluation of δ^2 requires $r(r-1)/2$ products of functions. Taking into account that δ^{s-1} is a polynomial in ε starting with ε^{s-1} , the evaluation of $\delta^s = \delta \cdot \delta^{s-1}$ requires $(r-s+2)(r-s+1)/2$ products of functions. Hence, the complete evaluation of $\delta^2, \dots, \delta^r$ requires $\sum_{s=2}^r (r-s+2)(r-s+1)/2$ products. This is roughly $r^3/6$ products. Having determined the powers of δ , the evaluation of the r.h.s. of the expression (37) requires $O(r^2)$ products, that we can neglect. We conclude that the substitution of variables requires $O(r^3)$ multiplications of functions.

Let us now come to the transformation via Lie series. The application of the operator $\exp(L_{\chi_s})$ to a function f requires r/s products. Taking into account that after the first step the operand is a polynomial of degree r in ε we must apply $\exp(L_{\chi_s})$ to r different functions, which requires at most r^2/s products of functions. Hence, the complete evaluation of the r.h.s. of (37) requires at most $r^2 \sum_{s=1}^r 1/s$ products. This amounts to $O(r^2 \ln r)$ multiplications of functions.

The conclusion is that for large r the Lie series algorithm is definitely more effective than the direct substitution of variables. Recalling that the algorithm of Lindstedt is based on a substitution of variables, and that all transformations in our formulation of Kolmogorov's algorithm are based on Lie series, we conclude that for a sufficiently large r the algorithm of Kolmogorov is more effective than Lindstedt's one.

One could ask: how large is "sufficiently large"? The considerations in this section do not allow me to formulate an answer. However, I can report an experimental result, so to say: in the calculation of the model problem discussed in sect. 4.2 Locatelli and I found that "sufficiently large" means $r = 5$.

References

- [1] Kolmogorov, A. N.: *Preservation of conditionally periodic movements with small change in the Hamilton function*, Dokl. Akad. Nauk SSSR, **98**, 527 (1954).
- [2] Thirring, W.: *Classical dynamical systems*, Springer Verlag, New York (1978).
- [3] Arnold, V. I.: *Proof of a theorem of A. N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian*, Usp. Mat. Nauk, **18**, 13 (1963); Russ. Math. Surv., **18**, 9 (1963).
- [4] Moser, J.: *On invariant curves of area-preserving mappings of an annulus*, Nachr. Akad. Wiss. Gött., II Math. Phys. Kl 1962, 1–20 (1962).

[‡] If all functions are trigonometric polynomials in q then the computational cost of a derivative is actually negligible. Conversely, multiplication is a quite long operation. One could perhaps reduce the time required by applying some methods related to the use of Fourier transforms, but this improvement takes place for both the Lindstedt and the Kolmogorov method, so that it is not relevant for our discussion.

- [5] Laskar, J. and Robutel, P: *Stability of the Planetary Three-Body Problem I. Expansion of the Planetary Hamiltonian*, Cel. Mech. **62**, 193–217, (1995).
- [6] Poincaré, H.: *Les méthodes nouvelles de la mécanique céleste*, Gauthier–Villars, Paris (1892).
- [7] Gallavotti, G.: *Twistless KAM tori*, Comm. Math. Phys. **164**, 145–156 (1994).
- [8] Gröbner, W.: *Die Lie-Reihen und Ihre Anwendungen*, Springer Verlag, Berlin (1960); it. transl.: *Le serie di Lie e le loro applicazioni*, Cremonese, Roma (1973).
- [9] A. Giorgilli: *Quantitative methods in classical perturbation theory*, in *From Newton to chaos*, A.E. Roy and B.D. Steves eds., Plenum Press, New York (1995).
- [10] Giorgilli, A. and Locatelli, U.: *Kolmogorov theorem and classical perturbation theory*, ZAMP **48**, 220–261 (1997).
- [11] Siegel, C. L.: *Iterations of analytic functions*, Annals of Math. **43**, 607–612 (1942).
- [12] Siegel, C. L.: *Über die Normalformen analytischer Differentialgleichungen in der Nähe einer Gleichgewichtslösung*, Nachr. Acad. Wiss. Göttingen, Math. Phys. Kl, **IIA** 21-30 (1952).
- [13] Eliasson, L. H.: *Generalization of an estimate of small divisors by Siegel*, in *Analysis et cætera*, P. Rabinowitz and E. Zehnder eds., Academic Press (1990).
- [14] Bruno, A.D.: *Analytic form of differential equations*, Trans. Moscow Math. Soc. **25**, 131–288 (1971).
- [15] Moser, J.: *Convergent series expansions for quasi-periodic motions*, Math. Ann. **169**, 136–176 (1967).
- [16] Eliasson, L. H.: *Absolutely convergent series expansion for quasi-periodic motions*, report 2–88, Dept. of Math., Univ. of Stockolm (1988).
- [17] Eliasson, L. H.: *Absolutely convergent series expansion for quasi-periodic motions*, MPEJ **2**, 1–33 (1996).
- [18] Eliasson, L. H.: *Hamiltonian system with normal form near an invariant torus*, in *Nonlinear dynamics*, G. Turchetti ed., World scientific (1989).
- [19] Gallavotti, G.: *Twistless KAM tori, quasi flat homoclinic intersections, and other cancellations in the perturbation series of certain completely integrable Hamiltonian systems. A review*, Reviews in Math. Phys. **6**, 343–411 (1994).
- [20] Gallavotti, G. and Gentile, G.: *Majorant series convergence for twistless KAM tori*, Ergodic Theory and Dynamical Systems **15**, 857–869 (1995).
- [21] Gentile, G. and Mastropietro, V.: *Tree expansion and multiscale analysis for KAM tori*, Nonlinearity **8**, 1159–1178 (1995).
- [22] Gentile, G. and Mastropietro, V.: *KAM theorem revisited*, Physica D **90**, 225–234 (1996).
- [23] Gentile, G. and Mastropietro, V.: *Methods of analysis of the Lindstedt series for KAM tori and renormalizability in classical mechanics. A review with some applications*, Reviews in Math. Phys **8**, 393–444 (1996).

- [24] Chierchia, L. and Falcolini, C.: *A direct proof of a theorem by Kolmogorov in Hamiltonian systems*, Ann. Sc. Norm. Sup. Pisa, Serie IV. **XXI** 541-593 (1994).
- [25] Chierchia, L. and Falcolini, C.: *Compensations in small divisor problems*, Comm. Math. Phys. **175**, 135-160, (1996).
- [26] Rüssmann, H.: *Non-degeneracy in the perturbation theory of integrable dynamical systems*, in *Number theory and dynamical systems*, Dodson, M.M., and Vickers, J.A.G., Eds., Cambr. Univ. Press (1989).
- [27] Rüssmann, H.: *On the frequencies of quasi periodic solutions of analytic nearly integrable Hamiltonian systems*, in *Seminar on Dynamical Systems*, S. Kuksin, V. Lazutkin and J. Pöschel eds., PNDLE 12, 160–183, Birkhäuser Verlag (1994).
- [28] Giorgilli, A., and Locatelli, U.: *Classical series expansions for quasiperiodic motions*, preprint (1997).