POINCARE'S NON - EXISTENCE THEOREM AND CLASSICAL PERTURBATION THEORY FOR NEARLY INTEGRABLE HAMILTONIAN SYSTEMS

GIANCARLO BENETTIN

Dipartimento di Fisica dell'Università di Padova
and Centro Interuniversitario di Struttura della Materia
Via Marzolo 8 – PADOVA (Italia)

LUIGI GALGANI

Dipartimento di Matematica dell'Università di Milano
Via Saldini 50 – MILANO (Italia)

ANTONIO GIORGILLI

Dipartimento di Fisica dell'Università di Milano
Via Celoria 16 – MILANO (Italia)

Abstract: *Classical perturbation theory is revisited in the light of Poincaré's theorem on the non existence of integrals of motion in nearly-integrable Hamiltonian systems.*

1. Introduction

As it is well known, at the turn of the century Poincaré [1] proved his fundamental theorem on the non-existence of integrals of motion in nearly-integrable Hamiltonian systems. Within apparently wide assumptions, this theorem states that generic perturbations modify quite significantly the basic features of integrable dynamical systems: indeed, integrals of motion disappear, and the topological character of trajectories is sensibly modified, at least in a dense subset of the phase space. On the one hand, this theorem had a quite strong impact on the development of classical mechanics as, beside its precise formulation, it was interpreted as asserting the general failure of perturbation theory for nearly-integrable Hamiltonian systems. On the other hand, because of the strong connections between absence of integrals of motion and ergodicity, the theorem, when conveniently generalized [2], was considered to support the idea that Hamiltonian dynamical systems (with at least three degrees of freedom) are in general ergodic [3], and thus suited for a statistical mechanical

approach.

Today, after the basic results by Birkhoff [4], Siegel [5,6], Kolmogorov [7] and Arnold [8], and more recently Nekhoroshev [9], it is clear that Poincaré's result is not as conclusive as it was supposed. Nevertheless, Poincaré's theorem still plays a central role in classical perturbation theory, being the starting point where the different branches of classical perturbation theory come from, and consequently a basic guide to understand its modern development.

The aim of this paper is precisely to illustrate the basic methods and results of classical perturbation theory in the light of Poincaré's theorem. The paper is organized as follows: first of all, in section 2, we will introduce and prove Poincaré's theorem, although in a reduced form. Then we will shortly recall, in section 3, the Lie method for canonical transformations close to the identity, which is in our opinion the most natural and efficient one in classical perturbation theory. Then, from section 4 to section 7, we will show how, by exploiting the possibilities still left open by Poincaré's negative result, one is naturally introduced to the different branches of classical perturbation theory. In particular, section 4 is concerned with Birkhoff's construction, section 5 is devoted to Siegel's diofantine condition, section 6 deals with perturbations having a finite number of Fourier components, while section 7 introduces, within the most elementary context, the basic ideas by Arnold (the ultraviolet cut-off) and by Kolmogorov (the weak normal form).

Up to this point, only finite order, or even first order perturbation theory is considered. Instead, section 8 is devoted to infinite order perturbation theory, and reports just a sketch of Kolmogorov theorem (in the recent formulation by Poeschel [10] and by Chierchia and Gallavotti [11]) and of Nekhoroshev theorem. Finally, section 9 illustrates the generalization of Poincaré's negative result accomplished by Fermi [2], and more recently extended in ref. [12].

The reader will understand the difficulties of compressing such a wide subject in a short space, and forgive us for the frequently informal exposition.

2. Poincaré's theorem

Let us consider a nearly-integrable Hamiltonian system with n degrees of freedom. Using action-angle variables $p = (p_1, \dots, p_n) \in \mathcal{B} \subset \mathbf{R}^n$, $q = (q_1, \dots, q_n) \in \mathbf{T}^n$, where $\mathcal{B}$ is an open subset of $\mathbf{R}^n$ and $\mathbf{T}^n$ the n-dimensional torus, the Hamiltonian is written

$$H(p, q, \varepsilon) = h(p) + \varepsilon f(p, q, \varepsilon) \ . \tag{2.1}$$

h and f are supposed to be regular or even analytic in all of the variables. For $\varepsilon = 0$ the system is integrable, and the phase space $\mathcal{W} = \mathcal{B} \times \mathbf{T}^n$ is foliated into invariant tori, where the motion is trivial. The natural basic question is whether in general, for ε small but not zero, this property is at least partially preserved. For example, one can look for a regular canonical transformation $(p, q) = C(p', q', \varepsilon)$, close to the identity, which transforms Hamiltonian (2.1) into an integrable one, i.e.

$$H'(p', q', \varepsilon) \equiv H(C(p', q', \varepsilon), \varepsilon) = h'(p', \varepsilon) \ . \tag{2.2}$$

A less ambitious purpose is to give H' the so-called (nonresonant) normal form up to order r, i.e.

$$H'(p', q', \varepsilon) = h'(p', \varepsilon) + \varepsilon^{r+1} f^{(r+1)}(p', q', \varepsilon) \ , \tag{2.3}$$

for $r \geq 1$. Clearly, such a form would mean that the system behaves as an integrable one, up to times of the order of $\varepsilon^{-(r+1)}$ (more precisely, up to times smaller than $T_r = \varepsilon^{-(r+1)} \| \partial f^{(r+1)} / \partial q' \|^{-1}$, $\|.\|$ denoting a convenient norm. An even less ambitious purpose is to obtain the following weaker normal form:

$$H'(p', q', \varepsilon) = h'(p', q', \varepsilon) + \varepsilon^{r+1} f^{(r+1)}(p', q', \varepsilon) \ , \tag{2.4}$$

where h' depends only on $m < n$ angles, or more generally on $m < n$ independent linear combinations of angles. Precisely, if $\mathcal{M} \subset \mathbf{Z}^n$ is a m-dimensional module (i.e. a m-dimensional "subspace" of $\mathbf{Z}^n$), we admit

$$h'(p', q', \varepsilon) = \sum_{k \in \mathcal{M}} h'_k(p', \varepsilon) e^{ik \cdot q} \ , \tag{2.5}$$

the dot denoting the ordinary scalar product between n-tuples. The previous case corresponds to $m = 0$. From this expression, it clearly follows

$$\frac{\partial h'}{\partial q'}(p', q', \varepsilon) = \sum_{k \in \mathcal{M}} c_k(p', q', \varepsilon) k \ , \qquad c_k = i h'_k(p', \varepsilon) e^{ik \cdot q} \ , \tag{2.6}$$

so that, up to time T_r, the variation of the actions is not arbitrary, being a linear combination of vectors of $\mathcal{M}$, and for any $\alpha \in \mathbf{R}$ orthogonal to $\mathcal{M}$, the quantity $\alpha \cdot p$ is an approximate integral of motion. The form (2.4), with h' given by (2.5), is called the resonant normal form, adapted to $\mathcal{M}$, up to order r.

At first sight, this program appears to be very promising: indeed, let us introduce a canonical transformation dependent on ε, for example by the standard mixed-variables method:

$$\begin{aligned} p &= p' + \varepsilon \frac{\partial S}{\partial q}(p', q, \varepsilon) \\ q' &= q + \varepsilon \frac{\partial S}{\partial p'}(p', q, \varepsilon) \\ S &= S_1(p', q) + \varepsilon S_2(p', q) + \dots \end{aligned} \tag{2.7}$$

(any other method could be equivalently used, in particular the Lie method which will be recalled in the next section). One easily recognizes that to accomplish the program sketched above, one has to solve, at any order $r \geq 1$ in ε, one and the same equation, precisely

$$\omega(p') \cdot \frac{\partial S_r}{\partial q}(p', q) + F_r(p', q) = h'_r(p', q) \ , \tag{2.8}$$

where S_r and h'_r are unknowns, while F_r is a known term, which, for $r = 1$, is nothing but $f(p', q, 0)$. Here ω denotes, as usual, the unperturbed angular frequency, i.e., $\omega = \partial h / \partial p$.

Equation (2.8) is the basic equation of classical perturbation theory; as a matter of fact, you are quite naturally led into the different branches of this theory, when you look for the possible assumptions you need, in order to be able to solve it. The starting point is, however, to establish that, in general, the above equation cannot be solved; this impossibility is indeed the heart of Poincaré's theorem, which (in a slightly reduced form) can be stated as follows:

Theorem 1 (Poincaré): *The following set of assumptions is not compatible:*

i) The unperturbed Hamiltonian $h(p)$ is strictly non-isocronous (or non-degenerate), i.e.

$$det\left(\frac{\partial^2 h}{\partial p \partial p}\right) \neq 0 \qquad \forall p \in \mathcal{B} \; ; \tag{2.9}$$

ii) The perturbation $f(p,q,0)$ has sufficiently many non-vanishing Fourier components: precisely, if

$$f(p,q,0) = \sum_{k \in \mathbf{Z}^n} f_k(p) e^{ik\cdot q} \; , \tag{2.10}$$

then for each $p \in \mathcal{B}$ and each $k \in \mathbf{Z}^n$ there exists at least a $\tilde{k} \in \mathbf{Z}^n$, $\tilde{k}$ parallel to k, such that $f_{\tilde{k}}(p) \neq 0$.

iii) There exists a canonical transformation $(p,q) = C(p',q',\varepsilon)$, defined (say) by (2.7), S being regular in $\mathcal{B}' \times \mathbf{T}^n \times I$, where $\mathcal{B}'$ is an open subset of $\mathcal{B}$ and I an interval containing the origin, which gives the new Hamiltonian $H'(p',q',\varepsilon) = H(C(p',q',\varepsilon),\varepsilon)$ the form

$$H'(p',q',\varepsilon) = h(p') + \varepsilon h_1'(p',q') + \varepsilon^2 f^{(2)}(p',q',\varepsilon) \; , \tag{2.11}$$

with

$$h_1'(p',q') = \sum_{k \in \mathcal{M}} h_{1k}'(p') e^{ik\cdot q} \; , \tag{2.12}$$

$\mathcal{M}$ being a module of dimension $m < n$.

Proof: by substituting (2.7) into (2.11) one finds, as already remarked, equation (2.8), with $r = 1$. One must then impose, as a necessary condition to assumption iii),

$$(ik \cdot \omega(p'))S_{1k}(p') + f_k(p') = 0 \qquad \forall k \notin \mathcal{M}, \quad \forall p' \in \mathcal{B}' \; , \tag{2.13}$$

where $S_{1k}, k \in \mathbf{Z}^n$, denote the Fourier coefficients of S_1.

Now, because of i), arbitrarily close to any $p' \in \mathcal{B}'$ there exists $\tilde{p} \in \mathcal{B}'$ and $\tilde{k} \in \mathbf{Z}^n$, such that $\tilde{k} \cdot \omega(\tilde{p}) = 0$. Moreover, one can easily arrange $\tilde{k} \notin \mathcal{M}$. It follows $f_{\tilde{k}}(\tilde{p}) = 0$ as well as $f_k(\tilde{p}) = 0$ for any k parallel to $\tilde{k}$. This is in conflict with assumption ii).

Remark: Poincaré's original proof is much more complicated and subtle; it also needs the assumption that everything be analytic, while here only differentiability is used. In fact, Poincaré proves more: essentially, he is able to exclude any integral of motion independent of energy, while here only those integrals are excluded, which for $\varepsilon \to 0$ become linear combinations of actions. To our purpose, this formulation is fairly sufficient, as it stresses deeply enough the basic difficulties of classical perturbation theory.

3. Canonical transformations by the Lie method

The different possibilities to escape Poincaré's difficulties will be considered in the next sections. However, we prefer to attack the basic existence theorems of classical perturbation theory, using the Lie method to generate canonical transformations close to the identity. This method, which in our opinion is the most natural and easy for classical perturbation theory, will be here shortly recalled at a purely formal level, i.e., disregarding problems of convergence (which are the counterpart of inversion problems in the standard mixed variables canonical transformations). For more details, see ref. [13-16]; in particular, in ref. [16] all the estimates necessary to ensure convergence are explicitly performed.

The idea of the Lie method is quite simple: as it is well known, if $\chi(p,q)$ is any Hamiltonian, defined in some open domain $\mathcal{D}$, and $C(p',q',t)$ denotes the solution of the corresponding Hamilton equations for initial datum (p',q'), then for any fixed t the change of variables $(p,q) = C(p',q',t)$ (which will be properly defined for sufficiently small t and $(p',q') \in \mathcal{D}'$, $\mathcal{D}'$ being a suitable subset of $\mathcal{D}$), is canonical and, for t small, close to the identity. Thinking to t as a "small parameter" ε, one thus defines a family of canonical transformations $(p,q) = C(p',q',\varepsilon)$ near the identity.

This canonical transformation may appear to be only implicitly defined. Instead, it is completely explicit (in particular, no inversions are needed) whenever it is used to transform functions. Indeed, let g be a regular function on $\mathcal{D}'$, and consider $(\mathcal{U}^\varepsilon g)(p',q') = g\big(C(p',q',\varepsilon)\big)$. One has then

$$\frac{d}{d\varepsilon}(\mathcal{U}^\varepsilon g)(p',q') = \{\chi, g\}\big(C(p',q',\varepsilon)\big) \tag{3.1}$$

and consequently, for any $r > 0$ (if χ and g are sufficiently regular)

$$\mathcal{U}^\varepsilon g = \sum_{j=1}^{r} \frac{\varepsilon^r}{r!} L_\chi^r g + O(\varepsilon^{r+1}) \; , \tag{3.2}$$

with $L_\chi = \{\chi, .\}$.

Let us see how this procedure works in our case, i.e. for Hamiltonian (2.1), at first order ($r = 1$). Performing one canonical transformation and retaining the terms of order one in ε, the new Hamiltonian $H'(p',q',\varepsilon) \equiv (\mathcal{U}^\varepsilon H)(p',q')$ takes the form

$$\begin{aligned} H'(p',q',\varepsilon) &= H(p',q',\varepsilon) + \varepsilon\{\chi, H\}(p',q',\varepsilon) + O(\varepsilon^2) \\ &= h(p') + \varepsilon f(p',q',0) + \varepsilon\{\chi, h\}(p',q') + O(\varepsilon^2) \; , \end{aligned} \tag{3.3}$$

and this is required to coincide up to order one with

$$h'(p',q',\varepsilon) = h(p') + \varepsilon h_1'(p',q') + O(\varepsilon^2) \; . \tag{3.4}$$

Thus one obtains $h'(p') = h(p')$ and for the unknown χ the equation

$$\{\chi, h\}(p', q') + f(p', q', 0) = h_1(p', q') \ , \tag{3.5}$$

i.e.,

$$\omega(p') \cdot \frac{\partial \chi}{\partial q}(p', q') + h'_1(p', q') = f(p', q', 0) \ , \tag{3.6}$$

which is the same as equation (2.8), with $\chi = -S$. (Notice however that the canonical transformations defined by the two methods are not identical, as they differ at higher orders in ε.)

To extend perturbation theory at any order r in ε, one could simply iterate the above procedure, setting $t = \varepsilon^2$, $t = \varepsilon^3$, and so on. A more efficient way is to consider a time-dependent Hamiltonian (generating function)

$$\begin{aligned} \chi(p, q, \varepsilon) &= \sum_{s=1}^{r} s\varepsilon^{s-1} \chi_s(p, q) \\ &= \frac{d}{d\varepsilon} \sum_{s=1}^{r} \varepsilon^s \chi_s(p, q) \end{aligned} \ , \tag{3.7}$$

defining there as before $\mathcal{C}(p', q', \varepsilon)$ to be the solution at "time" ε of the Hamilton equations corresponding to Hamiltonian $\chi(p, q, \varepsilon)$, with initial datum (p', q'). It could be seen [13–16] that for a generic regular function $g(p, q)$ the transformed function $\big(\mathcal{U}^\varepsilon g\big)(p, q) \equiv g\big(\mathcal{C}(p, q, \varepsilon)\big)$ is explicitly produced by the following algorithm, which generalizes (3.2):

$$\big(\mathcal{U}^\varepsilon g\big)(p, q) = \sum_{s=0}^{r} \varepsilon^s g^{(s)}(p, q) + O(\varepsilon^{r+1}) \ , \tag{3.5}$$

with

$$\begin{aligned} g^{(0)}(p, q) &= g(p, q) \\ g^{(s)}(p, q) &= \sum_{j=1}^{s} \frac{j}{s} (L_{\chi_j} g^{(s-j)})(p, q) \ . \end{aligned} \tag{3.6}$$

Notice that the primes previously appearing to distinguish the new variables have been dropped: this is possible and convenient because all of the functions are defined in the same domain, while no confusion is possible, as mixed variables are never used.

By applying this algorithm to Hamiltonian (2.1), the new Hamiltonian $H' = \mathcal{U}^\varepsilon H$ is straightforwardly constructed. As it could be easily seen, if one demands

$$\begin{aligned} H'(p, q, \varepsilon) &= h(p) + \sum_{s=1}^{r} h'_s(p, q) + \varepsilon^{r+1} f^{(r+1)}(p, q, \varepsilon) \\ h'_s(p, q) &= \sum_{k \in \mathcal{M}} h'_{sk}(p) e^{ik \cdot q} \ , \end{aligned} \tag{3.7}$$

then one must solve r equations of the form

$$\omega(p) \cdot \frac{\partial \chi_s}{\partial q}(p, q) + h'_s(p, q) = F_s(p, q) \tag{3.8}$$

for the unknowns χ_s and h'_s, while the known term F_s could be explicitely produced. Clearly, F_s is obtained from $f_1, \ldots, f_s$ as well as from the already determined $h'_1, \ldots, h'_{s-1}$, $\chi_1, \ldots, \chi_{s-1}$, via a finite number of Poisson brackets.

Let us stress that everything, with this algorithm, can be explicitly constructed, and that all of the necessary estimates, leading ultimately to the estimate of the remainder $f^{(r+1)}$ in (3.7), turn out to be quite easy, as shown for example in ref. [16].

4. Birkhoff's perturbation theory

The first idea to escape Poincaré's difficulties is to modify its first assumption. Although condition (2.9) should be considered (at least in a small $\mathcal{B}$) as being generic, there are relevant cases where it is not satisfied, like the Kepler problem, or the case of weakly coupled harmonic oscillators. The latter is indeed the case considered by Birkhoff [4] (after Whittaker [17] and Cherry [18]) who assumed, in place of (2.9),

$$h(p) = \omega \cdot p \ , \tag{4.1}$$

where $\omega = (\omega_1, \ldots, \omega_n) \in \mathbf{R}^n$ is now a fixed angular frequency.

Properly speaking, Birkhoff made use of the cartesian variables $x_j = \sqrt{2p_j} \cos q_j$, $y_j = \sqrt{2p_j} \sin q_j$, $j = 1, \ldots, n$ (the ordinary positions and momenta), and considered the Hamiltonian

$$\tilde{H}(x, y) = \sum_{j=1}^{n} \frac{\omega_j}{2} (x_j^2 + y_j^2) + \sum_{r=3}^{\infty} \tilde{f}_r(x, y) \ , \tag{4.2}$$

$\tilde{f}_r(x, y)$ being a homogeneous polynomial of degree r in $(x_1, \ldots, x_n)$, $(y_1, \ldots, y_n)$. In this expression the small parameter ε does not appear explicitly, but its analog, namely the strength of the perturbation, is given by the distance to the origin, as is formally seen by the standard technique of the "blowing up". Indeed, consider the ball $\mathcal{B}_\varepsilon \in \mathbf{R}^n$ defined by $p_j = \frac{1}{2}(x_j^2 + y_j^2) < \varepsilon$, $j = 1, \ldots, n$, and assume, only for simplicity, that the sum in (4.2) is restricted to r even. Then, turning back to action-angle variables, and performing a trivial rescaling of the actions via a factor ε, in order to report the domain $\mathcal{B}_\varepsilon$ to $\mathcal{B}_1$, Birkhoff's Hamiltonian is converted (up to a factor ε which can be reabsorbed in a rescaling of time) into a Hamiltonian of the form

$$H(p, q, \varepsilon) = \omega \cdot p + \sum_{r=1}^{\infty} \varepsilon^r f_r(p, q) \ , \tag{4.3}$$

where $f_r(p, q)$ is homogeneous polynomial of degree $2(r + 1)$ in $\cos q_1, \ldots, \cos q_n$, $\sin q_1, \ldots, \sin q_n$. So we have a standard Hamiltonian in action-angle variables of the

form (2.1), with $\mathcal{B}$ given by $|p_j| < 1$, $j = 1, \ldots, n$. We thus see by the way that Birkhoff's Hamiltonian violates also hypothesis ii) of Poincaré's theorem: indeed, all of the f_r's, in particular f_1, have only a finite number of Fourier components.

Following the developments by Moser [19], Gustavson [20], and Giorgilli and Galgani [13], let us see how classical perturbation theory does work with Hamiltonian (4.3). If we use the Lie method to generate canonical transformations, we are confronted, at any order r in ε, with the equation

$$\omega \cdot \frac{\partial \chi_r}{\partial q}(p,q) + h'_r(p,q) = F_r(p,q) \; ; \tag{4.4}$$

for r=1 it is $F_1(p,q) = f(p,q,0)$.

Suppose now ω satisfies just $m \leq n-1$ independent resonance relations, i.e. that the set

$$\mathcal{M} = \{k \in \mathbf{Z}^n; \; \omega \cdot k = 0\} \tag{4.5}$$

is a module of dimension m. Equation (4.4), for $r = 1$, is then easily seen to be solved by posing

$$\begin{aligned} h'_1(p,q) &= \sum_{k \in \mathcal{M}} f_{1k}(p) e^{ik \cdot q} \\ \chi_1(p,q) &= \sum_{k \notin \mathcal{M}} \frac{f_{1k}(p)}{ik \cdot \omega} e^{ik \cdot q} \; . \end{aligned} \tag{4.6}$$

Notice that the last sum is finite, so that there are no problems of convergence.

At higher orders in ε, the only delicate point is to recognize that F_r has always a finite number of Fourier components. This is true because, as remarked in the last section, at any order in ε, F_r is obtained via a finite number of Poisson brackets. In fact, it would not be difficult to see (for example, by meaking use of the cartesian coordinates x and y) that $F_r(p,q)$ is a homogeneous polynomial of degree $2(r+1)$ in $\cos q$ and $\sin q$. The procedure to solve equation (4.4) can then be iterated any number of times, thus accomplishing, to any order in ε, the program outlined in section 2, $\mathcal{M}$ being here imposed by the properties of the constant angular frequency ω appearing in the unperturbed Hamiltonian.

This is indeed the sketch of the proof of the following

Theorem 2 (Birkhoff): *Consider the Hamiltonian*

$$H(p,q,\varepsilon) = \omega \cdot p + \sum_{r=2}^{\infty} \varepsilon^{r-1} f_r(p,q) \; , \quad (p,q) \in \mathcal{B} \times \mathbf{Z}^n \; , \tag{4.7}$$

where $f_r(p,q)$ is a homogeneous polynomial of degree $2r$ in $\cos q$, $\sin q$, while $\omega = (\omega_1, \ldots, \omega_n)$ is a fixed angular frequency and $\mathcal{B} \in \mathbf{R}^n$ is the unit ball. Assume

$$\omega \cdot k \neq 0 \qquad \forall k \notin \mathcal{M} \; , \tag{4.8}$$

$\mathcal{M}$ being a module of dimension $m < n$.

Then, for any $r \geq 1$, and ε sufficiently small, there exists a canonical transformation generated by

$$\chi(p,q,\varepsilon) = \sum_{s=2}^{r} \varepsilon^{s-1} \hat{\chi}_s(p,q) \; , \tag{4.9}$$

$\hat{\chi}_s$ being a homogeneous polynomial of degree $2s$ in $\cos q$, $\sin q$, which gives the new Hamiltonian H' the form

$$H'(p,q,\varepsilon) = \omega \cdot p + \sum_{s=2}^{r} \varepsilon^{s-1} \hat{h}_s(p,q) + \varepsilon^{r+1} f^{(r+1)}(p,q,\varepsilon) \tag{4.10}$$

where $\hat{h}_s$ is also a homogeneous polynomial of degree $2s$ in $\cos q$, $\sin q$, satisfying

$$\hat{h}_s(p,q) = \sum_{k \in \mathcal{M}} \hat{h}_{sk}(p) e^{ik \cdot q} \; . \tag{4.10}$$

Remark: one can produce explicit expressions for both $\hat{h}_s$ and $\hat{\chi}_s$, $s = 1, \ldots, r$, and rigorous estimates for the remainder $f^{(r+1)}$.

5. An improvement to Birkhoff's program: Siegel's diofantine condition.

Let us consider again the basic equation of classical perturbation theory, namely

$$\omega \cdot \frac{\partial \chi_r}{\partial t} + h'_r = F_r \; . \tag{5.1}$$

As we have seen, Birkhoff's attitude against this equation is characterized by two assumptions: ω is fixed, while the known term F_r, at any r, has only finitely many Fourier components.

As pointed out by Siegel [5,6], the second condition is not necessary, as far as F_r is analytic, and ω satisfies the additional diofantine condition

$$|\omega \cdot k| \geq \gamma \, |k|^{-\eta} \qquad \forall k \in \mathbf{Z}^n \; , \tag{5.2}$$

where $|k| = |k_1| + \ldots + |k_n|$, while γ and η are any positive constants. As it is well known, for $\eta > n-1$ almost all frequencies satisfy this equation for some γ, although the complementary set, for any fixed γ and η, is open and dense in $\mathbf{R}^n$.

To see how these hypotheses allow solving equation (5.1), let us firstly consider the simpler equation for $\mathcal{G}(q)$

$$\omega \cdot \frac{\partial \mathcal{G}}{\partial q}(q) + \mathcal{F}(q) = 0 \; , \tag{5.3}$$

when the known term $\mathcal{F}(q)$ has zero average, i.e.

$$\mathcal{F}(q) = \sum_{\substack{k \in \mathbf{Z}^n \\ k \neq 0}} \mathcal{F}_k e^{ik \cdot q} \; , \tag{5.4}$$

and can be analitically continued for complex values of the angles, up to $|\Im q| \leq \xi$, where $|\Im q| = \max_{j\leq n} |\Im q_j|$, $\Im$ denoting the imaginary part of a complex number. Because of (5.4), equation (5.3) is formally solved by

$$\begin{aligned} \mathcal{G}(q) &= \sum_{\substack{k\in\mathbb{Z}^n \\ k\neq 0}} \mathcal{G}_k e^{ik\cdot q} \\ \mathcal{G}_k &= \frac{\mathcal{F}_k}{ik\cdot\omega} \ , \end{aligned} \tag{5.5}$$

where in virtue of (5.2) the denominators $ik\cdot\omega$ in particular never vanish.

Beside the formal level, it is quite easy to see that analyticity of $\mathcal{F}$ for $|\Im q| \leq \xi$, together with the diofantine condition (5.2), imply the convergence of the Fourier series for $\mathcal{G}(q)$, in the domain $|\Im q| < \xi$. Moreover, in any domain $|\Im q| \leq \xi - \delta$, where $\delta < \xi$ is any positive constant, $\mathcal{G}$ is bounded by

$$\begin{aligned} &\|\mathcal{G}\|_{\xi-\delta} \leq \|\mathcal{F}\|_\xi \gamma^{-1} C_n \delta^{-2n} \\ &\max_{j\leq n} \|\frac{\partial \mathcal{G}}{\partial q_j}\|_{\xi-\delta} \leq \|\mathcal{F}\|_\xi \gamma^{-1} C_n \delta^{-2n-1} \ , \end{aligned} \tag{5.6}$$

where for any function $\mathcal{F}(q)$, analytic for $|\Im q| \leq \xi$, we denote $\|\mathcal{F}\|_\xi = \sup_{|\Im q|\leq\xi} |\mathcal{F}(q)|$, and C_n is a suitable constant depending only on n. The proof relies on an elementary property, characteristic of analytic functions, which is here recalled as it will be useful later too:

Technical lemma (on Fourier estimates): *If $\mathcal{F}$ is analytic in the domain $|\Im q| \leq \xi$, then one has*

$$|\mathcal{F}_k| \leq \|\mathcal{F}\| e^{-\xi|k|} \ ; \tag{5.7}$$

conversely, if $|\mathcal{G}_k| \leq A e^{-\xi|k|}$, A being any positive constant, then $\mathcal{G}(q) = \sum_{k\in\mathbb{Z}^n} \mathcal{G}_k e^{ik\cdot q}$ is analytic in the domain $|\Im q| < \xi$, and for any $\delta < \xi$ it satisfies the estimates

$$\|\mathcal{G}\|_{\xi-\delta} \leq A D_n e^{-\delta|k|} \ , \tag{5.8}$$

D_n being another n-dependent constant.

For the details, see for example ref. [8,21,22].

These considerations allow directly to solve equation (4.4) for any r: indeed, considering for example the nonresonant case $m = 0$, it is clearly sufficient to set $h'_r(p)$ equal to the average of $F_r(p,q)$ over the angles, in order to be reported to equation (5.3). The estimates (5.6) can then be used to estimate the remainder of order ε^{r+1} in the expression of the new Hamiltonian.

This idea is easily generalized to resonant frequencies, and one can prove the following more general

Lemma 3 (Siegel): *Let $\mathcal{M}'$, $\dim \mathcal{M}' = m'$, be the resonant module for ω, defined by (4.5), and assume the generalized diofantine condition*

$$|\omega\cdot k''| \geq \gamma |k''|^{-\eta} \qquad \forall k'' \in \mathcal{M}'' \ , \tag{5.9}$$

where $\mathcal{M}''$ is a suitable module such that $\mathcal{M}' \oplus \mathcal{M}'' = \mathbf{Z}^n$, and γ, η are positive constants. Let $\mathcal{F}(q)$ be analytic for $|\Im q| \leq \xi$, with no Fourier components on $\mathcal{M}'$, i.e.

$$\mathcal{F}(q) = \sum_{\substack{k\in\mathbb{Z}^n \\ k\notin\mathcal{M}'}} \mathcal{F}_k e^{ik\cdot q} \ . \tag{5.10}$$

Then the equation for $\mathcal{G}(q)$:

$$\omega\cdot\frac{\partial \mathcal{G}}{\partial q}(q) + \mathcal{F}(q) = 0 \tag{5.11}$$

is solved for $|\Im q| < \xi$ by

$$\mathcal{G}(q) = \sum_{\substack{k\in\mathbb{Z}^n \\ k\notin\mathcal{M}'}} \frac{\mathcal{F}_k}{ik\cdot\omega} e^{ik\cdot q} \ , \tag{5.12}$$

and for any $\delta < \xi$ one has the estimates

$$\begin{aligned} &\|\mathcal{G}\|_{\xi-\delta} \leq \|\mathcal{F}\|_\xi \gamma^{-1} C_n \delta^{-2n} \\ &\max_{j\leq n} \|\frac{\partial \mathcal{G}}{\partial q_j}\|_{\xi-\delta} \leq \|\mathcal{F}\|_\xi \gamma^{-1} C_n \delta^{-2n-1} \ , \end{aligned} \tag{5.13}$$

with a suitable n-dependent constant C_n.

Let us notice that for $m' = 0$, i.e. $\mathcal{M}' = \{0\}$, this lemma reduces to the previously sketched result.

On the basis of this lemma, equation (5.1) can be solved even for resonant frequencies, if the "projection" of ω on a suitable $\mathcal{M}''$, according to (5.9), is diofantine, and $\mathcal{F}_r$ is analytic. Indeed, it is sufficient to pose $h'_r(p,q) = \sum_{k\in\mathcal{M}'} F_{rk}(p) e^{ik\cdot q}$, in order that the lemma can be applied to find χ_r.

Birkhoff's perturbation theory can then be carried on, for small ε, up to any order r in ε, even for resonant frequencies and perturbations with infinitely many Fourier components.

6. Perturbations with finitely many Fourier components, and any $\omega(p)$.

In this section we are going to discuss the second condition entering Poincaré's theorem, precisely the requirement that the perturbation $f(p,q,0)$ in Hamiltonian (2.1) has sufficiently many Fourier components. On the contrary, no hypotheses at all will be made on $\omega(p)$, which in particular can satisfy assumption i) of Poincaré's theorem.

As pointed out already by Cherry [18], there are no obstacles to solve an equation of the form

$$\omega(p) \cdot \frac{\partial \chi}{\partial q}(p,q) + F(p,q) = h'(p,q) \tag{6.1}$$

as far as $F(p,q)$ has finitely many Fourier components, say $F_k(p) = 0$ for $|k| > K$. The only delicate point is that such an equation must be solved differently in the different regions of the phase space, according to the different properties of $\omega(p)$. This idea, which was fully exploited only ten years ago by Nekhoroshev [9], is in fact quite simple. The first region one must consider, called $\mathcal{R}_0$, is the one where one can guarantee the non-resonance condition

$$|k \cdot \omega(p)| \geq \alpha_1 \qquad \text{for } k \neq 0,\ |k| \leq K \tag{6.2}$$

α_1 being a suitable constant. For p inside this region, equation (6.1) is solved by

$$\begin{aligned} h'(p,q) &= F_0(p) \\ \chi(p,q) &= \sum_{\substack{k \in \mathbb{Z}^n \\ k \neq 0}} \frac{F_k(p)}{ik \cdot \omega(p)} e^{ik \cdot q} \ . \end{aligned} \tag{6.3}$$

Notice that the sum is finite, and the denominators bounded away from zero.

To solve (6.1) for p outside this region, one must consider all possible moduli $\mathcal{M} \subset \mathbf{Z}^n$, of any dimension m between 1 and n, which can be generated by m independent vectors $(k^{(1)}, \ldots, k^{(m)})$ satisfying $|k^{(j)}| \leq K$, $j = 1, \ldots, m$. Given a sequence $\alpha_1 < \ldots < \alpha_n$, to each m-dimensional module $\mathcal{M}$ one associates the region $\mathcal{R}_{\mathcal{M}}$ where one has

$$|k^{(j)} \cdot \omega(p)| < \alpha_m \tag{6.4}$$

for at least a basis $(k^{(1)}, \ldots, k^{(m)})$ of $\mathcal{M}$, with $|k^{(1)}|, \ldots, |k^{(m)}| \leq K$, while at the same time it is also (for $m < n$)

$$|k \cdot \omega(p)| > \alpha_{m+1} \qquad \forall\, k \notin \mathcal{M}, \quad |k| \leq K \ . \tag{6.5}$$

These regions clearly constitute a covering of the action space (notice that, depending on $\omega(p)$, some regions, including $\mathcal{R}_0$, could be empty). For details on this construction, see the paper by Nekhoroshev [9], or ref. [23]. Inside a generic region $\mathcal{R}_{\mathcal{M}}$, equation (6.1) is conveniently solved by

$$\begin{aligned} h'(p,q) &= \sum_{k \in \mathcal{M}} F_k(p) e^{ik \cdot q} \\ \chi(p,q) &= \sum_{\substack{k \in \mathbb{Z}^n \\ k \notin \mathcal{M}}} \frac{F_k(p)}{ik \cdot \omega(p)} \ , \end{aligned} \tag{6.6}$$

which generalize (6.3).

To have a concrete example of this quite important construction, let us consider first order perturbation theory for a Hamiltonian representing a system of n weakly coupled rotators, precisely

$$H(p,q,\varepsilon) = \sum_{j=1}^{n} \frac{p_j^2}{2} + \varepsilon \sum_{1 \leq i < j \leq n} C_{ij} \cos(q_i - q_j) \ , \tag{6.7}$$

C_{ij} being real constants. For this Hamiltonian it is $\omega(p) = p$, while the perturbation has non-vanishing Fourier components only for $|k| \leq K = 2$. Models belonging to this class have been recently studied in ref. [24,25].

For such a system, the region $\mathcal{R}_0$ is defined, according to (6.2), by

$$|p_i \pm p_j| > \alpha_1 \qquad 1 \leq i < j \leq n \ , \tag{6.8}$$

while (6.3) turns into

$$\begin{aligned} h'(p,q) &= 0 \\ \chi(p,q) &= \sum_{1 \leq i < j \leq n} \frac{C_{ij}}{p_i - p_j} \sin(q_i - q_j) \ . \end{aligned} \tag{6.9}$$

Instead, if one considers, for example, the two-dimensional module generated by $k^{(1)} = (1, -1, 0 \ldots, 0)$ and $k^{(2)} = (1, 0, -1, \ldots, 0)$, the corresponding region $\mathcal{R}_{\mathcal{M}}$ is defined by

$$\begin{aligned} &|p_1 - p_2|,\ |p_1 - p_3| \leq \alpha_2 \\ &|p_1 + p_2|,\ |p_1 + p_3| > \alpha_3 \\ &|p_i \pm p_j| > \alpha_3 \qquad \text{for } (i,j) \neq (1,2),\ (1,3) \ . \end{aligned} \tag{6.10}$$

Inside this region, equation (6.1) is solved by

$$\begin{aligned} h'(p,q) &= C_{12} \cos(q_1 - q_2) + C_{13} \cos(q_1 - q_3) \\ \chi(p,q) &= \sum_{\substack{1 \leq i < j \leq n \\ (i,j) \neq (1,2),\ (1,3)}} \frac{C_{ij}}{p_i - p_j} \sin(q_i - q_j) \ . \end{aligned} \tag{6.11}$$

Essentially, it has been necessary to drop two terms from the expression of χ, in order to avoid small divisors; as a consequence, two terms of the perturbation have not been killed, and still survive in h'.

Let us see how, on the basis of these considerations, perturbation theory can be carried on up to any order r as far as the perturbation $f(p,q,\varepsilon)$ appearing in Hamitonian (2.1) has finitely many Fourier components. For simplicity, we introduce the unessential restriction that f be ε independent. As usual, we give here only a sketch of the construction; for the details, see ref. [16].

The essential point is the following: the property of the perturbation of having finitely many Fourier components is preserved at any r, due to the fact that F_r is constructed via Poisson brackets of functions which, in turn, have finitely many Fourier components. More precisely, it is not difficult to recognize, on the basis of elementary algebraic considerations, that one can always guarantee $F_{rk}(p) = 0$ for $|k| > K_r \equiv rK$, if $f_k(p) = 0$ for $|k| > K$. Thus, having in mind to construct perturbation theory up to a given order $\hat{r}$, one can introduce the above decomposition of the action space, using $K_{\hat{r}}$ in place of K. Within each region one is then guaranteed that the equation

$$\omega(p) \cdot \frac{\partial \chi_r}{\partial q}(p,q) + F_r(p,q) = h'_r(p,q) \tag{6.12}$$

can be solved by (6.6), for any $r \leq \hat{r}$.

This is sufficient as far as a formal construction is considered; beyond the formal level, one can prove the following

Theorem 4 (Giorgilli and Galgani): *Consider the Hamiltonian*

$$H(p,q,\varepsilon) = h(p) + \varepsilon f(p,q) \ , \tag{6.13}$$

and assume:

i) H is analytic for $(p,q) \in \mathcal{B} \times \mathbf{T}^n$;

ii) $f_k(p) = 0$ for $|k| > K$.

Then for any $r \geq 1$, if ε is sufficiently small, one can introduce the above decomposition of the action space into regions $\mathcal{R}_\mathcal{M}$, with reference to $K_r = rK$, and obtain in each region, via a suitable canonical transformation $C_\mathcal{M}(p,q,\varepsilon)$, the new Hamiltonian $H'(p,q,\varepsilon)$ in the adapted normal form

$$H'(p,q,\varepsilon) = h(p) + \sum_{s=1}^{r} \varepsilon^s h'_s(p,q) + \varepsilon^{r+1} f^{(r+1)}(p,q,\varepsilon) \ , \tag{6.14}$$

where

$$h'_s(p,q) = \sum_{k \in \mathcal{M}} h'_{sk}(p) e^{ik\cdot q} \ . \tag{6.15}$$

Remark: Notice that, by increasing r, the decomposition of $\mathcal{B}$ into resonant regions becomes finer and finer, and that h'_s, $s = 1, \ldots, r$, as well as $f^{(r+1)}$, is different for the different regions. Nevertheless, the regularity requirement, which is part of assumption iii) of Poincaré's theorem, is satisfied, at any fixed r, inside each resonant region.

7. Modifying the third of Poincaré's assumptions: Kolmogorov's and Arnold's methods

Up to now, we have seen how, by modifying either assumption i) or ii) of Poincaré's theorem, perturbation theory can be actually carried on up to any order r in ε. Here we will consider the third possibility, i.e., making weaker assumption iii). This can be done by following two basic ideas: either by looking for a weaker form of $h'(p,q,\varepsilon)$, or by accepting an essentially irregular dependence on ε. The first idea has been proposed by Kolmogorov [7], in the sketch of proof he wrote of his celebrated theorem (for a detailed proof exploiting Kolmogorov's idea, see for example ref. [22]); the second one is the so-called method of the ultraviolet cut-off, which was introduced by Arnold [9] in his own proof of Kolmogorov's theorem, and is a basic tool of classical perturbation theory.

Let us start from the latter one, and explain the basic idea in the most elementary situation. Consider equation (3.8) for $r = 1$, i.e.

$$\omega(p) \cdot \frac{\partial \chi_1}{\partial q}(p,q) + h'_1(p,q) = f(p,q) \tag{7.1}$$

(for simplicity f in (2.1) has been taken to be ε-independent). As already seen in section 5, if f is analytic for $|\Im q| \leq \xi$, then its coefficients decay exponentially with $|k|$. Then, for any $K > 0$, one can separate from f the "ultraviolet" part

$$f^{>K}(p,q) = \sum_{\substack{k \in \mathbb{Z}^n \\ |k| > K}} f_k(p) e^{ik\cdot q} \ , \tag{7.2}$$

which, as a consequence of (5.6), is easily seen to satisfy the estimate

$$\begin{aligned} &\|f^{>K}\|_{\xi-\delta} \leq B e^{-\frac{1}{2}\delta K} \\ &B = \left(\frac{1 + e^{-\frac{\xi}{2}}}{1 - e^{-\frac{\xi}{2}}} \right)^n \|f\|_\xi \ . \end{aligned} \tag{7.3}$$

In place of equation (7.1) one now solves

$$\omega(p) \cdot \frac{\partial \chi}{\partial q}(p,q) + h'_1(p,q) = f(p,q) - f^{>K}(p,q) \ , \tag{7.4}$$

i.e., one renounces to kill the ultraviolet part of the perturbation. But (7.4) has the same form as (7.1), with $\hat{f} = f - f^{>K}$ in place of f, so that, $\hat{f}$ having finitely many Fourier components, one can apply the method outlined in the previous section, and decompose the action space as there explained, according to the resonance properties of $\omega(p)$ with integers vectors k satisfying $|k| \leq K$. In each resonant region $\mathcal{R}_\mathcal{M}$ one then obtains

$$\begin{aligned} &H'(p,q,\varepsilon) = h(p) + \varepsilon h'_1(p,q) + \varepsilon f^{>K}(p,q) + \varepsilon^2 f^{(2)}(p,q,\varepsilon) \\ &h'_1(p,q) = \sum_{k \in \mathcal{M}} h'_{1k}(p) e^{ik\cdot q} \ , \end{aligned} \tag{7.5}$$

with a suitable $f^{(2)}$.

According to (7.3), if K is conveniently chosen as a function of δ and ε (a logaritmic dependence on ε^{-1} is sufficient), then the new disturbing term $\varepsilon f^{>K}$ can be made as small as one likes, say of order ε^2 as the remainder $f^{(2)}$ is. Hamiltonian (7.5) will then provide an integrable-like behavior on the appropriate time scale, exactly as Poincaré's Hamiltonian (2.11) does: however, the regularity of H' is now completely lost. Indeed, not only $f^{>K}$ depends discontinuously on ε (a fact which would be no dramatic, because it just concerns the remainder), but clearly, also the decomposition of the phase space, which strongly depends on K, will now acquire (even for fixed $r = 1$) a wild dependence on ε, as one certainly needs $K \to \infty$ for $\varepsilon \to 0$. As a result, the normalized Hamiltonian $h'(p, q, \varepsilon)$ itself will depend on ε in an essentially irregular way, against assumption iii) of Poincaré's theorem (notice that the only case where this does not happen, is that of fixed frequencies).

It could be seen that the method of the ultraviolet cut-off can be carried on up to any order r, as far as ε is sufficiently small, giving a Hamiltonian H' of the form

$$H'(p, q, \varepsilon) = h'(p, q, \varepsilon) + \varepsilon^{r+1} f_1^{(r+1)}(p, q, \varepsilon) + f_2^{(r+1)}(p, q, \varepsilon) \ , \tag{7.6}$$

where the two remainders are small for two different reasons: the former because it has ε^{r+1} in front of it; the latter because it contains ultraviolet parts, with a sufficiently large ε-dependent cut-off K.

Concerning Kolmogorov's proposal of looking for a weaker normal form, we limit ourselves to a very short exposition of the basic idea. Kolmogorov too looks, for $r = 1$, for a Hamiltonian of the form

$$H'(p, q, \varepsilon) = h'(p, q, \varepsilon) + \varepsilon^2 f^{(2)}(p, q, \varepsilon) \ ; \tag{7.7}$$

however, h' neither is integrable, nor it has (as in Poincaré's normal form) any integral of motion. Kolmogorov weak normal form is instead

$$h'(p, q, \varepsilon) = \omega(p^*) \cdot (p - p^*) + \sum_{1 \le i,j \le n} a_{ij}(p, q, \varepsilon)(p_i - p_i^*)(p_j - p_j^*) \ , \tag{7.8}$$

where the parameter $p^* \in \mathcal{B}$ has the only restriction that $\omega(p^*)$ satisfies Siegel's diofantine condition (5.2). As we have seen, this condition is satisfied by a large set of frequencies, and consequently, if condition i) of Poincaré's theorem is maintained, by a large set of values of p^*.

The interest of Hamiltonian (7.8), which in virtue of (7.7) governs the dynamics of our system up to times of order ε^{-2}, is quite transparent: indeed, the torus $\{p^*\} \times \mathbf{T}^n$ is immediately checked to be invariant, and consequently, because of the arbitrariness of p^*, the original Hamiltonian $H(p, q, \varepsilon)$ admits a large set of approximately invariant tori (i.e. invariant on that time scale).

Expression (7.8) for h' is achieved by Kolmogorov via an essential use of Siegel's lemma which, as we have seen, works whenever one has fixed diofantine frequencies and analytic functions. For small ε one is guaranteed that a true (i.e. non formal) canonical transformation is defined; the disadvantage of the method is that a different canonical transformation must be introduced for each different p^*.

8. The limit $r \to \infty$: Kolmogorov-Arnold's and Nekhoroshev's results

Let us recall what we have seen up to now. On the one hand, we have considered Poincaré's non-existence theorem, which definitely makes evident the impossibility of a naive perturbation theory in classical Hamiltonian dynamics. On the other hand, we have seen that many roads nevertheless remain open, which allow one to produce different forms of classical perturbation theory. As we have sometimes commented, one can work, for sufficiently small ε, up to any finite order r. The natural question then arises whether the limit $r \to \infty$ can be performed.

A first negative result comes from a theorem by Siegel [5]: for most perturbations (in a natural topology) and any fixed ε, Biskhoff's canonical transformation diverges. Positive results come instead, as it is well known, from the theorems of Kolmogorov [7], Arnold [8] and Nekhoroshev [9].

Although proceeding on different roads, Kolmogorov's and Arnold's methods lead to the same celebrated result: if $H(p, q, \varepsilon) = h(p) + \varepsilon f(p, q)$ is analytic, $h(p)$ is strictly non isocronous, and ε is sufficiently small, then invariant tori are abundant in phase space.

We recall here a quite recent formulation of this theorem, due to Poeschel [10] and independently to Chierchia and Gallavotti [11], which is the most powerful one and at the same time the most interesting for a comparison with Poincaré's theorem.

Theorem 5 (Poeschel; Chierchia and Gallavotti): *Consider the Hamiltonian*

$$H(p, q, \varepsilon) = h(p) + \varepsilon f(p, q) \ , \tag{8.1}$$

defined for $p \in \mathcal{B} \subset \mathbf{R}^n$ and $q \in \mathbf{T}^n$, and assume :

i) H is analytic for $(p, q) \in \mathcal{B} \times \mathbf{T}^n$;

ii) $\det\left(\dfrac{\partial h}{\partial p \partial p}\right) \neq 0$ for $p \in \mathcal{B}$;

iii) ε is sufficiently small.

Then there exist a canonical transformation $C(p, q, \varepsilon)$ and an integrable Hamiltonian $h'(p, \varepsilon)$ of class C^∞ in $\mathcal{B}_\varepsilon \times \mathbf{T}^n \times I$, where I is an interval around the origin and $\mathcal{B}_\varepsilon$ is a subset of $\mathcal{B}$ whose border is close to the border of $\mathcal{B}$, such that the new Hamiltonian $H'(p, q, \varepsilon) = H(C(p, q, \varepsilon), \varepsilon)$ satisfies the relation

$$H'(p, q, \varepsilon) \overset{\mathcal{B}_\varepsilon^*}{=} h'(p, \varepsilon) \ , \tag{8.2}$$

where $\mathcal{B}_\varepsilon^*$ is a closed set, whose measure is ε-close to the measure of $\mathcal{B}$, and $\overset{\mathcal{B}_\varepsilon^*}{=}$ denotes equality of the two members, as well as of all their derivatives, when $p \in \mathcal{B}_\varepsilon^*$.

Remark: it trivially follows from (8.2) that p is a constant of motion when it belongs to $\mathcal{B}_\varepsilon^*$; thus, there exist n integrals of motion for the original Hamiltonian (8.1), whenever the initial datum belongs to the image, through $\mathcal{C}$, of $\mathcal{B}_\varepsilon^* \times \mathbf{T}^n$. One can appreciate how thin is the separation between this result and Poincaré's non-existence theorem.

Let us come to Nekhoroshev's Theorem. This is in a sense more classical than Kolmogorov's theorem, as no "strange" sets like $\mathcal{B}_\varepsilon^*$ are involved; the proof is also more classical. While in Kolmogorov's theorem the limit $r \to \infty$ is studied at small but fixed ε, in Nekhoroshev's theorem one takes instead $r \to \infty$ and simultaneously $\varepsilon \to 0$: more precisely, one proves that r can be consistently chosen to be a convenient negative power of ε; as a consequence the remainder, which was of order ε^r, becomes exponentially small with ε, and correspondingly the time scale for which the system behaves as it were integrable becomes exponentially large.

As shown by Gallavotti [26] (see also ref. [27] and [16]), this method works straightforwardly in the case of fixed frequencies satisfying a diofantine condition, leading to the following

Theorem 6 (Nekhoroshev-Gallavotti): *Let $H(p,q,\varepsilon) = \omega \cdot p + \varepsilon f(p,q)$ be analytic for $(p,q) \in \mathcal{B} \times \mathbf{T}^n$, $\omega \in \mathbf{R}^n$ being a fixed angular frequency which, for suitable positive constants γ and η, satisfies the diofantine condition*

$$|\omega \cdot k| > \gamma |k|^{-\eta} \qquad \forall k \in \mathbf{Z}^n ; \quad k \neq 0 . \tag{8.3}$$

Then, if ε is sufficiently small, there exist constants P, T, α, β, such that any motion $\big(p(t),(q(t)\big)$, with $p(0)$ at distance at least $P\varepsilon^\alpha$ from the border of $\mathcal{B}$, satisfies the estimate

$$|p_j(t) - p_j(0)| < P\varepsilon^\alpha \qquad j = 1, \dots, n \tag{8.4}$$

for

$$|t| \leq T e^{-\left(\frac{1}{\varepsilon}\right)^\beta} . \tag{8.5}$$

For non isocronous systems, additional considerations of geometric nature are necessary. Indeed, at variance with Kolmogorov's case, here one wants to work in the whole of phase space. As already explained, in this case one must divide the action space into regions $\mathcal{R}_\mathcal{M}$, having well defined resonance properties, and work there separately. Clearly, this is possible for arbitrarily high times only if the orbit is somehow trapped inside these regions. Such a behavior is guaranteed by a geometric condition on $h(p)$, called by Nekhoroshev *steepness*, which is in fact nothing but a generalization of convexity. If we restrict ourselves to convex unperturbed Hamiltonians, the statement one can prove is the following

Theorem 7 (Nekhoroshev): *Consider the Hamiltonian*

$$H(p,q,\varepsilon) = h(p) + \varepsilon f(p,q) , \tag{8.6}$$

defined for $p \in \mathcal{B} \subset \mathbf{R}^n$ and $q \in \mathbf{T}^n$, and assume :

i) H is analytic for $(p,q) \in \mathcal{B} \times \mathbf{T}^n$;

ii) $\left(\dfrac{\partial h}{\partial p \partial p}\right)$ is positive definite in $\mathcal{B}$;

iii) ε is sufficiently small.

Then there exist constants P, T, α, β, such that any motion $\big(p(t),(q(t)\big)$, with $p(0)$ at distance at least $P\varepsilon^\alpha$ from the border of $\mathcal{B}$, satisfies the estimate

$$|p_j(t) - p_j(0)| < P\varepsilon^\alpha \qquad j = 1, \dots, n \tag{8.7}$$

for

$$|t| \leq T e^{-\left(\frac{1}{\varepsilon}\right)^\beta} . \tag{8.8}$$

For a more detailed statement of Nekhoroshev theorem, see the papers by Nekhoroshev [9], or ref. [23] and [27]. In the latter reference some further consequences of Nekhoroshev theorem are also drawn.

9. Extending Poincaré's non-existence result

As already remarked, the gap between Poincaré's non-existence theorem and theorem 5 is quite thin. In fact, the separation can be made even thiner, as Poincaré's result can be further extended, by a natural continuation of his arguments, as shown by Fermi [2,3] already in 1923.

The problem can be stated as follows: the presence of any regular integral of motion $I(p,q,\varepsilon)$ would imply the existence of a continuous foliation of the phase space into $(2n-1)$-dimensional invariant manifolds, of equation $I(p,q,\varepsilon) = I_0$, for I_0 in a suitable interval. The question posed by Fermi is now the following: is it possible that at least one of these manifolds — i.e., possibly even a single sheet — remains invariant for ε small but non zero? The answer, within the same assumptions of Poincaré, is negative, as far as the number of degrees of freedom exceeds two, and the manifold one is looking for is analytic. The necessity of $n > 2$ is easily understood: indeed, let ω^* be diofantine; then $\lambda\omega^*$, for $\lambda \geq 1$, is also diofantine. It follows (theorem 5) that the $(n+1)$-dimensional manifold given by $\omega(p) = \lambda\omega^*$, $\lambda \geq 1$ being a free parameter, is invariant. But for $n = 2$, it is $n+1 = 2n-1$, so that the reason why Fermi cannot exclude $(2n-1)$-dimensional invariant manifolds for $n = 2$, is just that they do exist (although, according to Poincaré's theorem, their union has empty interior). As already remarked in the introduction, the reason Fermi was so interested in the possible existence of these manifolds is the connection with the ergodic problem.

Proceeding along this direction, one can further extend this result [12], obtaining the following

Theorem 8 (Poincaré–Fermi): *Consider a nearly integrable Hamiltonian system with n degrees of freedom, with Hamiltonian*

$$H(p,q,\varepsilon) = h(p) + \varepsilon f(p,q) \ , \tag{9.1}$$

h and f being differentiable for $(p,q) \in \mathcal{B} \times \mathbf{T}^n$. Assume:

i) $\det\left(\dfrac{\partial^2 h}{\partial p \partial p}\right) \neq 0$ *for* $p \in \mathcal{B}$;

ii) the perturbation $f(p,q)$ has sufficiently many Fourier components, as in Poincaré's assumption ii);

iii) there exists a m-dimensional invariant manifold V, $m < 2n$, defined by $2n-m$ equations of the form

$$I_j(p,q,\varepsilon) = 0 \qquad j = 1,\dots,2n-m \ , \tag{9.2}$$

with $I_j(p,q,\varepsilon) = I_j^{(0)}(p) + \varepsilon I_j^{(1)}(p,q,\varepsilon)$, the gradients of $I_j^{(0)}$, $j = 1,\dots,2n-m$ and the gradient of $h(p)$ being linearly independent.

Then V must be exactly $(n+1)$-dimensional, and defined by an equation of the form $\omega(p) = \lambda\omega^*$, where λ is a free parameter (within a suitable range), and ω^* is a non-resonant angular frequency.

Remark: to exclude manifolds of dimension $m = 2n-1$, the assumption that grad $I^{(0)}$ and grad h be linearly independent can be replaced (as Fermi did) by analyticity of both H and I in all variables.

After this result, we can say that Poincaré's non-existence arguments, and modern existence theorems, essentially touch each other.

REFERENCES

[1] H. Poincaré: *Les Méthodes Nouvelles de la Méchanique Céleste*, Vol. 3 (Gautier–Villars, Paris, 1899).

[2] E. Fermi: *Nuovo Cimento* **26**, 105 (1923).

[3] E. Fermi: *Nuovo Cimento* **25**, 267 (1923); *Phys. Z.*, **24** (1923)

[4] G. D. Birkhoff: *Dynamical Systems*, (Am. Math. Soc., New York 1927).

[5] C. L. Siegel: *Ann. Math.*, **43**, 607 (1942).

[6] C. L. Siegel and J. K. Moser: *Lectures on Celestial Mechanics* (Springer Verlag, Hidelberg 1971).

[7] A. N. Kolmogorov: *Dokl. Akad. Nauk. SSSR* **98**, 527 (1954); English translation in G. Casati and G. Ford (Editors): *Lecture Notes in Physics* No. 93 (Springer Verlag, Berlin, 1979).

[8] V. I. Arnold: *Usp. Mat. Nauk* **18**, 13 (1963) [*Russ. Math. Surv.* **18**, 9 (1963)]; *Usp. Mat. Nauk* **18**, 91 (1963) [*Russ. Math. Surv.* **18**, 85 (1963)].

[9] N. N. Nekhoroshev: *Usp. Mat. Nauk* **32**, (1977) [*Russ. Math. Surv.* **32**, 1 (1977)]; *Trudy Sem. Petrows.* No. 5, 5 (1979).

[10] J. Poeschel: *Commun. Pure Appl. Math.* **35**, 653 (1982); *Celestial Mechanics* **28**, 133 (1982).

[11] L. Chierchia and G. Gallavotti: *Nuovo Cimento B* **67**, 277 (1982).

[12] G. Benettin, G. Ferrari, L. Galgani and A. Giorgilli: *Nuovo Cimento B* **72**, 137 (1982).

[13] A. Giorgilli and L. Galgani: *Celestial Mechanics* **17**, 267 (1978).

[14] J. Henrard: *Celestial Mechanics* **3**, 107 (1970); *Celestial Mechanics* **10**, 497 (1974).

[15] M. Rapaport: *Celestial Mechanics* **28**, 291 (1982).

[16] A. Giorgilli and L. Galgani: *Rigorous Estimates for the Series Expansions of Hamiltonian Perturbation Theory*, preprint.

[17] E. T. Whittaker: *Proc. Roy. Soc. Edinb.* **A37**, 95 (1916).

[18] T. M. Cherry: *Proc. Cambr. Phil. Soc.* **22**, 287, 325 and 510 (1924).

[19] J. Moser: *Lectures on Hamiltonian Systems*, in *Mem. Am. Math. Soc.* No 81, 1 (1968).

[20] F. Gustavson: *Astron. J.* **71**, 670 (1966).

[21] G. Gallavotti: *The Elements of Mechanics* (Springer Verlag, Berlin 1983).

[22] G. Benettin, L. Galgani, A. Giorgilli and J.-M. Strelcyn: *Nuovo Cimento B* **79**, 201 (1984).

[23] G. Benettin, L. Galgani, A. Giorgilli: *A Proof of Nekhoroshev's Theorem for the Stability Times in Nearly Integrable Hamiltonian Systems*, to appear in *Celestial Mechanics* .

[24] G. Benettin, L. Galgani, A. Giorgilli: *Classical Perturbation Theory for a System of Weakly Coupled Rotators*, and: *Numerical Investigation on a Chain of Weakly Coupled Rotators in the Light of Classical Perturbation Theory*, both to appear in *Nuovo Cimento B*.

[25] C. E. Wayne: *On the Elimination of Non-Resonant Harmonics* and *Bounds on Trajectories of a System of Weakly Coupled Rotators*, to appear in *J. Stat. Phys.*

[26] G. Gallavotti: lectures given at the 1984 *Les Houches* Summer School, to be published.

[27] G. Benettin and G. Gallavotti: *Exponential Estimates for the Stability times in Nearly Integrable Hamiltonian Systems*, to appear in *J. Stat. Phys.*

EQUIPARTITION PROBLEM IN THE THERMODYNAMIC LIMIT AND ULTRAVIOLET CATASTROPHE IN CLASSICAL FIELD THEORY

Roberto Livi[*], Marco Pettini°, Stefano Ruffo[*], Angelo Vulpiani"

[*] Dipartimento di Fisica dell'Università degli Studi di Firenze and Istituto Nazionale di Fisica Nucleare
Largo E. Fermi 2, I-50125 Firenze, Italy.

° Osservatorio Astrofisico di Arcetri and Gruppo Nazionale di Astronomia del Consiglio Nazionale delle Ricerche,
Largo E. Fermi 5, I-50125 Firenze, Italy.

" Dipartimento di Fisica, Università "La Sapienza" and Gruppo Nazionale di Struttura della Materia del Consiglio Nazionale delle Ricerche,
Piazzale A. Moro 2, I-00185 Roma, Italy.

(Dedicated to Massimo Sparpaglione, our friend)

ABSTRACT

It is shown that both the thermodynamic and the continuum limit do not necessarily imply the equipartition of energy. In the latter case equipartition is incompatible with analiticity properties of the equations of motion.
We study some applications to models of weakly nonlinear solids (Fermi-Pasta-Ulam model) and of a radiant cavity.

Since the first numerical experiment on weakly coupled oscillators performed by Fermi, Pasta and Ulam (F.P.U.) (1) it was evident that for hamiltonian systems at low energy there is not equipartition of the energy among the degrees of freedom and that ergodicity is broken. This result was in complete disagreement with the idea (usually assumed by the community of physicists) that the ergodic problem was substantially solved. This opinion was based on a wrong interpretation of a theorem of Poincaré and Fermi on the non existence of prime integrals independent from the energy in a generic hamiltonian system.

The scenario of the breakdown of ergodicity (and equipartition) has been confirmed theoretically (essentially by K.A.M. theorem) (2) and in a huge amount of numerical experiments (3). Roughly speaking in a generic hamiltonian system with N degrees of freedom there is a "critical" energy E_c so that for $E \lesssim E_c$ there is a prevalence of ordered orbits while for