

Effective Stability in Hamiltonian Systems in the light of Nekhoroshev's Theorem

Antonio Giorgilli

Dipartimento di Matematica dell'Università, Via Saldini 50, 20133 MILANO, Italy

Abstract.

The methods of classical perturbation theory are revisited in the light of a rigorous algebraic approach and of Nekhoroshev's theorem on stability over exponentially large times. The applications to the restricted three body problem and to a statistical model of a diatomic gas of identical molecules are illustrated, with the aim of giving good estimates for the size of the stability region and for the dependence on the number of degrees of freedom.

1. Introduction

The present lecture is concerned with the study of a near to integrable Hamiltonian system of differential equations. Such a problem is of interest in many fields of mathematical physics and astronomy. My plan is to illustrate recent rigorous results concerning such kind of systems, in the spirit of Nekhoroshev's theorem on stability over exponentially large times.

Let me recall that, according to Poincaré^[1], the fundamental problem of dynamics is the study of a canonical system with Hamiltonian

$$H(p, q, \varepsilon) = h(p) + \varepsilon f(p, q, \varepsilon) , \quad (1)$$

where $p = (p_1, \dots, p_n) \in \mathcal{G} \subset \mathbf{R}^n$, with \mathcal{G} an open set, and $q = (q_1, \dots, q_n) \in \mathbf{T}^n$ are action-angle variables, ε is a real (small) parameter and $H(p, q, \varepsilon)$ is assumed to be an analytic function of p , q and ε .

The dynamical evolution of the unperturbed system, i.e. the one with Hamiltonian $H(p, q, 0) = h(p)$, is a well known topic: the system admits n independent prime integrals, namely the actions p_1, \dots, p_n , and the flow in phase space is given by

$$p(t) = p^{(0)} , \quad q(t) = \omega(p^{(0)})t + q^{(0)} ,$$

where $\omega(p^{(0)}) = \frac{\partial h}{\partial p}(p^{(0)})$ are the unperturbed frequencies and $(p^{(0)}, q^{(0)})$ the initial data. So, the phase space is foliated into invariant tori, and the motion on a torus is either quasi periodic or periodic, according to the possible existence of resonance relations among the frequencies, namely relations like $k \cdot \omega = 0$ with $0 \neq k \in \mathbf{Z}^n$.

The naive approach of perturbation theory is essentially an attempt to extend the same plain picture to the perturbed system, namely the case $\varepsilon \neq 0$. From a technical viewpoint, one usually looks for a near to identity canonical transformation which removes the dependence of the transformed Hamiltonian on the angle variables. Fortunately, as was proven by Poincaré himself^[1], such a naive program fails. Indeed, he proved that under nondegeneracy conditions on the unperturbed Hamiltonian $h(p)$, namely $\det \left(\frac{\partial^2 h}{\partial p_j \partial p_l} \right) \neq 0$, the system (1) does not generally admit analytic prime integrals independent of the Hamiltonian.

On the other hand, it is well known that the series produced by perturbative methods, although generally non convergent, are very useful in practical applications. The relevance of such a fact is stressed by Poincaré, who also discusses the possibility that these series are in fact asymptotic expansions^[1]. A rigorous approach on this direction can already be found in some works of Moser^[2] and Littlewood^[3], and a general result is given by Nekhoroshev's theorem^[4]. Roughly speaking, one renounces to look for information on the dynamical evolution of the system for infinite times, being satisfied with results concerning a time interval like

$$T = T_* \exp\left(\frac{\varepsilon_*}{\varepsilon}\right)^a, \quad (2)$$

with suitable constants T_* , ε_* and a . Due to the exponential dependence on the inverse of the perturbative parameter ε , such a time interval can become very large, and possibly exceed any realistic time. In the very words of Littlewood^[3], “while not eternity, this is a considerable slice of it”.

Now, it is natural to raise the question whether such results can be effectively applied to physical systems, mainly in the case of statistical models, with a large number of degrees of freedom. A first problem here concerns the size of the constants T_* , ε_* and a in the exponential estimate (2): when explicitly computed, these constants turn out to be ridiculously small. For example, if one considers the solar system the natural perturbation parameter is the ratio between Jupiter's mass and the mass of the Sun; if one tries to apply the Nekhoroshev's result, as in the original formulation, one finds that the mass of Jupiter should be several orders of magnitude less than that of a proton. A second remark concerns the dependence of the same constants on the number of degrees of freedom: for example, still using the original Nekhoroshev's formulation, one has $a \sim 1/n^2$. This fact led some authors to conclude that Nekhoroshev's like results are not applicable to statistical systems, where the number n of degrees of freedom goes to infinity^[5].

Of course, one is not authorized to immediately conclude that Nekhoroshev's theorem is a wonderful mathematical result, but completely useless for physics. Indeed, the explicit values of the constants are estimated via a set of inequalities which are often far from being optimal, and could hardly be optimized without more specific assumptions on the physical system. The aim of the present lecture is precisely to discuss some recent theoretical improvements which give a positive, although partial answer to the question raised above.

2. Model problems

Let me illustrate the theory by making explicit reference to two physical examples: the stability of an elliptic equilibrium point, which is a classical topic, and the problem of relaxation times in statistical systems, which has recently been investigated^[6,7].

2.1. The stability of an elliptic equilibrium.

As a first model, let's consider a system of harmonic oscillators described by the Hamiltonian

$$H(x, y) = \frac{1}{2} \sum_{l=1}^n \omega_l (x_l^2 + y_l^2) + \sum_{s>0} H_s(x, y), \quad (3)$$

where H_s is a homogeneous polynomial of degree $s + 2$ in the canonical variables $(x, y) \in \mathbf{R}^{2n}$, and $\omega \in \mathbf{R}^n$ is the vector of the harmonic frequencies, which are assumed to be all nonvanishing. This is a classical problem, and can be considered as the simple version of the “fundamental problem of dynamics”. The reason is that the frequencies of the unperturbed

system are independent of the initial point, so that the formal approach is definitely simpler. Specific models of such a kind are the nonlinear chain, like the celebrated FPU model^[8], and the Lagrangian triangular equilibrium points L_4 and L_5 of the circular restricted problem of three bodies.

For the FPU model, namely a chain of $n + 2$ identical point masses on a straight line, connected with nonlinear springs and with fixed ends, the harmonic frequencies are given by $\omega_l = 2\alpha \sin \frac{l\pi}{2(n+1)}$, $1 \leq l \leq n$, (α being a constant).

For the Lagrangian points, denoting by m and M the masses of the primaries, and introducing the dimensionless parameter $\mu = m/(M + m)$, the triangular points turn out to be elliptic (and so linearly stable) for $0 < 27\mu(1 - \mu) < 1$, but, unlike the FPU model, the frequencies have different signs. For example, considering the point L_4 for the Sun–Jupiter case, and allowing the asteroid to move in space one gets $\mu \sim 0.95387 \times 10^{-3}$, $\omega_1 \sim 0.99676$, $\omega_2 \sim -0.80464 \times 10^{-1}$ and $\omega_3 = 1$.

In these models two problems naturally arise.

- i. *Stability of the equilibrium*: this is easily solved for the FPU model, since the Hamiltonian has a minimum in the equilibrium, but remains unsolved, up to now, for the L_4 point, because the harmonic frequencies have different sign, so that the equilibrium is a saddle point for the Hamiltonian; in the latter case strong results have been obtained in the framework of KAM theory, but definite conclusions can only be drawn in the planar case.
- ii. *Freezing of the harmonic actions*: since the pioneering work of Fermi, Pasta and Ulam a lot of numerical experiments produced evidence that, at least for low total energy, there is no equipartition of energy among the normal modes. Such a fact is relevant for the foundations of classical statistical mechanics.

A possible approach, which is valid for both these problems, consists in relaxing the definition of stability: instead of looking for a confinement of orbits (or for freezing of the actions) over an infinite time, one tries to prove that this happens up to a finite, but large time.

2.2. Relaxation times in statistical systems

The second model has been investigated in refs. [6] and [7]. One considers a canonical system with analytic Hamiltonian

$$H(p, x, \pi, \xi) = \hat{h}(p, x) + h_\omega(\pi, \xi) + f(p, x, \pi, \xi) , \quad (4)$$

where

$$h_\omega(\pi, \xi) = \frac{1}{2} \sum_{l=1}^{\nu} (\pi_l^2 + \omega_l^2 \xi_l^2) , \quad (\pi, \xi) \in \mathbf{R}^{2\nu}$$

is the Hamiltonian of a system of harmonic oscillators, $\hat{h}(p, x)$ is the Hamiltonian of a generic n -dimensional system, and $f(p, x, \pi, \xi)$ a coupling term which is assumed to be of order ξ , and so to vanish for $\xi = 0$. More specific models are the realization of physical constraints and a statistical system like a diatomic gas of identical molecules.

In the case of constraints the Hamiltonian $\hat{h}(p, x)$ describes the motion of the constrained system, while $h_\omega(\pi, \xi)$ describes the vibrations of the constraints. One is then interested in investigating the relations between the time evolution of the constrained system and that of the whole system in the limit $\omega \rightarrow \infty$.

The statistical model of a diatomic gas can be studied in the same spirit: the Hamiltonian $\hat{h}(p, x)$ describes the translational and rotational degrees of freedom, while $h_\omega(\pi, \xi)$ describes the internal vibrations. Assuming that the molecules interact via a regular short range

potential, one is still interested in understanding the dynamical evolution in the limit of large ω 's. The peculiarity of such a system with respect to that of constraints lies in the fact that the number ν of molecules is large, but all the frequencies are equal, so that the system is completely resonant.

The identification of a perturbative parameter in these systems proceeds as follows. Write $\omega = \lambda\Omega$ with Ω of the same order of the inverse of a typical time scale of the constrained system (for example the characteristic time for the collision of two molecules, which is non zero if the interaction potential is regular) and large λ ; then transform the variables according to $\pi = \pi'\sqrt{\lambda\Omega}$ and $\xi = \xi'/\sqrt{\lambda\Omega}$, and assume the total energy of the constraints (or of the internal vibrations) to be finite, so that the variables (π', ξ') turn out to be confined in a disk of size $1/\sqrt{\lambda}$. Then the Hamiltonian can be given the form, omitting primes,

$$H(p, x, \pi, \xi, \lambda) = \hat{h}(p, x) + \lambda h_\Omega(\pi, \xi) + \frac{1}{\lambda} f_\lambda(p, x, \pi, \xi) \quad (5)$$

with

$$h_\Omega(\pi, \xi) = \frac{1}{2} \sum_{l=1}^{\nu} \Omega_l (\pi_l^2 + \xi_l^2) \quad (6)$$

(here, a straightforward computation would give $\lambda^{-1/2}$ in front of f , but f itself turns out to be of order $\lambda^{-1/2}$, since it vanishes for $\xi = 0$). Thus, the problem looks similar to the fundamental problem of dynamics, corresponding to the Hamiltonian (1), since λ^{-1} plays the role of a small parameter; at the same time it resembles the problem of the elliptic equilibrium, due to the form of h_Ω . In fact, it is definitely different from both the models above, because \hat{h} is not restricted to represent an integrable system: one needs a perturbative scheme for a nonintegrable system.

The approach taken here consists in considering the whole system as composed of two separate subsystems $\hat{h}(p, x)$ and $h_\Omega(\pi, \xi)$, each with its own internal, possibly chaotic, dynamics, and in looking only for a bound on the energy exchange between these two subsystems over a large time scale, in the spirit of Nekhoroshev's theorem.

3. The perturbation scheme

My aim is now to describe a general perturbation scheme that can be applied to both the models above. The main tool will be an algebraic approach to perturbation theory based on explicit recursive formulae, like in the Lie transform method described in refs. [9–10]. Such a scheme is made rigorous by adding explicit estimates on the convergence properties of the series so produced; in doing this, particular attention is paid to obtaining good estimates, mainly for what concerns the dependence on the number of degrees of freedom. The general formulation of the present scheme can be found in ref. [11].

3.1. Algebraic framework

The first step consists in translating the concept of ‘‘perturbation order’’ into an algebraic structure. This is given by introducing a function space \mathcal{P} which is an algebra with respect to sum, scalar multiplication and product, and a sequence $\{\mathcal{P}_s\}_{s \geq 0}$ of subspaces of \mathcal{P} such that:

- i. \mathcal{P}_s is a linear space for $s \geq 0$, and $\mathcal{P} = \bigoplus_{s \geq 0} \mathcal{P}_s$
- ii. for any $f \in \mathcal{P}_s$ and $g \in \mathcal{P}_r$ one has $fg \in \mathcal{P}_{s+r}$
- iii. there exists $2n$ nonnegative integers $(a_1, \dots, a_n, b_1, \dots, b_n)$ such that for any $f \in \mathcal{P}_s$ one has $\frac{\partial f}{\partial p_j} \in \mathcal{P}_{s-a_j}$ and $\frac{\partial f}{\partial q_j} \in \mathcal{P}_{s-b_j}$ for $1 \leq j \leq n$. Here, p, q are the canonical coordinates.

The properties above essentially ask that the algebraic structure of the classes \mathcal{P}_s be compatible with the operations needed in order to build a perturbative scheme, namely sums, derivatives and Poisson brackets. Moreover, one can expand any function $f \in \mathcal{P}$ as $f = \sum_{s \geq 0} f_s$ with $f_s \in \mathcal{P}_s$, like in a power series development in a small parameter. The main, and important, difference with respect to the usual one-parameter scheme is that nonhomogeneous expansions are also allowed. Indeed, introducing for $s \geq 0$ the space $\mathcal{P}^{[s]} = \bigoplus_{l \geq s} \mathcal{P}_l$, one can consider an expansion like $f = \sum_{s \geq 0} f^{[s]}$ with $f^{[s]} \in \mathcal{P}^{[s]}$. Such an expansion is clearly not unique, but has the advantage that one is allowed to control the minimum order of each term. A more detailed control is possible by introducing for $0 \leq s \leq r$ the space $\mathcal{P}^{[s,r]} = \bigoplus_{s \leq l \leq r} \mathcal{P}_l$, so that also the maximum order of each term is bounded. Such a generalization is the technical tool which allows to build up a perturbation scheme for the model problem proposed in sect. 2.2. Indeed, in that case one has to deal with at least two parameters—the dimensionless quantity λ^{-1} and the size of the (π, ξ) variables. In such a case the straightforward application of the usual one-parameter expansion does not work. More generally, one can imagine cases when the identification of an expansion parameter is not trivial at all; a possible way out is to decide *a priori*, on the basis of heuristic considerations, *what is small and why*, and to check the consistency of that choice on the basis of rigorous estimates on the expansions. This is possible in the present general scheme.

In order to produce rigorous estimates one must take into account the domains where the functions are defined, and to introduce a suitable norm which allows to estimate the size of the functions.

The choice of the domains is the usual one: starting with a suitable subset M_0 of the phase space, one builds the domain

$$\mathcal{D}_R(M_0) = \bigcup_{(p,q) \in M_0} \Delta_R(p,q) \quad (7)$$

where $\Delta_R(p,q)$ is the polydisk

$$\Delta_R(p,q) = \{(p',q') \in \mathbf{C}^{2n} : |p_j - p'_j| \leq \varrho_j, |q_j - q'_j| \leq \sigma_j, 1 \leq j \leq n\}. \quad (8)$$

Here, $R \equiv (\varrho_1, \dots, \varrho_n, \sigma_1, \dots, \sigma_n)$ is a vector with positive components. For brevity, $\mathcal{D}_R(M_0)$ will be simply denoted by \mathcal{D}_R in what follows.

Having so defined the domain, one considers the function space $\mathcal{F}_R \subseteq \mathcal{P}$ of these functions which are analytic in the interior of \mathcal{D}_R and bounded on \mathcal{D}_R , and introduces a norm $\|\cdot\|_R$ on \mathcal{F}_R with the following properties:

- i. if $\mathcal{D}_{R'} \subseteq \mathcal{D}_R$ then $\|f\|_{R'} \leq \|f\|_R$;
- ii. for any $(p,q) \in \mathcal{D}_R$ one has $|f(p,q)| \leq \|f\|_R$;
- iii. there exists a positive constant C such that for any $\chi \in \mathcal{F}_{(1-d')R}$ with $0 \leq d' < 1$, for any $f \in \mathcal{F}_R$ and for any positive $d \leq 1 - d'$ one has

$$\|L_\chi f\|_{(1-d)R} \leq \frac{C}{(d+d')d} \|\chi\|_{(1-d')R} \|f\|_R ;$$

here, $L_\chi f$ denotes the Lie derivative of f , i.e., in fact, the Poisson bracket $\{\chi, f\}$. Let me add some comment. In order to bound the size of a function, the natural choice would be to use the supremum norm over the domain; this is indeed the usual method. However, when looking for good estimates, mainly with respect to the number n of degrees of freedom, this turns out to be a bad choice. The reason is that at some step of the perturbation procedure one necessarily has to consider an expansion of a function either in powers of the coordinates or in Fourier components, and to use general

theorems in order to get bounds on the coefficients of that expansion; then, after some operation on the coefficients one comes back to the supremum norm by summing up all the contributions. This, of course, introduces a very bad dependence on n . The only way out is to change the norm, still maintaining some compatibility properties with the supremum norm, like ii. above.

Let's come now to the examples. In the case of the elliptic equilibrium, the natural (and classical) choice is to identify \mathcal{P}_s with the linear space of homogeneous polynomials of degree $s + 2$ in the canonical variables, so that the Hamiltonian (3) is characterized by $H_s \in \mathcal{P}_s$. Choosing M_0 as the origin, the domain is naturally defined as a polydisk. Writing a polynomial $f \in \mathcal{P}_s$ as $f = \sum_{j,k} f_{jk} x^j y^k$, with $f_{jk} \in \mathbf{C}$, a suitable norm is then $\|f\|_R = \sum_{j,k} |f_{jk}| \varrho^j \sigma^k$. Then all the properties above are satisfied with $a_j = b_j = 1$ for $1 \leq j \leq n$ and with $C = 1$.

The model of constraints is more complex. Here, one must take into account the fact that the variables (π, ξ) play an active role in determining the perturbation order, since they are confined in a neighbourhood of the origin of size $\varepsilon = \lambda^{-1/2}$, while the variables (p, x) should be considered essentially as parameters. The algebraic framework is then built up by defining \mathcal{P}_s as the space of homogeneous polynomials of degree $s + 2$ in π, ξ, ε , whose coefficients are analytic functions of (p, x) . The domain has now the form $\mathcal{D}_R = \mathcal{G}_{R'} \times \Delta_{R''}$, where $\mathcal{G}_{R'}$ is the domain of the variables (p, x) and $\Delta_{R''}$ a polydisk centered on the origin for the variables (π, ξ) (here, I set $R = (R', R'')$, with analogous settings for ϱ and σ). Writing then $f \in \mathcal{P}_s$ as

$$f = \sum_{l+|j|+|k|=s+2} \varepsilon^l f_{jk}^{(l)}(p, x) \pi^j \xi^k ,$$

the norm is defined as

$$\|f\|_R = \sum_{l,j,k} \varepsilon^l \left| f_{jk}^{(l)} \right|_{R'} \varrho'^j \sigma''^k , \quad \left| f_{jk}^{(l)} \right|_{R'} = \sup_{(p,x) \in \mathcal{G}'_R} \left| f_{jk}^{(l)}(p, x) \right| .$$

Then, all the properties above are satisfied with $a_j = b_j = 0$ for the variables p, x , with $a_j = b_j = 1$ for the variables π, ξ , and with $C = 2$.

3.2. Near to identity canonical transformations

The next step consists in defining a near to identity canonical transformation via a recursive algebraic algorithm.

Considering a *generating sequence* $\chi = \{\chi_s\}_{s \geq 1}$ of functions $\chi_s \in \mathcal{P}^{[s]}$, one defines the operator

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

where

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

Such an operator turns out to be linear, invertible, and to preserve products and Poisson brackets, i.e. $T_\chi(f \cdot g) = (T_\chi f)(T_\chi g)$ and $T_\chi\{f, g\} = \{T_\chi f, T_\chi g\}$. Moreover, any near to identity canonical transformation $(p, q) = \mathcal{C}(p', q')$ can be given the form

$$p_l = T_\chi p'_l , \quad q_l = T_\chi q'_l , \quad 1 \leq l \leq n , \quad (9)$$

with a suitable generating sequence χ , and explicit recursive formulae can be found for the inverse operator T_χ^{-1} and for the generating sequence, φ say, of the composition of two transformations, T_χ and T_ψ say, so that $T_\varphi = T_\chi \circ T_\psi$.

Coming now to a rigorous viewpoint, one proves that, *under the hypothesis that there exist real constants $\beta \geq 0$ and $\Phi > 0$ such that $N_{R,R}(\chi_s) \leq \frac{\beta^{s-1}}{s}\Phi$, for any positive $d < 1/2$, and with the condition*

$$\left(\frac{2Ce^2\Phi}{d^2} + \beta \right) \leq 1$$

the canonical transformation defined by T_χ is analytic, and one has

$$\mathcal{D}_{(1-2d)R} \subset T_\chi(\mathcal{D}_{(1-d)R}) \subset \mathcal{D}_R .$$

The condition above on Φ , β and d turns out, in the examples discussed above, to be a condition on the size of both the parameter ε and the size R of the domain.

The present approach is more effective when compared with the classical one involving generating functions in mixed variables. Let me add two remarks to illustrate this fact.

A first remark is that the transformation is given by an explicit recursive expression: no inversion is needed, like in the classical method, and the definition can be easily translated into an explicit recursive algorithm. The same holds for the inverse transformation.

A second remark is concerned with the fact that, given a function $f(p, q)$, the canonical transformation (9) transforms it to $f'(p', q') = (f \circ T_\chi)(p', q')$: a standard result in the Lie series theory, namely the exchange theorem (see ref. [9]), states that $f'(p', q')$ coincides with $T_\chi f$, i.e. that $f \circ T_\chi = T_\chi f$. The actual consequence is that no substitution is needed in order to determine the transformed function f' , since it is given by an explicit algorithm. This is, according to Gröbner, the most fascinating aspect in the Lie transform theory.

Let me also stress that the present algebraic approach has some advantages with respect to the usual approach to Lie transforms, which makes essential use of the flow generated by a nonautonomous canonical system. Indeed, in the latter method one must choose a single perturbation parameter, to be identified with the time variable of the canonical flow, thus making not so natural the extension of the method to the case of many perturbation parameters. The usual approach could, of course, be recovered, but let me stress that only the algebraic algorithm is used, in fact, in practical applications, and that all the relevant properties can be proven by purely algebraic methods. The reference to a canonical flow looks like a psychological support more than like an essential mathematical tool.

3.3. Normal form of the Hamiltonian

The development of the theory now follows more closely the usual scheme. Let me make reference, for definiteness, to the Hamiltonian of the model of constraints. With the algebraic framework described in sect. 3.1, the Hamiltonian can be expanded in the form

$$H = H_0 + \sum_{s \geq 1} H_s ,$$

$$H_0 = h_\Omega(\pi, \xi) + \varepsilon^2 \hat{h}(p, x) , \quad H_s \in \mathcal{P}_{2s+2} ,$$

with $h_\Omega(\pi, \xi)$ and $\hat{h}(p, x)$ as in (6); one looks then for a truncated generating sequence $\chi^{(r)} = \{\chi_s\}_{s=1}^r$ which gives the Hamiltonian the form

$$H^{(r)} = Z_0 + \sum_{s=1}^r Z_s + \mathcal{R}^{(r)} ,$$

where $Z_0 = H_0$, Z_s is in normal form and $\mathcal{R}^{(r)} \in \mathcal{P}^{[2r+2]}$ is a nonnormalized remainder. By normal form it is simply meant here that $L_{h_\Omega} Z_s = 0$. This is a classical and well known problem, so let me stress only the relevant differences with the usual procedure. These differences are mainly concerned with two problems: first, the Poisson bracket between homogeneous functions necessarily produces a nonhomogeneous result, and, second, the unperturbed Hamiltonian H_0 is not integrable.

Trying to separate the homogeneous components is impractical, so, let's proceed as follows. By substituting the explicit definition of T_χ and the expansion above of the Hamiltonian in the expression of the normal form, and acting exactly as in the usual case of homogeneous expansions (for example the case of the elliptic equilibrium, as discussed in ref. [12]), one gets the infinite system of equations $L_{H_0} \chi_s + Z_s = \Psi_s$, with known Ψ_s , to be recursively solved for $s \geq 1$ with respect to χ_s and Z_s with the condition $L_{H_0} Z_s = 0$. Such a system can be consistently used in the nonhomogeneous case too, since, as is easily checked, the equation at order s involves only terms belonging to the space $\mathcal{P}^{[2s+2]}$, so that the minimum order of each term increases with s . At this point the second problem arises. Indeed, due to the lack of knowledge about \hat{h} , it is in general impossible to solve the equation above. The key point is instead to notice that $L_{H_0} \chi_s = L_{h_\Omega} \chi_s + L_{\varepsilon^2 \hat{h}} \chi_s$, and that, if $\chi_s \in \mathcal{P}^{[2s+2]}$, then one has $L_{h_\Omega} \chi_s \in \mathcal{P}^{[2s+2]}$, while $L_{\varepsilon^2 \hat{h}} \chi_s \in \mathcal{P}^{[2s+4]}$. Such an elementary remark allows to shift $L_{\varepsilon^2 \hat{h}} \chi_s$ to the next order, where it becomes a known term, so that the equation above takes the simpler form

$$L_{h_\Omega} \chi_s + Z_s = \Psi_s, \quad (10)$$

with known Ψ_s , and the (p, x) variables play essentially the role of parameters. The solution proceeds then, at least formally, as usual: the vector Ω determines a resonance module $\mathcal{M}_\Omega \in \mathbf{Z}^\nu$ defined as $\mathcal{M}_\Omega = \{k \in \mathbf{Z}^\nu : k \cdot \Omega = 0\}$; this module in turn determines a splitting $\Psi_s = \bar{\Psi}_s + \tilde{\Psi}_s$, where $\bar{\Psi}_s$ collects the resonant terms of Ψ_s and $\tilde{\Psi}_s$ the remaining ones; finally one puts $Z_s = \bar{\Psi}_s$ and uses $\tilde{\Psi}_s$ in order to determine χ_s .

From a rigorous viewpoint, this is not enough if one looks for good estimates. Indeed, as is well known, in solving the equation (10) above there appear small denominators of the form $k \cdot \Omega$ with $k \in \mathbf{Z}^\nu \setminus \mathcal{M}_\Omega$. A key point in removing, as far as possible, the dependence of the results on the number n of degrees of freedom, is that at any order one has to deal only with a finite number of small denominators. This is obtained by carefully controlling how the nonhomogeneity of the expansions propagate through the successive application of Poisson brackets. Indeed, one can check that at any order s of the perturbative procedure one has $\Psi_s \in \mathcal{P}^{[2s+2, 4s]}$, and so also Z_s and χ_s belong to the same space. This furnishes a bound on the maximum order of each term, and so ensures that only a finite number of small denominators do appear, in this case those with $0 < |k| \leq 4s$. So, in order to control the action of the small denominators, it is enough to determine a nonincreasing sequence $\{\alpha_s\}_{s \geq 1}$ of positive constants such that $|k \cdot \Omega| \geq \alpha_s$ for $0 < |k| \leq 4s$ and $k \notin \mathcal{M}_\Omega$.

Giving rigorous estimates for the generating sequence and the normal form is now a technical matter: the formal normalization algorithm is translated into a set of recursive estimates. The result is that, *assuming that there exist real positive constants E_0 , σ and E such that $\|\hat{h}\|_R \leq E_0$ and $\|H_s\|_R \leq \sigma^{s-1} E$, one proves that for any positive $d < 1$ one has $\|\chi_s\|_{(1-d)R} \leq \frac{\beta^{s-1}}{s} \Phi$ with $\beta = C_0 r / \alpha_r$ and $\Phi = 4E / \alpha_r$; the constant C_0 depends on E_0 , E , σ and R . Thus, the generating sequence satisfies the conditions for the convergence of the canonical transformation stated in sect. 3.2, provided the constants β and Φ are small enough. This can be satisfied if R' is chosen to be of order ε , namely if the initial condition are such that the energy of the subsystem h_Ω is small. Setting now $d = 1/4$, and recalling that $\varepsilon^2 = \lambda^{-1}$ for the original Hamiltonian (5) of the model of constraints, one gets the*

final result that *under the hypotheses above there exist real positive constants λ_* , C_1 , C_2 and C_3 depending on E_0 , E , σ and R such that for any $\lambda > \lambda_*$ the normalized Hamiltonian is analytic in the domain $\mathcal{D}_{R/2}$, and one has the estimates*

$$\begin{aligned} |T_\chi h_\Omega - h_\Omega| &< C_1 \lambda^{-2} \\ |T_\chi \hat{h} - \hat{h}| &< C_2 \frac{\lambda^{-2}}{\alpha_r} \\ |\mathcal{R}^{(r)}| &< C_3 \left(\frac{r}{\alpha_r}\right)^r \left(\frac{\lambda_*}{\lambda}\right)^r. \end{aligned} \tag{11}$$

Here, the dependence on the normalization order r has been put into evidence.

The estimates above have been adapted to the case of constraints, but essentially the same results are obtained in the case of the elliptic equilibrium, with the same dependence on the normalization order r . The size of the constants β and Φ in this case depends only on the size of the domain, which is a polydisk centered on the origin. A detailed proof for the case of constraints can be found in ref. [7]; the case of elliptic equilibrium has been investigated in ref. [12].

3.4. Exponential estimates of Nekhoroshev type

Forget now, for a moment, the remainder $\mathcal{R}^{(r)}$, and consider the dynamical evolution of the truncated Hamiltonian $Z = Z_0 + \sum_{s=1}^r Z_s$. Let $I = I_0 + I_1 + \dots$ be a prime integral for Z (in general, given $I_0 = \sum_{l=1}^\nu \mu_l (\pi_l^2 + \xi_l^2)$ with $\mathcal{M}_\Omega \perp \mu \in \mathbf{R}^\nu$, then $I = T_\chi I_0$ is a prime integral; for example, $T_\chi h_\Omega$ is such an integral). Then, if one observes the dynamical evolution of I_0 one sees that its value changes over a time scale of order λ , i.e. a quite short time; however, since I is constant, this change in time is bounded by $|I - I_0|$, evaluated over the domain where the motion takes place. If one can guarantee that this domain is contained in $\mathcal{D}_{R/2}$, then the change of I_0 is bounded for all times, being of order λ^{-1} . Such a change is due to the *deformation* of coordinates induced by the near to identity canonical transformation.

Taking now into account the whole Hamiltonian $H^{(r)}$, one sees that I is an approximate integral whose time derivative is of the same order λ^{-r} of the remainder $\mathcal{R}^{(r)}$. Thus, superimposed to the deformation, there may be a diffusion (the Arnold diffusion) due to the *noise* induced by the remainder. Such a diffusion is slow, since it can become of the same order λ^{-1} of the deformation only after a time interval of order λ^{r-1} , but its possible existence does not allow to draw any conclusion about the motion for infinite times.

The bound above on the noise has an evident shortcoming: it depends on the normalization order r , which is clearly an extraneous element introduced by the perturbation scheme. Thus, one is naturally led to remove such an element by looking for a choice of r , say r_{opt} , which minimizes the size of the remainder.

A general method is the following. Recalling the usual diophantine theory, choose the sequence $\{\alpha_s\}_{s \geq 1}$ as $\alpha_s = \gamma s^{-\tau}$, with suitable constants $\gamma > 0$ and $\tau \geq 0$ (it is known that for $\tau > \nu - 1$ such a condition is satisfied by a set of Ω 's of large measure). Replace now this value in the bound (11) of the remainder, so that

$$|\mathcal{R}^{(r)}| < C' r^{r(\tau+1)} \left(\frac{\lambda_*}{\lambda}\right)^r,$$

with a suitable constant C' , and minimize it by setting

$$r = r_{\text{opt}} = \left[\frac{1}{e} \left(\frac{\lambda}{\lambda_*} \right)^{\frac{1}{\tau+1}} \right] \quad (12)$$

(here, $[\cdot]$ denotes the integer part). Thus, the optimal normalization order r_{opt} is determined as a function of the perturbation parameter λ . By substituting this value in the estimate for the remainder one gets

$$|\mathcal{R}_\lambda| < \mathcal{A}\lambda \exp \left[-\frac{\tau+1}{e} \left(\frac{\lambda}{\lambda_*} \right)^{\frac{1}{\tau+1}} \right], \quad (13)$$

where \mathcal{R}_λ denotes the remainder when the normalization order is chosen according to (12), and \mathcal{A} is a constant.

The conclusion is that *the change in time of a prime integral is of the same order of the deformation up to a time $|t| \leq \min(T, T_0)$, where*

$$T = T_* \lambda^{-1} \exp \left[\frac{\tau+1}{e} \left(\frac{\lambda}{\lambda_*} \right)^{\frac{1}{\tau+1}} \right], \quad (14)$$

and T_0 is the escape time of the orbit from the domain $\mathcal{D}_{R/2}$, where the estimates hold. The time T_0 can usually be determined by independent considerations; for example, it is actually infinite if the motion takes place on a compact surface of constant energy which is contained in $\mathcal{D}_{R/2}$.

4. Application to the model problems

Let me now conclude by coming back to the specific problems illustrated in sect. 2, and quoting some recent results. As said in the introduction, the aim is to show that the application of the previous scheme to these models can substantially improve the original Nekhoroshev's results, at least in some cases.

4.1. Nonlinear chains

Assuming that the frequencies ω are nonresonant, the normalized Hamiltonian admits n prime integrals of the form $\Phi^{(l)} = I_l + \Phi_1^{(l)} + \dots$, which are perturbations of the harmonic actions $I_l = \frac{1}{2} \sum_l \omega_l (x_l^2 + y_l^2)$. If the initial data $I_l(0)$ are small enough, then one has

$$|I_l(t) - I_l(0)| < C I_{\max}^{3/2},$$

with a suitable constant C and $I_{\max} = \max_l(I_l(0))$, up to a time

$$T = T_* \exp \left[\frac{\tau+1}{e} \left(\frac{I_*}{I_{\max}} \right)^{\frac{1}{2(\tau+1)}} \right].$$

Unfortunately, here one has $\tau > n - 1$, which is a bad dependence on n ; however, this result has not been optimized.

Such results can be obtained by the methods described in the present lecture. A substantially simpler, but unfortunately less general, method can be found in ref. [13]. An extension to the case of an infinite system of coupled oscillators, but with a particular class of initial conditions, can be found in ref. [14].

4.2. The Lagrangian point L_4

Considering the spatial, circular restricted three body problem in the Sun–Jupiter case, one can look for a radius R_0 such that an asteroid starting in a neighbourhood of radius R_0 of the Lagrangian point L_4 is confined to a neighbourhood of radius $2R_0$ for a time T of the order of the age of the universe.

Due to the particular simplicity of the problem, one can improve the choice of the optimal normalization order by explicitly computing the constant α_r in (11), and performing optimization by computer. This gives for R_0 a value of a few kilometers, which is not far from being a realistic result. Let me stress, in such a connection, that the use of the computer is only limited to the evaluation of the constants which appear in the theory.

4.3. The diatomic gas

In this model, the only prime integral for the normalized Hamiltonian is h_Ω . Such a result is in a sense optimal, since, due to the complete resonance, essentially no bound can be put on the exchange of energy among the internal vibration of the molecules. So, the best we can do is to prove that the subsystem of the translational and rotational degrees of freedom of the molecules and the subsystem of the internal vibrations evolve in fact independently, each with its own internal, possibly chaotic dynamics, and that a significant transfer of energy between these two subsystems takes an exponentially large time.

The results of sect. 3.4 allow to conclude, in this case, that the energy exchange is of order λ^{-1} up to a time

$$T = T_* \exp\left(\frac{\lambda}{\lambda_*}\right).$$

Note that the exponent a which appears in (2), and is of order $1/n$ in the case of the nonlinear chain, here is exactly 1. This is precisely due to the resonance. Indeed, due to the fact that $\Omega_1 = \dots = \Omega_\nu$, one has that the expression $k \cdot \Omega$ either vanishes, but the corresponding term goes into the normal form, or is a multiple of Ω_1 , so that there are no small denominators at all. The price paid is the fact that no information is available on the internal evolution of the subsystem of the vibrations.

A dependence on n still remains in the constants T_* and λ_* , which turn out to be typically of order $1/n^2$. However one should note that this is too pessimistic, since the perturbation scheme, being global in (a subset of) the phase space, is developed as if all the molecules were simultaneously colliding at any time: a foolish picture. It seems quite reasonable to hope that such a dependence can be removed in the framework of a statistical approach, which for example excludes the initial data leading to multiple collisions. Some more indications can be found in ref. [7].

- [1] Poincaré, H.: *Les méthodes nouvelles de la mécanique céleste*, Gauthier–Villars, Paris (1892).
- [2] J. Moser: *Stabilitätsverhalten kanonischer differentialgleichungssysteme*, Nachr. Akad. Wiss. Göttingen, Math. Phys. K1 IIa, nr.6 (1955), 87–120.
- [3] J. E. Littlewood: *On the equilateral configuration in the restricted problem of three bodies*, Proc. London Math. Soc.(3) **9** (1959), 343–372;
J. E. Littlewood: *The Lagrange configuration in celestial mechanics*, Proc. London Math. Soc.(3) **9** (1959), 525–543.
- [4] Nekhoroshev N. N.: *Exponential estimate of the stability time of near–integrable Hamiltonian systems*, Russ. Math. Surveys **32** N.6, 1–65 (1977).
Nekhoroshev N. N.: *Exponential estimate of the stability time of near–integrable Hamiltonian systems, II*, Trudy Sem. Petrovs. N.5, 5–50 (1979).

- Benettin, G., Galgani, L. and Giorgilli, A.: *A proof of Nekhoroshev's theorem for the stability times in nearly integrable Hamiltonian systems*, Cel. Mech **37**, 1–25 (1985).
- Benettin, G. and Gallavotti, G.: *Stability of motion near resonances in quasi integrable Hamiltonian systems*, J. Stat. Phys. **44**, 293–338 (1986).
- Gallavotti, G.: *Quasi-integrable mechanical systems*, in K. Osterwalder and R. Stora (eds.), *Les Houches, Session XLIII* (1984).
- [5] Zaslavski, H. M.: *Stochasticity of dynamical systems*, Nauka, Moscow (1984).
- Chirikov, B. V.: Phys.Rep. **52**, 263 (1979).
- [6] Benettin, G., Galgani, L. and Giorgilli, A.: *Realization of holonomic constraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory, part I*, Comm. Math. Phys., **113**, 87–103 (1987).
- [7] Benettin, G., Galgani, L. and Giorgilli, A.: *Realization of holonomic constraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory, part II*, Comm. Math. Phys, to appear.
- [8] Fermi, E., Pasta, J., and Ulam, S.: *Los Alamos Report No. LA-1940* (1955), later published in *Fermi, E.: Collected Papers* (Chicago, 1965), and *Lect. Appl. Math.* **15**, 143 (1974).
- [9] Gröbner, W.: *Die Lie-Reihen und Ihre Anwendungen*, VEB Deutscher Verlag der Wissenschaften (1967).
- [10] Hori, G.: *Theory of general perturbation with unspecified canonical variables*, Publ. Astron. Soc. of Japan **23** 567–587 (1966).
- Deprit, A.: *Canonical transformations depending on a small parameter*, Cel. Mech. **1**, 12–30 (1969).
- [11] Colombo, R. and Giorgilli, A.: *A rigorous algebraic approach to Lie transforms in Hamiltonian mechanics*, preprint (1988).
- [12] Giorgilli, A., Delshams, A., Fontich, E., Galgani, L. and Simó, C.: *Effective stability for a Hamiltonian system near an elliptic equilibrium point, with an application to the restricted three body problem*, J. Diff. Eqs., **77** N.1, 167–198 (1989).
- [13] Giorgilli, A.: *Rigorous results on the power expansions for the integrals of a Hamiltonian system near an elliptic equilibrium point*, Ann. I.H.P, **48** N.4, 423–439 (1989).
- [14] Benettin, G., Fröhlich, J. and Giorgilli, A.: *A Nekhoroshev-type theorem for Hamiltonian systems with infinitely many degrees of freedom*, Comm. Math. Phys., to appear.

This article was processed by the author using the T_EX Macropackage from Springer-Verlag.