Journal of Statistical Physics, Vol. 101, Nos. 3/4, 2000

Estimates for Non-Resonant Normal Forms in Hamiltonian Perturbation Theory

Enrico Valdinoci[1]

Received April 13, 1999; final June 28, 2000

We make a remark about an estimate of the rest for the non-resonant action-angle normal forms and exhibit a simple example suggesting the optimality of this estimate when there are no small divisors. Given a polynomial perturbation of degree P and an integer k, calling γ the size of the small denominators up to order k, we prove that the kth order remainder is bounded by $(2/\varepsilon_0)^{k+1}$ with $\varepsilon_0 = \text{const}\, \gamma^2/(kP^2)$. Thus, fixing the degree of the perturbation, if γ is independent of k (i.e., if there are no small divisors), we obtain a rest bounded by $(\text{const}\, k)^{k+1}$. These estimates are also applied to the case in which the small divisors are absent, and they are conjectured to be optimal in this context. To support this idea we present a simplified model problem with no small denominators, formally related to the above calculations, and we show that it indeed has factorial divergence of its Birkhoff series. We also obtain Nekhoroshev's Theorem for harmonic oscillators. We hope that our simple approach makes more accessible to a general audience this important (although quite technical) topic.

KEY WORDS: Perturbation theory; stability of Hamiltonian systems; Birkhoff series; method of majorants; harmonic oscillators.

1. INTRODUCTION AND SETUP

We study the perturbation theory of an integrable system expressed in action-angle variables

$$H(I, \varphi) = h(I) + \varepsilon f(I, \varphi) \tag{1}$$

[1] Department of Mathematics, The University of Texas at Austin, Austin, Texas 78712-1082; e-mail: enrico@math.utexas.edu

0022-4715/00/1100-0905$18.00/0

whose equations of motion are $\dot{I} = -\varepsilon \partial_\varphi f$ and $\dot{\varphi} = \partial_I H$. We consider ε, $|\varepsilon| \leqslant \varepsilon_0$ as a small parameter. As usual, we denote by $\mathbb{T}^N$ the N-dimensional torus, and we define

$$B_r^N(I') \equiv \{I \in \mathbb{C}^N \text{ s.t. } |I - I'| \leqslant r\}, \qquad \text{and}$$
$$\mathbb{T}_\zeta^N \equiv \{\varphi \in \mathbb{C} \text{ s.t. } \Re\varphi \in \mathbb{T} \text{ and } \max_{1 \leqslant i \leqslant N} |\Im\varphi| \leqslant \zeta\}$$

We require h and f to be real analytic for I in a neighborhood of $I_0 \in \mathbb{R}^N$ (say $B_{\rho^*}^N(I_0)$, with $0 < \rho^* < 1$) and f to be a trigonometric polynomial in φ of degree P. For convenience we fix $0 < \xi < 1$ and consider $\mathbb{T}_{2\xi}^N$ as the domain in the φ's.

We now fix an integer $k \geqslant 2$. We call

$$\gamma = \gamma(k) \equiv \min\{1, \rho^* \xi / \kappa_0, \min_{\substack{0 < |n_1| + \cdots + |n_N| \leqslant kP \\ I \in B_{\rho^*}^N(I_0)}} |h'(I) \cdot n|\}$$

the size of the "small denominators," where κ_0 is a suitable number (say $\kappa_0 = 32$). We assume that the system (1) is *non-resonant* up the kPth order, i.e., $\gamma > 0$.

Following a standard procedure in Hamiltonian perturbation theory we find a change of variables that puts the system into one which depends only on the actions until the kth order in ε. More precisely, we recall the following classic theorem, due to Birkhoff:[2]

Theorem (Birkhoff). Consider the system (1) under the above assumptions. Then, if ε_0 is small enough,[3] there exists a real analytic canonical transformation $(I, \varphi) = \Phi_k(y, x)$, ε-close to the identity, and suitable[4] $\rho, \rho^\#, \xi^*, \xi^\#$, such that

$$B_{\rho^\#}^N(I_0) \times \mathbb{T}_{\xi^\#}^N \subseteq \Phi_k(B_\rho^N(y_0) \times \mathbb{T}_{\xi^*}^N) \subseteq B_{\rho^*}^N(I_0) \times \mathbb{T}_\xi^N \tag{2}$$

and

$$\mathscr{H}(y, x) \equiv H \circ \Phi_k(y, x) = h(y) + \sum_{j=1}^k \varepsilon^j g_j(y) + \varepsilon^{k+1} R_k(y, x; \varepsilon) \tag{3}$$

for some real analytic functions g_j.

[2] See, for example, refs. 5 and 1. The method of "averaging" and dividing the motion into a slow evolution and rapid oscillations was also used by Gauss, Lagrange and Laplace in their studies of secular planetary motions.

[3] See (7) and (8) for the choice of ε_0.

[4] We choose the superscript of ρ and ξ in this (apparently asymmetric) way for future convenience, since the variable of the functions χ_j in (4) will be y and φ.

The representation (3) is often called "normal form," the new variables (y, x) are called "normal coordinates," and R_k is called the "rest." The study of the existence of transformations reducing a dynamical system to a simple form was already in ref. 24 and has been investigated by many authors. For a detailed survey on several normal forms in Hamiltonian systems, see ref. 10. See also ref. 8 where a theory of normal forms for differential equations is considered and Section 2 of ref. 9 for some applications to the Hamiltonian case.

Unless differently stated, we will adopt the name of "constants" to denote quantities which do not depend on ε, k and P, but may depend on N, ρ^*, ξ, h and f. We will reserve the symbol κ_i to denote suitable "pure numbers" (like 2, 4, etc.).

The statement of Birkhoff's Theorem can be made more explicit, saying that g_1 is the average of f in the angles and that the transformation Φ_k is obtained by a generating function of the type $\mathscr{G}(y, \varphi) = y \cdot \varphi + \sum_{j=1}^{k} \varepsilon^j \chi_j(y, \varphi)$, corresponding to the transformation

$$I = y + \sum_{j=1}^{k} \varepsilon^j \partial_\varphi \chi_j, \qquad x = \varphi + \sum_{j=1}^{k} \varepsilon^j \partial_y \chi_j \tag{4}$$

The χ_j's can be computed recursively as follows. Denote the Fourier coefficients by "$\hat{\cdot}$,"

$$\begin{aligned} \chi_1(y, \varphi) &= \sum_{\substack{n \in \mathbb{Z}^N \\ 0 < |n_1| + \cdots + |n_N| \leqslant P}} \frac{i\hat{f}_n(y)}{h'(y) \cdot n} e^{in \cdot \varphi}, \\ \chi_j(y, \varphi) &= \sum_{\substack{n \in \mathbb{Z}^N \\ 0 < |n_1| + \cdots + |n_N| \leqslant jP}} \frac{i\hat{\Gamma}_n^{(j)}(y)}{h'(y) \cdot n} e^{in \cdot \varphi} \end{aligned} \tag{5}$$

for all $2 \leqslant j \leqslant k$, where, denoting by $[\,\cdot\,]_j$ the jth order in power series of ε,

$$\Gamma^{(j)} \equiv \left[h\left(y + \sum_{m=1}^{j-1} \varepsilon^m \partial_\varphi \chi_m \right) \right]_j + \left[f\left(y + \sum_{m=1}^{j-1} \varepsilon^m \partial_\varphi \chi_m, \varphi \right) \right]_{j-1} \tag{6}$$

The function g_j, for $2 \leqslant j \leqslant k$, is the average of $\Gamma^{(j)}$ in the angles. Sending k to ∞, we obtain a formal series $\sum_{j \geqslant 1} \varepsilon^j \chi_j$, called *Birkhoff series*.[5]

[5] Other approaches to this problem are also possible. See for instance ref. 2, in which Birkhoff series of generating functions are replaced by a convenient iteration of Lie transforms. Indeed, the bound on ε_0 in formula (16) of our note can be seen as a slight improvement, in the non-resonance case with polynomial perturbation of a fixed degree, of the one given in formula (24) of ref. 2. Indeed, the latter requires ε_0 to be smaller than γ^2/k^{N+2}, while our formula (16) allows ε_0 to be of the size of γ^2/k.

It is well known that Birkhoff series are in general divergent. For example one can show that the two degrees of freedom Hamiltonian

$$\varpi I_1 + I_2 + \varepsilon \left(I_1 + \sum_{n \in \mathbb{Z}^2} e^{-|n_1| - |n_2|} \cos n \cdot \varphi \right)$$

has divergent Birkhoff series, where ϖ is the golden number (see ref. 14).

In Proposition 2.5 we will show that this divergence is, in the worst case, *factorial*, i.e., the size of χ_j will be controlled by $P^{2(j-1)}(j-1)!/\gamma^{2j-1}$.

We now discuss the condition over ε_0 to be required in Birkhoff's Theorem. First, we must assume that

ε_0 is small enough that the transformation Φ_k, implicitly defined by (4), is a diffeomorphism for (y, x) in a suitable domain $B^N_\rho(y_0) \times \mathbb{T}^N_{\chi^*}$, satisfying (2) (7)

Moreover, we require that

$$\sum_{j=1}^{k} \varepsilon_0^j \sup_{B^N_\rho(y_0) \times \mathbb{T}^N_\xi} |\partial_\varphi \chi_j| \leqslant \rho^*/4 \tag{8}$$

This condition ensures that relations such as (6) are well defined, and that $B^N_{2\rho}(y_0) \subseteq B^N_{\rho^*/2}(I_0)$, provided that $\rho \leqslant \rho^*/8$ (we will use it in Proposition 2.4; Proposition 2.6 will allow us to choose $\rho = \rho^*/8$).

The general estimate stated in our paper is Theorem 2.7, in which we prove that

Theorem. Set $\varepsilon_0 \equiv c^* \gamma^2/(kP^2)$, where c^* is a suitable constant. Then, the following bound on the rest R_k of the Birkhoff normal form (3) holds:

$$\sup_{B^N_\rho(y_0) \times \mathbb{T}^N_{\xi^*},\ |\varepsilon| \leqslant \varepsilon_0/2} |R_k| \leqslant \sup_{B^N_{\rho^*}(I_0) \times \mathbb{T}^N_\xi} |H| \cdot \left(\frac{2}{\varepsilon_0} \right)^{k+1}$$

Moreover the following bound on the size of the jth order in ε of the new Hamiltonian holds:

$$\sup_{B^N_\rho(y_0) \times \mathbb{T}^N_{\xi^*}} |g_j| \leqslant \sup_{B^N_{\rho^*}(I_0) \times \mathbb{T}^N_\xi} |H| \cdot \left(\frac{1}{\varepsilon_0} \right)^j$$

Similar estimates are already present in the literature. For instance, one can compare our Proposition 2.5 with Section 6 of ref. 15 and with the Main

Proposition of ref. 3. In the case of harmonic oscillators with (σ, N)-Diophantine frequency[6] with polynomial perturbation, our Proposition 2.5 is indeed a slight improvement of formula (6.6) of ref. 15 and formula (6.11) of ref. 3. Namely, for polynomial perturbation with a fixed degree, the above mentioned papers give a bound on $\partial_\varphi \chi_k$ of the type $(\text{const})^k\, k!\, k^{2Nk+k}$, while Proposition 2.5 here leads indeed to the bound $(\text{const})^k\, k!\, k^{2Nk-N}$.

At variance with the proofs mentioned above, we will make use only of the traditional method of majorants (see the following Definition 2.1 and, for example, refs. 18 and 26). In spite of its very classical flavor, this method (as far as we know) has not been extensively used in this context. In our paper, the use of the method of majorants will provide a very short proof, in which only elementary computations are involved.

Also, the way the above result is stated and proved here, directly allows a comparison with the case of no small divisors (which is of independent interest: see ref. 22). Indeed, we believe that the estimates above are probably not too far from being sharp even when the small divisors are absent, as the simplified example in Section 3 will show. We consider the Hamilton–Jacobi equation for the perturbation of a simple one-dimensional harmonic oscillator $\omega I + 2(1-I)\,\varepsilon \cos\varphi$. Since the effect of the operator $\omega\partial_\varphi$ is just multiplying for a small denominator, a reasonable example with no small denominators could be obtained substituting 1 instead of $\omega\partial_\varphi$. We easily prove that this simplified model shows a factorial growth of Birkhoff series. The same procedure of considering easier problems with no small denominators was used, in a different setting, in ref. 6.

As remarked in refs. 15, 16, 3, and 25 the estimates developed in our paper are enough to prove *Nekhoroshev's Theorem in the case of harmonic oscillators*, obtaining exponentially long times of stability for systems of the type $H(I, \varphi) = \omega \cdot I + \varepsilon f(I, \varphi)$, where ω is a Diophantine vector and f is an analytic periodic perturbation. For comments about the relevance of the case of the harmonic oscillators, see page 294 of ref. 3.

With more refined techniques it is possible to extend the validity of these estimates even to the resonant case, providing the so called "analytic part" of Nekhoroshev's Theorem, which was firstly proved, under suitable conditions of "steepness," in ref. 23. See for example: ref. 25 with an iterative finite KAM algorithm; refs. 2 and 17; refs. 20 and 21 which consider perturbation theory around periodic orbits. We believe that the estimates developed in our paper, although not new in the literature, are easier than the ones usually developed in the proofs of Nekhoroshev's Theorem, but they remain essentially optimal in the case of no small

[6] See page 13 for the definition of Diophantine vector.

denominators. Unfortunately, it is not easy to extend our method near the resonances, since in this case differentiation with respect to the actions appears in the definition of the transformation, making it difficult to invent a good majorant problem.

We also stress the fact that the problem of the optimality of the exponents in Nekhoroshev's Theorem (as stated in refs. 25 and 20) is related but not equivalent to the one discussed here in Section 3, since the latter refers to the case of no small divisors. For a more detailed discussion about the optimality of refs. 25 and 20 and a conjecture of Chirikov, see refs. 11 and 4.

We remark that the problem of divergent Birkhoff series with no small denominators is a classic topic in literature: see, for example, refs. 22 and 12. Also perturbation theory around periodic orbits (as the one used in ref. 20 for Nekhoroshev's Theorem) leads to series with no small divisors, since the denominators are in this case controlled by the inverse of the period. Estimates similar to the one in our paper arise in computing center manifolds too, and in this context they turn out to be sharp. We refer to refs. 27, 28, 19, 13 and references therein for a more complete discussion of center manifolds.

Also the study of the properties of divergent series is a classic topic in the theory of "resurgent analysis" and "resummation:" see, for example, refs. 7 and 28. The sketch of our prove is the following: instead of solving the recurrence (5)–(6), we will consider in Lemma 2.3 a "majorant problem," i.e., a problem whose solution "dominates" the solution of (5)–(6). In detail: we will consider the functional equation (11), showing that it admits a solution $\Psi = \sum \varepsilon^j \Psi_j$, and we will prove that the Fourier coefficients of the Ψ_j's control, up to a factorial, the ones of the χ_j's defined by (5)–(6).

2. ESTIMATE ON THE REST

Definition 2.1. Consider the formal Taylor–Fourier expansions of two functions f and g:

$$f(z, \varphi) = \sum_{n \in \mathbb{Z}^N,\, m \in \mathbb{N}^M} f_{n,m} z^m e^{in \cdot \varphi} \qquad \text{and}$$

$$g(z, \varphi) = \sum_{n \in \mathbb{Z}^N,\, m \in \mathbb{N}^M} g_{n,m} z^m e^{in \cdot \varphi}$$

where z is an M-dimensional vector and φ is an N-dimensional angle. We say that g is a *majorant* of f (and we briefly write $f \prec g$) if $|f_{n,m}| \leqslant g_{n,m}$.

We denote

$$\bar{f}(z, \varphi) \equiv \sum_{n \in \mathbb{Z}^N, m \in \mathbb{N}^M} |f_{n,m}| \, z^m e^{in\cdot\varphi}.$$

We will use afterwards the following easy results, the proof of which is elementary:

Lemma 2.2. For all $n, m \in \mathbb{N}$, $m \geqslant 2$,

$$\max\{j_1! \cdot \cdots \cdot j_p!,\, j_1,..., j_p \in \mathbb{N},\, j_1 + \cdots + j_p = n\} = n! \tag{9}$$

$$\begin{aligned} &\max\{j_1! \cdot \cdots \cdot j_p!,\, j_1,..., j_p \in \mathbb{N}, \\ &\qquad 1 \leqslant j_1,..., j_p \leqslant m-1,\, j_1 + \cdots + j_p = m\} = (m-1)! \end{aligned} \tag{10}$$

The following lemma formulates the majorant problem. We denote, for all $x \in \mathbb{C}$, $\vec{x} \equiv (\underbrace{x,..., x}_{N \text{ times}})$. We call η_0 an auxiliary parameter, and consider by now $|\varepsilon| \leqslant \eta_0$:

Lemma 2.3. Take $y \in B^N_{\rho^*/2}(I_0)$. Define,[7] for all $z \in B^N_{\rho^*/2}(0)$,

$$Q_y(z) \equiv h(y+z) - h(y) - h'(y) \cdot z,$$

$$F_y(z, \varphi) \equiv f(y+z, \varphi)$$

Set $\mathscr{E} \equiv \{\varepsilon \in \mathbb{C}, |\varepsilon| \leqslant \eta_0\}$. Fix $R \leqslant \rho^*/(2\sqrt{N})$ such that

$$\sup_{|y-I_0| \leqslant \rho^*/2,\, |z|_\infty \leqslant R} |\partial_z \bar{Q}_y| \leqslant \gamma/(4\sqrt{N})$$

Then, there exists a unique $\Psi = \Psi_y(\varphi, \varepsilon)$ real analytic for $(y, \varphi, \varepsilon) \in B^N_{\rho^*/2}(I_0) \times \mathbb{T}^N_{2\xi} \times \mathscr{E}$, such that $\Psi_y(\varphi, 0) = 0$, $\sup_{B^N_{\rho^*/2}(I_0) \times \mathbb{T}^N_{2\xi} \times \mathscr{E}} |\Psi| \leqslant R$, and

$$\gamma \Psi - \bar{Q}_y(\vec{\Psi}) - \varepsilon \bar{F}_y(\vec{\Psi}, \varphi) = 0 \tag{11}$$

provided that $\varepsilon \in \mathscr{E}$ and

$$\eta_0 \leqslant \min\left\{\frac{\gamma R}{2 \sup_{y \in B^N_{\rho^*/2}(y_0),\, \varphi \in \mathbb{T}^N_{2\xi}} |\bar{F}_y(0, \varphi)|}, \frac{\gamma}{4\sqrt{N} \sup_{y \in B^N_{\rho^*/2}(I_0),\, |z|_\infty \leqslant R,\, \varphi \in \mathbb{T}^N_{2\xi}} |\partial_z \bar{F}_y|}\right\} \tag{12}$$

[7] Beware that the subscript "$\cdot y$", in this paper, does *not* mean derivative with respect to y.

Proof. Denote $\|\cdot\| \equiv \sup_{B^N_{\rho^*/2}(I_0) \times \mathbb{T}^N_{2\xi} \times \mathscr{E}} |\cdot|$, $\mathscr{B} \equiv \{u \in C^\omega(B^N_{\rho^*/2}(I_0) \times \mathbb{T}^N_{2\xi} \times \mathscr{E})$ s.t. $\|u\| \leqslant R\}$ and, $\forall u \in \mathscr{B}$,

$$\mathscr{F}[u](y, \varphi, \varepsilon) \equiv \gamma^{-1} \bar{Q}_y(\vec{u}(y, \varphi, \varepsilon)) + \gamma^{-1} \varepsilon \bar{F}_y(\vec{u}(y, \varphi, \varepsilon), \varphi)$$

one sees that $\mathscr{F}$ is a contraction over $\mathscr{B}$ endowed with the norm $\|\cdot\|$. Notice that the condition $R \leqslant \rho^*/(2\sqrt{N})$ assures that $\bar{Q}_y(\vec{u})$ and $\bar{F}_y(\vec{u}, \varphi)$ are well defined. ▮

Remark. It is easy to check that the hypotheses of the lemma above are fulfilled if we choose $R \equiv c_1 \gamma$ and $\eta_0 \equiv c_2 \gamma^2$, for suitable constants c_1 and c_2.

We now consider the ε-expansion of the function Ψ built in the previous lemma: $\Psi_y(\varphi, \varepsilon) = \sum_{j \geqslant 1} \varepsilon^j \psi_j(y, \varphi)$.

Note that, by definition, $\forall r \geqslant 2$,

$$\Gamma^{(r)}(y, \varphi) = \left[Q_y \left(\sum_{j=1}^{r-1} \varepsilon^j \partial_\varphi \chi_j \right) \right]_r + \left[F_y \left(\sum_{j=1}^{r-1} \varepsilon^j \partial_\varphi \chi_j, \varphi \right) \right]_{r-1} \tag{13}$$

We remark that in the previous lemma (as well as in the next proposition) it is convenient to think y as a parameter, and we do majorant theory in z and φ only. So if

$$Q_y(z) = \sum_{p \in \mathbb{N}^N} Q_p(y)\, z^p, \qquad F_y(z, \varphi) = \sum_{p \in \mathbb{N}^N,\, n \in \mathbb{Z}^N} F_{n,p}(y)\, z^p e^{in \cdot \varphi}$$

we have

$$\bar{Q}_y(z) = \sum_{p \in \mathbb{N}^N} |Q_p(y)|\, z^p, \qquad \bar{F}_y(z, \varphi) = \sum_{p \in \mathbb{N}^N,\, n \in \mathbb{Z}^N} |F_{n,p}(y)|\, z^p e^{in \cdot \varphi}$$

In the same way, fixing y in the next proposition, we do majorant theory for χ and Ψ only with respect to the variable φ.

The following Proposition 2.4 will show that Ψ_j is a majorant, up to a factorial, of χ_j. From this, Proposition 2.5 will deduce the claimed bound on the coefficients of Birkhoff series.

Proposition 2.4. Assume $\rho \leqslant \rho^*/8$. Fix $y \in B^N_{2\rho}(y_0) \subseteq B^N_{\rho^*/2}(I_0)$. Let Ψ be the function defined in the previous lemma. Then, for all $j \in \mathbb{N}$,

$$\chi_j \prec P^{2(j-1)}(j-1)!\, \Psi_j$$

[i.e., $|(\widehat{\chi_j})_n(y)| \leqslant P^{2(j-1)}(j-1)!\, (\widehat{\Psi_j})_n(y)$.]

Proof. Writing the ε-expansion of Ψ in (11) one obtains

$$\gamma \sum_{j\geqslant 1} \varepsilon^j \Psi_j - \bar{Q}_y\left(\sum_{j\geqslant 1} \varepsilon^j \vec{\Psi}_j\right) - \varepsilon \bar{F}_y\left(\sum_{j\geqslant 1} \varepsilon^j \vec{\Psi}_j, \varphi\right) = 0 \tag{14}$$

Taking the first order, we get $\gamma\Psi_1 - \bar{f}(y, \varphi) = 0$; so, by definition of χ_1, $\Psi_1 = \bar{f}/\gamma \succ \chi_1$. This is the inductive basis of our, statement. For the inductive step we suppose to have proven the result for all $1 \leqslant r-1$ and we take the rth order of (14):

$$\Psi_r = \frac{1}{\gamma}\left(\left[\bar{Q}_y\left(\sum_{j=1}^{r-1} \varepsilon^j \vec{\Psi}_j\right)\right]_r + \left[\bar{F}_y\left(\sum_{j=1}^{r-1} \varepsilon^j \vec{\Psi}_j, \varphi\right)\right]_{r-1}\right) \tag{15}$$

So that, by (9), (10), (13), the definition of χ_r and the fact that $\partial_\varphi \chi_j \prec jP\vec{\bar{\chi}}_j$:

$$\begin{aligned}
\chi_r &\prec \frac{1}{\gamma}\,\Gamma^{(r)} \prec \frac{1}{\gamma}\left(\left[\bar{Q}_y\left(\sum_{j=1}^{r-1} \varepsilon^j\, \overline{\frac{\partial \chi_j}{\partial \varphi}}\right)\right]_r + \left[\bar{F}_y\left(\sum_{j=1}^{r-1} \varepsilon^j\, \overline{\frac{\partial \chi_j}{\partial \varphi}}, \varphi\right)\right]_{r-1}\right)\\
&\prec \frac{1}{\gamma}\left(\left[\bar{Q}_y\left(\sum_{j=1}^{r-1} \varepsilon^j Pj\vec{\bar{\chi}}_j\right)\right]_r + \left[\bar{F}_y\left(\sum_{j=1}^{r-1} \varepsilon^j Pj\vec{\bar{\chi}}_j, \varphi\right)\right]_{r-1}\right)\\
&\prec \frac{1}{\gamma}\left(\left[\bar{Q}_y\left(\sum_{j=1}^{r-1} \varepsilon^j P^{2j-1} j!\, \vec{\chi}_j\right)\right]_r + \left[\bar{F}_y\left(\sum_{j=1}^{r-1} \varepsilon^j P^{2j-1} j!\, \vec{\chi}_j, \varphi\right)\right]_{r-1}\right)\\
&\prec \frac{1}{\gamma}\left(P^{2(r-1)} \max_{\substack{1\leqslant j_1,\dots,j_p\leqslant r-1\\ j_1+\cdots+j_p=r\\ 2\leqslant p\leqslant r}} j_1!\cdots j_p! \left[\bar{Q}_y\left(\sum_{j=1}^{r-1} \varepsilon^j \vec{\Psi}_j\right)\right]_r\right.\\
&\qquad \left. + P^{2(r-1)-1} \max_{\substack{1\leqslant j_1,\dots,j_p\leqslant r-1\\ j_1+\cdots+j_p=r-1\\ 1\leqslant p\leqslant r-1}} j_1!\cdots j_p! \left[\bar{F}_y\left(\sum_{j=1}^{r-1} \varepsilon^j \vec{\Psi}_j\right)\right]_{r-1}\right)\\
&\prec \frac{1}{\gamma}\, P^{2(r-1)}(r-1)! \left(\left[\bar{Q}_y\left(\sum_{j=1}^{r-1} \varepsilon^j \vec{\Psi}_j\right)\right]_r + \left[\bar{F}_y\left(\sum_{j=1}^{r-1} \varepsilon^j \vec{\Psi}_j, \varphi\right)\right]_{r-1}\right)\\
&= P^{2(r-1)}(r-1)!\, \Psi_r
\end{aligned}$$

ending the proof. ▮

Lemma 2.3 and Proposition 2.4 imply the following

Proposition 2.5. There exists a constant c such that

$$\sup_{B^N_{2\rho}(y_0)\times\mathbb{T}^N_\xi} |\chi_j| \leqslant c\,\frac{P^{2(j-1)}(j-1)!\,\gamma}{\eta_0^j} \qquad \text{and}$$

$$\sup_{B^N_{2\rho}(y_0)\times\mathbb{T}^N_\xi} |\partial_\varphi \chi_j| \leqslant c\,\frac{P^{2j-1}j!\,\gamma}{\eta_0^j}$$

Remark. Following the notations of ref. 28, pp. 12–14, the previous proposition states that, in the case of no small denominator (i.e., γ independent of k), the Birkhoff series $\sum \varepsilon^j\chi_j$ are in the *Gevrey class* G_1, and the corresponding Borel transform $\sum \varepsilon^j\chi_j/j!$ is defined for $|\varepsilon|$ small enough.

We are now in position to choose ε_0, ρ, $\rho^\#$, ξ^* and $\xi^\#$ in such a way to fulfill the requirements of Birkhoff's Theorem. Proposition 2.5 and an easy contraction argument lead to the following:

Proposition 2.6. Define

$$\varepsilon_0 \equiv \frac{c^*\gamma^2}{kP^2}, \qquad \rho \equiv \frac{\rho^*}{\kappa_1}, \qquad \rho^\# \equiv \frac{\rho^*}{\kappa_2}, \qquad \xi^* \equiv \frac{\xi}{\kappa_3}, \qquad \xi^\# \equiv \frac{\xi}{\kappa_4} \tag{16}$$

where c^* is a suitable constant, and κ_i are suitable numbers. Then conditions (7) and (8) are satisfied.

Proof. By Proposition 2.5 and the Cauchy inequality, there exists a constant $\tilde{c}$ such that

$$\sup_{B^N_{2\rho}(y_0)\in\mathbb{T}^N_\xi} |\partial_\varphi\chi_j| \leqslant \gamma\left(\frac{\tilde{c}jP^2}{\eta_0}\right)^j, \qquad \sup_{B^N_{\rho}(y_0)\times\mathbb{T}^N_\xi} |\partial_y\chi_j| \leqslant \frac{\gamma}{\rho}\left(\frac{\tilde{c}jP^2}{\eta_0}\right)^j,$$

$$\text{and} \qquad \sup_{B^N_{\rho}(y_0)\times\mathbb{T}^N_\xi} |\partial^2_{y\varphi}\chi_j| \leqslant \frac{\gamma}{\rho}\left(\frac{\tilde{c}jP^2}{\eta_0}\right)^j \tag{17}$$

So, choosing c^* to make $\tilde{c}kP^2\varepsilon_0/\eta_0 \leqslant 1/2$,

$$\sum_{j=1}^k \varepsilon_0^j \sup_{B^N_{\rho}(y_0)\times\mathbb{T}^N_\xi} |\partial_\varphi\chi_j| \leqslant \gamma\sum_{j=1}^k\left(\frac{1}{2}\right)^j < \gamma < \frac{\rho^*}{4}$$

proving (8).

We now show that, with the above choice of ε_0, ρ and χ^*, for any $(y, x) \in B^N_\rho(y_0) \times \mathbb{T}^N_{\xi*}$ there exists a unique $(I, \varphi) \in B^N_{\rho*}(I_0) \times \mathbb{T}^N_\xi$ verifying (4). Given $(y, x) \in B^N_\rho(y_0) \times \mathbb{T}^N_{\xi*}$, define $\mathscr{C}(\varphi) \equiv x - \sum_{j=1}^k \varepsilon^j \partial_y \chi_j(y, \varphi)$. By (17), $\mathscr{C}$ is a contraction on $\mathbb{T}^N_\xi$, so there exists a unique $\varphi_* \in \mathbb{T}^N_\xi$ such that $\mathscr{C}(\varphi_*) = \varphi_*$. Setting $I_* = y + \sum_{j=1}^k \varepsilon^j \partial_\varphi \chi_j(y, \varphi_*)$, we have that $I_* \in B_{\rho^*}(I_0)$ by (17). This proves the existence.

For the uniqueness, if $(I, \varphi), (I', \varphi') \in B^N_{\rho*}(I_0) \times \mathbb{T}^N_\xi$ verify:

$$I = y + \sum_{j=1}^k \varepsilon^j \partial_\varphi \chi_j(y, \varphi), \quad I' = y + \sum_{j=1}^k \varepsilon^j \partial_\varphi \chi_j(y, \varphi'), \quad \mathscr{C}(\varphi), \mathscr{C}(\varphi') = \varphi'$$

then it must be $\varphi = \varphi'$ and so $I = I'$.

In the same way, using the contraction $\widetilde{\mathscr{C}}(y) \equiv I - \sum_{j=1}^k \varepsilon^j \partial_\varphi \chi_j(y, \varphi)$ on $B^N_\rho(y_0)$, one sees that $\forall (I, \varphi) \in B^N_{\rho\sharp}(I_0) \times \mathbb{T}^N_{\xi\sharp}$ there exists a unique $(y, x) \in B_\rho(y_0) \times \mathbb{T}^N_{\xi*}$ verifying (4). ∎

The estimates above are summarized in the following

Theorem 2.7. With the choices in (16), the following bound on the rest R_k of the Birkhoff normal form (3) holds:

$$\sup_{y \in B^N_\rho(y_0),\, x \in \mathbb{T}^N_{\xi*},\, |\varepsilon| \leqslant \varepsilon_0/2} |R_k| \leqslant \sup_{B^N_{\rho*}(I_0) \times \mathbb{T}^N_\xi} |H| \cdot \left(\frac{2}{\varepsilon_0}\right)^{k+1} \tag{18}$$

Moreover, the following bond on the size of the jth order in ε of the new Hamiltonian holds:

$$\sup_{y \in B^N_\rho(y_0),\, x \in \mathbb{T}^N_{\xi*}} |g_j| \leqslant \sup_{B^N_{\rho*}(I_0) \times \mathbb{T}^N_\xi} |H| \cdot \left(\frac{1}{\varepsilon_0}\right)^j \tag{19}$$

Proof. Fixed $|\varepsilon| \leqslant \varepsilon_0/2$, writing the ε-expansion of $\mathscr{H}$ we have that there exist ε^*, $|\varepsilon^*| \leqslant \varepsilon_0/2$ such that

$$\mathscr{H}(y, x; \varepsilon) = \sum_{j=0}^k \varepsilon^j \mathscr{H}_j(y, x) + \frac{\varepsilon^{k+1}}{(k+1)!} \frac{\partial^{k+1} \mathscr{H}}{\partial \varepsilon^{k+1}}(y, x; \varepsilon^*)$$

From Birkhoff's Theorem one has $\mathscr{H}_0 = h$ and $\mathscr{H}_j = (1/j!)(\partial^j \mathscr{H}/\partial \varepsilon^j)(y, x; 0) = g_j$ for $1 \leqslant j \leqslant k$ and $R_k(y, x; \varepsilon) \equiv [1/(k+1)!](\partial^{k+1} \mathscr{H}/\partial \varepsilon^{k+1})(y, x; \varepsilon^*)$. Then apply the Cauchy Estimate. ∎

3. A SIMPLE MODEL WITH NO SMALL DENOMINATORS BUT WITH DIVERGING BIRKHOFF SERIES, SUGGESTING THE "SHARPNESS" OF THE PREVIOUS ESTIMATES

The estimates in Proposition 2.4 state that the divergence of the formal series $\sum \varepsilon^j \chi_j$ is *not only due to the small divisors*, but it is, in the worst cases, *factorial*. To suggest the optimality of this estimate we present a *simplified model*: we consider the perturbation of an *harmonic oscillator*

$$H(i, \varphi) = \omega I + \varepsilon f(I, \varphi)$$

with $I \in \mathbb{R}$, $\varphi \in S^1$ and $f = (e^{i\varphi} + e^{-i\varphi})(1 - I) = 2\cos\varphi \cdot (1 - I)$. Looking for a transformation generated by $\mathcal{G}(y, \varphi) = y\varphi + \chi(y, \varphi) = y\varphi + \sum_{j \geqslant 1} \chi_j(y, \varphi)$ in order to "integrate" the system, i.e., to find a new Hamiltonian depending only on the action variables, we obtain the "Hamilton–Jacobi equation"

$$\omega \partial_\varphi \chi + \varepsilon f(y + \partial_\varphi \chi, \varphi) = 0 \tag{20}$$

Since the effect of the operator $\omega \partial_\varphi$ is "multiplying for a small denominator," instead of (20) we consider the following simplified problem, in which we suppress the small divisors:

$$u + \varepsilon f(y + \partial_\varphi u, \varphi) = 0 \tag{21}$$

where $u(y, \varphi) = \sum_{j \geqslant 1} \varepsilon^j u_j(y, \varphi)$ is the unknown function. The formal solution of (21) is

$$u_j = \begin{cases} \sigma_j(y - 1) \displaystyle\sum_{k=0}^{(j-1)/2} c_{j,k} \cos((j - 2k)\varphi) & \text{if } j \text{ is odd} \\ \sigma_j(y - 1) \displaystyle\sum_{k=0}^{(j-2)/2} c_{j,k} \sin((j - 2k)\varphi) & \text{if } j \text{ is even} \end{cases}$$

for suitable $c_{j,k} \geqslant 0$ with $c_{j,0} = 2(j-1)!$ and $\sigma_j = \pm 1$. This shows that, if j is odd, u_j evaluated in $y = 0$ and $\varphi = 0$ grows up factorially. This example shows that, even in problems with no small divisors, the Birkhoff series can diverge with a factorial growth.

4. NEKHOROSHEV'S THEOREM FOR HARMONIC OSCILLATORS WITH DIOPHANTINE FREQUENCIES

We now apply the previous estimates to obtain exponentially long times of stability for perturbations of harmonic oscillators with "strongly non-resonant" frequencies. As customary, we denote by $(I(t), \varphi(t))$ the solution of the Hamilton equations with initial value $(I(0), \varphi(0))$.

The following result can be easily derived by an optimization argument and a "cut-off" in the perturbation similar to the one in ref. 3. More precisely: given a periodic perturbation, we will split it into two terms, and the first term will be a trigonometric polynomial of an appropriate degree: the degree of this polynomial is chosen in order to optimize our estimate.

Theorem 4.1. Consider the system $H(I, \varphi) = \omega \cdot I + \varepsilon f(I, \varphi)$, where $I \in B^N_{\rho^*}(I_0)$, $\varphi \in \mathbb{T}^N_{2\xi}$, f is real analytic and periodic in φ. Assume that $\omega \in \mathbb{R}^N$ is (σ, τ)-Diophantine, i.e., there exist two constants $\sigma, \tau > 0$, such that $|\omega \cdot n| \geqslant \sigma\, |n|^{-\tau}$, $\forall n \in \mathbb{Z}^N - \{0\}$.

Then, we have stability in the actions for exponentially long times, i.e., there exist constants $c_i > 0$, $i = 0,..., 6$, such that

$$|I(t) - I(0)| \leqslant c_0 \varepsilon^{c_1} \qquad \forall |t| \leqslant c_2 \varepsilon^{c_3} \exp \frac{c_4}{\varepsilon^{c_5}} \tag{22}$$

for any initial data satisfying $|I(0) - I_0| \leqslant \rho^*/c_6$.

The constants c_1, c_3 and c_5 depend only on τ, the others may also depend on ρ^*, ξ, H and σ.

Proof. Take $P \equiv 1/\varepsilon^{1/(3\tau)}$ and k as the integer part of $c^\#/(\varepsilon P^{2\tau+2})^{1/(2r+1)}$ with $c^\# \equiv 1/(C_2 e)^{1/(2r+1)}$, standing the C_i's for suitable positive constants as follows. Write $f = f_{\leqslant P} + f_{>P}$, where $f_{\leqslant P} \equiv \sum_{|n| \leqslant P} \hat{f}_n(I)\, e^{in \cdot \varphi}$.

The size of $f_{>P}$ is controlled by $C_1 e^{-C_1/\varepsilon^{1/(3\tau)}}$. We call $(y(0), x(0))$ the point in normal coordinates corresponding to $(I(0), \varphi(0))$. By the Diophantine condition, we have that $\gamma \geqslant C_0/(kP)^\tau$. Looking at the system in normal form (3), recalling Theorem 2.7, we have

$$\begin{gathered} |y(t) - y(0)| \leqslant (\sup |H| + 1)\, k^{\tau+1} \varepsilon \leqslant C_3 \varepsilon^{c_1} \leqslant \rho, \\ \forall |t| \leqslant \min \left\{ \frac{k^{\tau+1}\varepsilon}{(C_2 \varepsilon P^{2\tau+2} k^{2\tau+1})^{k+1}}, \frac{k^{\tau+1}}{C_1 e^{-C_1/\varepsilon^{1/(3\tau)}}} \right\} \end{gathered} \tag{23}$$

By Proposition (2.5), $\forall\, |y - y_0| \leqslant \rho$ and $\varphi \in \mathbb{T}^N_\xi$,

$$\left| \sum_{j=1}^{k} \varepsilon^j \partial_\varphi \chi_j \right| \leqslant \frac{\gamma}{P} \sum_{j=1}^{k} \left(\frac{\tilde{c} k \varepsilon P^2}{\eta_0} \right)^j \leqslant \frac{2\gamma}{P} \cdot \frac{\tilde{c} k \varepsilon P^2}{\eta_0} \leqslant C_4 \varepsilon^{c_1} \tag{24}$$

So (22) follows by (23) and (24). Notice that, with the definitions in (16), C_2 depends on c^* [explicitly one can assume $C_2 \equiv 2/(c^* C_0^2)$], so that $\varepsilon_0 \geqslant c^* C_0^2 (C_2 e)\, \varepsilon \geqslant 2\varepsilon$. ∎

ACKNOWLEDGMENTS

It is a pleasure to thank Prof. Luigi Chierchia and Prof. Rafael de la Llave for their strong help and encouragement. This work has been partially supported by Indam.

REFERENCES

1. V. I. Arnol'd (ed.), *Dynamical Systems III*, Encyclopaedia of Mathematical Sciences, Vol. 3 (New York, Springer-Verlag, 1985).
2. G. Benettin, L. Galgani, and A. Giorgilli, A proof of Nekhoroshev's Theorem for the stability times in nearly integrable Hamiltonian systems, *Celestial Mechanics* **37**(1):1–25 (1985).
3. G. Benettin and G. Gallavotti, Stability of motions near resonances in quasi-integrable Hamiltonian systems, *J. Stat. Phys.* **44**:293–338 (1986).
4. U. Bessi, L. Chierchia, and E. Valdinoci, Upper bounds on Arnold diffusion time via mather theory, *Journal des Math. Pures et Appl.*, to appear.
5. G. D. Birkhoff, Dynamical systems, *Amer. Math. Soc.* (New York, 1927).
6. F. Bonetto, G. Gallavotti, G. Gentile, and V. Mastropietro, Lindstedt series, ultraviolet divergences and Moser's Theorem, *Ann. Scuola Norm. Sup. Pisa Cl. Sci. IV* **26**(3):545–503 (1998).
7. E. Borel, *Les séries divergentes*, Gauthier Villars Imprimeur Libraire, Paris, 1901.
8. A. D. Bryuno, Analytic form of differential equations, *Trans. Moscow Math. Soc.* **25**:131–288 (1971).
9. A. D. Bryuno, Analytic form of differential equations II, *Trans. Moscow Math. Soc.* **26**:199–239 (1972).
10. A. D. Bryuno, Normalization of a Hamiltonian system near an invariant cycle or torus, *Russian Math. Surveys* **44**(2):53–89 (1989).
11. J. Cresson, Thesis, 1998.
12. R. Cushman, Examples of nonintegrable analytic Hamiltonian vector fields with no small divisors, *Trans. Amer. Math. Soc.* **238**:45–55 (1978).
13. R. de la Llave, A. Jorba, work in progress.
14. G. Gallavotti, A criterion of integrability for perturbed nonresonant harmonic oscillators. "Wick ordering" of the perturbations in classical mechanics and invariance of the frequency spectrum, *Comm. Math. Phys.* **87**(3):365–383 (1982).
15. G. Gallavotti, Quasi-integrable mechanical systems, in *Les Houches, Session XLIII–1984* (Elsevier Sc. Publ., 1986).
16. A. Giorgilli and L. Galgani, Rigorous estimates for the series expansion of Hamiltonian perturbation theory, *Celestial Mechanics* **37**(2):95–112 (1985).
17. A. Giorgilli and A. Morbidelli, Invariant KAM torn and global stability for Hamiltonian systems, *ZAMP* **48**(1):102–134 (1997).
18. F. John, *Partial Differential Equations* (New York, Springer-Verlag, 1982).
19. A. Jorba, *A Methodology for the Numerical Computation of Normal Forms*, *Centre Manifolds and First Integrals of Hamiltonian Systems*, mp_arc 98-6.
20. P. Lochak, Canonical perturbation theory via simultaneous approximation, *Russian Math. Surveys* **47**(6):57–133 (1992).
21. P. Lochak, A. I. Neistadt, and L. Niederman, *Stability of nearly integrable convex Hamiltonian systems over exponentially long times*, Seminar on Dynamical Systems, Progr. Nonlinear Differential Equations Appl. 12, Birkhauser, Basel, pp. 15–34, 1994.

22. J. Moser, On the integrability of area preserving Cremona mappings near an elliptic fixed point, *Bol. Soc. Mat. Mexicana* **5**(2):176–180 (1960).
23. N. N. Nekhoroshev, An exponential estimate of the time of stability of nearly integrable Hamiltonian systems, *Russian Math. Surveys* **32**(6):176–180 (1977).
24. H. Poincaré, *Sur les propriétés des fonctions définies par les équations aux différences partielles*, Thèse, Paris, 1879.
25. J. Pöschel, Nekhoroshev estimates for quasi-convex Hamiltonian systems, *Math. Zeitsc.* **213**(2):187–216 (1993).
26. C. L. Siegel and J. Moser, *Lectures on Celestial Mechanics* (New York, Springer-Verlag, 1971).
27. J. Sijbrand, Properties of center manifolds, *Trans Amer. Math. Soc.* **289**(2):431–469 (1985).
28. B. Sternin and V. Shatalov, *Borel–Laplace Transform and Asymptotic Theory* (CRC Press, New York, 1996).
29. A. Vanderbauwhede, Centre Manifolds, Normal Forms and Elementary Bifurcations, in *Dynamics Reported*, Vol. 2 (Wiley, 1989), pp. 89–169.