 Istituto Lombardo (Rend. Scienze) 146, 133-160 (2012)

ON A THEOREM OF LYAPOUNOV

Nota del s.c. ANTONIO GIORGILLI (*)

(Adunanza del 3 maggio 2012)

SUNTO. Si mostra che un sistema Hamiltoniano nell'intorno di un punto di equilibrio, sotto condizione che gli autovalori soddisfino delle condizioni di non–risonanza del tipo di Melnikov, ammette una forma normale che rende evidente l'esistenza di una varietà invariante (locale) a due dimensioni sulla quale si hanno soluzioni note. Nel caso di un autovalore puramente immaginario tali soluzioni formano una famiglia periodica a due parametri che costituisce la continuazione naturale di un modo normale. Questo secondo risultato è stato dimostrato in precedenza da Lyapounov. In questo lavoro si completa quello di Lyapounov dimostrando la convergenza della trasformazione dell'Hamiltoniana a forma normale e rimuovendo le restrizione che gli autovalori siano puramente immaginari.

***

ABSTRACT. It is shown that a Hamiltonian system in the neighbourhood of an equilibrium may be given a special normal form in case the eigenvalues of the linearized system satisfy non–resonance conditions of Melnikov's type. The normal form possesses a two dimensional (local) invariant manifold on which the solutions are known. If the eigenvalue is pure imaginary then these solutions are the natural continuation of a normal mode of the linear system. The latter result was first proved by Lyapounov. The present paper completes Lyapounov's result in that the convergence of the transformation of the Hamiltonian to a normal form is proven and the condition that the eigenvalues be pure imaginary is removed.

(*) Dipartimento di Matematica, Via Saldini 50, 20133 Milano, Italy.
E-mail: antonio.giorgilli@unimi.it

1. Introduction

Consider a canonical system of differential equations in a neighbourhood of an equilibrium, with Hamiltonian

$$H(x,y) = H_0(x,y) + H_1(x,y) + \dots, \quad (x,y) \in \mathbb{C}^{2n}, \tag{1}$$

where the unperturbed quadratic part of the Hamiltonian is

$$H_0(x,y) = \sum_{j=1}^{n} \lambda_j x_j y_j, \quad (\lambda_1, \dots, \lambda_n) \in \mathbb{C}^n, \tag{2}$$

and $H_s(x,y)$ for $s \geq 1$ is a homogeneous polynomial of degree $s+2$. The form (2) is a typical one for the quadratic part of a Hamiltonian system in the neighbourhood of an equilibrium, as under quite general conditions the system may be given that form via a (complex) linear canonical transformation (see, e.g., [10] or [12], § 15).

The Hamiltonian is assumed to be analytic in some neighborhood of the origin of $\mathbb{C}^{2n}$. Moreover λ_1 will be assumed to satisfy at least the first of the following non–resonance conditions:

(i) *First Melnikov's condition:*

$$\lambda_\nu - k\lambda_1 \neq 0 \quad \text{for } k \in \mathbb{Z} \text{ and } \nu = 2, \dots, n\,. \tag{3}$$

(ii) *Second Melnikov's condition:*

$$\lambda_\nu \pm \lambda_{\nu'} - k\lambda_1 \neq 0 \quad \text{for } k \in \mathbb{Z} \text{ and } \nu, \nu' = 2, \dots, n\,, \tag{4}$$

the case $\nu' = \nu$ being included.

In [10] Lyapounov proved that if $\lambda_1 = i\omega_1$ is pure imaginary and the non resonance condition (i) above is satisfied then there exists a two parameter family of solutions of the form

$$x_j = \varphi_j(\xi_1, \eta_1), \quad y_j = \psi_j(\xi_1, \eta_1) \tag{5}$$

written as convergent power series in the arguments

$$\xi_1 = \overset{\circ}{\xi}_1 e^{ita_1(\overset{\circ}{\zeta})}, \quad \eta_1 = \overset{\circ}{\eta}_1 e^{-ita_1(\overset{\circ}{\zeta})}, \tag{6}$$

where $a_1(\overset{\circ}{\zeta}) = \lambda_1 + \dots$ is a convergent power series in $\overset{\circ}{\zeta}_1 = \overset{\circ}{\xi}_1 \overset{\circ}{\eta}_1$. In the case $n = 1$ this actually describes all solutions of the system. A proof in case all λ's are pure imaginary is reported in [12].

The proof of the theorem is worked out by the authors quoted above by expanding the solution in the form (5) and proceeding by comparison of coefficients. From a formal viewpoint the statement above looks equivalent to the existence of a canonical transformation that gives the system (1) a suitable normal form, making the Hamiltonian to depend at least quadratically on $x_2, \dots, x_n, y_2, \dots, y_n$. A formal construction giving such a normal form can be easily produced. However, proving the convergence of the normalization procedure seems to be more difficult. The aim of this paper is precisely to produce a proof of convergence of the transformation to normal form.

I will actually give two different statements that can be proved with the same method. The first one is

THEOREM 1.1. *With the nonresonance hypothesis (i) above (first Melnikov's condition) on $\lambda_1, \dots, \lambda_n$, there exists a canonical, near the identity transformation in the form of a power series convergent in a neighbourhood of the origin, which gives the Hamiltonian* (1) *the normal form*

$$H(x, y) = H_0(x, y) + \Gamma(x_1 y_1) + F(x, y), \tag{7}$$

where $H_0(x, y)$ as in (1)*, $\Gamma(x_1 y_1)$ depends only on the product $x_1 y_1$, and $F(x, y)$ is at least quadratic in $x_2, \dots, x_n, y_2, \dots, y_n$*

The existence of the Lyapounov orbits for λ_1 pure imaginary is evident from the normal form: just put initially $x_2 = \dots = x_n = y_2 = \dots = y_n = 0$, which defines a local invariant two dimensional manifold on which the dynamics is generated by the Hamiltonian $\lambda_1 x_1 y_1 + \Gamma(x_1 y_1)$. The advantage of the normal form is that it allows also to

investigate the dynamics in the neighbourhood of the orbits so found. To this end the following statement may be even more useful.

THEOREM 1.2. *With the nonresonance hypotheses (i) and (ii) above (first and second Melnikov's conditions) on $\lambda_1, \dots, \lambda_n$, there exists a canonical, near the identity transformation in the form of a power series convergent in a neighbourhood of the origin, which gives the Hamiltonian* (1) *the normal form*

$$H(x,y) = H_0(x,y) + \Gamma(x_1y_1, \dots, x_ny_n) + F(x,y), \tag{8}$$

where $H_0(x,y)$ as in (1)*, $\Gamma(x,y)$ contains only monomials $x_1^j y_1^j x_\nu y_\nu$ with a positive integer j and with $\nu = 2, \dots, n$, and $F(x,y)$ is at least cubic in $x_2, \dots, x_n, y_2, \dots, y_n$*

This requires a stronger non–resonance condition. However this normal form may be more convenient if one is interested in the stability of a Lyapounov orbit. Indeed, let all λ's be pure imaginary, say $\lambda_j = i\omega_j$, and write the Hamiltonian restricted to the invariant manifold $x_2 = \dots = x_n = y_2 = \dots = y_n = 0$ in action–angle variables by transforming $x_1 = \sqrt{p}\, e^{iq}$, $y_1 = -i\sqrt{p}\, e^{-iq}$. Thus one gets the Hamiltonian $\omega_1 p + \Gamma(p)$, which represents a non linear oscillator, with orbits written as $p(t) = p_*$, $q(t) = q(0) + \Omega(p_*)t$, where $\Omega(p_*) = \omega_1 + O(p_*)$ is a fixed frequency. By a translation $p' = p - p_*$ the Hamiltonian may be reexpanded (omitting primes) as

$$H(q,p,x,y) = \Omega p + \sum_{j=2}^{n} \lambda_j x_j y_j + H_1 + H_2 + \dots$$

where H_s is a homogeneous polynomial of degree $s+2$ in $p^{1/2}, x_2, \dots, y_n$ with coefficients periodically depending on q. The dynamics of the latter Hamiltonian may be investigated with known methods from perturbation theory. The advantage with respect to the normal form of theorem 1.1 is that the quadratic part of the Hamiltonian is independent of the angle q.

The proof is based on a previous work by the author [5] concerning the construction of the normal form in a case investigated by Cherry [2]

and Moser [11]. It must be stressed that this problem does not involve small divisors. Rather, the possible source of divergence is due to the use of Cauchy's estimates for the derivatives required by the normalization algorithm. The global effect of accumulation of derivatives is controlled with a technique introduced by the author and U. Locatelli in order to achieve a proof of KAM theorem using classical expansions in a perturbation parameter (see [6, 7, 4, 8]).

2. FORMAL ALGORITHM

Reducing the Hamiltonian to a normal form is a quite general problem which may be solved in a number of different ways. Moreover, the concept of "normal form" may assume a quite general meaning, depending on what one is looking for. Here I state the algorithm in a general form, using the method of composition of Lie series.

2.1 The algorithm for the normal form

Write the Hamiltonian after r normalization steps as

$$H^{(r)}(x,y) = H_0(x,y) + Z_1(x,y) + \ldots + Z_r(x,y) + \sum_{s>r} H_s^{(r)}(x,y), \quad (9)$$

where $Z_1(x,y), \ldots, Z_r(x,y)$ are in normal form, whatever it means, and are homogeneous polynomials of degree $3, \ldots, r+2$. For $r = 0$ the Hamiltonian (1) is already in the wanted form, with no functions Z.

Assume that the Hamiltonian has been given a normal form (9) up to order $r-1$, so that $H^{(r-1)}$ is known. The generating function χ_r and the normal form Z_r are determined by solving the equation

$$L_{H_0}\chi_r + Z_r = H_r^{(r-1)}. \quad (10)$$

where the common notation $L_\varphi \cdot := \{\cdot, \varphi\}$ has been used. The solution of this equation depends on what is meant by "normal form". At a formal level, any choice thatallows to solve the equation above for Z_r and χ_r is

acceptable. Assume for a moment that a method of solution has been found. Then the transformed Hamiltonian is expanded as

$$H^{(r)}_{sr+m} = \frac{1}{s!}L^s_{\chi_r} Z_m + \sum_{p=0}^{s-1} \frac{1}{p!} L^p_{\chi_r} H^{(r-1)}_{(s-p)r+m}$$

$$\text{for } r \geq 2\,,\ s \geq 1 \text{ and } 1 \leq m < r,$$

$$H^{(r)}_{sr} = \frac{1}{(s-1)!} L^{s-1}_{\chi_r} \left(\frac{1}{s} Z_r + \frac{s-1}{s} H^{(r-1)}_r \right) + \sum_{p=0}^{s-2} \frac{1}{p!} L^p_{\chi_r} H^{(r-1)}_{(s-p)r}$$

$$\text{for } r \geq 1 \text{ and } s \geq 2. \tag{11}$$

The justification of the algorithm requires only some straightforward calculation, and is deferred to appendix A.

Thus, the problem is how to solve the equation (10) for the generating function χ_r and the normal form Z_r. Let me make some general considerations.

Let $\mathcal{P}_s$ denote the linear space of homogeneous polynomials of degree s in the complex variables x, y. Let also $\mathcal{P} = \bigcup_{s\geq 0} \mathcal{P}_s$, so that a formal power series is an element of $\mathcal{P}$. A basis in $\mathcal{P}$ is given by the monomials $x^j y^k := x_1^{j_1} \cdot \ldots \cdot x_n^{j_n} y_1^{k_1} \cdot \ldots \cdot y_n^{k_n}$, where j, k are integer vectors with non–negative components. The linear operator L_{H_0} maps every space $\mathcal{P}_s$ into itself. If, due to the choice of the coordinates, the unperturbed Hamiltonian H_0 has the form (2) then the operator L_{H_0} is diagonal, since

$$L_{H_0} x^j y^k = \langle j - k, \lambda \rangle x^j y^k.$$

The kernel and the range of L_{H_0} are defined as usual, namely $\mathcal{N} = L^{-1}_{H_0}(0)$, the inverse image of the null vector in $\mathcal{P}$, and $\mathcal{R} = L_{H_0}(\mathcal{P})$. Both $\mathcal{N}$ and $\mathcal{R}$ are actually subspaces of the same space $\mathcal{P}$, and it turns out that they are complementary subspaces, i.e., $\mathcal{N} \cap \mathcal{R} = \{0\}$, the null vector, and $\mathcal{N} \oplus \mathcal{R} = \mathcal{P}$. A consequence of the properties above is

that L_{H_0} restricted to the subspace $\mathcal{R}$ is uniquely inverted, i.e., the equation $L_{H_0}\chi = \psi$ with $\psi \in \mathcal{R}$ admits an unique solution χ satisfying the condition $\chi \in \mathcal{R}$. That unique solution will be written as $\chi = L_{H_0}^{-1}\psi$, i.e., $L_{H_0}^{-1}$ is defined as the inverse of L_{H_0} restricted to $\mathcal{R}$. It's easy to identify the subspaces $\mathcal{N}$ and $\mathcal{R}$ using the coordinates. Thanks to the diagonal form of L_{H_0} one has

$$\begin{aligned} \mathcal{N} &= \text{span}\left\{x^j y^k \ : \ \langle j-k, \lambda\rangle = 0\right\}, \\ \mathcal{R} &= \text{span}\left\{x^j y^k \ : \ \langle j-k, \lambda\rangle \neq 0\right\}. \end{aligned} \tag{12}$$

Given $\psi \in \mathcal{R}$ and writing $\psi = \sum_{j,k} \psi_{j,k} x^j y^k$, with $\psi_{j,k} = 0$ for $x^j y^k \in \mathcal{N}$, one has

$$L_{H_0}^{-1}\psi = \sum_{j,k} \frac{\psi_{j,k}}{\langle j-k, \lambda\rangle} x^j y^k. \tag{13}$$

In view of the general considerations above we can conclude that the choice of a normal form is subjected to the constraint that in equation (10) we have $H_r^{(r-1)} - Z_r \in \mathcal{R}$. The simplest choice is to ask also $Z_r \in \mathcal{N}$, i.e., to set Z_r to be the projection of $H_r^{(r-1)}$ on the subspace $\mathcal{N}$. This is known indeed as Birkhoff's normal form.

2.2 Normal form for Lyapounov's orbits

I come now to show that the construction of the normal form of theorem 1.1 is formally consistent. Consider the disjoint subsets of $\mathbb{Z}^n$

$$\begin{aligned} \mathcal{K}^\sharp &= \{k \in \mathbb{Z}^n \ : \ k_2 = \ldots = k_n = 0\}, \\ \mathcal{K}^\natural &= \{k \in \mathbb{Z}^n \ : \ |k_2| + \ldots + |k_n| = 1\}, \\ \mathcal{K}^\flat &= \{k \in \mathbb{Z}^n \ : \ |k_2| + \ldots + |k_n| > 1\}. \end{aligned} \tag{14}$$

One has $\mathbb{Z}^n = \mathcal{K}^\sharp \cup \mathcal{K}^\natural \cup \mathcal{K}^\flat$, of course. Considering only integer vectors $j,\ k$ with non–negative components, introduce the subspaces of $\mathcal{P}$

$$\begin{aligned} \mathcal{P}^\sharp &= \text{span}\left\{ x^j y^k \ :\ j+k \in \mathcal{K}^\sharp \right\} \\ \mathcal{P}^\natural &= \text{span}\left\{ x^j y^k \ :\ j+k \in \mathcal{K}^\natural \right\} \\ \mathcal{P}^\flat &= \text{span}\left\{ x^j y^k \ :\ j+k \in \mathcal{K}^\flat \right\} \end{aligned} \tag{15}$$

These subspaces are clearly disjoint, and moreover one has $\mathcal{P} = \mathcal{P}^\sharp \oplus \mathcal{P}^\natural \oplus \mathcal{P}^\flat$. Finally, let $\mathcal{N}^\sharp = \mathcal{N} \cap \mathcal{P}^\sharp$ and $\mathcal{R}^\sharp = \mathcal{R} \cap \mathcal{P}^\sharp$, and define the subspaces $\mathcal{Z}$ and $\mathcal{W}$ of $\mathcal{P}$ as

$$\mathcal{Z} = \mathcal{N}^\sharp \oplus \mathcal{P}^\flat, \quad \mathcal{W} = \mathcal{R}^\sharp \oplus \mathcal{P}^\natural. \tag{16}$$

It is an easy matter to check that $\mathcal{Z} \cap \mathcal{W} = \{0\}$ and $\mathcal{Z} \oplus \mathcal{W} = \mathcal{P}$. The construction of bases for $\mathcal{Z}$ and $\mathcal{W}$ is quite straightforward: a monomial $x^j y^k$ belongs to $\mathcal{Z}$ in either case ($j+k \in \mathcal{K}^\sharp$ and $\langle j-k, \lambda \rangle = 0$) or ($j+k \in \mathcal{K}^\flat$); else it belongs to $\mathcal{W}$. The hypothesis (i) on λ (first Melnikov's condition) formulated at the beginning of the introduction means that the non-resonance condition

$$\langle k, \lambda \rangle \neq 0 \quad \text{for } 0 \neq k \in \mathcal{K}^\sharp \cup \mathcal{K}^\natural \tag{17}$$

is satisfied. This implies $\mathcal{W} \subset \mathcal{R}$, so that for every $\psi \in \mathcal{W}$ the unique solution $\chi = L_{H_0}^{-1}\psi$, $\chi \in \mathcal{W}$ of the equation $L_{H_0}\chi = \psi$ exists. With this setting, the equation

$$L_{H_0}\chi + Z = \Psi, \tag{18}$$

with Ψ known, admits a straightforward solution. Split $\Psi = \Psi_{\mathcal{Z}} + \Psi_{\mathcal{W}}$ with $\Psi_{\mathcal{Z}} \in \mathcal{Z}$ and $\Psi_{\mathcal{W}} \in \mathcal{W}$; such a decomposition exists and is unique, because $\mathcal{Z}$ and $\mathcal{W}$ are complementary subspaces. Then set $Z = \Psi_{\mathcal{Z}}$, and determine $\chi = L_{H_0}^{-1}\Psi_{\mathcal{W}}$ according to (13).

2.3 Adding the second Melnikov's condition

With minor changes one can also prove that the normal form of theorem 1.2 can be constructed. Let

$$\begin{aligned}\mathcal{K}^\sharp &= \{k \in \mathbb{Z}^n \ : \ k_2 = \ldots = k_n = 0\}\,,\\ \mathcal{K}^\natural &= \{k \in \mathbb{Z}^n \ : \ |k_2| + \ldots + |k_n| = 1, 2\}\,,\\ \mathcal{K}^\flat &= \{k \in \mathbb{Z}^n \ : \ |k_2| + \ldots + |k_n| > 2\}\,.\end{aligned} \tag{19}$$

One has again $\mathbb{Z}^n = \mathcal{K}^\sharp \cup \mathcal{K}^\natural \cup \mathcal{K}^\flat$, of course. The subspaces of $\mathcal{P}$ are defined again as in (15), although they turn out to be different in view of the differences in the sets $\mathcal{K}$. Finally, let $\mathcal{N}^\sharp = \mathcal{N} \cap \mathcal{P}^\sharp$, $\mathcal{R}^\sharp = \mathcal{R} \cap \mathcal{P}^\sharp$, $\mathcal{N}^\natural = \mathcal{N} \cap \mathcal{P}^\natural$ and $\mathcal{R}^\natural = \mathcal{R} \cap \mathcal{P}^\natural$, and define the subspaces $\mathcal{Z}$ and $\mathcal{W}$ of $\mathcal{P}$ as

$$\mathcal{Z} = \mathcal{N}^\sharp \oplus \mathcal{N}^\natural \oplus \mathcal{P}^\flat, \quad \mathcal{W} = \mathcal{R}^\sharp \oplus \mathcal{R}^\natural. \tag{20}$$

The difference with respect to the previous case is just that now $\mathcal{N}^\natural$ is not empty, because it contains all monomials of the form $(x_1 y_1)^j \times x_\nu y_\nu$ with $\nu = 2, \ldots, n$ and positive j. This forces the change in the definition of the subspaces $\mathcal{Z}$ and $\mathcal{W}$. However, the properties $\mathcal{Z} \cap \mathcal{W} = \{0\}$ and $\mathcal{Z} \oplus \mathcal{W} = \mathcal{P}$ remain true. Furthermore, in view of the second Melnikov's condition, also the property $\mathcal{W} \subset \mathcal{R}$ holds true, so that for every $\psi \in \mathcal{W}$ the unique solution $\chi = L_{H_0}^{-1}\psi$, $\chi \in \mathcal{W}$ of the equation $L_{H_0}\chi = \psi$ exists.

3. QUANTITATIVE ESTIMATES

Pick a real vector $R \in \mathbb{R}^n$ with positive components. and consider the domain

$$\Delta_R = \{(x, y) \in \mathbb{C}^n \ : \ |x_j| \leq R_j \,, \ |y_j| \leq R_j \,, \ 1 \leq j \leq n\}\,, \tag{21}$$

namely a polydisk which is the product of disks of radii $R_1, \ldots, R_n$ in the planes of the complex coordinates $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$, respectively. Let also

$$\Lambda = \min_{1 \leq j \leq n} R_j. \tag{22}$$

The norm $\|f\|_R$ in the polydisk Δ_R is defined as

$$\|f\|_R = \sum_{|j+k|=r} |f_{j,k}| R^{j+k}. \tag{23}$$

A family of polydisks $\Delta_{\delta R}$ of radii δR, with $0 < \delta \leq 1$ will be considered below. With a minor abuse the simplified notation $\|\cdot\|_\delta$ in place of $\|\cdot\|_{\delta R}$ will be used.

The main result of this section is

LEMMA 3.1. *Let the Hamiltonian $H^{(0)}$ satisfy $\|H^{(0)}_s\|_1 \leq h^{s-1}E$ for $s \geq 1$, with some constants $h \geq 0$ and $E > 0$. Let $0 < d < 1/2$. Then there exist positive constants β and G depending on E, h, Λ, d and on $\lambda_1, \ldots, \lambda_n$ such that*

$$\|\chi_r\|_{1-d} \leq \beta^{r-1} G \quad \text{for all } r \geq 1.$$

The rest of this section is devoted to the proof. Some technical calculations are deferred to appendix B.

3.1 An arithmetic lemma

The following lemma will play a crucial role in the proof of lemma 3.1.

LEMMA 3.2. *Let $\lambda \in \mathbb{C}^n$ be such that λ_1 satisfies the non-resonance condition* (3)*. Then there exists a positive γ such that the inequality*

$$|\langle k, \lambda \rangle| \geq |k| \gamma$$

holds true for all non–zero $k \in \mathcal{K}^\sharp \cup \mathcal{K}^\natural$ defined as in (14)*.*

COROLLARY 3.3. *Let in addition the non-resonance condition* (4) *be satisfied. Then the same statement holds true for all non–zero $k \in \mathcal{K}^\sharp \cup \mathcal{K}^\natural$ defined as in* (19)*.*

The proof of the corollary is just a trivial modification of the PROOF of LEMMA 3.2. For $k \in \mathcal{K}^\sharp$ the claim is obvious, since $|\langle k, \lambda \rangle| = |k_1 \lambda_1|$. So, let $k \in \mathcal{K}^\natural$. Set $\vartheta = \max\big(|\lambda_2|, \ldots, |\lambda_n|\big)$ (for the corollary

maximize also over $|2\lambda_\nu|$ and $|\lambda_\nu \pm \lambda_{\nu'}|$ with $\nu' \neq \nu$). Pick an integer $N \geq 1 + 2\vartheta$ and set

$$\delta = \min_{\substack{k \in \mathcal{K}^\natural \\ |k| \leq N}} \big|\langle k, \lambda\rangle\big|, \quad \gamma = \min\left(\frac{\delta}{N}, \frac{|\lambda_1|}{2}\right);$$

in view of the non-resonance condition (17) one has $\delta > 0$. Then the claim of the lemma holds true with the given value of γ. Indeed, let $k \in \mathcal{K}^\natural$, so that $|k_1| = |k| - 1$. If $|k| \leq N$ then $\big|\langle k, \lambda\rangle\big| \geq \delta \geq N\gamma \geq |k|\gamma$. If $|k| > N$ use $\vartheta \leq (N-1)\delta/2$, which follows from the choice of N, and evaluate

$$\begin{aligned} \big|\langle k, \lambda\rangle\big| &\geq \big|k_1\lambda_1\big| - \vartheta \geq \big(|k| - 1\big)\delta - \tfrac{(N-1)}{2}\,\delta \\ &\geq \frac{|k| - 1}{2}\,\delta + \frac{N}{2}\,\delta - \frac{(N-1)}{2}\,\delta = |k|\frac{\delta}{2} \geq |k|\gamma. \end{aligned}$$

∎

3.2 Generalized Cauchy estimates

Here I refer to the more restrictive hypotheses of theorem 1.2, and in particular to the spaces $\mathcal{P}$ defined as in sect. 2.3. However, the same arguments with very little simplifications apply also to the setting of sect. 2.2, which applies to theorem 1.1. I will insert short comments in parentheses concerning the latter case, where appropriate.

The estimates in this section strongly depend on a suitable splitting of all functions over the subspaces $\mathcal{P}^\sharp$, $\mathcal{P}^\natural$ and $\mathcal{P}^\flat$. At a formal level, it is useful to keep in mind the following table concerning the Poisson bracket:

$\{\cdot,\cdot\}$	$\mathcal{P}^\sharp$	$\mathcal{P}^\natural$	$\mathcal{P}^\flat$
$\mathcal{P}^\sharp$	$\mathcal{P}^\sharp$	$\mathcal{P}^\natural$	$\mathcal{P}^\flat$
$\mathcal{P}^\natural$	$\mathcal{P}^\natural$	$\mathcal{P}^\sharp \oplus \mathcal{P}^\natural \oplus \mathcal{P}^\flat$	$\mathcal{P}^\natural \oplus \mathcal{P}^\flat$
$\mathcal{P}^\flat$	$\mathcal{P}^\flat$	$\mathcal{P}^\natural \oplus \mathcal{P}^\flat$	$\mathcal{P}^\flat$

(24)

(For the subspaces defined as in sect.2.2 just remove $\mathcal{P}^\natural$ from the central case corresponding to the Poisson bracket between functions in $\mathcal{P}^\natural$.) In

view of the transformation formulæ (11) the situation to be considered is the following. A generating function $\chi \in \mathcal{W} \cap \mathcal{P}_r$ with some $r \geq 1$ is given in the form $\chi = L_{H_0}^{-1}\psi$, with known $\psi \in \mathcal{W} \cap \mathcal{P}_r$. Since $\mathcal{W} = \mathcal{R}^\sharp \oplus \mathcal{R}^\natural$ one has $\chi = \chi^\sharp + \chi^\natural$, with an obvious meaning of the notation. The operator L_χ may be applied either to a generic function $f = f^\sharp + f^\natural + f^\flat \in \mathcal{P}_s$ with $s \geq r$ or to a function in normal form $Z = Z^\sharp + Z^\natural + Z^\flat \in \mathcal{Z} \cap \mathcal{P}_m$ with $0 < m < r$ (in theorem 1.1 one has $Z^\natural = 0$). In particular one has

$$Z^\sharp = \sum_{j>1} z_j x_1^j y_1^j, \quad Z^\natural = \sum_{j>0,\, 2\leq\nu\leq n} z_{j,\nu} x_1^j y_1^j x_\nu y_\nu,$$

due to the non–resonance conditions on λ. For some non–negative δ', δ'', δ satisfying $0 \leq \max(\delta', \delta'') < \delta \leq 1/2$ the norms $\|\psi\|_{1-\delta'}$, $\|f\|_{1-\delta''}$ and $\|Z\|_{1-\delta''}$ are assumed to be known, and one looks for an estimate of the Lie derivative in a domain $\Delta_{(1-\delta)R}$. The following estimates will be used in the rest of the paper.

(i) The generating function χ is estimated by

$$\|\chi\|_{1-\delta'} \leq \frac{1}{\gamma} \|\psi\|_{1-\delta'}, \tag{25}$$

with γ as in lemma 3.2.

(ii) The general estimate for the Lie derivative of a generic function f is

$$\left\|L_\chi f\right\|_{1-\delta} \leq \frac{1}{(\delta-\delta')(\delta-\delta'')\Lambda^2} \|\chi\|_{1-\delta'} \|f\|_{1-\delta''} \tag{26}$$

with Λ as in (22). Denoting by $\left(L_{\chi^\natural} f^\flat\right)^\natural$ the projection of $L_{\chi^\natural} f^\flat$ over $\mathcal{P}^\natural$ one has

$$\left\|\left(L_{\chi^\natural} f^\flat\right)^\natural\right\|_{1-\delta} \leq \frac{4}{(\delta-\delta'')\Lambda^2} \|\chi\|_{1-\delta'} \|f\|_{1-\delta''}. \tag{27}$$

(iii) For a function Z in normal form one has

$$\left\|L_\chi(Z^\sharp + Z^\natural)\right\|_{1-\delta} \leq \frac{1}{(\delta-\delta'')\gamma\Lambda^2} \|\psi\|_{1-\delta'} \|Z\|_{1-\delta''}. \tag{28}$$

I recall the reader's attention on the missing denominator $\delta - \delta'$ in (27) and (28). This is crucial for the convergence proof. For, working out the convergence proof requires a quite accurate control of the accumulation of the divisors $\delta - \delta'$, $\delta - \delta''$ that appear in the generalized Cauchy estimates for derivatives. The scheme in the next section is specially devised in order to allow such a control.

The proof of (25) is a straightforward consequence of the definition of the norm and of (13). For, the denominators are uniformly estimated from below by γ, in view of lemma 3.2.

The proof of the estimates (26), (27) and (28) is a purely technical matter, and is deferred to appendix B.

3.3 Recursive estimates

The aim of this section is to obtain estimates for the norms of the generating functions and of the transformed Hamiltonians, at every step of the normalization procedure.

Consider a sequence of boxed domains $\Delta_{(1-\delta_r)R}$, where $\{\delta_r\}_{r\geq 1}$ is a monotonically increasing sequence of positive numbers converging to some $d < 1/2$. Let also $\delta_0 = 0$, and $d_r = \delta_r - \delta_{r-1}$ for $r \geq 1$, so that $d_r < 1$ for all positive r. The purpose is to look for estimates of the norms of the generating function χ_r and of the normal form Z_r in the polydisk $\Delta_{(1-\delta_{r-1})R}$, and of the functions $H_s^{(r)}$ in the domain $\Delta_{(1-\delta_r)R}$.

Let $\mathcal{J}_{r,s}$ for $1 < r < s$ be the set of integer arrays defined as

$$\mathcal{J}_{r,s} = \Big\{ J = \{j_1, \dots j_k\} \,:\, j_m \in \{1, \dots, r\}\,,\ 1 \leq k \leq 2(s-1)\,, \\ \sum_{m=1}^{k} \log_2 j_m \leq 2(s - 1 - \log_2 s) \Big\}. \tag{29}$$

Let also $\mathcal{J}_{0,s} = \emptyset$ for $s \geq 1$. Recalling that $\{d_r\}_{r\geq 1}$ is a sequence of positive numbers not exceeding 1 define the sequence $\{T_{r,s}\}_{0\leq r<s}$ as

$$T_{0,s} = 1, \quad T_{r,s} = \max_{J\in\mathcal{J}_{r,s}} \prod_{j\in J} d_j^{-1}. \tag{30}$$

The following properties will be used below: for $0 \le r \le r' < s$ one has

$$T_{r,s} \quad \le T_{r',s}, \tag{31}$$

$$\frac{1}{d_r^2} T_{r-1,r} T_{r',s} \quad \le T_{r',r+s}. \tag{32}$$

Checking (31) is easy: for $r = 0$ use $d_l \le 1$ for $l \ge 1$; for $r > 0$ use the inclusion relation $\mathcal{J}_{r,s} \subset \mathcal{J}_{r',s}$ for $r < r'$. In order to prove (32) remark that by definition one has

$$\begin{aligned} \frac{1}{d_r^2} T_{r-1,r} T_{r',s} &= \frac{1}{d_r^2} \max_{J \in \mathcal{J}_{r-1,r}} \prod_{j \in J} d_j^{-1} \max_{J' \in \mathcal{J}_{r',s}} \prod_{j' \in J'} d_{j'}^{-1} \\ &= \max_{J \in \mathcal{J}_{r-1,r}} \max_{J' \in \mathcal{J}_{r',s}} \prod_{j \in \{r,r\} \cup J \cup J'} d_j^{-1}. \end{aligned}$$

It is enough to prove that $\{r, r\} \cup J \cup J' =: \tilde{J} \in \mathcal{J}_{r',r+s}$. First check that

$$\begin{aligned} \#\big(\tilde{J}\big) &= 2 + \#(J) + \#(J') \le 2 + 2(r-1) + 2(s-1) \\ &= 2(r+s-1). \end{aligned}$$

On the other hand, since $1 \le j \le r-1$ for all $j \in J$ and $1 \le j' \le r'$ for all $j' \in J'$, one also has $1 \le \tilde{j} \le r'$ for all $\tilde{j} \in \tilde{J}$. Finally, evaluate

$$\begin{aligned} \sum_{\tilde{j} \in \tilde{J}} \log_2 \tilde{j} &= 2 \log_2 r + \sum_{j \in J} \log_2 j + \sum_{j' \in J'} \log_2 j' \\ &\le 2 \log_2 r + 2(r - 1 - \log_2 r) \\ &\quad + 2(s - 1 - \log_2 s) \\ &\le 2\big[r + s - 1 - (1 + \log_2 s)\big] \\ &\le 2\big[r + s - 1 - \log_2(r+s)\big], \end{aligned}$$

where the elementary inequality $1 + \log_2 s = \log_2 2 + \log_2 s = \log_2(2s) > \log_2(r+s)$ has been used (recall that $r \le r' < s$). Hence, $\tilde{J} \in \mathcal{J}_{r',r+s}$, as claimed.

I shall also use the numerical sequence $\{\mu_{r,s}\}_{r\geq 0, s\geq 0}$ defined as

$$\begin{aligned} &\mu_{0,0} = 0\,, \quad \mu_{0,s} = 1 \quad \text{for } s > 0, \\ &\mu_{r,s} = \sum_{0\leq rp<s} \mu_{r-1,r}^{p} \mu_{r-1,s-rp} \quad \text{for } r > 0 \text{ and } s \geq 0. \end{aligned} \tag{33}$$

The recursive estimates are collected in

LEMMA 3.4. *Let the Hamiltonian $H^{(0)}$ satisfy $\|H_s^{(0)}\|_1 \leq h^{s-1}E$ for some constants $h \geq 0$ and $E > 0$. Let $d_0 = 1$ and $\{d_r\}_{r\geq 1}$ be an arbitrary sequence of positive numbers satisfying $\sum_{r\geq 1} d_r = d$ with $d < 1$. Let also $\delta_0 = 0$ and $\delta_r = d_1 + \ldots + d_r$. Then for $s > r \geq 1$ the following estimates hold true:*

$$\|\chi_r\|_{1-\delta_{r-1}} \leq \mu_{r-1,r} T_{r-1,r} C^{r-1} \frac{E}{\gamma}, \tag{34}$$

$$\|Z_r\|_{1-\delta_{r-1}} \leq \mu_{r-1,r} T_{r-1,r} C^{r-1} \frac{E}{d_{r-1}}, \tag{35}$$

$$\|Z_r^{\sharp} + Z_r^{\natural}\|_{1-\delta_{r-1}} \leq \mu_{r-1,r} T_{r-1,r} C^{r-1} E, \tag{36}$$

$$\|H_s^{(r)}\|_{1-\delta_r} \leq \mu_{r,s} T_{r,s} C^{s-1} \frac{E}{d_r}, \tag{37}$$

$$\|H_s^{(r),\sharp} + H_s^{(r),\natural}\|_{1-\delta_r} \leq \mu_{r,s} T_{r,s} C^{s-1} E, \tag{38}$$

where

$$C = h + \frac{4e^2 E}{\gamma \Lambda^2}, \tag{39}$$

and $\mu_{r,s}$ and $T_{r,s}$ are the sequences defined by (30) *and* (33).

Remark that (34), (36) and (38) differ from (35) and (37), respectively, only because a divisor d_r is missing.

PROOF. By induction. For $r = 0$ only (37) and (38) are meaningful, and hold true in view of $d_0 = \mu_{0,s} = T_{0,s} = 1$ and of $h < C$. The induction consists in first proving that if (37) and (38) hold true up to $r - 1$ then (34), (35) and (36) are true for r; next proving that if (34), (35) and (36) hold true up to r then (37) and (38) are true for r.

Let $r > 0$ and put $r-1$ in place of r and r in place of s in (37) and (38). Recalling that only $H_r^{(r-1),\sharp} + H_r^{(r-1),\natural}$ is used in order to determine χ_r use the definition of the norm, the form of the solution of eq. (10) discussed in sect. 2, and the estimate (25). This immediately shows that (34), (35) and (36) are true for r provided (37) and (38) hold true for $r-1$. Coming to (37) and (38) and recalling the recursive definitions (11) there are only two kinds of terms to be estimated, namely $\frac{1}{s!} L^s_{\chi_r} Z_m$ for $1 \le m < r$ and $\frac{1}{p!} L^p_{\chi_r} H^{(r-1)}_{(s-p)r+m}$ for $0 \le p \le s$ and $0 \le m < r$. For, remarking that Z_r and $H_r^{(r-1)}$ are estimated by exactly the same quantity it is safe to estimate $\|H^{(r)}_{sr}\|$ by replacing Z_r with $H_r^{(r-1)}$ in the second of (11). This is tantamount to extending the sum in the second of (11) to $p = s-1$ and making it identical with the sum in the first of (11), with $m = 0$.

Denote $\varphi_s = L^s_{\chi_r} Z_m$, where $r > 1$, and split $\varphi_s = \varphi_s^\sharp + \varphi_s^\natural + \varphi_s^\flat$. I claim

$$\|\varphi_s\|_{1-\delta_r} \le s! \left(\tfrac{e}{d_r \Lambda}\right)^{2(s-1)} \|\chi_r\|^{s-1}_{1-\delta_{r-1}} \tfrac{D}{d_r}, \tag{40}$$

$$\|\varphi_s^\sharp + \varphi_s^\natural\|_{1-\delta_r} \le s! \left(\tfrac{2e}{d_r \Lambda}\right)^{2(s-1)} \|\chi_r\|^{s-1}_{1-\delta_{r-1}} D, \tag{41}$$

for $s \ge 1$, where

$$D = \mu_{r-1,r} \mu_{m-1,m} T_{r,r+m} C^{r+m-1} E \tag{42}$$

The proof proceeds by induction. Let $s = 1$. By the general estimate (26) one has

$$\|\varphi_1\|_{1-\delta_r} \le \frac{1}{d_r d_m \Lambda^2} \|\chi_r\|_{1-\delta_{r-1}} \|Z_m\|_{1-\delta_{m-1}}.$$

Using (34) and (35) one gets

$$\|\varphi_1\|_{1-\delta_r} \le \frac{1}{d_r} \mu_{m-1,m} \mu_{r-1,r} \frac{1}{d_{m-1} d_m} T_{m-1,m} T_{r-1,r} C^{r+m-2} \frac{E^2}{\gamma \Lambda^2},$$

so that (40) immediately follows from (31), (32) and (39).

Still keeping $s = 1$, (41) is obtained by remarking that the contributions to $\varphi_1^\sharp + \varphi_1^\natural$ come only from $L_{\chi_r}(Z_m^\sharp + Z_m^\natural)$ and $\big(L_{\chi_r^\natural} Z_m^\flat\big)^\natural$. Proceeding as above, from (28) and (27) one gets

$$\begin{aligned}\big\|\varphi_1^\sharp + \varphi_1^\natural\big\|_{1-\delta_r} \le{}& \frac{1}{\gamma d_m \Lambda^2}\big\|H_r^{(r-1),\sharp} + H_r^{(r-1),\natural}\big\|_{1-\delta_{r-1}}\big\|Z_m\big\|_{1-\delta_{m-1}}\\ &+\frac{4}{d_m\Lambda^2}\big\|\chi_r\big\|_{1-\delta_{r-1}}\big\|Z_m\big\|_{1-\delta_{m-1}}.\end{aligned}$$

Then (41) for $s = 1$ follows from (35), (38) for $r - 1$ and (39). Remark that the divisor d_r does not appear here.

Let now $s > 1$, and assume that (40) be true up to $s - 1$. Recalling that the divisor d_r due to the generalized Cauchy estimates is arbitrary, replace d_r with $\frac{s-1}{s} d_r$ in the estimates (40) and (41) for φ_{s-1}, thus getting

$$\begin{aligned}&\big\|\varphi_{s-1}\big\|_{1-\delta_r+d_r/s}\\ &\qquad\le (s-1)!\left(\frac{s}{s-1}\right)^{2s-3}\left(\frac{e}{d_r\Lambda}\right)^{2(s-2)}\|\chi_r\|_{1-\delta_{r-1}}^{s-2}\,\frac{D}{d_r},\\ &\big\|\varphi_{s-1}^\sharp + \varphi_{s-1}^\natural\big\|_{1-\delta_r+d_r/s}\\ &\qquad\le (s-1)!\left(\frac{s}{s-1}\right)^{2s-4}\left(\frac{2e}{d_r\Lambda}\right)^{2(s-2)}\|\chi_r\|_{1-\delta_{r-1}}^{s-2}\,D.\end{aligned}\tag{43}$$

Consider first the estimate (41). Remarking that the contributions to $\varphi_s^\sharp + \varphi_s^\natural$ come only from $L_{\chi_r}\big(\varphi_{s-1}^\sharp + \varphi_{s-1}^\natural\big)$ and $\big(L_{\chi_r^\natural}\varphi_{s-1}^\flat\big)^\natural$, use (26) and (27) to estimate

$$\begin{aligned}\big\|\varphi_s^\sharp + \varphi_s^\natural\big\|_{1-\delta_r} &\le \big\|L_{\chi_r}\big(\varphi_{s-1}^\sharp + \varphi_{s-1}^\natural\big)\big\|_{1-\delta_r} + \big\|\big(L_{\chi_r^\natural}\varphi_{s-1}^\flat\big)^\natural\big\|_{1-\delta_r}\\ &\le \frac{s}{d_r^2\Lambda^2}\big\|\chi_r\big\|_{1-\delta_{r-1}}\big\|\varphi_{s-1}^\sharp + \varphi_{s-1}^\natural\big\|_{1-\delta_r+d_r/s}\\ &\quad+\frac{2s}{d_r\Lambda^2}\big\|\chi_r\big\|_{1-\delta_{r-1}}\big\|\varphi_{s-1}\big\|_{1-\delta_r+d_r/s}\end{aligned}$$

Replacing (43) in the latter expression one gets

$$\left\|\varphi_s^\sharp+\varphi_s^\flat\right\|_{1-\delta_r}\leq\frac{s!}{d_r^2\Lambda^2}\left(\frac{s}{s-1}\right)^{2s-3}\left(\frac{2e}{d_r\Lambda}\right)^{2(s-2)}\left\|\chi_r\right\|_{1-\delta_{r-1}}^{s-1}D,$$

so that (41) follows from the trivial inequality $\left(\frac{s}{s-1}\right)^{s-1}<e$. The estimate (40) is checked with a similar calculation, just taking into account that (26) must be used in order to estimate $L_{\chi_r}\varphi_{s-1}$. This produces an extra divisor d_r with respect to the calculation above.

Finally, replace (34) and (42) in (40) and (41). Using also (39), one gets

$$\begin{aligned}\left\|\varphi_s\right\|_{1-\delta_r}&\leq s!\,\mu_{r-1,r}^{s-1}\left(\frac{1}{d_r^2}\,T_{r-1,r}\right)^{s-1}T_{r,r+m}C^{sr+m-1}\frac{E}{d_r}.\\ \left\|\varphi_s^\sharp+\varphi_s^\flat\right\|_{1-\delta_r}&\leq s!\,\mu_{r-1,r}^{s-1}\left(\frac{1}{d_r^2}\,T_{r-1,r}\right)^{s-1}T_{r,r+m}C^{sr+m-1}E.\end{aligned}$$

Using $s-1$ times the inequalities (31) and (32) one easily gets

$$\left(\frac{1}{d_r^2}\,T_{r-1,r}\right)^{s-1}T_{r,r+m}\leq\left(\frac{1}{d_r^2}\,T_{r-1,r}\right)^{s-2}T_{r,2r+m}\leq\ldots\leq T_{r,sr+m}.$$

Thus one concludes

$$\begin{aligned}&\frac{1}{s!}\left\|L_{\chi_r}^s Z_m\right\|_{1-\delta_r}\\ &\qquad\leq\mu_{r-1,r}^s\mu_{m-1,m}T_{r,sr+m}C^{sr+m-1}\frac{E}{d_r},\end{aligned}\tag{44}$$

$$\begin{aligned}&\frac{1}{s!}\left\|\left(L_{\chi_r}^s Z_m\right)^\sharp+\left(L_{\chi_r}^s Z_m\right)^\flat\right\|_{1-\delta_r}\\ &\qquad\leq\mu_{r-1,r}^s\mu_{m-1,m}T_{r,sr+m}C^{sr+m-1}E.\end{aligned}\tag{45}$$

The estimate for $\frac{1}{p!}L^p_{\chi_r}H^{(r-1)}_{(s-p)r+m}$ is a minor *variazione* of the scheme above. Only the first step must be omitted. E.g., set $\varphi_p = L^p_{\chi_r}H^{(r-1)}_{(s-p)r+m}$ and proceed as follows. Using (37) for $r-1$ get

$$\left\|\varphi_0\right\|_{1-\delta_{r-1}} \leq \mu_{r-1,sr+m}T_{r,sr+m}C^{sr+m-1}E;$$

this starts the induction on p. Then proceed for $p > 0$ as above. The conclusion is

$$\begin{aligned}
&\frac{1}{p!}\ \left\|L^p_{\chi_r}H^{(r-1)}_{(s-p)r+m}\right\|_{1-\delta_r} \leq \mu^p_{r-1,r}\mu_{r-1,(s-p)r+m}T_{r,sr+m}C^{sr+m-1}\frac{E}{d_r},\\
&\frac{1}{p!}\ \left\|\left(L^p_{\chi_r}H^{(r-1)}_{(s-p)r+m}\right)^\sharp + \left(L^p_{\chi_r}H^{(r-1)}_{(s-p)r+m}\right)^\natural\right\|_{1-\delta_r}\\
&\qquad\qquad \leq \mu^p_{r-1,r}\mu_{r-1,(s-p)r+m}T_{r,sr+m}C^{sr+m-1}E.
\end{aligned}$$

Collecting the latter equations together with (44) and (45), and referring to the transformation formulæ (11) it is now an easy matter to verify that (37) and (38) hold true provided the sequence $\mu_{r,s}$ for $0 < r < s$ is defined as

$$\begin{aligned}
\mu_{0,s} &= 1 && \text{for } s > 0,\\
\mu_{r,sr+m} &= \mu^s_{r-1,r}\mu_{m-1,m} + \sum_{p=0}^{s-1}\mu^p_{r-1,r}\mu_{r-1,(s-p)r+m}\\
&&& \text{for } r \geq 2,\ s \geq 1,\ 1 \leq m < r.\\
\mu_{r,sr} &= \sum_{p=0}^{s-1}\mu^p_{r-1,r}\mu_{r-1,(s-p)r} && \text{for } r \geq 1,\ s \geq 2.
\end{aligned} \tag{46}$$

This looks quite different from (33). However, I claim that (33) is just a harmless extension of (46). Indeed, just redefine the indexes by replacing $sr + m$ with s, also accepting $s \geq 0$, which removes the implicit restriction $s > r$. This is harmless, because for $s \leq r$ one gets for (33) $\mu_{r,s} = \mu_{r-1,s}$. Therefore, in the second line one can replace $\mu_{m-1,m} = \mu_{r-1,m}$ and include it into the sum. This completes the proof of lemma 3.4. ∎

3.4 Completion of the proof of lemma 3.1

The statement of the lemma concerns only the sequence of generating functions, that are estimated by (34). The completion of the proof rests on a suitable choice of the sequence $\{d_r\}_{r\geq 1}$, that was left arbitrary, and on a suitable estimate of the sequence $\{\mu_{r,s}\}_{s\geq r\geq 0}$. As a matter of fact, only the diagonal elements of the latter sequence need to be estimated, because the estimate for the generating functions in lemma 3.4 involves only $\mu_{r-1,r} = \mu_{r,r}$.

First, pick a value for d, with $0 < d < 1/2$, and set

$$d_r = \frac{b}{r^2}, \quad b = \frac{6d}{\pi^2},$$

so that $\sum_{r\geq 1} d_r = d$ in view of $\sum_{r\geq 1} 1/r^2 = \pi^2/6$. The immediate consequence is that

$$T_{r,s} \leq \left(\frac{16}{b^2}\right)^{s-1}. \tag{47}$$

For, use the definition (30) and recall the definition (29) of $\mathcal{J}_{r,s}$; then let $J \in \mathcal{J}_{r,s}$ and evaluate

$$\prod_{j\in J} \frac{1}{d_j} \leq \frac{1}{b^{2(s-1)}} \prod_{j\in J} j^2,$$

because $\#(J) \leq 2(s-1)$. On the other hand one has

$$\log_2 \prod_{j\in J} j^2 = 2\sum_{j\in J} \log_2 j \leq 4(s-1).$$

This proves (47)

Coming to the sequence (33), the problem is to show that $\mu_{r-1,r} \leq \eta^{r-1}$ for some positive η. For, only $\mu_{r-1,r}$ enters the estimate (34). By separating the term $p = 0$ in the sum one gets

$$\begin{aligned}\mu_{r,s} &= \mu_{r-1,s} + \mu_{r-1,r} \sum_{0\leq q<s-r} \mu_{r-1,r}^q \mu_{r-1,s-r-rq} \\ &= \mu_{r-1,s} + \mu_{r-1,r}\mu_{r,s-r}.\end{aligned}$$

Putting $r-1$ in place of r and r in place of s in the latter formula one gets

$$\begin{aligned}\mu_{r-1,r} &= \mu_{r-2,r} + \mu_{r-2,r-1}\mu_{r-1,1} \\ &= \mu_{r-3,r} + \mu_{r-3,r-2}\mu_{r-2,2} + \mu_{r-2,r-1}\mu_{r-1,1} \\ &\dots \\ &\leq \mu_{0,r} + \mu_{0,1}\mu_{1,r-1} + \dots \mu_{r-2,r-1}\mu_{r-1,1} \\ &\leq \mu_{0,1}\mu_{r-2,r-1} + \dots + \mu_{r-2,r-1}\mu_{0,1}.\end{aligned}$$

The last inequality requires a justification. Just use

$$\mu_{0,r} < \mu_{1,r} < \dots < \mu_{r-1,r} = \mu_{r,r} = \dots,$$

which is an immediate consequence of the definition. Then, for $r-j \geq j$ substitute $\mu_{r-j,j} = \mu_{j-1,j}$, and for $r-j < j$ use $\mu_{r-j,j} < \mu_{j-1,j}$, and so also $\mu_{r-j,j} + 1 \leq \mu_{j-1,j}$ which gets rid of the extra term $\mu_{0,r} = 1$.

Thus, the sequence $\{\nu_r\}_{r\geq 1}$ defined as

$$\nu_1 = 1, \quad \nu_r = \sum_{j=1}^{r-1} \nu_j \nu_{r-j} \quad \text{for } r > 1 \tag{48}$$

is a majorant of $\{\mu_{r-1,r}\}_{r\geq 1}$. This is known as the Catalan's sequence, and one has

$$\nu_r = \frac{2^{r-1}(2r-3)!!}{r!} \leq 4^{r-1}, \tag{49}$$

where the common notation $(2n+1)!! = 1 \cdot 3 \cdot \dots \cdot (2n+1)$ has been used.

Thus, we conclude that $\mu_{r-1,r} \leq 4^{r-1}$. Inserting the latter inequality and (47) in (34) the statement of lemma 3.1 follows.

4. Proof of theorem 1.1

Having established the estimate of lemma 3.1 on the sequence of generating functions it is now a standard matter to complete the proof of theorem 1.1. Hence this section will be less detailed with respect to the previous ones.

The situation to be dealt with is the following. An infinite sequence $\{\chi_r\}_{r\geq 1}$ of generating functions is given, with $\chi_r \in \mathcal{P}_{r+2}$ (a homogeneous polynomial of degree $r+2$) satisfying $\|\chi_r\|_R \leq \beta^{r-1}G$ for some real vector R with positive components and some positive β

and G. Define a corresponding sequence of canonical transformations $(x^{(r-1)}, y^{(r-1)}) = \exp(L_{\chi_r})(x^{(r)}, y^{(r)})$. By composition one also constructs a sequence $\{\mathcal{C}^{(r)}\}_{r\geq 0}$ of canonical transformations $(x^{(0)}, y^{(0)}) = \mathcal{C}^{(r)}(x^{(r)}, y^{(r)})$ recursively defined as

$$\mathcal{C}^{(0)} = \mathrm{Id}, \quad \mathcal{C}^{(r)} = \exp(L_{\chi_r}) \circ \mathcal{C}^{(r-1)},$$

Id denoting the identity operator. The problem is to prove the following statements.

(i) Every near the identity canonical transformation defined via the exponential operator $\exp(L_{\chi_r})$ is expressed as a power series which is convergent in a polydisk $\Delta_{\varrho R}$ for some positive ϱ.

(ii) For any function $f(x^{(r-1)}, y^{(r-1)})$ analytic in $\Delta_{\varrho R}$ the transformed function is analytic in the same polydisk, and moreover

$$f(x^{(r-1)}, y^{(r-1)})\Big|_{(x^{(r-1)},y^{(r-1)})=\exp(L_{\chi r})(x^{(r)},y^{(r)})} = \big[\exp(L_{\chi_r})f\big](x^{(r)}, y^{(r)}).$$

(iii) The sequence $\big\{\mathcal{C}^{(r)}\big\}_{r\geq 0}$ of canonical transformations converges for $r \to \infty$ to a canonical transformation $\mathcal{C}^{(\infty)}$ which is analytic in a polydisk $\Delta_{(1-d)\varrho R}$ for some positive $d < 1/2$.

(iv) For any function f analytic in $\Delta_{\varrho R}$ the sequence recursively defined as $f^{(0)} = f$, $f^{(r)} = \exp(L_{\chi_r})f^{(r-1)}$ converges for $r \to \infty$ to a function $f^{(\infty)}$ that is analytic in $\Delta_{(1-d)\varrho R}$, and moreover one has

$$f^{(\infty)} = f \circ \mathcal{C}^{(\infty)}.$$

The statement (i) actually reduces to Cauchy's proof of the existence and uniqueness of the local solution of an analytic system of differential equations. For, the transformation defined via the exponential operator is the time–one canonical flow induced by the Hamiltonian vector field generated by χ_r. The statement (ii) actually claims that the substitution of variables in a function f may be effectively replaced by the application of the exponential operator to f; this is indeed the basis of the algorithm for constructing the normal form used in sect. 2. A detailed proof of both

these statements may be found, e.g., in [9]; however, the reader may be able to reconstruct the proof by following the hints in [3].

The proof of (iii) rests on the following remarks. In the polydisk $\Delta_{\varrho R}$ one has $|\chi_r(x,y)| \le \varrho^{r+2}\,\|\chi_r\|_{\varrho R}$; this, in turn, implies that $\left|x^{(r)} - x^{(r-1)}\right| \sim \beta^{r-1}\varrho^{r+2}$ and $\left|y^{(r)} - y^{(r-1)}\right| \sim \beta^{r-1}\varrho^{r+2}$. The geometric bound on the latter quantities implies that $\sum_{r>1}\left|x^{(r)} - x^{(r-1)}\right|$ and $\sum_{r\ge1}\left|y^{(r)} - y^{(r-1)}\right|$ behave like geometric series, i.e., converge for ϱ small enough. Thus, the claim follows from Weierstrass theorem. Finally, the statement (iv) follows from (ii) being true for all $r > 0$, which implies that both sequences $f^{(r)} = \mathcal{C}^{(r)}f$ and $f \circ \mathcal{C}^{(r)}$ converge to the same limit. This concludes the proof of theorem 1.1.

APPENDIX A: JUSTIFICATION OF THE NORMALIZATION ALGORITHM

Justifying the normalization algorithm of sect. 2 is just matter of rearranging terms in the expansion of $\exp(L_{\chi_r})H^{(r-1)}$. Considering first H_0 and $H_r^{(r-1)}$ together, one has

$$\begin{aligned}\exp(L_{\chi_r})\left(H_0 + H_r^{(r-1)}\right) = H_0 + &L_{\chi_r}H_0 + \sum_{s\ge2}\frac{1}{s!}L_{\chi_r}^s H_0 \\ &+H_r^{(r-1)} + \sum_{s\ge1}\frac{1}{s!}L_{\chi_r}^s H_r^{(r-1)}.\end{aligned}$$

Here, H_0 is the first term in the transformed Hamiltonian $H^{(r)}$ in (9). In view of (10) one has $L_{\chi_r}H_0 + H_r^{(r-1)} = Z_r$, which kills the unwanted term $H_r^{(r-1)}$ and replaces it with the normalized term Z_r. The two sums may be collected and simplified by calculating

$$\begin{aligned}\sum_{s\ge2}\frac{1}{s!}L_{\chi_r}^s\, H_0 + \sum_{s\ge1}\frac{1}{s!}L_{\chi_r}^s H_r^{(r-1)} &\\ = \sum_{s\ge2}\frac{1}{(s-1)!}L_{\chi_r}^{s-1}\left[\frac{1}{s}\left(L_{\chi_r}H_0 + H_r^{(r-1)}\right) + \frac{s-1}{s}H_r^{(r-1)}\right] &\\ = \sum_{s\ge2}\frac{1}{(s-1)!}L_{\chi_r}^{s-1}\left(\frac{1}{s}Z_r + \frac{s-1}{s}H_r^{(r-1)}\right).&\end{aligned}$$

Here, both $L_{\chi_r}^{s-1} Z_r$ and $L_{\chi_r}^{s-1} H_r^{(r-1)}$ are homogeneous polynomials of degree $sr+2$, that are added to $H_{sr}^{(r)}$ in the second of (11).

Proceed now by transforming the functions $Z_1, \ldots, Z_{r-1}$ that are already in normal form. Recall that no such term exists for $r = 1$. For $r > 1$ calculate

$$\exp(L_{\chi_r}) Z_m = Z_m + \sum_{s \geq 1} \frac{1}{s!} L_{\chi_r}^s Z_m, \quad \text{for } 1 \leq m < r.$$

The term Z_m is copied into $H^{(r)}$ in (9). The term $L_{\chi_r}^s Z_m$ is a homogeneous polynomial of degree $sr + m + 2$ that is added to $H_{sr+m}^{(r)}$ in the first of (11).

Finally, consider all terms $H_s^{(r-1)}$ with $s > r$, that may be written as $H_{lr+m}^{(r-1)}$ with $l \geq 1$ and $0 \leq m < r$, the case $l = 1\,,\ m = 0$ being excluded. One gets

$$\exp(L_{\chi_r}) H_{lr+m}^{(r-1)} = \sum_{p \geq 0} \frac{1}{p!} L_{\chi_r}^p H_{lr+m}^{(r-1)}$$

where $L_{\chi_r}^p H_{lr+m}^{(r-1)}$ is a homogeneous polynomial of degree $(p+l)r+m$. Collecting all homogeneous terms with $m = 0$, $l \geq 2$ and $p + l = s \geq 2$ one gets $\sum_{p=0}^{s-2} \frac{1}{p!} L_{\chi_r}^p H_{(s-p)r}^{(r-1)}$, that is added to $H_{sr}^{(r)}$ in the second of (11). Similarly, collecting all homogeneous terms with $0 < m < r,\ l \geq 1$ and $p+l = s \geq 1$ one gets $\sum_{p=0}^{s-1} \frac{1}{p!} L_{\chi_r}^p H_{(s-p)r+m}^{(r-1)}$, that is added to $H_{sr+m}^{(r)}$ in the first of (11). The latter case does not occur for $r = 1$. This completes the justification of the formal algorithm.

APPENDIX B: TECHNICAL CALCULATIONS

The aim is to check the estimates (26), (27) and (28). Write, generically, $\chi = \sum_{j,k} c_{j,k} x^j y^k$ and $f = \sum_{j,k} f_{j,k} x^j y^k$. Then compute

$$L_\chi f = \sum_{j,k,j',k'} \sum_{l=1}^{n} \frac{j_l' k_l - j_l k_l'}{x_l y_l} c_{j,k} f_{j',k'} x^{j+j'} y^{k+k'}. \tag{50}$$

Using the definition of norm evaluate

$$\begin{aligned}
&\|L_\chi f\|_{1-\delta} \\
&\le \sum_{j,k,j',k'} \sum_{l=1}^{n} \frac{|j'_l k_l - j_l k'_l|}{R_l^2} |c_{j,k}|\, |f_{j',k'}| (1-\delta)^{|j+k|+|j'+k'|-2} R^{j+k+j'+k'} \\
&\le \frac{1}{\Lambda^2} \sum_{j,k} \sum_{j',k'} \sum_{l=1}^{n} \left|j'_l k_l - j_l k'_l\right|\, |c_{j,k}| \left((1-\delta') - (\delta-\delta')\right)^{|j+k|-1} R^{j+k} \\
&\quad \times |f_{j',k'}| \left((1-\delta'') - (\delta-\delta'')\right)^{|j'+k'|-1} R^{j'+k'}.
\end{aligned}$$

If f is a generic function, then in view of $|j_l| \le |j+k|$ and $k_l \le |j+k|$ one has

$$\sum_{l=1}^{n} |j'_l k_l - j_l k'_l| < |j+k| \sum_{l=1}^{n} |j'_l + k'_l| = |j+k| \cdot |j'+k'|. \tag{51}$$

Replacing in the estimate above and using the elementary inequality

$$m(\lambda - x)^{m-1} < \frac{\lambda^m}{x} \quad \text{for } 0 < x < \lambda \text{ and } m \ge 1 \tag{52}$$

one gets

$$\begin{aligned}
\|L_\chi f\|_{1-\delta} &\le \frac{1}{\Lambda^2} \sum_{j,k} |c_{j,k}|\, |j+k| \left((1-\delta') - (\delta-\delta')\right)^{|j+k|-1} R^{j+k} \\
&\quad \times \sum_{j',k'} |f_{j',k'}|\, |j'+k'| \left((1-\delta'') - (\delta-\delta'')\right)^{|j'+k'|-1} R^{j'+k'} \\
&\le \frac{1}{(\delta-\delta')(\delta-\delta'')\Lambda^2} \sum_{j,k} |c_{j,k}|\, (1-\delta')^{|j+k|} R^{j+k} \\
&\quad \times \sum_{j',k'} |f_{j',k'}| (1-\delta'')^{|j'+k'|} R^{j'+k'},
\end{aligned} \tag{53}$$

from which (26) immediately follows in view of the definition of the norm.

In order to prove (27) recall that $\chi^\natural$ contains only monomials $c_{j,k}x^j y^k$ with $j+k \in \mathcal{K}^\natural$. The projection $\left(L_{\chi^\natural} f^\flat\right)^\natural$ is just part of the general expression (50). In particular the value $l=1$ in the sum must be discarded because the resulting monomials belong to $\mathcal{P}^\flat$. Moreover, for $k \in \mathcal{K}^\natural$ one has $\sum_{l=2}^n |j_l + k_l| \leq 2$. Thus the estimate (51) may be replaced by

$$\sum_{l=2}^n |j'_l k_l - j_l k'_l| < 2 \sum_{l=2}^n |j'_l + k'_l| = 2|j' + k'|. \tag{54}$$

Hence the inequality (52) must be used only for the term involving $|j'+k'|$, and there is no need to introduce the divisor $\delta - \delta'$. Use instead $1/(1-\delta) < 2$ in view of $\delta < 1/2$.

Coming finally to (28), replace f in the general expression (50) by $Z^\sharp + Z^\natural = \sum_{\nu \in \mathcal{K}^\sharp \cup \mathcal{K}^\sharp} z_{\nu,\nu} x^\nu y^\nu$. Recall also that the coefficients $c_{j,k}$ of χ have the form $c_{j,k} = \frac{\psi_{j,k}}{\langle k-j,\lambda\rangle}$, in view of (13). Then (51) may be replaced by

$$\sum_l |\nu_l (j_l - k_l)| \leq |\nu| \sum_l |j_l - k_l| \leq |\nu|\,|j-k| \tag{55}$$

On the other hand, by lemma 3.2 one has

$$|c_{j,k}| \leq \frac{|\psi_{j,k}|}{|j-k|\gamma},$$

so that the factor $|j-k|$ in (55) is compensated by the divisor here. This removes the need to introduce the divisor $\delta - \delta'$ in the rest of the estimates. Use instead $(1-\delta)^{|j+k|-1} \leq 2(1-\delta)^{|j+k|}$, which holds true

in view of $\delta < 1/2$. Then (53) is replaced by

$$\begin{aligned}
&\|L_\chi(Z^\sharp + Z^\natural)\|_{1-\delta} \\
&\quad \le \frac{1}{\Lambda^2}\sum_{j,k}\frac{2}{\gamma}|\psi_{j,k}|\,(1-\delta)^{|j+k|}R^{j+k} \\
&\qquad \times \sum_{\nu\in\mathcal{K}^\sharp}|z_{\nu,\nu}|\,|\nu|\big((1-\delta'')-(\delta-\delta'')\big)^{2|\nu|-1}R^{2\nu} \\
&\quad \le \frac{1}{(\delta-\delta'')\gamma\Lambda^2}\sum_{j,k}|\psi_{j,k}|\,(1-\delta')^{|j+k|}R^{j+k} \\
&\qquad \times \sum_{\nu\in\mathcal{K}^\sharp}|z_{\nu,\nu}|(1-\delta'')^{2|\nu|}R^{2\nu}.
\end{aligned}$$

Thus (28) follows in view of the definition of the norm.

REFERENCES

[1] G.D. Birkhoff: *Dynamical systems*, New York (1927).

[2] T.M. Cherry: *On the solutions of Hamiltonian systems in the neighborhood of a singular point*, Proc. London Math. Soc., Ser. 2, **27**, 151–170 (1926).

[3] A. Giorgilli: *Quantitative methods in classical perturbation theory*, proceedings of the Nato ASI school "From Newton to chaos: modern techniques for understanding and coping with chaos in N–body dynamical systems", A.E. Roy e B.D. Steves eds., Plenum Press, New York (1995).

[4] A. Giorgilli: *Classical constructive methods in KAM theory*, PSS, **46**, 1441–1451 (1998).

[5] A. Giorgilli: *Unstable equilibria of Hamiltonian systems*, Disc. and Cont. Dynamical Systems, Vol. 7, N. 4, 855–871 (2001).

[6] A. Giorgilli and U. Locatelli: *Kolmogorov theorem and classical perturbation theory*, ZAMP **48**, 220–261 (1997).

[7] A. Giorgilli and U. Locatelli: *On classical series expansions for quasi–periodic motions*, MPEJ **3** N. 5 (1997).

[8] A. Giorgilli and U. Locatelli: *A classical self–contained proof of Kolmogorov's theorem on invariant tori*, in *Hamiltonian systems with three or more degrees of freedom*, Carles Simó ed., NATO ASI series C, Vol. 533, Kluwer Academic Publishers, Dordrecht–Boston–London (1999).

[9] W. Gröbner: *Die Lie–Reihen und Ihre Anwendungen*, VEB Deutscher Verlag der Wissenschaften (1967).

[10] A.M. Lyapunov: *The General Problem of the Stability of Motion* (In Russian), Doctoral dissertation, Univ. Kharkov (1892). French translation in: *Problème général de la stabilité du mouvement*, Annales de la Faculté des Sciences de Toulouse, deuxième série, Tome IX, 203–474 (1907). Reprinted in: Ann. Math. Study, Princeton University Press, n. 17, (1949).

[11] J.K. Moser: *On the generalization of a theorem of A. Liapounoff*, Comm. Pure Appl. Math. **11**, 257–271 (1958).

[12] C.L. Siegel and J.K. Moser: *Lectures in Celestial Mechanics*, Springer–Verlag, Berlin Heidelberg New York (1971).