

UNSTABLE EQUILIBRIA OF HAMILTONIAN SYSTEMS

ANTONIO GIORGILLI

*Dipartimento di Matematica e Applicazioni, Via degli Arcimboldi 8,
20126 — Milano, Italy.*

Abstract. It is shown that a Hamiltonian system in the neighbourhood of an equilibrium may be given a special normal form in case four of the eigenvalues of the linearized system are of the form $\lambda_1, -\lambda_1, \lambda_2, -\lambda_2$, with λ_1 and λ_2 independent over the reals, i.e., $\lambda_1/\lambda_2 \notin \mathbf{R}$. That is, for a real Hamiltonian system and concerning the variables x_1, y_1, x_2, y_2 the equilibrium is of either type *center–saddle* or *complex–saddle*. The normal form exhibits the existence of a four–parameter family of solutions which has been previously investigated by Moser. This paper completes Moser’s result in that the convergence of the transformation of the Hamiltonian to a normal form is proven.

1. Introduction

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

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

where the unperturbed quadratic part of the Hamiltonian is

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

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], § 15.

The Hamiltonian is assumed to be analytic in some neighborhood of the origin of \mathbf{C}^{2n} . Moreover λ_1, λ_2 will be assumed to be independent over the reals, and to satisfy the non–resonance condition $\lambda_\nu \neq k_1 \lambda_1 + k_2 \lambda_2$ for all integers k_1, k_2 and $\nu \geq 3$. Independence over the reals of λ_1 and λ_2 means that the ratio λ_1/λ_2 is not a real number. Concerning the variables x_1, y_1, x_2, y_2 this covers the case of a real Hamiltonian system with an equilibrium of either type *center–saddle* or *complex–saddle*.

In a 1958 paper^[9] Moser proved that under the conditions above there exists a four parameter family of solutions of the form

$$(3) \quad x_j = \varphi_j(\xi_1, \xi_2, \eta_1, \eta_2), \quad y_j = \psi_j(\xi_1, \xi_2, \eta_1, \eta_2)$$

written as convergent power series in the arguments

$$(4) \quad \xi_l = \overset{\circ}{\xi}_l e^{ta_l(\overset{\circ}{\zeta})}, \quad \eta_l = \overset{\circ}{\eta}_l e^{-ta_l(\overset{\circ}{\zeta})}, \quad l = 1, 2,$$

where $a_1(\overset{\circ}{\zeta}) = \lambda_1 + \dots$, $a_2(\overset{\circ}{\zeta}) = \lambda_2 + \dots$ are convergent power series in $\overset{\circ}{\zeta}_l = \overset{\circ}{\xi}_l \overset{\circ}{\eta}_l$. In the case $n = 2$ this actually describes all solutions of the system. The case $n = 2$ had been previously investigated by Cherry^[2]. However, according to Moser, Cherry's paper is not error free.

From a formal viewpoint the statement above looks equivalent to the existence of a canonical transformation that gives the system (1) the normal form

$$(5) \quad \Gamma(\xi, \eta) = Z(\zeta_1, \zeta_2) + R(\xi, \eta),$$

where $Z(\zeta_1, \zeta_2) = \lambda_1 \xi_1 \eta_1 + \lambda_2 \xi_2 \eta_2 + \dots$ is a power series, $\zeta_1 = \xi_1 \eta_1$, $\zeta_2 = \xi_2 \eta_2$, and $R(\xi, \eta)$ is at least quadratic in $\xi_3, \dots, \xi_n, \eta_3, \dots, \eta_n$. For, if one sets $a_l = \frac{\partial Z}{\partial \zeta_l}$ then (4) together with $\xi_3 = \dots = \xi_n = \eta_3 = \dots = \eta_n = 0$ is a four parameter family of solutions for the system (5), which via the canonical transformation produces a corresponding family of solutions of the form (3) for the system (1).

However, two remarks are in order. The first remark is that the formal algorithm leading to the power series (3) does not use the canonical structure of the equations, and so it has no direct relation with the existence of a canonical transformation reducing the Hamiltonian to a normal form. The second remark is that neither the series (3) nor the canonical transformation are uniquely determined: a full discussion concerning the relations among different determinations is reported in Moser's paper. What Moser actually proves is that there exists a *particular* determination of the series in (3) which is convergent in a neighbourhood of the origin.

It seems then interesting to investigate whether a convergent normal form (5) may be constructed via a convergent canonical transformation. This question is proposed, indeed, in Moser's paper. However, his conclusion is that to establish directly the convergence of the generating function W of the canonical transformation seems to be impossible, since W occurs in the argument of the unknown function Γ .

The aim of this paper is precisely to produce a proof of convergence of the transformation to normal form:

Theorem 1: *With the hypotheses above 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 (5).*

The proof is worked out by using the composition of Lie series as a basic tool for constructing canonical transformation (see, e.g., [3]). This makes the algorithm explicit, thus avoiding the difficulty pointed out by Moser. 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 here 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^{[4][5][6][7]}.

2. Formal algorithm

The algorithm based on composition of Lie series may be stated as follows. Write the Hamiltonian after r normalization steps as

$$(6) \quad H^{(r)}(x, y) = H_0(x, y) + Z_1(x, y) + \dots + Z_r(x, y) + \sum_{s>r} H_s^{(r)}(x, y) ,$$

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

Assume that the Hamiltonian has been given a normal form (6) 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

$$(7) \quad L_{H_0} \chi_r + Z_r = H_r^{(r-1)} .$$

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

$$(8) \quad H_{sr+m}^{(r)} = \frac{1}{s!} L_{\chi_r}^s Z_m + \sum_{p=0}^{s-1} \frac{1}{p!} L_{\chi_r}^p H_{(s-p)r+m}^{(r-1)} \quad \text{for } r \geq 2, s \geq 1 \text{ and } 1 \leq m < r ,$$

$$H_{sr}^{(r)} = \frac{1}{(s-1)!} L_{\chi_r}^{s-1} \left(\frac{1}{s} Z_r + \frac{s-1}{s} H_r^{(r-1)} \right) + \sum_{p=0}^{s-2} \frac{1}{p!} L_{\chi_r}^p H_{(s-p)r}^{(r-1)}$$

for $r \geq 1$ and $s \geq 2$.

Justifying this algorithm 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

$$\begin{aligned} \exp(L_{\chi_r})(H_0 + H_r^{(r-1)}) &= H_0 + L_{\chi_r} H_0 + \sum_{s \geq 2} \frac{1}{s!} L_{\chi_r}^s H_0 \\ &\quad + H_r^{(r-1)} + \sum_{s \geq 1} \frac{1}{s!} L_{\chi_r}^s H_r^{(r-1)} . \end{aligned}$$

Here, H_0 is the first term in the transformed Hamiltonian $H^{(r)}$ in (6). In view of (7) 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

$$\begin{aligned} &\sum_{s \geq 2} \frac{1}{s!} L_{\chi_r}^s H_0 + \sum_{s \geq 1} \frac{1}{s!} L_{\chi_r}^s H_r^{(r-1)} \\ &= \sum_{s \geq 2} \frac{1}{(s-1)!} L_{\chi_r}^{s-1} \left[\frac{1}{s} \left(L_{\chi_r} H_0 + H_r^{(r-1)} \right) + \frac{s-1}{s} H_r^{(r-1)} \right] \\ &= \sum_{s \geq 2} \frac{1}{(s-1)!} L_{\chi_r}^{s-1} \left(\frac{1}{s} Z_r + \frac{s-1}{s} H_r^{(r-1)} \right) . \end{aligned}$$

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 (8).

Proceed now by transforming the functions Z_1, \dots, 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 (6). 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 (8).

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 (8). 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 (8). The latter case does not occur for $r=1$. This completes the justification of the formal algorithm.

Thus, the problem is how to solve the equation (7) for the generating function and the normal form. A suitable method is the following. 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 $x^j y^k := x_1^{j_1} \cdot \dots \cdot x_n^{j_n} y_1^{k_1} \cdot \dots \cdot y_n^{k_n}$, 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 $\mathcal{N} = L_{H_0}^{-1}(0)$, 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 a unique solution χ satisfying the condition $\chi \in \mathcal{R}$. That unique solution will be denoted by $\chi = L_{H_0}^{-1} \psi$. 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

$$(9) \quad \begin{aligned} \mathcal{N} &= \text{span} \{ x^j y^k : \langle j - k, \lambda \rangle = 0 \}, \\ \mathcal{R} &= \text{span} \{ x^j y^k : \langle j - k, \lambda \rangle \neq 0 \}. \end{aligned}$$

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

$$(10) \quad L_{H_0}^{-1} \psi = \sum_{j,k} \frac{\psi_{j,k}}{\langle j-k, \lambda \rangle} x^j y^k .$$

The case $n = 2$ is rather simple. A function $Z \in \mathcal{P}$ is said to be in normal form if $Z \in \mathcal{N}$, i.e., if $L_{H_0} Z = 0$. This is the well known Birkhoff's normal form (see [1]). The case $n > 2$ together with the particular requirements for the normal form in sect. 1 requires a more complicated setting. Consider the disjoint subsets of \mathbf{Z}^n

$$(11) \quad \begin{aligned} \mathcal{K}^\sharp &= \{k \in \mathbf{Z}^n : k_3 = \dots = k_n = 0\} , \\ \mathcal{K}^\natural &= \{k \in \mathbf{Z}^n : |k_3| + \dots + |k_n| = 1\} , \\ \mathcal{K}^\flat &= \{k \in \mathbf{Z}^n : |k_3| + \dots + |k_n| > 1\} . \end{aligned}$$

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

$$(12) \quad \begin{aligned} \mathcal{P}^\sharp &= \text{span} \{x^j y^k : j+k \in \mathcal{K}^\sharp\} \\ \mathcal{P}^\natural &= \text{span} \{x^j y^k : j+k \in \mathcal{K}^\natural\} \\ \mathcal{P}^\flat &= \text{span} \{x^j y^k : j+k \in \mathcal{K}^\flat\} \end{aligned}$$

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

$$(13) \quad \mathcal{Z} = \mathcal{N}^\sharp \oplus \mathcal{P}^\flat , \quad \mathcal{W} = \mathcal{R}^\sharp \oplus \mathcal{P}^\natural .$$

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 hypotheses on λ formulated at the beginning of the introduction mean that the non-resonance condition

$$(14) \quad \langle k, \lambda \rangle \neq 0 \quad \text{for } 0 \neq k \in \mathcal{K}^\sharp \cup \mathcal{K}^\natural$$

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

$$(15) \quad L_{H_0} \chi + Z = \Psi ,$$

with Ψ known, admits a simple 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 (10).

3. Estimates of the generating functions

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

$$(16) \quad \Delta_R = \{(x, y) \in \mathbf{C}^n : |x_j| \leq R_j, |y_j| \leq R_j, 1 \leq j \leq n\} ,$$

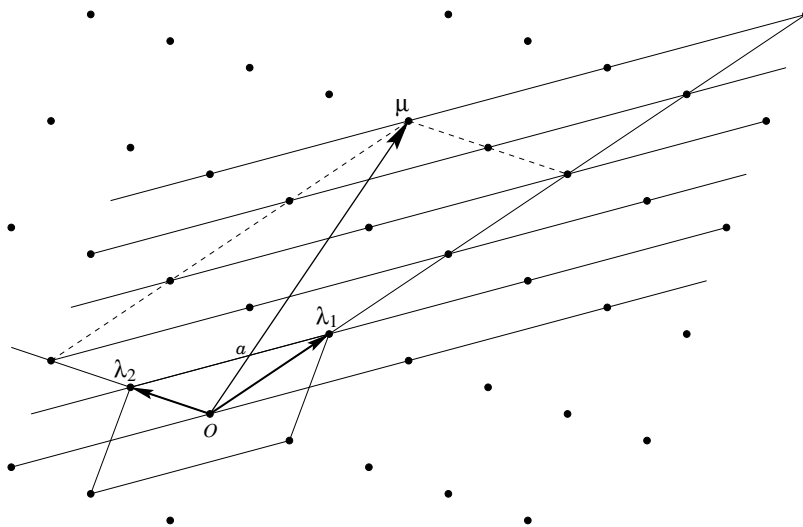


Figure 1. Illustrating the proof of lemma ???. The segment from the origin to $\mu = 3\lambda_1 + 2\lambda_2$ intersects the perimeter of the parallelogram with vertices in $\lambda_1, \lambda_2, -\lambda_1, -\lambda_2$ at the point a . Using Thales theorem it is immediately seen that $\mu = 5a$.

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

$$(17) \quad \Lambda = \min_{1 \leq j \leq n} R_j .$$

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

$$(18) \quad \|f\|_R = \sum_{|j+k|=r} |f_{j,k}| R^{j+k} .$$

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 1: *Let the Hamiltonian $H^{(0)}$ satisfy $\|H_s^{(0)}\|_1 \leq h^{s-1}E$ 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 $\lambda_1, \dots, \lambda_n$ such that*

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

The rest of this section contains all the technicalities that contribute to the proof.

3.1 An arithmetic lemma

The following lemma will play a crucial role in the proof of lemma 1.

Lemma 2: *Let $\lambda \in \mathbf{C}^n$ be such that λ_1, λ_2 are independent over the reals, and the non-resonance condition (14) be satisfied. Then there exists a positive γ such that the inequality*

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

holds true for all non-zero $k \in \mathcal{K}^\sharp \cup \mathcal{K}^\natural$.

Proof. Let first $n = 2$. Since λ_1, λ_2 are independent over the reals the set of points $k_1\lambda_1 + k_2\lambda_2$ with k_1, k_2 arbitrary integers is a discrete lattice in the complex plane (see fig. 1). Consider the parallelogram with vertices in $\lambda_1, \lambda_2, -\lambda_1, -\lambda_2$. Let $\mu = k_1\lambda_1 + k_2\lambda_2$ be any point of the lattice different from the origin, i.e., $|k| \neq 0$, and let a be the intersection of the segment from the origin to μ with the perimeter of the parallelogram. Such a point exists because μ can not belong to the interior of the parallelogram, by construction. Then $|\mu| = (|k_1| + |k_2|)|a|$, by Thales theorem. Taking δ as the distance of the perimeter of the parallelogram from the origin the inequality $|k_1\lambda_1 + k_2\lambda_2| \geq (|k_1| + |k_2|)\delta$ holds true for all integers k_1, k_2 .

Let now $n > 2$. Set $\vartheta = \max(|\lambda_3|, \dots, |\lambda_n|)$. Pick an integer $N \geq 1 + 2\vartheta/\delta$ and set

$$\delta' = \min_{\substack{k \in \mathcal{K}^\natural \\ |k| \leq N}} |\langle k, \lambda \rangle| ;$$

in view of the non-resonance condition (14) one has $\delta' > 0$. Then the claim of the lemma holds true with, e.g.,

$$\gamma = \min \left(\frac{\delta'}{N}, \frac{\delta}{2} \right) .$$

Indeed, if $k \in \mathcal{K}^\sharp$ one has just to consider the case $n = 2$, with no further comment. So, let $k \in \mathcal{K}^\natural$, so that $|k_1| + |k_2| = |k| - 1$. If $|k| \leq N$ then $|\langle k, \lambda \rangle| \geq \delta' \geq N\gamma \geq |k|\gamma$. If $|k| > N$ use $\vartheta \leq (N - 1)\delta/2$, which follows from the choice of N , and evaluate

$$\begin{aligned} |\langle k, \lambda \rangle| &\geq |k_1\lambda_1 + k_2\lambda_2| - \vartheta \geq (|k| - 1)\delta - \frac{(N - 1)}{2} \delta \\ &\geq \frac{|k| - 1}{2} \delta + \frac{N}{2} \delta - \frac{(N - 1)}{2} \delta = |k| \frac{\delta}{2} \geq |k|\gamma . \end{aligned}$$

Q.E.D.

3.2 Generalized Cauchy estimates

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}^b . 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}^b
\mathcal{P}^\sharp	\mathcal{P}^\sharp	\mathcal{P}^\natural	\mathcal{P}^b
\mathcal{P}^\natural	\mathcal{P}^\natural	$\mathcal{P}^\sharp \oplus \mathcal{P}^b$	$\mathcal{P}^\natural \oplus \mathcal{P}^b$
\mathcal{P}^b	\mathcal{P}^b	$\mathcal{P}^\natural \oplus \mathcal{P}^b$	\mathcal{P}^b

The table is meaningful only for $n > 2$. For $n = 2$ all functions belong to \mathcal{P}^\sharp .

In view of the transformation formulæ (8) 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 $\mathcal{W} = \mathcal{R}^\sharp \oplus \mathcal{P}^\natural$ one has $\chi = \chi^\sharp + \chi^\natural$, with an obvious meaning of the notation. The operator L_χ may be applied to either a generic function $f = f^\sharp + f^\natural + f^b \in \mathcal{P}_s$ with $s \geq r$ or a function in normal form $Z = Z^\sharp + Z^b \in \mathcal{Z} \cap \mathcal{P}_m$ with $0 < m < r$. In particular one has $Z^\sharp = \sum_{\nu \in \mathcal{K}^\sharp} z_{\nu,\nu} x^\nu y^\nu$, due to the non-resonance condition on λ . For some non-negative d', d'', d satisfying $0 \leq \max(d', d'') < d \leq 1/2$ the norms $\|\psi\|_{1-d'}$, $\|f\|_{1-d''}$ and $\|Z\|_{1-d''}$ are assumed to be known, and one looks for an estimate of the Lie derivative in a domain $\Delta_{(1-d)R}$. The following estimates will be used in the rest of the paper.

(i) The generating function χ is estimated by

$$(20) \quad \|\chi\|_{1-d'} \leq \frac{1}{\gamma} \|\psi\|_{1-d'} \quad ,$$

with γ as in lemma 2.

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

$$(21) \quad \|L_\chi f\|_{1-d} \leq \frac{1}{(d-d')(d-d'')\Lambda^2} \|\chi\|_{1-d'} \|f\|_{1-d''}$$

with Λ as in (17). Denoting by $(L_{\chi^\natural} f^b)^\natural$ the projection of $L_{\chi^\natural} f^b$ over \mathcal{P}^\natural one has

$$(22) \quad \|(L_{\chi^\natural} f^b)^\natural\|_{1-d} \leq \frac{2}{(d-d'')\Lambda^2} \|\chi\|_{1-d'} \|f\|_{1-d''} \quad .$$

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

$$(23) \quad \|L_\chi Z^\sharp\|_{1-d} \leq \frac{1}{(d-d'')\gamma\Lambda^2} \|\psi\|_{1-d'} \|Z\|_{1-d''} \quad .$$

The missing denominator $d-d'$ in (22) and (23) is crucial for the convergence proof.

The proof of (20) is a straightforward consequence of the definition of the norm and of (10). For, the denominators are uniformly estimated from below by γ , in view of lemma 2.

Coming to the estimates (21), (22) and (23), 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

$$(24) \quad L_\chi f = \sum_{j,k,j',k'} \sum_{l=1}^n \frac{j_l' k_l - j_l k_l'}{x_l y_l} c_{j,k} f_{j',k'} x^{j+j'} y^{k+k'} \quad .$$

Using the definition of norm evaluate

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

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

$$(25) \quad \sum_{l=1}^n |j'_l k_l - j_l k'_l| < |j+k| \sum_{l=1}^n |j'_l + k'_l| = |j+k| \cdot |j' + k'| .$$

Replacing in the estimate above and using the elementary inequality[†]

$$(26) \quad m(\lambda - x)^{m-1} < \frac{\lambda^m}{x} \quad \text{for } 0 < x < \lambda \text{ and } m \geq 1$$

one gets

$$(27) \quad \begin{aligned} \|L_\chi f\|_{1-d} &\leq \frac{1}{\Lambda^2} \sum_{j,k} |c_{j,k}| |j+k| ((1-d') - (d-d'))^{|j+k|-1} R^{j+k} \\ &\quad \times \sum_{j',k'} |f_{j',k'}| |j'+k'| ((1-d'') - (d-d''))^{|j'+k'|-1} R^{j'+k'} \\ &\leq \frac{1}{(d-d')(d-d'')\Lambda^2} \sum_{j,k} |c_{j,k}| (1-d')^{|j+k|} R^{j+k} \\ &\quad \times \sum_{j',k'} |f_{j',k'}| (1-d'')^{|j'+k'|} R^{j'+k'} , \end{aligned}$$

from which (21) immediately follows in view of the definition of the norm.

In order to prove (22) remark that χ^\natural contains only monomials $c_{j,k} x^j y^k$ with $j+k \in \mathcal{K}^\natural$. Moreover, the projection $(L_{\chi^\natural} f^\natural)^\natural$ is just part of the general expression (24). In particular all terms of the Poisson bracket containing derivatives with respect to the variables x_1, x_2, y_2, y_2 must be discarded, because the resulting monomials belong to \mathcal{P}^b . Moreover, exactly one of $j_3, k_3, \dots, j_n, k_n$ has the value 1, which implies $\sum_{l=3}^n |j_l + k_l| = 1$. This means that the estimate (25) may be replaced by

$$(28) \quad \sum_{l=3}^n |j'_l k_l - j_l k'_l| < \sum_{l=1}^n |j'_l + k'_l| = |j' + k'| .$$

Hence the inequality (26) must be used only for the term involving $|j' + k'|$, while there is no need to introduce the divisor $d - d'$. Use instead $1/(1-d) < 2$ in view of $d < 1/2$.

Coming finally to (23), replace f in the general expression (24) by $Z^\sharp = \sum_{\nu \in \mathcal{K}^\sharp} z_{\nu,\nu} x^\nu y^\nu$. 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 (10). Then (25) may be replaced by

$$(29) \quad \sum_{l=1,2} |\nu_l(j_l - k_l)| \leq |\nu| \sum_{l=1,2} |j_l - k_l| \leq |\nu| |j - k|$$

[†] The function $x(\lambda - x)^{m-1}$ in the interval $0 \leq x \leq \lambda$ has a maximum for $x = \lambda/m$, and the inequality follows from $x(\lambda - x)^{m-1} \leq \frac{\lambda}{m} \left(\lambda - \frac{\lambda}{m}\right)^{m-1} < \frac{\lambda^m}{m}$.

On the other hand, by lemma 2 one has

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

so that the factor $|j-k|$ in (29) is compensated by the divisor here. This removes the need to introduce the divisor $d-d'$ in the rest of the estimates. Use instead $(1-d)^{|j+k|-1} \leq 2(1-d)^{|j+k|}$, which holds true in view of $d < 1/2$. Then (27) is replaced by

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

Checking (23) is now straightforward matter.

Working out the convergence proof requires a quite accurate control of the accumulation of the divisors $d-d'$, $d-d''$ 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.

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 the elements of which are positive integers not exceeding r , the length of the array being at most $2(s-1)$, and satisfying the *selection rule* that the sum of the base 2 logarithms of the array's elements does not exceed $2(s-1 - \log_2 s)$. Formally:

$$(30) \quad \mathcal{J}_{r,s} = \left\{ J = \{j_1, \dots, j_k\} : j_m \in \{1, \dots, r\}, 1 \leq k \leq 2(s-1), \right. \\ \left. \sum_{m=1}^k \log_2 j_m \leq 2(s-1 - \log_2 s) \right\}.$$

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

$$(31) \quad T_{0,s} = 1, \quad T_{r,s} = \max_{J \in \mathcal{J}_{r,s}} \prod_{j \in J} d_j^{-1}.$$

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

$$(32) \quad T_{r,s} \leq T_{r',s} ,$$

$$(33) \quad \frac{1}{d_r^2} T_{r-1,r} T_{r',s} \leq T_{r',r+s} .$$

Checking (32) is easy: for $r = 0$ use $d_l \leq 1$ for $l \geq 1$; for $r > 0$ use the inclusion relation $\mathcal{J}_{r,s} \subset \mathcal{J}_{r',s}$ for $r < r'$. In order to prove (33) 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 $\{r,r\} \cup J \cup J' =: \tilde{J} \in \mathcal{J}_{r',r+s}$. Now, the number of elements of \tilde{J} satisfies

$$\#(\tilde{J}) = 2 + \#(J) + \#(J') \leq 2 + 2(r-1) + 2(s-1) = 2(r+s-1) .$$

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

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

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 $r \leq r' < s$). Hence, $\tilde{J} \in \mathcal{J}_{r',r+s}$, as claimed.

Finally, define a numerical sequence $\{\mu_{r,s}\}_{r \geq 0, s \geq 0}$ as follows:

$$(34) \quad \begin{aligned} \mu_{0,0} &= 0, \quad \mu_{0,s} = 1 \quad \text{for } s > 0, \\ \mu_{r,sr+m} &= \sum_{p=0}^s \mu_{r-1,r}^p \mu_{r-1,(s-p)r+m} \quad \text{for } s \geq 0, 0 \leq m < r . \end{aligned}$$

The recursive estimates are collected in

Lemma 3: *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 $\delta_r = d_1 + \dots + d_r$. Then for $s > r \geq 1$ the following estimates hold true:*

$$(35) \quad \|\chi_r\|_{1-\delta_{r-1}} \leq \mu_{r-1,r} T_{r-1,r} C^{r-1} \frac{E}{\gamma} ,$$

$$(36) \quad \|Z_r\|_{1-\delta_{r-1}} \leq \mu_{r-1,r} T_{r-1,r} C^{r-1} \frac{E}{d_{r-1}} ,$$

$$(37) \quad \|Z_r^\sharp\|_{1-\delta_{r-1}} \leq \mu_{r-1,r} T_{r-1,r} C^{r-1} E ,$$

$$(38) \quad \|H_s^{(r)}\|_{1-\delta_r} \leq \mu_{r,s} T_{r,s} C^{s-1} \frac{E}{d_r} ,$$

$$(39) \quad \|H_s^{(r),\sharp} + H_s^{(r),\natural}\|_{1-\delta_r} \leq \mu_{r,s} T_{r,s} C^{s-1} E ,$$

where

$$(40) \quad C = h + \frac{4e^2 E}{\gamma \Lambda^2} ,$$

and $\mu_{r,s}$ and $T_{r,s}$ are the sequences defined above.

Remark that (35), (37) and (39) differ from (36) and (38), respectively, only because a divisor d_r is missing. This is the key of the convergence proof.[‡]

Proof. By induction. For $r = 0$ only (38) and (39) 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 (38) and (39) hold true up to $r - 1$ then (35), (36) and (37) are true for r ; next proving that if (35), (36) and (37) hold true up to r then (38) and (39) are true for r .

Let $r > 0$ and put $r - 1$ in place of r and r in place of s in (38) and (39). 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. (7) discussed in sect. 2, and the estimate (20). This immediately shows that (35), (36) and (37) are true for r provided (38) and (39) hold true for $r - 1$. It remains to prove that (38) and (39) hold true also for r as a consequence of (35), (36) and (37) and of the recursive definitions (8). Actually there are only two kinds of terms to be estimated, namely $\frac{1}{s!} L_{\chi_r}^s Z_m$ for $1 \leq m < r$ and $\frac{1}{p!} L_{\chi_r}^p H_{(s-p)r+m}^{(r-1)}$ for $0 \leq p \leq s$ and $0 \leq m < r$. For, remarking that Z_r and $H_r^{(r-1)}$ are estimated by exactly the same quantity it is safe to estimate $\|H_{sr}^{(r)}\|$ by replacing Z_r with $H_r^{(r-1)}$ in the second of (8). This is tantamount to extending the sum in the second of (8) to $p = s - 1$ and making it identical with the sum in the first of (8), with $m = 0$.

Denote $\varphi_s = L_{\chi_r}^s Z_m$ (which, by the way, exists only for $r > 1$), and split $\varphi_s = \varphi_s^\sharp + \varphi_s^\natural + \varphi_s^\flat$. I claim

$$(41) \quad \|\varphi_s\|_{1-\delta_r} \leq s! \left(\frac{e}{d_r \Lambda} \right)^{2(s-1)} \|\chi_r\|_{1-\delta_{r-1}}^{s-1} \frac{D}{d_r} ,$$

$$(42) \quad \|\varphi_s^\sharp + \varphi_s^\natural\|_{1-\delta_r} \leq s! \left(\frac{2e}{d_r \Lambda} \right)^{2(s-1)} \|\chi_r\|_{1-\delta_{r-1}}^{s-1} D ,$$

[‡] The reader may be a little puzzled by this claim. However, he or she may get convinced by paying attention to the accumulation of the divisors d_r in the estimates. Using (38) in order to estimate χ_r causes an extra divisor d_r to appear in (20). This happens at every step, thus causing a divisor $d_1 d_2 \cdots d_r$ to appear in the estimate of χ_r . A moment's thought will allow the reader to realize that this causes the estimate of the norm of χ_r to grow faster than geometrically with r , due to $\sum_r d_r = d$. The dangerous accumulation of divisors above does not occur because only (39) is used in order to estimate the norm of χ_r , due to $\chi_r \in \mathcal{W} \subset \mathcal{P}^\sharp \oplus \mathcal{P}^\natural$.

for $s \geq 1$, where

$$(43) \quad D = \mu_{r-1,r} \mu_{m-1,m} T_{r,r+m} C^{r+m-1} E$$

The proof proceeds by induction. Let $s = 1$. By the general estimate (21) one has

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

Using (35) and (36) one gets

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

so that (41) immediately follows from (32), (33) and (40)

Still keeping $s = 1$, (42) is obtained by remarking that the contributions to $\varphi_1^\sharp + \varphi_1^\natural$ come only from $L_{\chi_r} Z_m^\sharp$ and $(L_{\chi_r^\natural} Z_m^\natural)^\natural$. Proceeding as above, from (23) and (22) one gets

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

Then (42) for $s = 1$ follows from (36), (39) for $r - 1$ and (40). Remark that the divisor d_r does not appear here.

Let now $s > 1$, and assume that (41) 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 (41) and (42) for φ_{s-1} , thus getting

$$(44) \quad \begin{aligned} \|\varphi_{s-1}\|_{1-\delta_r+d_r/s} &\leq (s-1)! \left(\frac{s}{s-1}\right)^{2s-3} \left(\frac{e}{d_r \Lambda}\right)^{2(s-2)} \|\chi_r\|_{1-\delta_{r-1}}^{s-2} \frac{D}{d_r}, \\ \|\varphi_{s-1}^\sharp + \varphi_{s-1}^\natural\|_{1-\delta_r+d_r/s} &\leq (s-1)! \left(\frac{s}{s-1}\right)^{2s-4} \left(\frac{2e}{d_r \Lambda}\right)^{2(s-2)} \|\chi_r\|_{1-\delta_{r-1}}^{s-2} D. \end{aligned}$$

Let me consider first the estimate (42) — which is the most delicate one. Remarking that the contributions to $\varphi_s^\sharp + \varphi_s^\natural$ come only from $L_{\chi_r}(\varphi_{s-1}^\sharp + \varphi_{s-1}^\natural)$ and $(L_{\chi_r^\natural} \varphi_{s-1}^\natural)^\natural$, use (21) and (22) to estimate

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

Replacing (44) in the latter expression one gets

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

so that (42) follows from the trivial inequality $\left(\frac{s}{s-1}\right)^{s-1} < e$. The estimate (41) is checked with a similar calculation, just taking into account that only (21) may 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 (35) and (43) in (41) and (42). Using also (40), one gets

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

Using $s-1$ times the inequalities (32) and (33) one easily gets

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

Thus one concludes

$$(45) \quad \frac{1}{s!} \|L_{\chi_r}^s Z_m\|_{1-\delta_r} \leq \mu_{r-1,r}^s \mu_{m-1,m} T_{r,sr+m} C^{sr+m-1} \frac{E}{d_r} ,$$

$$(46) \quad \frac{1}{s!} \|(L_{\chi_r}^s Z_m)^\sharp + (L_{\chi_r}^s Z_m)^\natural\|_{1-\delta_r} \leq \mu_{r-1,r}^s \mu_{m-1,m} T_{r,sr+m} C^{sr+m-1} E .$$

The estimate for $\frac{1}{p!} L_{\chi_r}^p H_{(s-p)r+m}^{(r-1)}$ is a minor *variazione* of the scheme above. Only the first step must be omitted. E.g., set $\varphi_p = L_{\chi_r}^p H_{(s-p)r+m}^{(r-1)}$ and proceed as follows. Using (38) for $r-1$ get

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

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

$$(47) \quad \frac{1}{p!} \|L_{\chi_r}^p H_{(s-p)r+m}^{(r-1)}\|_{1-\delta_r} \leq \mu_{r-1,r}^p \mu_{r-1,(s-p)r+m} T_{r,sr+m} C^{sr+m-1} \frac{E}{d_r} ,$$

$$(48) \quad \begin{aligned} \frac{1}{p!} \|(L_{\chi_r}^p H_{(s-p)r+m}^{(r-1)})^\sharp + (L_{\chi_r}^p H_{(s-p)r+m}^{(r-1)})^\natural\|_{1-\delta_r} \\ \leq \mu_{r-1,r}^p \mu_{r-1,(s-p)r+m} T_{r,sr+m} C^{sr+m-1} E . \end{aligned}$$

Collecting (45), (46), (47) and (48) and referring to the transformation formulæ (8) it is now an easy matter to verify that (38) and (39) hold true provided the sequence $\mu_{r,s}$ for $0 < r < s$ is defined as

$$(49) \quad \begin{aligned} \mu_{0,s} &= 1 && \text{for } s > 0 , \\ \mu_{r,sr+m} &= \mu_{r-1,r}^s \mu_{m-1,m} + \sum_{p=0}^{s-1} \mu_{r-1,r}^p \mu_{r-1,(s-p)r+m} && \text{for } r \geq 2, s \geq 1, 1 \leq m < r . \\ \mu_{r,sr} &= \sum_{p=0}^{s-1} \mu_{r-1,r}^p \mu_{r-1,(s-p)r} && \text{for } r \geq 1, s \geq 2 . \end{aligned}$$

This looks quite different from (34). However, I claim that (34) is just a harmless extension of (49) that covers the cases $s = 0$ and $s = 1, m = 0$. For, putting $s = 0$ in (34) we get $\mu_{r,m} = \mu_{r-1,m}$ for $0 \leq m < r$; in particular, for $m = 0$ this implies $\mu_{r,0} = \dots = \mu_{0,0} = 0$ for all $r \geq 0$. On the other hand, putting $s = 1, m = 0$ in (34) we get $\mu_{r,r} = \mu_{r-1,r} + \mu_{r-1,r}\mu_{r-1,0} = \mu_{r-1,r}$ in view of $\mu_{r-1,0} = 0$. Hence in the second of (49) we may replace $\mu_{m-1,m}$ with $\mu_{r-1,m} = \dots = \mu_{m,m} = \mu_{m-1,m}$. That is, we extend the sum up to $p = s$. On the other hand, adding the term $p = s$ in the third of (49) is harmless, because we add $\mu_{r,r}^s \mu_{r-1,0} = 0$. This completes the proof of lemma 3. *Q.E.D.*

3.4 Completion of the proof of lemma 1

The statement of the lemma concerns only the sequence of generating functions, that are estimated by (35). 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 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

$$(50) \quad T_{r,s} \leq \left(\frac{16}{b^2} \right)^{s-1}.$$

For, use the definition (31) and recall the definition (30) 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 (50)

Coming to the sequence (34), the problem is to show that $\mu_{r,r} \leq \eta^{r-1}$ for some positive η . Start by proving the following properties:

$$(51) \quad \mu_{r,0} < \dots < \mu_{r-1,r} = \mu_{r,r} = \mu_{r+1,r} = \dots \quad \text{for } r \geq 1,$$

$$(52) \quad \mu_{1,s} = s \quad \text{for } s \geq 0,$$

$$(53) \quad \mu_{r,s} = \mu_{r-1,s} + \mu_{r-1,r}\mu_{r,s-r} \leq \mu_{r-1,s} + \mu_{r,r}\mu_{s-r,s-r} \quad \text{for } s \geq r \geq 2,$$

$$(54) \quad \mu_{r,r} \leq r + \sum_{j=2}^{r-1} \mu_{j,j}\mu_{r-j,r-j} \quad \text{for } r \geq 3.$$

Properties (51) and (52) are straightforward consequences of the definition. In order to check (53) write $s = kr + m$ with $0 \leq m < r$ and calculate

$$\begin{aligned} \mu_{r,kr+m} &= \mu_{r-1,kr+m} + \sum_{p=1}^k \mu_{r-1,r}^p \mu_{r-1,(k-p)r+m} \\ &= \mu_{r-1,kr+m} + \underbrace{\mu_{r-1,r} \sum_{p=0}^{k-1} \mu_{r-1,r}^p \mu_{r-1,(k-1-p)r+m}}_{=\mu_{r,(k-1)r+m}} ; \end{aligned}$$

this gives the equality, and the inequality follows from (51). Property (54) is a repeated application of (53), followed by use of (52). For, calculate

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

this is nothing but (54) since $\mu_{1,r} = r$.

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

$$(55) \quad \begin{aligned} \nu_1 &= 1 , \quad \nu_2 = 2 , \\ \nu_r &= r + \sum_{j=2}^{r-1} \nu_j \nu_{r-j} \quad \text{for } r \geq 3 \end{aligned}$$

is a majorant of $\{\mu_{r,r}\}_{r \geq 1}$. An easy remark is that $\{\nu_r\}_{r \geq 1}$ resembles the well known sequence $\{\lambda_r\}_{r \geq 1}$ defined as

$$(56) \quad \begin{aligned} \lambda_1 &= 1 , \\ \lambda_r &= \sum_{j=1}^{r-1} \lambda_j \lambda_{r-j} \quad \text{for } r \geq 2 . \end{aligned}$$

Indeed it is an easy matter to see that $\nu_r \leq a^{r-1} \lambda_r$ for some positive constant a (e.g., $a = 2$). Hence it is enough to study the latter sequence. It is known that

$$(57) \quad \lambda_r = \frac{2^{r-1}(2r-3)!!}{r!} \leq 4^{r-1} ,$$

where the standard notation $(2n+1)!! = 1 \cdot 3 \cdot \dots \cdot (2n+1)$ has been used. As a check, let the function $g(z)$ be defined as $g(z) = \sum_{r \geq 1} \lambda_r z^r$, so that $\lambda_r = g^{(r)}(0)/r!$. Then it is immediate to check that the recursive definition (56) is equivalent to the equation $g = z + g^2$. By repeated differentiation of the latter equation one readily finds

$$g' = \frac{1}{1-2g} , \dots , g^{(r)} = \frac{2^{r-1}(2r-3)!!}{(1-2g)^{2^{r-1}}}$$

(check by induction). From this, (57) follows.

Thus, $\mu_{r,r} \leq \zeta^{r-1}$ for some positive ζ . Inserting the latter inequality and (50) in (35) the statement of lemma 1 follows.

4. Proof of theorem 1

Having established the estimate of lemma 1 on the sequence of generating functions it is now a standard matter to complete the proof of theorem 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

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

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

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

- (iii) The sequence $\{\mathcal{C}^{(r)}\}_{r \geq 0}$ of canonical transformations converges for $r \rightarrow \infty$ 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 $r \rightarrow \infty$ 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 [8]; 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 $|\chi_r(x, y)| \leq \varrho^{r+2} \|\chi_r\|_{\varrho R}$; this, in turn, implies that $|x^{(r)} - x^{(r-1)}| \sim \beta^{r-1} \varrho^{r+2}$ and $|y^{(r)} - y^{(r-1)}| \sim \beta^{r-1} \varrho^{r+2}$. The geometric bound on the latter quantities implies that $\sum_{r>1} |x^{(r)} - x^{(r-1)}|$ and $\sum_{r>1} |y^{(r)} - y^{(r-1)}|$ 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.

Acknowledgements. This work has been completed while the author was visiting the *Observatoire de la côte d'Azur*, Nice (France). I want to express my gratitude to C. Froeschlé for his kind invitation.

References

- [1] Birkhoff, G. D.: *Dynamical systems*, New York (1927).
- [2] Cherry, T. M.: *On the solutions of Hamiltonian systems in the neighborhood of a singular point*, Proc. London Math. Soc., Ser. 2, **27**, 151–170 (1926).
- [3] Giorgilli, A.: *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] Giorgilli, A. and Locatelli, U.: *Kolmogorov theorem and classical perturbation theory*, ZAMP **48**, 220–261 (1997).
- [5] Giorgilli, A. and Locatelli, U.: *On classical series expansions for quasi–periodic motions*, MPEJ **3** N. 5 (1997).
- [6] Giorgilli: *Classical constructive methods in KAM theory*, PSS, **46** 1441–1451 (1998).
- [7] Giorgilli, A. and Locatelli, U.: *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).
- [8] Gröbner, W.: *Die Lie–Reihen und Ihre Anwendungen*, VEB Deutscher Verlag der Wissenschaften (1967).
- [9] Moser, J.: *On the generalization of a theorem of A. Liapounoff*, Comm. Pure Appl. Math. **11**, 257–271 (1958).
- [10] Siegel, C.L. and Moser, J.K.: *Lectures in Celestial Mechanics*, Springer–Verlag, Berlin Heidelberg New York (1971).