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

[9] A. Giorgilli and L. Galgani, *Rigorous estimates for the series expansions of Hamiltonian perturbation theory*, Cel. Mech. **37**, 95-112 (1985).

[10] A. Giorgilli, A. Delshams, E. Fontich, L. Galgani and C. Simó: *Effective stability for a Hamiltonian system near an elliptic equilibrium point, with an application to the restricted three body problem*, J. Diff. Eqs., in print.

[11] A. Giorgilli and L. Galgani *Formal integrals for an autonomous Hamiltonian system near an equilibrium point*, Cel. Mech. **17**, 267-280 (1978).

[12] W. Grobner, *Serie di Lie e loro applicazioni*, Cremonese, Roma (1973).

References

[1] G.Benettin, L. Galgani and A. Giorgilli: *Realization of holonomic constraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory, part I*, Comm. Math. Phys., **113**, 87–103, (1987).

[2] N.N. Nekhoroshev: *Exponential estimate of the stability time of near–integrable Hamiltonian systems*, Russ. Math. Surveys **32** N.6 (1977), 1–65.
N.N. Nekhoroshev: *Exponential estimate of the stability time of near–integrable Hamiltonian systems, II*, Trudy Sem. Petrovs. N.5 (1979), 5–50 (in russian).
G. Benettin, L. Galgani and A. Giorgilli: *A Proof of Nekhoroshev's Theorem for the Stability times in Nearly Integrable Hamiltonian Systems*, Cel. Mech **37**, 1-25 (1985);
G. Benettin and G. Gallavotti: *Stability of Motion Near Resonances in Quasi Integrable Hamiltonian Systems*, J. Stat. Phys. **44**, 293–338 (1986).
G. Gallavotti: *Quasi-integrable Mechanical Systems* in K. Osterwalder and R. Stora (eds.), *Les Houches, Session XLIII* (1984).

[3] J.H. Jeans: Phil. Magazine **6**, 279, (1903);
J.H. Jeans: Phil. Magazine **10**, 91, (1905);

[4] L. D. Landau: Phys. Z. Sowjetunion **1**, 81 (1932); **2**, 46 (1932);
D. Rapp: Journ. Chem. Phys. **32**, 735 (1960).

[5] G. Benettin, L. Galgani and A. Giorgilli: *Exponential law for the equipartition times among translational and vibrational degrees of freedom*, Phys. Lett. A, **120** N.1, 23–27 (1987).

[6] E.T. Whittaker: *On the adelphic integral of the differential equations of dynamics*, Proc. Roy Soc. Edinburgh, Sect. A, **37**, 95–109, (1916);
E.T. Whittaker: *A treatise on the analytical dynamics of particles and rigid bodies*, London, pp. 432 ff. (1970).

[7] T.M. Cherry: *On integrals developable about a singular point of a Hamiltonian system of differential equations*, Proc. Camb. Phil. Soc. **22**, 325–349, (1924);
T.M. Cherry: *On integrals developable about a singular point of a Hamiltonian system of differential equations, II*, Proc. Camb. Phil. Soc. **22**, 510–533, (1924);

lemma 9.6 with $\frac{d}{s+1}$ in place of d and $2d$ in place of $d+d'$, and get

$$N_{R-2d}(\{I_\mu,h_s\}) \le \frac{s+1}{2d^2}\|I_\mu\|_R\left(\frac{s+1}{s}\right)^{2(s-1)}\frac{2^{2s}\mathcal{F}}{s}\left(\frac{2e^2\Phi}{d^2}+\beta\right)^{s-1}$$
$$< \frac{2^{2s-1}e^2}{d^2}\mathcal{F}\left(\frac{2e^2\Phi}{d^2}+\beta\right)^{s-1}\|I_\mu\|_R \ .$$

A similar computation using (12.6) gives

$$N_{R-2d}(\{I_\mu,\hat h_s\}) < \frac{(s+1)e^2}{2d^2}\|I_\mu\|_R\frac{2^{2s}\Phi}{d^2}\left(\frac{2e^2\Phi}{d^2}+\beta\right)^{s-1}N_{R-d}(\hat h)$$
$$N_{R-2d}(\{I_\mu,H_{l,s-l}\}) < \frac{(s-l+1)e^2}{2d^2}\|I_\mu\|_R\frac{2^{2s-l+2}\Phi}{d^2}\left(\frac{2e^2\Phi}{d^2}+\beta\right)^{s-l-1}N_{R-d}(H_l) \ .$$

We can now proceed as in the computation of the estimate (12.5), (12.7) and (12.8), and get

$$\left|\{I_\mu,h^{(m)}\}(p',x',\pi',\xi',\varepsilon)\right| < \frac{2^4e^2E}{d^2}\delta^4\left(\frac{\delta\overline{\delta}}{\delta_{*r}}\right)^{2m}\|I_\mu\|_R$$
$$\left|\{I_\mu,\hat h^{(m)}\}(p',x',\pi',\xi',\varepsilon)\right| < \frac{2^4e^2E}{\alpha_r d^4}(m+2)E_0\delta^4\left(\frac{\delta\overline{\delta}}{\delta_{*r}}\right)^{2m}\|I_\mu\|_R$$
$$\left|\{I_\mu,\mathcal{H}^{(m)}\}(p',x',\pi',\xi',\varepsilon)\right| < \frac{2^{10}e^2E}{\alpha_r d^4}(m+1)E\delta_{*r}^2\delta^4\left(\frac{\delta\overline{\delta}}{\delta_{*r}}\right)^{2m}\|I_\mu\|_R \ ,$$

and collect all these inequalities to get, for $r>1$,

$$\left|\{I_\mu,\mathcal{R}^{(r)}\}(p',x',\pi',\xi',\varepsilon)\right|$$
$$< \frac{2^4e^2E}{d^2}\delta^4\left(\frac{\delta\overline{\delta}}{\delta_{*r}}\right)^{2r}\|I_\mu\|_R\left(1+\frac{(r+1)\delta_{*r}^2E_0}{\alpha_r d^2}+\frac{2^5(r+1)\delta_{*r}^2E}{\alpha_r d^2}\right) \ ,$$

so that (7.9) follows by again noting that the last factor, via to the explicit expression of δ_{*r}, does not exceed 2. Here too the estimate for $r=1$ should be separately computed by using $\mathcal{R}^{(1)}=h^{(1)}+\mathcal{H}^{(1)}$. The lemma is thus proven.

so that one has

$$\begin{aligned}\left|\mathcal{H}^{(m)}(p',x',\pi',\xi',\varepsilon)\right| &< \frac{2^{10}E^2}{\alpha_r d^2}\delta_{*r}^2\delta^4\sum_{s>m}\left(\frac{\delta\overline{\delta}}{\delta_{*r}}\right)^{2(s-1)} \\ &\leq \frac{2^{11}E^2}{\alpha_r d^2}\delta_{*r}^2\delta^4\left(\frac{\delta\overline{\delta}}{\delta_{*r}}\right)^{2m}\end{aligned} \tag{12.8}$$

In order to get (7.6) we recall now that the remainder $\mathcal{R}^{(r)}$ is defined for $r>1$ as $\mathcal{R}^{(r)} = h^{(r)} + \varepsilon^2\hat{h}^{(r-1)} + \mathcal{H}^{(r)}$, (see (6.2) and (6.4)), so that from (12.7) and (12.8) one obtains

$$\left|\mathcal{R}^{(r)}(p',x',\pi',\xi',\varepsilon)\right| < 2^5E\delta^4\left(\frac{\delta\overline{\delta}}{\delta_{*r}}\right)^{2r}\left(\frac{1}{r+1}+\frac{\delta_{*r}^2E_0}{\alpha_r d^2}+\frac{2^6\delta_{*r}^2E}{\alpha_r d^2}\right) ,$$

and (7.6) follows from the trivial inequality $\left(\frac{1}{r+1}+\frac{\delta_{*r}E_0}{\alpha_r d^2}+\frac{2^6\delta_{*r}E}{\alpha_r d^2}\right) < \frac{2}{r+1}$, which in turn is obtained via the very definition of δ_{*r}. For $r=1$ the remainder reduces to $\mathcal{R}^{(1)} = h^{(1)} + \mathcal{H}^{(1)}$, and the estimate is trivially checked.

In order to conclude the proof of theorem 7.1 we must prove (7.5). To this end, we use the bound (6.11) on the norms of Z_s, and again transform to real variables by lemma 9.2. This gives, for $1\leq s\leq r$,

$$N_{R-d}(Z_s) \leq \frac{2^{2s+2}\beta^{s-1}}{s}E < \frac{2^4}{s}\delta_{*r}^{-2(s-1)}E ,$$

so that in the domain $\mathcal{D}_{R-2d,\varrho}$ one has

$$\left|Z^{(r)}(p',x',\pi',\xi',\varepsilon)\right| < 2^5E\delta^4\sum_{s=1}^{r}\left(\frac{\delta\overline{\delta}}{\delta_{*r}}\right)^{2(s-1)} ,$$

and (7.5) follows by simply extending the sum to infinity. This concludes the proof of theorem 7.1.

Next, we come to the

Proof of corollary 7.2. That I_μ is a prime integral of the normalized part of the Hamiltonian (7.2) follows from the characterization of the normal form with respect to the module $\mathcal{M}$ given in sect. 2, and from the condition $\mu\perp\mathcal{M}$. Indeed, using complex variables one immediately sees that $\{I_\mu, Z^{(r)}\}=0$. The bound (7.8) has already been proven above. So, we only need to prove the bound on the time derivative $\dot{I}_\mu$.

Using $\dot{I}_\mu = \{I_\mu, \mathcal{R}^{(r)}\}$ and the expression of the remainder, namely $\mathcal{R}^{(r)} = h^{(r)} + \varepsilon^2\hat{h}^{(r-1)} + \mathcal{H}^{(r)}$, we are led to look for bounds on $\{I_\mu, h_s\}$, $\{I_\mu, \hat{h}_s\}$ and $\{I_\mu, H_{l,s-l}\}$. To this end we use the bound (12.4) on the norm of h_s with $\frac{s}{s+1}d$ in place of d and

$\delta\overline{\delta} \leq \delta_{*r}/\sqrt{2}$, on the domain $\mathcal{D}_{R-2d,\varrho}$ one has

$$\begin{aligned} \left| h^{(m)}(p', x', \pi', \xi', \varepsilon) \right| &\leq 2^4 E \delta^4 \sum_{s>m} \frac{1}{s} \left(\frac{\delta\overline{\delta}}{\delta_{*r}} \right)^{2(s-1)} \\ &< \frac{2^5 E}{m+1} \delta^4 \left(\frac{\delta\overline{\delta}}{\delta_{*r}} \right)^{2m} \end{aligned} \tag{12.5}$$

Inequality (7.3) then follows by just recalling the exchange theorem, i.e. that $h_\Omega(\pi, \xi)$ is changed by the canonical transformation $T_{\chi^{(r)}}$ to $h_\Omega(\pi', \xi') + h^{(0)}(p', x', \pi', \xi', \varepsilon)$.

By similar computations one proves the inequalities (7.4), (7.6) and (7.8). Indeed, denote $T_{\chi^{(r)}} I_\mu = \sum_{s\geq 0} I_s$, $T_{\chi^{(r)}} \hat{h} = \sum_{s\geq 0} \hat{h}_s$ and $T_{\chi^{(r)}} \sum_{s\geq 1} H_s = \sum_{s\geq 1} \sum_{l=1}^{s} H_{l,s-l}$. By lemma 10.2 one has that I_s and $\hat{h}_s$ are of class $\mathcal{P}_{2s+2,4s}$ and $H_{l,s-l}$ is of class $\mathcal{P}_{2s+2,4s-2l+2}$. Again, we use lemma 10.3 in order to get the estimates for the norms of I_s, $\hat{h}_s$ and $H_{l,s-l}$ in complex variables, and lemma 9.2 to transform back to real variables, so that we get

$$\begin{aligned} N_{R-2d}(I_s) &\leq \frac{2^{2s}\Phi}{d^2} \left(\frac{2e^2\Phi}{d^2} + \beta \right)^{s-1} \|I_\mu\|_{R-d} \\ N_{R-2d}(\hat{h}_s) &\leq \frac{2^{2s}\Phi}{d^2} \left(\frac{2e^2\Phi}{d^2} + \beta \right)^{s-1} N_{R-d}(\hat{h}) \\ N_{R-2d}(H_{l,s-l}) &\leq \frac{2^{2s-l+2}\Phi}{d^2} \left(\frac{2e^2\Phi}{d^2} + \beta \right)^{s-l-1} N_{R-d}(H_l) \end{aligned} \tag{12.6}$$

Proceeding now as above, one gets

$$\begin{aligned} \left| I^{(m)}(p', x', \pi', \xi', \varepsilon) \right| &\leq \frac{2^5 E}{\alpha_r d^2} \|I_\mu\|_{R-d} \delta^4 \left(\frac{\delta\overline{\delta}}{\delta_{*r}} \right)^{2m} \\ \left| \hat{h}^{(m)}(p', x', \pi', \xi', \varepsilon) \right| &\leq \frac{2^5 E}{\alpha_r d^2} E_0 \delta^4 \left(\frac{\delta\overline{\delta}}{\delta_{*r}} \right)^{2m} . \end{aligned} \tag{12.7}$$

For $m = 0$, these inequalities immediately give (7.8) and (7.4). Denoting now $\mathcal{H}^{(m)} = \sum_{s>m} \sum_{l=1}^{s} H_{l,s-l}$, and recalling that in complex variables one has by hypothesis $N_R(H_l) \leq (2\sigma)^{l-1} 4E$, one gets

$$\left| \mathcal{H}^{(m)}(p', x', \pi', \xi', \varepsilon) \right| \leq \frac{2^9 E^2}{\alpha_r d^2} \delta^4 \sum_{s>m} \sum_{l=1}^{s} \delta_{*r}^{-2(s-l-1)} (4\sigma)^{l-1} \delta^{2(s-1)} \overline{\delta}^{2(s-l)} .$$

Using now the definition (7.1) of δ_{*r}, and using also $\overline{\delta} > 1$ one easily computes $\sum_{l=1}^{s} \delta_{*r}^{-2(s-l-1)} (4\sigma)^{l-1} \overline{\delta}^{2(s-l)} < \delta_{*r}^{-2(s-2)} \overline{\delta}^{2(s-1)} \sum_{l=1}^{s} (1/16)^{l-1} < 2\delta_{*r}^{-2(s-2)} \overline{\delta}^{2(s-1)}$,

12. Proof of the theorem on the estimate of the remainder

According to sect. 7, we now collect the results of the previous sections and give the estimates for the Hamiltonian in normal form up to the order r. We start with the

Proof of theorem 7.1. Starting with the Hamiltonian $H(p,x,\pi,\xi,\varepsilon)$ in real variables, as in the hypotheses of the theorem, we first perform a transformation to complex variables via the canonical transformation (2.8), so that, omitting the primes, we recover an Hamiltonian of the same form with $\varepsilon^2\hat{h}(p,x)$ of class $\mathcal{P}_{2,2}$ (in fact it is unchanged) and H_s of class $\mathcal{P}_{2s+2,2s+2}$ on the same domain $\mathcal{G}_R$, while h_Ω is changed to the form required by theorem 6.1. From the hypotheses of the theorem and lemma 9.2, the norms of the new Hamiltonian change according to

$$N_R(\varepsilon^2\hat{h}) \le E_0 \ , \qquad N_R(H_s) \le \gamma^{s-1}\mathcal{F} \ , \qquad s \ge 1 \ , \tag{12.1}$$

with

$$\gamma = 2\sigma \ , \quad \mathcal{F} = 4E \ , \tag{12.2}$$

while the norm of h_Ω, or more generically the norm of I_μ, is unchanged, as is easily computed. By substitution of these values in (6.12) we obtain for the generating sequence the bound (6.11), as required by the theorems on canonical transformations, with

$$\Phi = \frac{4E}{\alpha_r} \ , \quad \beta = \frac{12(4E+E_0)(r-1)}{\alpha_r d^2} + 4\sigma \ . \tag{12.3}$$

Since the bounds on the norms of the generating sequence are given in a domain $\mathcal{D}_{R-d,\varrho}$, we can apply the theorems on canonical transformations only on a smaller domain, that we choose to be $\mathcal{D}_{R-2d,\varrho}$; the quantity δ_* in (5.6) turns out to be larger than δ_{*r} given by (7.1). This proves the consistency of the canonical transformation.

Coming now to the bounds, consider first the inequality (7.3). As in sect. 10, denote $T_{\chi^{(r)}} h_\Omega = \sum_{s\ge 0} h_s$, and also $h^{(m)} = \sum_{s>m} h_s$. By lemma 10.2 one has that h_s is of class $\mathcal{P}_{2s+2,4s}$, and lemma 11.3 gives the estimates for the norms in complex variables. By lemma 9.2 one can then produce the norms in real variables, namely

$$N_{R-2d}(h_s) \le \frac{2^{2s}\mathcal{F}}{s}\left(\frac{2e^2\Phi}{d^2}+\beta\right)^{s-1} \ . \tag{12.4}$$

By defining now $\delta_{*r}^2 = \frac{1}{4}\left(\frac{2e^2\Phi}{d^2}+\beta\right)^{-1}$, and substituting the expressions (12.3) for Φ and β, the definition (7.1) of δ_{*r} is immediately obtained; moreover, by the hypothesis form

$$\begin{aligned} h_1 &= -L_{h_\Omega}\chi_1 \\ h_s &= \sum_{j=1}^{s-1} \frac{j}{s} L_{\chi_j} h_{s-j} - L_{h_\Omega}\chi_s \ , \quad 2 \le s \le r \ , \\ h_s &= \sum_{j=1}^{r} \frac{j}{s} L_{\chi_j} h_{s-j} \ , \quad s > r \end{aligned} \tag{11.16}$$

(the second line exists only for $r > 1$), and recall that, by (6.4), $L_{h_\Omega}\chi_s = Z_s - \Psi_s$ for $1 \le s \le r$. The proof now closely follows the line of the proof of lemmas 10.2 and 10.3. Precisely, we can build a sequence $\{C_s\}_{s\ge1}$ such that $N_{R-2d}(h_s) \le C_s$ via the recursive definition

$$\begin{aligned} C_1 &= N_{R-2d}(Z_1 - \Psi_1) \\ C_s &= \frac{2e^2}{sd^2} \sum_{j=1}^{s-1} j(s-j) N_{R-d}(\chi_j) C_{s-j} + N_{R-2d}(Z_s - \Psi_s) \ , \quad 2 \le s \le r \\ C_s &= \frac{2e^2}{sd^2} \sum_{j=1}^{r} j(s-j) N_{R-d}(\chi_j) C_{s-j} \ , \quad s > r \ , \end{aligned}$$

analogous to that of lemma 10.2 for the sequence I_s. Here a factor $s-j$ appears in the sum instead of $s-j+1$ as in (10.14) because h_s can be computed by at most $s-1$ Poisson brackets, since $L_{h_\Omega}\chi_s$ is already known and bounded in $\mathcal{D}_{R-d}$. From this, using $N_{R-d}(\chi_j) \le \frac{\beta^{j-1}}{j}\Phi$ and $N_{R-2d}(Z_s - \Psi_s) \le N_{R-d}(\Psi_s) \le \frac{\beta^{s-1}}{s}\mathcal{F}$ one gets $N_{R-2d}(h_s) \le \tilde{C}_s$, with

$$\begin{aligned} \tilde{C}_1 &= \mathcal{F} \\ \tilde{C}_s &= \frac{2e^2\Phi}{sd^2} \sum_{j=1}^{s-1} (s-j)\beta^{j-1}\tilde{C}_{s-j} + \frac{\beta^{s-1}}{s}\mathcal{F} \ , \quad 2 \le s \le r \\ \tilde{C}_s &= \frac{2e^2\Phi}{sd^2} \sum_{j=1}^{r} (s-j)\beta^{j-1}\tilde{C}_{s-j} \ , \quad s > r \ , \end{aligned}$$

which are the analogous of (10.18). Proceeding now as in the proof of lemma 10.3 one gets

$$\tilde{C}_s \le \frac{s-1}{s} \left(\frac{2e^2\Phi}{d^2} + \beta \right) \tilde{C}_{s-1} \ , \quad s > 1 \ ,$$

and this immediately gives (11.15), so that the lemma is proven.

are known, one easily computes

$$g^{(s)}=\sum_{j=1}^{s-1}\sum_{k=1}^{2s-3}\sum_{l=s}^{2s-2}\beta_{j,k,l}\left[\frac{jCf^{j+1}}{(1-Cg)^{k+1}(1-\gamma z)^{l+2}}\right.$$
$$\left.+\frac{kCf^{j+1}}{(1-Cg)^{k+2}(1-\gamma z)^{l+2}}+\frac{l\gamma f^{j}}{(1-Cg)^{k}(1-\gamma z)^{l+1}}\right] .$$

This would allow us to explicitly compute the coefficients $\alpha_{j,k,l}$. However, we can avoid these horrendous computations, and note that $g^{(s)}(0)=\sum_{j,k,l}\alpha_{j,k,l}$, i.e.

$$\begin{aligned}g^{(s)}(0)&=\sum_{j=1}^{s-1}\sum_{k=1}^{2s-3}\sum_{l=s}^{2s-2}\beta_{i,k,l}\left[(j+k)C+l\gamma\right]\\&\le(s-1)(3C+2\gamma)\sum_{j,k,l}\beta_{j,k,l} .\end{aligned}$$

This allows us to obtain the final estimate

$$\eta_s\le\frac{(3C+2\gamma)^{s-1}}{s} ,$$

so that, by using the explicit expressions of C given by (11.14) and of $\tilde{C}$ given by (11.11), we obtain (11.12), and the lemma is proven.

We can now conclude the proof of the theorem. Indeed, by (11.3) and (11.8) one has $N_{R-d}(Z_s)\le\tilde{\eta}_s\mathcal{F}$ and $N_{R-d}(\chi_s)\le\tilde{\eta}_s\frac{\mathcal{F}}{\alpha_r}$, and using $\tilde{\eta}_s\le\eta_s$ we immediately obtain the bound (6.11) with β and Φ given by (6.12), so that the theorem 6.1 is proven.

Before concluding this section we prove the following

Lemma 11.3: *The sequence $\{h_s\}_{s\ge1}$ defined by $T_{\chi^{(r)}}h_\Omega$ is bounded by*

$$N_{R-2d}(h_s)\le\frac{1}{s}\left(\frac{2e^2\Phi}{d^2}+\beta\right)^{s-1}\mathcal{F} . \tag{11.15}$$

Such lemma will be useful in bounding the remainder, since the estimate given here is definitely better than the one given in lemma 10.3 for a generic function I_μ.

Proof. Recall that the terms of the generating sequence are bounded on the common domain $\mathcal{D}_{R-d,\varrho}$, and look for bounds of h_s in $\mathcal{D}_{R-2d,\varrho}$. To this end, write h_s in the

Next, we define two new sequences $\{\vartheta_s\}_{0\le s\le r-1}$ and $\{\eta_s\}_{1\le s\le r}$ by

$$\begin{aligned}
\vartheta_0 &= 1 \\
\eta_1 &= 1 \\
\vartheta_s &= \frac{C}{s}\sum_{l=1}^{s} l\eta_l\vartheta_{s-l}\ , && 1\le s\le r \\
\eta_s &= \frac{C}{s}\sum_{l=1}^{s-1} l\eta_l\eta_{s-l} + \frac{1}{s}\sum_{l=1}^{s} l\gamma^{l-1}\vartheta_{s-l}\ , && 2\le s\le r
\end{aligned} \tag{11.13}$$

with

$$C = \tilde{C}(\mathcal{F}+E_0)\ . \tag{11.14}$$

It is an easy matter to check that one has $\zeta_s \le \vartheta_s$ and $\tilde{\eta}_s \le \eta_s$, so we now look for an estimate of the sequence $\{\eta_s\}$. To this end, we forget the limitations on s in (11.13), and consider the sequences ϑ_s and η_s as being defined for any s. Thus, we can look for two functions f and g of the complex variable z

$$\begin{aligned}
f(z) &= \sum_{s\ge 0}\vartheta_s z^s \\
g(z) &= \sum_{s\ge 1}\eta_s z^s\ ,
\end{aligned}$$

so that η_s can be evaluated via the s–th derivative of $g(z)$, being $\eta_s = g^{(s)}(0)/s!$. Eq. (11.13) then gives

$$\begin{aligned}
f' &= Cfg' \\
g' &= \frac{f}{(1-Cg)(1-\gamma z)^2}\ ,
\end{aligned}$$

and the conditions $\vartheta_0 = \eta_1 = 1$ above can be translated into the equivalent conditions $g(0) = 0$ and $f(0) = g'(0) = 1$. Here, use was also made of the identity $\sum_{s\ge 1} s\gamma^{s-1}z^{s-1} = (1-\gamma z)^{-2}$. We note now that, by induction, the s–th derivative of $g(z)$ has the general form

$$g^{(s)} = \sum_{j=1}^{s}\sum_{k=1}^{2s-1}\sum_{l=s}^{2s}\alpha_{j,k,l}\frac{f^j}{(1-Cg)^k(1-\gamma z)^l}\ ,$$

where $\alpha_{j,k,l}$ are nonnegative coefficients. Indeed, this is true for $s=1$, and by assuming that the coefficients $\beta_{j,k,l}$ of

$$g^{(s-1)} = \sum_{j=1}^{s-1}\sum_{k=1}^{2s-3}\sum_{l=s}^{2s-2}\beta_{j,k,l}\frac{f^j}{(1-Cg)^k(1-\gamma z)^l}$$

and

$$
\begin{aligned}
\tilde\zeta_{l,s} &= \frac{\tilde C\mathcal{F}}{s}\sum_{j=1}^{s} j\tilde\eta_j\tilde\zeta_{l,s-j}\ , && 1\le l<r\ ,\quad 1\le s\le r-l\ ,\\
\zeta_s &= \frac{\tilde C\mathcal{F}}{s}\sum_{j=1}^{s} j\tilde\eta_j\zeta_{s-j}\ , && 1\le s\le r-1\ ,\\
\tilde\eta_s &= \frac{\tilde C\mathcal{F}}{s}\sum_{l=1}^{s-1} l\tilde\eta_l\tilde\eta_{s-l} + \frac{(s-1)\tilde C E_0}{s}\tilde\eta_{s-1}\\
&\quad +\frac{1}{s}\sum_{l=1}^{s} l\tilde\zeta_{l,s-l} + \frac{E_0}{s\mathcal{F}}\zeta_{s-1}\ , && 2\le s\le r\ ,
\end{aligned}
\tag{11.10}
$$

with

$$
\tilde C = \frac{4(r-1)}{\alpha_r d^2}\ . \tag{11.11}
$$

In fact, we are only interested in bounding the sequence $\{\tilde\eta_s\}_{1\le s\le r}$. This is given by the following

Lemma 11.2: *The sequence $\{\tilde\eta_s\}_{1\le s\le r}$ defined by (11.9) and (11.10) is bounded by*

$$
\tilde\eta_s \le \frac{1}{s}\left[\frac{12(\mathcal{F}+E_0)(r-1)}{\alpha_r d^2} + 2\gamma\right]^{s-1}\ . \tag{11.12}
$$

Proof. First, it is immediately checked that one has $\tilde\zeta_{l,s}=\gamma^{l-1}\zeta_s$, so that, using also $\tilde\eta_1=1$, we can write the sequence (11.10) in the simpler form

$$
\begin{aligned}
\zeta_0 &= 1\ ,\\
\tilde\eta_1 &= 1\ ,\\
\zeta_s &= \frac{\tilde C\mathcal{F}}{s}\sum_{l=1}^{s} l\tilde\eta_l\zeta_{s-l}\ , && 1\le s\le r-1\ ,\\
\tilde\eta_s &= \frac{\tilde C\mathcal{F}}{s}\sum_{l=1}^{s-1} l\tilde\eta_l\tilde\eta_{s-l} + \frac{(s-1)\tilde C E_0}{s}\tilde\eta_1\tilde\eta_{s-1}\\
&\quad +\frac{1}{s}\sum_{l=1}^{s} l\gamma^{l-1}\zeta_{s-l} + \frac{E_0}{s\mathcal{F}}\zeta_{s-1}\ , && 2\le s\le r\ .
\end{aligned}
$$

Then we look for sequences $\{\tilde\eta_s\}_{1\le s\le r}$, $\{\zeta_s\}_{0\le s\le r-1}$, and $\{\tilde\zeta_{l,s}\}_{1\le l\le r, 0\le s\le r-l}$ such that

$$\begin{aligned} N_{R-d_s}(\Psi_s) &\le \tilde\eta_s \mathcal{F} \ , \\ N_{R-d_{s+1}}(\hat h_s) &\le \zeta_s E_0 \ , \\ N_{R-d_{l+s}}(H_{l,s}) &\le \tilde\zeta_{l,s}\mathcal{F} \ . \end{aligned} \tag{11.8}$$

Here we can take $\eta_1 = \zeta_0 = 1$ and $\tilde\zeta_{l,0} = \gamma^{l-1}$. In order to determine such sequences, we put them in the recursive formula (11.5), and use also the definition (5.2) of T_χ and lemma 9.6 to get

$$\begin{aligned} N_{R-d_{l+s}}(H_{l,s}) &\le \sum_{j=1}^{s} \frac{2j}{s(d_{l+s}-d_j)(d_{l+s}-d_{l+s-j})\alpha_j}\tilde\eta_j\tilde\zeta_{l,s-j}\mathcal{F}^2 \ , \\ N_{R-d_{s+1}}(\hat h_s) &\le \sum_{j=1}^{s} \frac{2j}{s(d_{s+1}-d_j)(d_{s+1}-d_{s-j})\alpha_j}\tilde\eta_j\zeta_{s-j}\mathcal{F}E_0 \ , \\ N_{R-d_s}(\Psi_s) &\le \sum_{l=1}^{s-1} \frac{2l}{s(d_s-d_l)(d_s-d_{s-l})\alpha_l}\tilde\eta_l\tilde\eta_{s-l}\mathcal{F}^2 \\ &\quad + \frac{s-1}{s(d_s-d_{s-1})d_s\alpha_{s-1}}\tilde\eta_{s-1}\mathcal{F}E_0 \\ &\quad + \sum_{l=1}^{s}\frac{l}{s}\tilde\zeta_{l,s-l}\mathcal{F} + \frac{1}{s}\zeta_{s-1}E_0 \ . \end{aligned}$$

Recalling now the definition (11.6) of d_s, and using the inequality

$$\left(\sqrt{s-1}-\sqrt{j-1}\right)\left(\sqrt{s-1}-\sqrt{s-j-1}\right) \ge \frac{1}{2}$$

for $1 \le j \le s-1$, we immediately get

$$\frac{1}{(d_s-d_j)(d_s-d_{s-j})} \le \frac{2(r-1)}{d^2} \ ;$$

recall moreover that one has $\alpha_j \ge \alpha_r$ for $1 \le j \le r$, so that the sequences above can be defined by

$$\begin{aligned} \tilde\zeta_{l,0} &= \gamma^{l-1} \ , \qquad 1 \le l \le r \ , \\ \zeta_0 &= 1 \ , \\ \tilde\eta_1 &= 1 \ , \end{aligned} \tag{11.9}$$

This is seen as follows. Using the second and third of (6.4) for h_l, write (6.6) in the form

$$\begin{aligned}\Psi_s = &\sum_{l=1}^{s-1} \frac{l}{s} L_{\chi_l} Z_{s-l} + \sum_{l=1}^{s} \frac{l}{s} H_{l,s-l} + \frac{1}{s}\varepsilon^2 \hat{h}_{s-1} + \frac{s-1}{s} L_{\chi_{s-1}} \varepsilon^2 \hat{h}_0 \\ &- \sum_{l=1}^{s-1} \frac{l}{s} L_{\chi_l} \varepsilon^2 \hat{h}_{s-l-1} - \sum_{l=1}^{s-1} \frac{l}{s} L_{\chi_l} \sum_{m=1}^{s-l} H_{m,s-l-m} \\ &+ \sum_{l=1}^{s} \frac{s-l}{s} H_{l,s-l} + \frac{s-1}{s} \varepsilon^2 \hat{h}_{s-1} \ .\end{aligned}$$

The first four terms give exactly the expression (11.5), so we should only check that the remaining terms vanish. To this end, use the explicit expression of $\hat{h}_{s-1}$ and $H_{l,s-l}$ given by the definition (5.2) of T_χ, and compute the latter two terms as

$$\frac{s-1}{s}\varepsilon^2 \hat{h}_{s-1} = \sum_{l=1}^{s-1} \frac{l}{s} L_{\chi_l} \varepsilon^2 \hat{h}_{s-l-1}$$

and

$$\begin{aligned}\sum_{l=1}^{s} \frac{s-l}{s} H_{l,s-l} &= \sum_{l=1}^{s-1} \frac{s-l}{s} \sum_{m=1}^{s-l} \frac{m}{s-l} L_{\chi_m} H_{l,s-l-m} \\ &= \sum_{m=1}^{s-1} \frac{m}{s} L_{\chi_m} \sum_{l=1}^{s-m} H_{l,s-l-m} \ ,\end{aligned}$$

and check that these terms compensate the negative terms in the expression above for Ψ_s.

We are now ready to obtain the estimates on the generating sequence. Recalling the result of lemma 11.1, namely the inequalities (11.3), we note that we just have to estimate the norms of Ψ_s, for $1 \le s \le r$, via the recursive formula (11.5). To this end, we fix d, with $0 < d < R$, and introduce a sequence $0 = d_1 < \ldots < d_r = d$ by

$$d_s = \sqrt{\frac{s-1}{r-1}} \ , \quad 1 \le s \le r \ . \tag{11.6}$$

Then we look for a recursive estimate of Ψ_s in the domain $\mathcal{G}_{R-d_s}$, so that Ψ_r will be estimated in $\mathcal{G}_{R-d}$. Starting with $s = 1$, one has, by (11.5), $N_R(\Psi_1) = N_R(H_1) \le \mathcal{F}$, and so also

$$N_R(Z_1) \le \mathcal{F} \ , \quad N_R(\chi_1) \le \frac{\mathcal{F}}{\alpha_1} \ . \tag{11.7}$$

resonant part) all the monomials $\pi^j \xi^k$ with $j-k \in \mathcal{M}_\Omega$, and in $\tilde{\Psi}_s$ (the nonresonant part) the remaining ones; then, the equation above splits into $Z_s = \overline{\Psi}_s$, which defines the normalized part of the Hamiltonian, and $L_{h_\Omega}\chi_s = \tilde{\Psi}_s$, which can be uniquely solved by the method above, and gives the generating sequence.

In fact, we can do something better, by including into the normal form Z_s more resonant terms than in the previous solution, precisely also terms such that $(j-k)\cdot\Omega$ is very small, although not vanishing. This is done by considering a larger resonance module $\mathcal{M}$ such that $\mathcal{M} \supset \mathcal{M}_\Omega$, and by performing the splitting of Ψ_s by the condition $(j-k) \in \mathcal{M}$ instead of $(j-k) \in \mathcal{M}_\Omega$ as above.

Coming now to the estimates on the solution of the eq. (6.5), we can prove the following

Lemma 11.1: *For a known function Ψ_s of class $\mathcal{P}_{2s+2,4s}$ on the domain $\mathcal{G}_R$ and with reference to a resonance module $\mathcal{M} \supset \mathcal{M}_\Omega$, eq. (6.5) admits a solution Z_s and χ_s of class $\mathcal{P}_{2s+2,4s}$ such that one has the bounds*

$$\begin{aligned} N_R(Z_s) &\le N_R(\Psi_s) \ , \\ N_R(\chi_s) &\le \frac{1}{\alpha_s} N_R(\Psi_s) \ , \end{aligned} \tag{11.3}$$

where α_s satisfies (6.10).

Proof. We write Ψ_s in the form (4.9), i.e. $\Psi_s = \sum_{2s+2\le l+m\le 4s} \varepsilon^l \Psi_s^{(l,m)}$ with $\Psi_s^{(l,m)} \in \Pi_m$. Then, denoting by $Z_s^{(l,m)}$ and $\chi_s^{(l,m)}$ the solutions of eq. (6.5) in Π_m, using the fact that $\{\alpha_s\}_{s\ge 1}$ is a nonincreasing sequence, and by the definition (4.13) of the norm, one finds

$$\|Z_s^{(l,m)}\|_R \le \|\Psi_s^{(l,m)}\|_R \ , \qquad \|\chi_s^{(l,m)}\|_R \le \frac{1}{\alpha_s}\|\Psi_s^{(l,m)}\|_R \ .$$

Finally, the statement follows from the definition (4.14) of the norm, so that the lemma is proven.

Before coming to the estimates on the generating sequence, let us obtain for Ψ_s a more suitable form than (6.6). We claim that (6.5) and (6.6) can be written as

$$L_{h_\Omega}\chi_s + Z_s = \Psi_s \ , \tag{11.4}$$

where

$$\begin{aligned} \Psi_1 &= H_1 \ , \\ \Psi_s &= \sum_{l=1}^{s-1} \frac{l}{s} L_{\chi_l} Z_{s-l} + \frac{s-1}{s} L_{\chi_{s-1}} \varepsilon^2 \hat{h}_0 + \sum_{l=1}^{s} \frac{l}{s} H_{l,s-l} + \frac{1}{s}\varepsilon^2 \hat{h}_{s-1} \ . \end{aligned} \tag{11.5}$$

which implies (10.23). So, in order to prove (5.8) for f it is enough to prove that, for any l, one has

$$\left(T_\chi f^{(l)}\right)(p',x',\pi',\xi',\varepsilon) = f^{(l)}(T_\chi(p',x',\pi',\xi',\varepsilon)) \ ,$$

i.e., that it holds for polynomials. This is immediately obtained by the properties of T_χ of being linear and preserving products. So, we have proven (5.8) in a polydisk $D_{R-d}(\overline{p},\overline{x}) \times \Delta_{\varrho(R-d)}$ for any $(\overline{p},\overline{x}) \in \mathcal{G}$, and this implies that the same holds on $\mathcal{G}_{R-d} \times \Delta_{\varrho(R-d)} = \mathcal{D}_{R-d,\varrho}$. This concludes the proof.

11. Proof of the theorem on the generating sequence

We determine the generating sequence $\chi^{(r)}$ and the normal form $Z^{(r)}$ by recursively solving the eq. (6.5). In fact, this is the general problem of solving an equation of the form

$$L_{h_\Omega} f = g \ , \tag{11.1}$$

g being a known function of class $\mathcal{P}_{\Lambda,K}$, with the condition that f also be of class $\mathcal{P}_{\Lambda,K}$. Let us discuss in some detail how such equation is solved.

Let us first consider $g \in \Pi_\Lambda$; the extension of the result to the class $\mathcal{P}_{\Lambda,K}$ is immediate, by the linearity of the operator L_{h_Ω}. We can write $g(p,x,\pi,\xi) = \sum_{jk} g_{jk}(p,x)\pi^j\xi^k$, with known coefficients $g_{jk}(p,x)$. Assuming for f the same form with unknown coefficients $f_{jk}(p,x)$, and taking h_Ω in complex variables as in (2.7), we compute

$$L_{h_\Omega} f = i\sum_{jk}(j-k)\cdot\Omega f_{jk}(p,x)\pi^j\xi^k \ .$$

Then, the solution of the equation (11.1) above in Π_Λ exists and is given by

$$f_{jk}(p,x) = -\frac{i}{(j-k)\cdot\Omega} g_{jk}(p,x) \ , \tag{11.2}$$

provided the polynomial expansion of $g(p,x,\pi,\xi)$ in π,ξ contains no monomials $\pi^j\xi^k$ such that $(j-k)\cdot\Omega = 0$, i.e such that $(j-k) \in \mathcal{M}_\Omega$, the resonance module associated to h_Ω defined by (2.10). This solution is unique up to a term $\overline{g}(p,x,\pi,\xi) \in \Pi_\Lambda$ such that $L_{h_\Omega}\overline{g} = 0$, i.e. such that the polynomial expansion of $\overline{g}(p,x,\pi,\xi)$ in π,ξ contains only monomials $\pi^j\xi^k$ with $j-k \in \mathcal{M}_\Omega$. However, we will not need to introduce such arbitrary terms, and the solution will be made unique by simply taking $\overline{g}(p,x,\pi,\xi) = 0$.

Coming back to eq. $L_{h_\Omega}\chi_s + Z_s = \Psi_s$, the obvious solution is to consider the polynomial expansion of Ψ_s in π,ξ, and to split Ψ_s into $\overline{\Psi}_s + \tilde{\Psi}_s$, by putting in $\overline{\Psi}_s$ (the and (5.6) of δ, $\overline{\delta}$ and δ_*, one has, for $(p',x',\pi',\xi')\in\mathcal{D}_{R-d,\varrho}$,

$$\begin{aligned}|T_\chi f(p',x',\pi',\xi',\varepsilon)| &\le \frac{2\Phi}{d^2}N_R(f)\sum_{r\ge0}\left(\frac{2e^2\Phi}{d^2}+\beta\right)^{r-1}\delta^{2r+\Lambda}\overline{\delta}^{2r+K-\Lambda}\\ &\le \delta^\Lambda\overline{\delta}^{K-\Lambda}N_R(f)\sum_{r\ge0}\left(\frac{2e^2\Phi}{d^2}+\beta\right)^{r}(\delta\overline{\delta})^{2r}\\ &\le \delta^\Lambda\overline{\delta}^{K-\Lambda}\left(1-\frac{\delta^2\overline{\delta}^2}{\delta_*^2}\right)^{-1}N_R(f)\ .\end{aligned}\tag{10.20}$$

On the other hand, it is an easy matter to see that if $|f(p,x,\pi,\xi,\varepsilon)|\le a$ on the domain $\mathcal{D}_{R,\varrho}$, then there exists a positive constant C' such that $N_R(f)\le C'a$, so that

$$|T_\chi f(p',x',\pi',\xi',\varepsilon)|<Ca\ ,\quad C=2\delta^\Lambda\overline{\delta}^{K-\Lambda}C'\ .\tag{10.21}$$

The actual size of C' is not relevant in what follows. Coming to the formula (5.8), namely the exchange theorem, denote by $D_\alpha(\overline{p},\overline{x})$ a polydisk of radius α in the variables (p,x) with center in $(\overline{p},\overline{x})$, and consider the power series expansion of $f(p,x,\pi,\xi,\varepsilon)$ in the polydisk $D_R(\overline{p},\overline{x})\times\Delta_{\varrho R}$, with $(\overline{p},\overline{x})\in\mathcal{G}$. This means that we can write

$$f(p,x,\pi,\xi,\varepsilon)=\sum_{l\ge0}f^{(l)}(p,x,\pi,\xi,\varepsilon)\ ,\tag{10.22}$$

$f^{(l)}$ being an homogeneous polynomial in the variables p,x,π,ξ, and also a function of class $\mathcal{P}_{\Lambda,K}$, and such series convergent in the domain $D_R(\overline{p},\overline{x})\times\Delta_{\varrho R}$, where f is analytic by hypothesis. We prove now that at every point (p',x',π',ξ') of the polydisk $D_{R-d}(\overline{p},\overline{x})\times\Delta_{\varrho(R-d)}$ one has

$$\big(T_\chi f\big)(p',x',\pi',\xi',\varepsilon)=\sum_{l\ge0}\big(T_\chi f^{(l)}\big)(p',x',\pi',\xi',\varepsilon)\ .\tag{10.23}$$

Indeed, the convergence of (10.22) means that for any $\eta>0$ there exists $\overline{L}=\overline{L}(\eta)$ such that for $L\ge\overline{L}$ one has $\big|f(p,x,\pi,\xi,\varepsilon)-\sum_{l\le L}f^{(l)}(p,x,\pi,\xi,\varepsilon)\big|<\eta$ for $(p,x,\pi,\xi)\in D_R\times\Delta_{\varrho R}$, and we can use the fact that $f-\sum_{l\le L}f^{(l)}$ is a function of class $\mathcal{P}_{\Lambda,K}$, the linearity of T_χ and (10.21) to compute

$$\begin{aligned}&\left|(T_\chi f)(p',x',\pi',\xi',\varepsilon)-\sum_{l\le L}\left(T_\chi f^{(l)}\right)(p',x',\pi',\xi',\varepsilon)\right|\\ &\qquad=\left|\left[T_\chi\left(f-\sum_{l\le L}f^{(l)}\right)\right](p',x',\pi',\xi',\varepsilon)\right|<C\eta\ ,\end{aligned}$$

It is now an easy matter to check that the latter two terms do not exceed $\beta\tilde{B}_{r-1}$, so that one has

$$\tilde{B}_r \le \left(\frac{2e^2\Phi}{d^2} + \beta\right) \tilde{B}_{r-1} \ .$$

Moreover, it is immediately seen that this holds also for $r = 2$, so that one has

$$\tilde{B}_r \le \left(\frac{2e^2\Phi}{d^2} + \beta\right)^{r-1} \tilde{B}_1 \ ,$$

and the last of (10.16) follows from the definition (10.18) of $\tilde{B}_1$. The remaining inequalities in (10.16) are obtained in essentially the same way. The lemma is thus proven.

We can now come to the

Proof of theorem 5.1. As in lemma 10.2, denote by z any of the canonical coordinates p, x and by ζ any of the coordinates π, ξ, and recall that z_r is of class $\mathcal{P}_{2r+2,4r}$ and ζ_r of class $\mathcal{P}_{2r+1,4r-1}$. In order to investigate the convergence of the series defining $T_\chi z$ and $T_\chi \zeta$ we follow the procedure illustrated at the end of sect. 4, making use of the explicit estimates on the norms given by lemma 10.3 and the definitions (5.5) and (5.6) of δ, $\overline{\delta}$ and δ_*, and get

$$\begin{aligned} |T_\chi z - z| &\le \frac{\delta^4\Phi}{d} \sum_{r\ge 1} \left(\frac{2e^2\Phi}{d^2} + \beta\right)^{r-1} \delta^{2(r-1)}\overline{\delta}^{2(r-1)} \\ &\le \frac{\delta^4\Phi}{d} \left(1 - \frac{\delta^2\overline{\delta}^2}{\delta_*^2}\right)^{-1} \ , \\ |T_\chi \zeta - \zeta| &\le \frac{\delta^3\Phi}{d} \sum_{r\ge 1} \left(\frac{2e^2\Phi}{d^2} + \beta\right)^{r-1} \delta^{2(r-1)}\overline{\delta}^{2(r-1)} \\ &\le \frac{\delta^3\Phi}{d} \left(1 - \frac{\delta^2\overline{\delta}^2}{\delta_*^2}\right)^{-1} \ , \end{aligned}$$

provided $\delta\overline{\delta} < \delta_*$. This shows that the canonical transformation defined by T_χ is analytic. Using now the conditions $\varrho \ge \varepsilon/e^2$ and $\delta\overline{\delta} \le \delta_*/\sqrt{2}$ one easily checks that the inequalities (5.7) hold, and moreover that $\frac{2\delta^4\Phi}{d} < d$ and $\frac{2\delta^3\Phi}{d} < \varrho d$, so that $\mathcal{D}_{R-2d,\varrho} \subset T_\chi(\mathcal{D}_{R-d,\varrho}) \subset \mathcal{D}_{R,\varrho}$. This concludes the proof.

We come now to the

Proof of theorem 5.2. First, we prove that $T_\chi f$ is analytic in $\mathcal{D}_{R-d,\varrho}$. Still applying the procedure illustrated at the end of sect. 4, by lemma 10.3 and the definitions (5.5)

Then, for any positive $d \le R$ and for $r \ge 1$ one has the bounds

$$
\begin{aligned}
N_{R-d}(z_r) &\le \left(\frac{2e^2\Phi}{d^2}+\beta\right)^{r-1} \frac{\Phi}{d}\ ,\\
N_{R-d}(\zeta_r) &\le \left(\frac{2e^2\Phi}{d^2}+\beta\right)^{r-1} \frac{\Phi}{d}\ ,\\
N_{R-d}(I_r) &\le \left(\frac{2e^2\Phi}{d^2}+\beta\right)^{r-1} \frac{\Phi}{d^2}\|I_\mu\|_R\ ,\\
N_{R-d}(\hat{h}_r) &\le \left(\frac{2e^2\Phi}{d^2}+\beta\right)^{r-1} \frac{\Phi}{d^2}|\hat{h}|_R\ ,\\
N_{R-d}(f_r) &\le \left(\frac{2e^2\Phi}{d^2}+\beta\right)^{r-1} \frac{2\Phi}{d^2}N_R(f)\ ,
\end{aligned}
\tag{10.16}
$$

Proof. Using the hypothesis on χ_l in the recursive formula (10.13) one immediately has, by lemma 10.2,

$$
\begin{aligned}
&N_{R-d}(z_r)\le \tilde{G}_r\ ,\quad N_{R-d}(\zeta_r)\le \tilde{\Gamma}_r\ ,\quad N_{R-d}(I_r)\le \tilde{C}_r\ ,\\
&N_{R-d}(\hat{h}_r)\le \tilde{D}_r\ ,\quad N_{R-d}(f_r)\le \tilde{B}_r\ ,
\end{aligned}
\tag{10.17}
$$

with

$$
\begin{aligned}
\tilde{G}_1 &= \frac{\Phi}{rd}\ , & \tilde{G}_r &= \frac{2e^2\Phi}{rd^2}\sum_{j=1}^{r}(r-j+1)\beta^{j-1}\tilde{G}_{r-j} + \frac{\Phi}{rd}\beta^{r-1}\ ,\\
\tilde{\Gamma}_1 &= \frac{\Phi}{d}\ , & \tilde{\Gamma}_r &= \frac{2e^2\Phi}{rd^2}\sum_{j=1}^{r}(r-j+1)\beta^{j-1}\tilde{\Gamma}_{r-j} + \frac{\Phi}{rd}\beta^{r-1}\ ,\\
\tilde{C}_1 &= \frac{\Phi}{d^2}\|I_\mu\|_R\ , & \tilde{C}_r &= \frac{2e^2\Phi}{rd^2}\sum_{j=1}^{r}(r-j+1)\beta^{j-1}\tilde{C}_{r-j} + \frac{\Phi}{rd^2}\beta^{r-1}\|I_\mu\|_R\ ,\\
\tilde{D}_1 &= \frac{\Phi}{d^2}|\hat{h}|_R\ , & \tilde{D}_r &= \frac{2e^2\Phi}{rd^2}\sum_{j=1}^{r}(r-j+1)\beta^{j-1}\tilde{D}_{r-j} + \frac{\Phi}{rd^2}\beta^{r-1}|\hat{h}|_R\ ,\\
\tilde{B}_1 &= \frac{2\Phi}{d^2}N_R(f)\ , & \tilde{B}_r &= \frac{2e^2\Phi}{rd^2}\sum_{j=1}^{r}(r-j+1)\beta^{j-1}\tilde{B}_{r-j} + \frac{2\Phi}{rd^2}\beta^{r-1}N_R(f)\ ,
\end{aligned}
\tag{10.18}
$$

Consider now the last of (10.18), and take $r \ge 3$. By isolating the term $j=1$ in the sum, and replacing j by $j+1$ in the remaining ones we obtain

$$
\begin{aligned}
\tilde{B}_r &= \frac{2e^2\Phi}{d^2}\tilde{B}_{r-1}\\
&\quad + \frac{2\beta e^2\Phi}{rd^2}\sum_{j=1}^{r-2}(r-j)\beta^{j-1}\tilde{B}_{r-1-j} + \frac{2\Phi}{rd^2}\beta^{r-1}N_R(f)\ .
\end{aligned}
\tag{10.19}
$$

most r operators L_{χ_j}, and c_α a suitable set of coefficients. Thus, from lemma 10.1 on the repeated application of Poisson brackets, one has $N_{R-d}(E_\alpha f) \le A_\alpha$, where A_α can be explicitly computed via the recursive formula (10.10) once the actual factorization of E_α is known. So, one has $N_{R-d}(f_r) \le \sum_\alpha |c_\alpha| A_\alpha$. We look now for a recursive bound of the latter quantity. If we assume that the actual factorization of $f_s = \sum_\beta E_\beta f$ is known for $1 \le s < r$, which is obviously true for $r = 2$, then the definition (10.15) immediately gives the factorization of f_r (notice that the set of indexes β actually depends on s). So, in order to obtain a recursive bound, we assume that the constants A_β are known, and that $\sum_\beta |c_\beta| A_\beta \le B_s$, with known constants B_s, and use lemma 10.1 to evaluate a single term in the factorization of $L_{\chi_j} f_{r-j} = \sum_\beta c_\beta L_{\chi_j} E_\beta f$. This gives

$$N_{R-d}(L_{\chi_j} E_\beta f) \le \frac{2e^2(r-j+1)}{d^2} N_R(\chi_j) A_\beta \ .$$

Here use has been made of the fact that the only relevant informations needed in applying the recursive formula (10.10) are the number of Poisson brackets in the actual factorization of E_β, in our case at most $r-j$, and the norm of χ_j. The contribution of $L_{\chi_j} f_{r-j}$ to $\sum_\alpha |c_\alpha| A_\alpha$ can then be estimated by simply performing the sum over the set of indexes β, which is a trivial task, since A_β is multiplied by a coefficient which does not depend on β, and gives

$$\begin{aligned} N_{R-d}(L_{\chi_j} f_{r-j}) &\le \frac{2e^2(r-j+1)}{d^2} N_R(\chi_j) \sum_\beta |c_\beta| A_\beta \\ &\le \frac{2e^2(r-j+1)}{d^2} N_R(\chi_j) B_{r-j} \end{aligned}$$

Finally, we perform the sum over j which appears in (10.15), and add the estimate for $L_{\chi_r} f$, and obtain $\sum_\alpha |c_\alpha| A_\alpha \le B_r$, with B_r defined by (10.14). The lemma is thus proven.

We specialize now the contents of lemma 10.2 to the case, assumed in theorem 5.1, in which the generating sequence $\{\chi_l\}_{l\ge1}$ is characterized by the condition that $N_R(\chi_l) \le \frac{\beta^{l-1}}{l}\Phi$ for some real positive constants β and Φ. We have the

Lemma 10.3: *Consider the generating sequence $\{\chi_l\}_{l\ge1}$, the functions z, ζ, $I_\mu(\pi,\xi)$, $\hat h(p,x)$ and f, and the sequences $\{z_r\}_{r\ge0}$, $\{\zeta_r\}_{r\ge0}$, $\{I_r\}_{r\ge0}$, $\{\hat h_r\}_{r\ge0}$, and $\{f_r\}_{r\ge0}$ as in lemma 10.2. Assume moreover that there exist real positive constants β and Φ such that $N_R(\chi_l) \le \frac{\beta^{l-1}}{l}\Phi$ for $l \ge 1$.*

fined by

$$
\begin{aligned}
G_1 &= \frac{1}{d} N_R(\chi_1) \ , \\
\Gamma_1 &= \frac{1}{d} N_R(\chi_1) \ , \\
C_1 &= \frac{1}{d^2} N_R(\chi_1) \|I_\mu\|_R \ , \\
D_1 &= \frac{1}{d^2} N_R(\chi_1) |\hat{h}|_R \ , \\
B_1 &= \frac{2}{d^2} N_R(\chi_1) N_R(f) \ ,
\end{aligned}
\tag{10.13}
$$

and

$$
\begin{aligned}
G_r &= \frac{2e^2}{rd^2} \sum_{j=1}^{r} j(r-j+1) N_R(\chi_j) G_{r-j} + \frac{1}{d} N_R(\chi_r) \\
\Gamma_r &= \frac{2e^2}{rd^2} \sum_{j=1}^{r} j(r-j+1) N_R(\chi_j) \Gamma_{r-j} + \frac{1}{d} N_R(\chi_r) \\
C_r &= \frac{2e^2}{rd^2} \sum_{j=1}^{r} j(r-j+1) N_R(\chi_j) C_{r-j} + \frac{1}{d^2} N_R(\chi_r) \|I_\mu\|_R \\
D_r &= \frac{2e^2}{rd^2} \sum_{j=1}^{r} j(r-j+1) N_R(\chi_j) D_{r-j} + \frac{1}{d^2} N_R(\chi_r) |\hat{h}|_R \\
B_r &= \frac{2e^2}{rd^2} \sum_{j=1}^{r} j(r-j+1) N_R(\chi_j) B_{r-j} + \frac{2}{d^2} N_R(\chi_r) N_R(f)
\end{aligned}
\tag{10.14}
$$

Proof. That z_r is of class $\mathcal{P}_{2r+2,4r}$, ζ_r is of class $\mathcal{P}_{2r+1,4r-1}$, I_r is of class $\mathcal{P}_{2r+2,4r}$, $\hat{h}_r$ is of class $\mathcal{P}_{2r+4,4r+2}$ and f_r is of class $\mathcal{P}_{2r+\Lambda,4r+K}$ easily follows from lemma 10.1 So, we concentrate on the bounds.

Writing the recursive definition (5.2) in the equivalent form

$$
f_1 = L_{\chi_1} f \ , \quad f_r = \sum_{j=1}^{r-1} \frac{j}{r} L_{\chi_j} f_{r-j} + L_{\chi_r} f \ , \quad r > 1 \tag{10.15}
$$

we immediately find G_1, Γ_1, C_1, D_1 and B_1 from lemma 10.1. The same holds for the last term in the r.h.s of the recursive definition of G_r, Γ_r, C_r, D_r and B_r. So, we only need to discuss the first term in the r.h.s. of (10.14), which is the really recursive part of such formulae. This has the same form for all the sequences, so we only consider the sequence for B_r, i.e. the last one. Looking at (10.15) one immediately sees that $f_r = \sum_\alpha c_\alpha E_\alpha f$ for a suitable set of indices α, where each E_α is a composition of at

Such inequality can be applied s times, and gives the explicit estimate

$$N_{R-d}(\eta_s) \le s^2 A_1 \prod_{l=2}^{s} \frac{2s^2}{ld^2} N_R(\vartheta_l) \ , \tag{10.11}$$

which holds for any $s > 0$. It is now immediate to check that (10.10) holds for $s = 2$. To prove it for $s \ge 3$ we use induction. Write the r.h.s of (10.11) by isolating the term $l = s$ in the product, and compute

$$\begin{aligned} &\frac{2s}{d^2} N_R(\vartheta_s) \cdot s^2 A_1 \prod_{l=2}^{s-1} \frac{2s^2}{ld^2} N_R(\vartheta_l) \\ &\quad = \left(\frac{s}{s-1}\right)^{2(s-1)} \frac{2s}{d^2} N_R(\vartheta_s) \cdot s^2 A_1 \prod_{l=2}^{s-1} \frac{2(s-1)^2}{ld^2} N_R(\vartheta_l) \\ &\quad < \frac{2se^2}{d^2} N_R(\vartheta_s) A_{s-1} \ . \end{aligned}$$

Here the inequality $\left(\frac{s}{s-1}\right)^{2(s-1)} < e^2$ was used. This concludes the proof of the lemma.

We can now come back to the operator T_χ, and prove the

Lemma 10.2: *For a given generating sequence $\{\chi_s\}_{s\ge1}$ with χ_s of class $\mathcal{P}_{2s+2,4s}$ on a domain $\mathcal{G}_R$, and for functions z, ζ, $I_\mu(\pi,\xi)$, $\hat{h}(p,x)$ and $f(p,x,\pi,\xi,\varepsilon)$ as in lemma 10.1, with $h_\Omega(\pi,\xi)$ of class $\mathcal{P}_{2,2}$, $\hat{h}(p,x)$ of class $\mathcal{P}_{0,0}$ and $f(p,x,\pi,\xi,\varepsilon)$ of class $\mathcal{P}_{\Lambda,K}$ on the same domain $\mathcal{G}_R$, the sequences $\{z_r\}_{r\ge0}$, $\{\zeta_r\}_{r\ge0}$, $\{I_r\}_{r\ge0}$, $\{\hat{h}_r\}_{r\ge0}$, and $\{f_r\}_{r\ge0}$ defining the transformed functions under the operator T_χ have the following properties for $0 < d < R$ and $r \ge 1$:*

$$\begin{array}{llll} i. & z_r \text{ is of class } \mathcal{P}_{2r+2,4r} \ , & \text{and } N_{R-d}(z_r) \le G_r \ , & \\ ii. & \zeta_r \text{ is of class } \mathcal{P}_{2r+1,4r-1} \ , & \text{and } N_{R-d}(\zeta_r) \le \Gamma_r \ , & \\ iii. & I_r \text{ is of class } \mathcal{P}_{2r+2,4r} \ , & \text{and } N_{R-d}(I_r) \le C_r \ , & (10.12) \\ iv. & \hat{h}_r \text{ is of class } \mathcal{P}_{2r+2,4r} \ , & \text{and } N_{R-d}(\hat{h}_r) \le D_r \ , & \\ v. & f_r \text{ is of class } \mathcal{P}_{2r+\Lambda,4r+K} \ , & \text{and } N_{R-d}(f_r) \le B_r \ . & \end{array}$$

The sequences $\{G_r\}_{r\ge1}$, $\{\Gamma_r\}_{r\ge1}$, $\{C_r\}_{r\ge1}$, $\{D_r\}_{r\ge1}$ and $\{B\}_{r\ge1}$ are recursively de-

domain $\mathcal{G}_R$. *Let* g *be any of the functions* z, ζ, $I_\mu(\pi,\xi)$, $\hat{h}(p,x)$ *and* $f(p,x,\pi,\xi,\varepsilon)$ *above, and consider the sequence* $\{\eta_s\}_{s\geq 0}$ *of functions recursively defined by*

$$\eta_0 = g \ , \quad \eta_s = L_{\vartheta_s}\eta_{s-1} \ . \tag{10.7}$$

Then, for $s \geq 1$ *one can determine three sequences* $\{M_s\}_{s\geq 1}$, $\{N_s\}_{s\geq 1}$ *and* $\{A_s\}_{s\geq 1}$ *such that* η_s *is of class* $\mathcal{P}_{M_s,N_s}$, *and, for any positive* $d < R$, *one has the estimate*

$$N_{R-d}(\eta_s) \leq A_s \ . \tag{10.8}$$

The sequences $\{M_s\}_{s\geq 1}$, $\{N_s\}_{s\geq 1}$ *and* $\{A_s\}_{s\geq 1}$ *are recursively defined as follows:*

$$\begin{array}{llll}
\text{i. if } g = z, & \text{then } M_1 = \Lambda_1 \ , & N_1 = K_1 \ , & A_1 = \frac{1}{d} N_R(\vartheta_1) \\
\text{ii. if } g = \zeta, & \text{then } M_1 = \Lambda_1 - 1 \ , & N_1 = K_1 - 1 \ , & A_1 = \frac{1}{d} N_R(\vartheta_1) \\
\text{iii. if } g = I_\mu, & \text{then } M_1 = \Lambda_1 \ , & N_1 = K_1 \ , & A_1 = \frac{1}{d^2} N_R(\vartheta_1) \|I_\mu\|_R \\
\text{iv. if } g = \hat{h}, & \text{then } M_1 = \Lambda_1 \ , & N_1 = K_1 \ , & A_1 = \frac{1}{d^2} N_R(\vartheta_1) |\hat{h}|_R \\
\text{v. if } g = f, & \text{then } M_1 = \Lambda_1 + \Lambda - 2 \ , & N_1 = K_1 + K \ , & A_1 = \frac{2}{d^2} N_R(\vartheta_1) N_R(f)
\end{array} \tag{10.9}$$

and, for $s \geq 2$,

$$\begin{aligned}
M_s &= M_{s-1} + \Lambda_s - 2 \ , \quad N_s = N_{s-1} + K_s \ , \\
A_s &= \frac{2se^2}{d^2} N_R(\vartheta_s) A_{s-1} \ .
\end{aligned} \tag{10.10}$$

Proof. The fact that η_s is of class $\mathcal{P}_{M_s,N_s}$ immediately follows from the definition of M_s and N_s, and from the algebraic properties enunciated in sect. 4. Moreover, for $s = 1$ the inequalities (10.9) are trivial consequences of lemmas 9.4 and 9.6, taking into account that the Poisson brackets $\{\vartheta_1, z\}$ and $\{\vartheta_1, \zeta\}$ are nothing but derivatives of ϑ_1, and recalling that I_μ and $\hat{h}$ depend only on π, ξ and p, x respectively. So, we fix $s > 1$ and d, and introduce the sequence of domains $\mathcal{G}_{R-l\tilde{d}}$, with $\tilde{d} = d/s$, for $0 \leq l \leq s$. We start with $N_{R-\tilde{d}}(\eta_1) \leq s^2 A_1$, where A_1 is known by (10.9) and the factor s^2 has been introduced because the bound holds in the domain $\mathcal{G}_{R-\tilde{d}}$, and, for $2 \leq l \leq s$, we use lemma 9.6 with $(l-1)\tilde{d}$ in place of d' and $\tilde{d}$ in place of d, and obtain

$$N_{R-l\tilde{d}}(\eta_l) \leq \frac{2}{l\tilde{d}^2} N_R(\vartheta_l) N_{R-(l-1)\tilde{d}}(\eta_{l-1})$$

iii. f has the form $I_\mu(\pi,\xi)=\sum_{l=1}^{\nu}\frac{\mu_l}{2}\left(\pi_l^2+\xi_l^2\right)$, with $0\neq\mu\in\mathbf{R}^\nu$. In such case the transformed function will be denoted by

$$T_\chi I_\mu=\sum_{r\geq 0} I_r\ . \tag{10.3}$$

This is similar to considering a function of the variables π,ξ only. In particular the unperturbed Hamiltonian h_Ω has exactly such form, with $\mu=\Omega$.

iv. f coincides with $\hat h(p,x)$. In such case the transformed function will be denoted by

$$T_\chi\hat h=\sum_{r\geq 0}\hat h_r\ . \tag{10.4}$$

This is similar to considering a function of the variables p,x only, i.e. a function of class $\mathcal{P}_{0,0}$.

v. f is a generic function $f(p,x,\pi,\xi,\varepsilon)$ of class $\mathcal{P}_{\Lambda,K}$ for some $K\geq\Lambda\geq 0$. The transformed function will be denoted by

$$T_\chi f=\sum_{r\geq 0} f_r\ . \tag{10.5}$$

Here, the sequences $\{z_r\}_{r\geq0}$, $\{\zeta_r\}_{r\geq0}$, $\{I_r\}_{r\geq0}$, $\{\hat h_r\}_{r\geq0}$, and $\{f_r\}_{r\geq0}$ are recursively defined as in (5.2)., i.e.

$$\begin{aligned}
z_0&=z\ ,\quad & z_r&=\sum_{j=1}^{r}\frac{j}{r}L_{\chi_j}z_{r-j}\ ,\\
\zeta_0&=\zeta\ ,\quad & \zeta_r&=\sum_{j=1}^{r}\frac{j}{r}L_{\chi_j}\zeta_{r-j}\ ,\\
I_0&=I\ ,\quad & I_r&=\sum_{j=1}^{r}\frac{j}{r}L_{\chi_j}I_{r-j}\ ,\\
\hat h_0&=\hat h\ ,\quad & \hat h_r&=\sum_{j=1}^{r}\frac{j}{r}L_{\chi_j}\hat h_{r-j}\ ,\\
f_0&=f\ ,\quad & f_r&=\sum_{j=1}^{r}\frac{j}{r}L_{\chi_j}f_{r-j}\ .
\end{aligned} \tag{10.6}$$

Before coming to the estimates on the sequences $\{z_r\}_{r\geq0}$, $\{\zeta_r\}_{r\geq0}$, $\{I_r\}_{r\geq0}$, $\{\hat h_r\}_{r\geq0}$, and $\{f_r\}_{r\geq0}$, we generalize the lemma 9.6 by giving the estimate for the repeated application of the Poisson bracket.

Lemma 10.1: *Let $\{\Lambda_l\}_{l\geq1}$ and $\{K_l\}_{l\geq1}$, with $K_l\geq\Lambda_l\geq 2$, be sequences of integers, and $\{\vartheta_l\}_{l\geq1}$ a sequence of functions, with ϑ_l of class $\mathcal{P}_{\Lambda_l,K_l}$ on a common*

Lemma 9.6: *For two functions f of class $\mathcal{P}_{\Lambda,K}$ on the domain $\mathcal{G}_R$ and f' of class $\mathcal{P}_{\Lambda',K'}$ on the domain $\mathcal{G}_{R-d'}$ with $0 \le d' < R$, and for any positive $d < R - d'$ one has*

$$\begin{aligned} N_{R-d'-d}(\{f,f'\}_{p,x}) &\le \frac{1}{d(d'+d)} N_R(f) N_{R-d'}(f') \\ N_{R-d'-d}(\{f,f'\}_{\pi,\xi}) &\le \frac{1}{d(d'+d)} N_R(f) N_{R-d'}(f') \\ N_{R-d'-d}(\{f,f'\}) &\le \frac{2}{d(d'+d)} N_R(f) N_{R-d'}(f') \end{aligned} \tag{9.10}$$

Proof. Using the form (4.9) for f and f', compute

$$\{f,f'\}_{p,x} = \sum_{\Lambda \le l+m \le K} \sum_{\Lambda' \le l'+m' \le K'} \varepsilon^{l+l'} \{f^{(l,m)}, f'^{(l',m')}\}_{p,x} \ ;$$

then, by using the definition (4.14) of the norm and lemma 9.5, one has

$$\begin{aligned} N_{R-d'-d}(\{f,f'\}_{p,x}) &\le \sum_{l,m} \sum_{l',m'} \left\| \{f^{(l,m)}, f'^{(l',m')}\}_{p,x} \right\|_{R-d'-d} \\ &\le \frac{1}{d(d'+d)} \sum_{l,m} \left\| f^{(l,m)} \right\|_R \cdot \sum_{l',m'} \left\| f'^{(l'm')} \right\|_{R-d'} \ , \end{aligned}$$

and the first of (9.10) follows from the definition of the norm. The second inequality is deduced by essentially the same computation, and the third one immediately follows from the previous ones by $\{f,f'\} = \{f,f'\}_{p,x} + \{f,f'\}_{\pi,\xi}$, so that the lemma is proven.

10. Proof of the theorems on canonical transformations

In order to prove the theorems on canonical transformations of sect. 5 we must estimate the application of the operator T_χ to a function $f(p,x,\pi,\xi,\varepsilon)$. We shall consider the following cases.

i. f is one of the canonical coordinates p, x. In such case we shall denote by z any of these coordinates, and the transformation will be denoted by

$$T_\chi z = \sum_{r \ge 0} z_r \ . \tag{10.1}$$

ii. f is one of the canonical coordinates π, ξ. In such case we shall denote by ζ any of these coordinates, and the transformation will be denoted by

$$T_\chi \zeta = \sum_{r \ge 0} \zeta_r \ . \tag{10.2}$$

Lemma 9.5: *For two functions $f \in \Pi_\Lambda$ on the domain $\mathcal{G}_R$ and $f' \in \Pi_{\Lambda'}$ on the domain $\mathcal{G}_{R-d'}$ with $0 \le d' < R$, and for any positive $d < R - d'$ one has*

$$\begin{aligned}
\left\| \{f, f'\}_{p,x} \right\|_{R-d'-d} &\le \frac{1}{d(d'+d)} \|f\|_R \|f'\|_{R-d'} \ , \\
\left\| \{f, f'\}_{\pi,\xi} \right\|_{R-d'-d} &\le \frac{1}{d(d'+d)} \|f\|_R \|f'\|_{R-d'} \ .
\end{aligned} \tag{9.9}$$

Proof. Using the form (4.7) for f and f', compute

$$\{f, f'\}_{p,x} = \sum_{jkj'k'} \{f_{jk}, f'_{j'k'}\}_{p,x} \pi^{j+j'} \xi^{k+k'} \ ;$$

then, by the definition (4.13) of the norm and by (9.8), one has

$$\begin{aligned}
\|\{f, f'\}_{p,x}\|_{R-d'-d} &\le (R - d' - d)^{\Lambda+\Lambda'} \sum_{jkj'k'} \left|\{f_{jk}, f'_{j'k'}\}_{p,x}\right|_{R-d'-d} \\
&\le \frac{1}{d(d'+d)} \left[R^\Lambda \sum_{jk} |f_{jk}|_R \right] \cdot \left[(R-d')^{\Lambda'} \sum_{j'k'} \left|f'_{j'k'}\right|_{R-d'} \right] \ ,
\end{aligned}$$

and the first of (9.9) follows from the definition of the norm. In order to prove the second one compute

$$\{f, f'\}_{\pi,\xi} = \sum_{jkj'k'} f_{jk} f'_{j'k'} \sum_{l=1}^{\nu} \frac{j_l k'_l - j'_l k_l}{\pi_l \xi_l} \pi^{j+j'} \xi^{k+k'} \ ;$$

then, by the definition (4.13) of the norm, one has

$$\begin{aligned}
\left\| \{f, f'\}_{\pi,\xi} \right\|_{R-d'-d} &\le (R - d' - d)^{\Lambda+\Lambda'-2} \sum_{jkj'k'} |f_{jk}|_R \left|f'_{j'k'}\right|_{R-d'} \sum_l (j_l k'_l + j'_l k_l) \\
&\le \left[\Lambda (R - d' - d)^{\Lambda-1} \sum_{jk} |f_{jk}|_R \right] \cdot \left[\Lambda' (R - d' - d)^{\Lambda'-1} \sum_{j'k'} \left|f'_{j'k'}\right|_{R-d'} \right] \\
&\le \left[\frac{R^\Lambda}{d'+d} \sum_{jk} |f_{jk}|_R \right] \cdot \left[\frac{(R-d')^{\Lambda'}}{d} \sum_{j'k'} \left|f'_{j'k'}\right|_{R-d'} \right] \ ,
\end{aligned}$$

and the second of (9.9) follows from the definition of the norm. Here the inequalities $\sum_{l=1}^{\nu} (j_l k'_l + j'_l k_l) \le \Lambda \sum_l (k'_l + j'_l) \le \Lambda\Lambda'$ and $s(R-\delta)^{s-1} \le \frac{R^s}{\delta}$ for $0 < \delta < R$ and $s \ge 1$ were used. The lemma is thus proven.

Such a result is extended to functions of class $\mathcal{P}_{\Lambda,K}$ by the following i.e. the first of (9.6); using now the trivial inequality $\Lambda(R-d)^{\Lambda-1} \le \frac{R^\Lambda}{d}$ for $0<d<R$ and $\Lambda \ge 1$, one has

$$\left\| \frac{\partial f}{\partial \pi_l} \right\|_{R-d} \le (R-d)^{\Lambda-1} \sum_{j,k} j_l \, |f_{jk}|_R \le \Lambda (R-d)^{\Lambda-1} \sum_{j,k} |f_{jk}|_R \le \frac{1}{d} R^\Lambda \sum_{j,k} |f_{jk}|_R \ ,$$

i.e., the third of (9.6). The derivatives with respect to x and ξ are bounded in exactly the same way, so that the lemma is proven.

Such a result is extended to functions of class $\mathcal{P}_{\Lambda,K}$ by the

Lemma 9.4: *For a function f of class $\mathcal{P}_{\Lambda,K}$ on the domain $\mathcal{G}_R$ one has, for $0<d<R$,*

$$\begin{aligned} N_{R-d}\left(\frac{\partial f}{\partial p_l}\right) &\le \frac{1}{d} N_R(f) \ , \quad N_{R-d}\left(\frac{\partial f}{\partial x_l}\right) \le \frac{1}{d} N_R(f) \ , \qquad 1 \le l \le n \ , \\ N_{R-d}\left(\frac{\partial f}{\partial \pi_l}\right) &\le \frac{1}{d} N_R(f) \ , \quad N_{R-d}\left(\frac{\partial f}{\partial \xi_l}\right) \le \frac{1}{d} N_R(f) \ , \qquad 1 \le l \le \nu \ . \end{aligned} \tag{9.7}$$

Proof. Using the form (4.9) for f, compute

$$\frac{\partial f}{\partial p_l} = \sum_{\Lambda \le j+m \le K} \varepsilon^j \frac{\partial f^{(j,m)}}{\partial p_l} \ ;$$

then, by the definition (4.14) of the norm and by lemma 9.3, one has

$$N_{R-d}\left(\frac{\partial f}{\partial p_l}\right) = \sum_{\Lambda \le j+m \le K} \left\| \frac{\partial f^{(j,m)}}{\partial p_l} \right\|_{R-d} \le \frac{1}{d} \sum_{\Lambda \le j+m \le K} \left\| f^{(j,m)} \right\|_R \ ,$$

and the first of (9.7) follows. The remaining inequalities are proven in exactly the same way. The lemma is thus proven.

Concerning the Poisson bracket, we first consider two analytic bounded functions $\varphi(p,x)$ on the domain $\mathcal{G}_R$ and $\varphi'(p,x)$ on the domain $\mathcal{G}_{R-d'}$, with $0 \le d' < R$. In such case we use the bound

$$|\{\varphi, \varphi'\}|_{R-d'-d} \le \frac{|\varphi|_R \, |\varphi'|_{R-d'}}{d(d'+d)} \ , \tag{9.8}$$

where $0 < d < R - d'$. This is nothing but a trivial generalization of lemma 2 in the first part of the paper.

Using such results we prove the

Lemma 9.2: *For a function $f(p, x, \pi, \xi, \varepsilon)$ of class $\mathcal{P}_{\Lambda,K}$ with $0 \le \Lambda \le K$ on the domain $\mathcal{G}_R$ the transformed function $\tilde{f}(p, x, \tilde{\pi}, \tilde{\xi}, \varepsilon)$ under the canonical transformation (2.8) is of class $\mathcal{P}_{\Lambda,K}$, and its norm is bounded by*

$$N_R(\tilde{f}) \le 2^{\frac{K}{2}} N_R(f) \ . \tag{9.3}$$

Conversely, for a function $\tilde{g}(p, x, \tilde{\pi}, \tilde{\xi}, \varepsilon)$ of class $\mathcal{P}_{\Lambda,K}$ on the domain $\mathcal{G}_R$ the transformed function $g(p, x, \pi, \xi, \varepsilon)$ under the inverse of the canonical transformation (2.8) also is of class $\mathcal{P}_{\Lambda,K}$, and its norm is bounded by

$$N_R(g) \le 2^{\frac{K}{2}} N_R(\tilde{g}) \ . \tag{9.4}$$

The proof is a straightforward application of lemma 9.1, by only taking into account the fact that f and g are polynomials of degree $\le K$ in the variables π, ξ and $\tilde{\pi}, \tilde{\xi}$ respectively.

We come now to the technical lemmas related to derivatives and Poisson brackets. In estimating the derivatives of the coefficients in the polynomial expansions in π, ξ we make use of Cauchy's inequality. Precisely, for an analytic bounded function $\varphi(p, x)$ on a domain $\mathcal{G}_R$ and for any positive $d < R$ Cauchy's inequality reads

$$\left|\frac{\partial \varphi}{\partial p_l}\right|_{R-d} \le \frac{|\varphi|_R}{d} \ , \quad \left|\frac{\partial \varphi}{\partial x_l}\right|_{R-d} \le \frac{|\varphi|_R}{d} \ . \tag{9.5}$$

Using those inequalities we can prove the

Lemma 9.3: *For a function $f \in \Pi_\Lambda$ on the domain $\mathcal{G}_R$ one has, for $0 < d < R$,*

$$\begin{aligned} \left\|\frac{\partial f}{\partial p_l}\right\|_{R-d} &\le \frac{1}{d}\|f\|_R \ , \quad \left\|\frac{\partial f}{\partial x_l}\right\|_{R-d} \le \frac{1}{d}\|f\|_R \ , \qquad 1 \le l \le n \ , \\ \left\|\frac{\partial f}{\partial \pi_l}\right\|_{R-d} &\le \frac{1}{d}\|f\|_R \ , \quad \left\|\frac{\partial f}{\partial \xi_l}\right\|_{R-d} \le \frac{1}{d}\|f\|_R \ , \qquad 1 \le l \le \nu \ . \end{aligned} \tag{9.6}$$

Proof. Using the form (4.7) for f, compute

$$\frac{\partial f}{\partial p_l} = \sum_{j,k} \frac{\partial f_{jk}}{\partial p_l} \pi^j \xi^k \ , \qquad \frac{\partial f}{\partial \pi_l} = \sum_{j,k} f_{jk} \frac{j_l}{\pi_l} \pi^j \xi^k \ ,$$

so that, by the definition (4.13) of the norm and the Cauchy's inequality (9.5), one has

$$\left\|\frac{\partial f}{\partial p_l}\right\|_{R-d} \le (R-d)^\Lambda \sum_{j,k} \left|\frac{\partial f_{jk}}{\partial p_l}\right|_{R-d} \le \frac{R^\Lambda}{d} \sum_{j,k} |f_{jk}|_R \ ,$$

Part C – Technical sections.

9. Technical lemmas

We prove here some technical lemmas used in the paper.

First, we consider the canonical transformation (2.8) to complex variables (π', ξ'). The norm of a function $f \in \Pi_\Lambda$ is changed according to the following

Lemma 9.1: *For a function $f(p,x,\pi,\xi) \in \Pi_\Lambda$ on the domain $\mathcal{G}_R$ the transformed function $\tilde{f}(p,x,\tilde{\pi},\tilde{\xi})$ under the canonical transformation (2.8) also belongs to Π_Λ, and its norm is bounded by*

$$\left\|\tilde{f}\right\|_R \le 2^{\frac{\Lambda}{2}} \|f\|_R \ . \tag{9.1}$$

Conversely, for a function $\tilde{g}(p,x,\tilde{\pi},\tilde{\xi}) \in \Pi_\Lambda$ on the domain $\mathcal{G}_R$ the transformed function $g(p,x,\pi,\xi)$ under the inverse of the canonical transformation (2.8) also belongs to Π_Λ, and its norm is bounded by

$$\|g\|_R \le 2^{\frac{\Lambda}{2}} \|\tilde{g}\|_R \ . \tag{9.2}$$

Proof. Take $f = \sum_{|j+k|=\Lambda} f_{jk}(p,x)\pi^j\xi^k$, and perform the substitution

$$\begin{aligned}
\tilde{f} &= \sum_{j,k} i^{|k|} 2^{-\frac{|j+k|}{2}} f_{jk}(p,x)(\tilde{\pi}+i\tilde{\xi})^j(\tilde{\pi}-i\tilde{\xi})^k \\
&= \sum_{j,k} i^{|k|} 2^{-\frac{|j+k|}{2}} f_{jk}(p,x) \prod_{l=1}^{\nu} \sum_{s=0}^{j_l} \binom{j_l}{s} i^{j_l-s} \sum_{r=0}^{k_l} \binom{k_l}{r} i^{k_l-r} \tilde{\pi}_l^{s+r} \tilde{\xi}_l^{j_l+k_l-s-r} \ ,
\end{aligned}$$

so that $\tilde{f} \in \Pi_\Lambda$, and using the definition (4.13) of the norm one has

$$\begin{aligned}
\left\|\tilde{f}\right\|_R &\le R^\Lambda \sum_{j,k} 2^{-\frac{|j+k|}{2}} |f_{jk}|_R \prod_{l=1}^{\nu} \sum_{s=0}^{j_l} \binom{j_l}{s} \sum_{r=0}^{k_l} \binom{k_l}{r} \\
&= R^\Lambda \sum_{j,k} 2^{\frac{|j+k|}{2}} |f_{jk}|_R \\
&= 2^{\frac{\Lambda}{2}} R^\Lambda \sum_{j,k} |f_{jk}|_R \ ,
\end{aligned}$$

so that (9.1) follows. The statement about the inverse transformation is proven by essentially the same computations, and this concludes the proof of the lemma.

Such result is immediately extended to a function of class $\mathcal{P}_{\Lambda,K}$ by the following and replacing E by $\hat{E}(p,x)$, as illustrated in sect. 2. The only point where the global constant E is needed is the choice of the optimal normalization order $r_{\rm opt}$ in sect. 7 (otherwise we would not obtain one and the same canonical transformation on the whole domain $\mathcal{D}_{R,\varrho_*}$). The proof then immediately follows.

For what concerns corollary 2.2, (2.20) follows from (7.17) by using $\|I_\mu\|_{R-d} = |\mu|\,(R-d)^2 = \frac{9}{16}\,|\mu|\,R^2$, and $\frac{9}{16}\mathcal{B} < 2e^{-\tau}$; (2.21) follows from (7.18) by substituting the explicit expressions for $\mathcal{A}$, d and $\|I_\mu\|_R$. This proves corollary 2.2.

Let's now come to corollaries 2.3 and 2.4. By the estimate (2.21), which holds in the domain $\mathcal{D}_{\frac{1}{2}R,\varrho}$, one immediately gets, for $|t| < \min(T_0, T)$,

$$|I_\mu(\pi'(t),\xi'(t)) - I_\mu(\pi'(0),\xi'(0))| < 2^6\lambda^{-2}\,|\mu|\,|\Omega|^{-1}\,E\frac{|t|}{T}\ , \tag{8.5}$$

and such estimate can be made similar to (2.20) by using the inequalities $\lambda > e^{\tau+1}\lambda_*$, $\lambda_*^{-1} < 2^{-9}\gamma R^2/E_0$ and $\gamma < |\Omega|$, from which (2.22) easily follows. The first of (2.25) is nothing but (8.5) with $\mu = \Omega$, and the second one follows from the energy conservation, by using

$$\begin{aligned}\left|\hat{h}(p'(t),x'(t)) - \hat{h}(p'(0),x'(0))\right| &< |\lambda h_\Omega(\pi'(t),\xi'(t)) - \lambda h_\Omega(\pi'(0),\xi'(0))| \\ &\quad + 2\,|Z(p',x',\pi',\xi',\lambda)| + 2\,|\mathcal{R}(p',x',\pi',\xi',\lambda)|\ ,\end{aligned}$$

and using the estimates (2.19) for Z and $\mathcal{R}$. This proves corollary 2.3. Next, (2.26) and the first of (2.27) are obtained by summing up the deformation due to the canonical transformation, which is estimated by (2.19) and (2.20), to the change above due to the noise of the remainder; the second of (2.27) is still obtained via the conservation of energy in the original Hamiltonian (2.16). This concludes the proof of the corollaries.

and this is the bound to be used in evaluating the norm of H_s in (4.6), by just computing the number of terms in the sum, which, since $|j|=l$ and $|k|=s$, does not exceed $\binom{\nu+l-1}{l}\binom{\nu+s-1}{s}\le\nu^{s+l}$. So, by the definition (4.14) of the norm, one has

$$N_R(H_s)<\sum_{l=0}^{L}\left[\frac{\nu R}{\eta}\right]^{s+l}F=\sigma^{s-1}E$$

provided one takes

$$\sigma=\frac{\nu R}{\eta}\ ,\quad E=\frac{1-\sigma^{L+1}}{1-\sigma}\sigma F\ . \tag{8.3}$$

These general estimates of the values of σ and E can be used in theorem 2.1.

Better values of these constants, mainly for what concerns the dependence on ν, can instead be found if one knows more about the perturbation f. For example, consider the case of a diatomic gas of ν identical molecules. Then $f(p,x,\pi,\xi)$, as well as $\hat{h}(p,x)$, can be written as the sum of $\nu(\nu-1)/2$ terms due to the interaction between pairs of molecules. Thus F, as well as E_0, turns out to be of order ν^2, while the number of terms in the sum (8.2) turns out not to exceed 2^{s+l}. Moreover, since the interaction does not depend on the momenta π, one has $K=0$, so that the values of σ and E can be replaced by

$$\sigma=\frac{2R}{\eta}\ ,\quad E=\sigma F_0\nu(\nu-1)\ , \tag{8.4}$$

with constants η and F_0 independent of ν.

The proof of theorem 2.1 is now just a matter of straightforward computation. First take $\varepsilon=\varrho=\delta$, and recall that $\varepsilon=\lambda^{-1/2}$, so that it is natural to assume also $\delta\le 1$, which in turn gives $\overline{\delta}=1$. Then take $d=R/4$, so that the canonical transformation is defined in the domain $\mathcal{D}_{\frac{1}{2}R,\varrho}$, and the image of $\mathcal{D}_{\frac{1}{2}R,\varrho}$ contains $\mathcal{D}_{\frac{1}{4}R,\varrho}$ and is contained in $\mathcal{D}_{\frac{3}{4}R,\varrho}$. With these settings one computes $\lambda_*=\delta_{*1}^{-2}$. Finally, the estimates for $Z(p',x',\pi',\xi',\lambda)$ and $\mathcal{R}(p',x',\pi',\xi',\lambda)$ are multiplied by λ, because the Hamiltonian was previously divided by the same factor. The inequalities (2.19) then follow from (7.15) by just substituting the explicit expressions (7.16) for $\mathcal{A}$ and $\mathcal{B}$ and the trivial estimate $\mathcal{B}<2^2e^{-\tau}$, which in turn follows from the definition of δ_{*1}, and this proves theorem 2.1.

Let us now add a few words on the very minor changes which are needed in order to prove the local version of the theorem, namely theorem 2.1'. We just notice that all theorems given in sects. 5 to 7 admit a local formulation, which is obtained by simply considering domains of the form $\mathcal{D}_{R,\varrho_*}(p,x)$ in place of the full domain $\mathcal{D}_{R,\varrho_*}$, *and one has the bound*

$$|I_\mu(p',x',\pi',\xi',\varepsilon)-I_\mu(p,x,\pi,\xi,\varepsilon)|<\mathcal{B}\frac{E}{E_0}\|I_\mu\|_{R-d}\delta^2\left(\frac{\delta}{\delta_{*1}}\right)^{\frac{2}{\tau+1}}\ , \tag{7.17}$$

with $\mathcal{B}$ *given by (7.16). Moreover, the time derivative of* I_μ *taking into account the full system (7.14) is bounded in the domain* $\mathcal{D}_{R-2d,\varrho}$ *by*

$$\left|\dot I_\mu(p',x',\pi',\xi',\varepsilon)\right|<\frac{\mathcal{A}e}{d^2}\delta^4\exp\left[-\frac{\tau+1}{e}\left(\frac{\delta_{*1}}{\delta\bar\delta}\right)^{\frac{2}{\tau+1}}\right]\|I_\mu\|_R \tag{7.18}$$

with $\mathcal{A}$ *defined by (7.16).*

8. Connection with the main results

We show now how the results of the previous section can be applied to the models discussed in sect. 2. Coming back to the Hamiltonian (1.1), let $f(p,x,\pi,\xi)$ be an analytic function of its variables in a domain of the form $\mathcal{G}_R\times\tilde\Delta_\eta$, where $\mathcal{G}_R$ is defined by (2.12), and $\tilde\Delta_\eta$ has the form

$$\tilde\Delta_\eta=\left\{(\pi,\xi):|\pi_j|\le\eta\Omega_j^{-1/2},|\xi_j|\le\eta\Omega_j^{1/2},1\le j\le\nu\right\}\ , \tag{8.1}$$

Ω being defined as in (2.1). Let now $F=\sup_{\mathcal{G}_R\times\tilde\Delta_\eta}|f(p,x,\pi,\xi)|$, and consider the power series development (4.1), which can be given the more explicit form

$$f^{(s,l)}(p,x,\pi,\xi)=\sum_{|j|=l}\sum_{|k|=s}f^{(s,l)}_{j,k}(p,x)\pi^j\xi^k\ .$$

Then, by using Cauchy's inequality, in the domain $\mathcal{G}_R$ one has the bound

$$\left|f^{(s,l)}_{j,k}(p,x)\right|\le\frac{F}{\eta^{s+l}}\Omega^{\frac{k-j}{2}}\ .$$

Performing the transformation (2.2) we now get

$$f^{(s,l)}(p,x,\pi',\xi')=\lambda^{\frac{l-s}{2}}\sum_{j,k}\Omega^{\frac{j-k}{2}}f^{(s,l)}_{jk}(p,x)\pi'^j\xi'^k\ ,$$

so that, by ignoring the power of λ which does not contribute to the norm, and noting that the factor $\Omega^{\frac{j-k}{2}}$ cancels out, just due to the initial choice of $\tilde\Delta_\eta$ in (8.1), the coefficient of $\pi'^j\xi'^k$ is bounded in $\mathcal{G}_R$ by F/η^{s+l}. Since $f^{(s,l)}\in\Pi_{s+l}$, we immediately get

$$\left\|f^{(s,l)}\right\|_R\le\sum_{j,k}\left(\frac{R}{\eta}\right)^{s+l}F\ , \tag{8.2}$$

be a resonance module, and assume

$$|k \cdot \Omega| \geq \gamma\, |k|^{-\tau} \;\; \text{for } k \in \mathbf{Z}^\nu \setminus \mathcal{M}$$

with real constants $\gamma > 0$ *and* $\tau \geq 0$. *Consider a domain* $\mathcal{D}_{R,\varrho} = \mathcal{G}_R \times \Delta_{\varrho R}$ *with* $\Delta_{\varrho R}$ *defined by (2.14) and with* $\varrho \geq \varepsilon/e^2$, *and define* $\delta = \max(\varrho, \varepsilon)$ *and* $\overline{\delta} = \max(1, \delta)$. *Then for any positive* $d < R/3$ *and for*

$$\delta\overline{\delta} \leq e^{-\frac{\tau+1}{2}} \delta_{*1} \ , \tag{7.13}$$

where δ_{*1} *is defined by (7.11), one can find a real analytic, near to identity canonical transformation* $(p, x, \pi, \xi) = \mathcal{C}_\varepsilon(p', x', \pi', \xi')$ *from* $\mathcal{D}_{R-2d,\varrho}$ *to* $\mathcal{D}_{R,\varrho}$, *and such that* $\mathcal{D}_{R-3d,\varrho} \subset \mathcal{C}_\varepsilon(\mathcal{D}_{R-2d,\varrho}) \subset \mathcal{D}_{R-d,\varrho}$, *which puts the Hamiltonian into the form*

$$\begin{aligned} H(p', x', \pi', \xi', \varepsilon) &= h_\Omega(\pi', \xi') + \varepsilon^2 \hat{h}(p', x') \\ &\quad + Z(p', x', \pi', \xi', \varepsilon) + \mathcal{R}(p', x', \pi', \xi', \varepsilon) \ , \end{aligned} \tag{7.14}$$

where $Z(p', x', \pi', \xi', \varepsilon)$ *is in normal form with respect to the resonance module* $\mathcal{M}$. *Moreover in the domain* $\mathcal{D}_{R-2d,\varrho}$ *one has the bounds*

$$\begin{aligned} &|h_\Omega(\pi', \xi') - h_\Omega(\pi, \xi)| < 2^5 E \delta^4 \\ &\left|\hat{h}(p', x') - \hat{h}(p, x)\right| < \mathcal{B} E \delta^2 \left(\frac{\delta}{\delta_{*1}}\right)^{\frac{2}{\tau+1}} \\ &\left|Z^{(r)}(p', x', \pi', \xi', \varepsilon)\right| < 2^5 E \delta^4 \\ &\left|\mathcal{R}^{(r)}(p', x', \pi', \xi', \varepsilon)\right| < \mathcal{A} \delta^4 \left(\frac{\delta\overline{\delta}}{\delta_{*1}}\right)^{\frac{2}{\tau+1}} \exp\left[-\frac{\tau+1}{e}\left(\frac{\delta_{*1}}{\delta\overline{\delta}}\right)^{\frac{2}{\tau+1}}\right] \ , \end{aligned} \tag{7.15}$$

where

$$\mathcal{A} = 2^6 e^{\tau+2} E \ , \quad \mathcal{B} = \frac{2^5 e^{-\tau} E_0}{\gamma d^2} \delta_{*1}^2 \ . \tag{7.16}$$

In fact, the condition (7.13), which follows from (7.12) by requiring $r_{\rm opt} \geq 1$, introduces, besides the threshold δ_{*1}, an effective threshold, in the sense that above the latter one the perturbation procedure is useless.

It is now straightforward to remove the dependence on r also from the estimate given in corollary 7.2, thus obtaining the following

Corollary 7.4: *The normalized part* $h_\Omega(\pi', \xi') + \varepsilon^2 \hat{h}(p', x') + Z(p', x', \pi', \xi', \varepsilon)$ *of the Hamiltonian (7.14) admits* $\nu - \dim \mathcal{M}$ *independent prime integrals of the form (7.7),* The existence of integrals of the form above is a classical result which follows from the characterization of the normal form given in sect. 2. The proofs of the bound (7.9) and of theorem 7.1 are deferred to sect. 12.

Our aim is now, in the spirit of Nekhoroshev's theory, to remove the normalization order r from our theory. Thus, along the lines of ref. [9], we look for an optimal normalization order $r_{\rm opt}$ which minimizes the bound (7.6) on the remainder $\mathcal{R}^{(r)}$.

In order to do that, we must give an explicit expression for α_r. As described in sect. 2, we use the diophantine bound (2.11) for the small denominators, and define

$$\alpha_r = \gamma r^{-\tau} \tag{7.10}$$

with $\gamma > 0$ and $\tau \geq 0$. Moreover, since the explicit expression (7.1) of δ_{*r} is too complicated for an analytical optimization, we use the inequality

$$\delta_{*r} \geq \frac{\delta_{*1}}{\sqrt{r^{\tau+1}}} \ , \quad \delta_{*1} = \frac{1}{4}\sqrt{\frac{\gamma d^2}{2e^2 E + 3(4E+E_0) + \gamma d^2 \sigma}} \ . \tag{7.11}$$

The quantity δ_{*1} introduced here plays the role of a threshold above which no perturbation procedure can be done.

By substituting these simpler expressions in (7.6), one is thus led to look for the minimum with respect to r of $r^{(\tau+1)r}\left(\frac{\delta\overline{\delta}}{\delta_{*1}}\right)^{2r}$, and so to choose $r_{\rm opt}$ as the integer satisfying the inequality

$$\frac{1}{e}\left(\frac{\delta_{*1}}{\delta\overline{\delta}}\right)^{\frac{2}{\tau+1}} - 1 < r_{\rm opt} \leq \frac{1}{e}\left(\frac{\delta_{*1}}{\delta\overline{\delta}}\right)^{\frac{2}{\tau+1}} \ . \tag{7.12}$$

This choice is not in contrast with the condition $\delta\overline{\delta} \leq \delta_{*r}/\sqrt{2}$ of theorem 7.1, since in (7.11) one has

$$\delta_{*r_{\rm opt}} > \frac{\delta_{*1}}{\sqrt{r_{\rm opt}^{\tau+1}}} > e^{\frac{\tau+1}{2}}\delta\overline{\delta} \ .$$

Finally, we substitute $r_{\rm opt}$ everywhere in (7.10), (7.4) and (7.6) in order to completely remove the dependence of the bounds on r. This proves the

Theorem 7.3: *Consider the Hamiltonian* $H(p,x,\pi,\xi,\varepsilon) = \sum_{s\geq 0} H_s(p,x,\pi,\xi,\varepsilon)$ *with* $H_0(p,x,\pi,\xi,\varepsilon) = h_\Omega(\pi,\xi) + \varepsilon^2\hat{h}(p,x)$, $h_\Omega(\pi,\xi) = \frac{1}{2}\sum_{l=1}^{\nu}\Omega_l(\pi_l^2+\xi_l^2)$, H_s *of class* $\mathcal{P}_{2s+2,2s+2}$ *on the domain* $\mathcal{G}_R$ *defined by (2.12), and* ε *a real parameter. Assume that* $N_R(\hat{h}) \leq E_0$ *and* $N_R(H_s) \leq \sigma^{s-1}E$ *for positive* σ, E_0 *and* E. *With reference to the vector* Ω *of the frequencies of* h_Ω *and to the module* $\mathcal{M}_\Omega$ *related to it, let* $\mathcal{M} \supset \mathcal{M}_\Omega$

and define $\delta = \max(\varrho, \varepsilon)$ *and* $\overline{\delta} = \max(1, \delta)$.
Then, for any positive $d < R/3$ *and any integer* $r \geq 1$, *and for* $\delta\overline{\delta} \leq \delta_{*r}/\sqrt{2}$, *with*

$$\delta_{*r} = \frac{1}{4}\left[\frac{2e^2E + 3(4E+E_0)(r-1) + \alpha_r d^2\sigma}{\alpha_r d^2}\right]^{-\frac{1}{2}} \tag{7.1}$$

there exists a real analytic, near to identity canonical transformation $T_{\chi^{(r)}}$ *from* $\mathcal{D}_{R-2d,\varrho}$ *to* $\mathcal{D}_{R-d,\varrho}$, *with* $T_{\chi^{(r)}}(\mathcal{D}_{R-2d,\varrho}) \supset \mathcal{D}_{R-3d,\varrho}$, *which puts the Hamiltonian in normal form up to order* r *with respect to the module* $\mathcal{M}$, *i.e.*

$$\begin{aligned} H^{(r)}(p',x',\pi',\xi',\varepsilon) &= h_\Omega(\pi',\xi') + \varepsilon^2\hat{h}(p',x') \\ &\quad + Z^{(r)}(p',x',\pi',\xi',\varepsilon) + \mathcal{R}^{(r)}(p',x',\pi',\xi',\varepsilon)\ , \end{aligned} \tag{7.2}$$

and one has the bounds

$$|h_\Omega(\pi',\xi') - h_\Omega(\pi,\xi)| < 2^5 E\delta^4 \tag{7.3}$$

$$\left|\hat{h}(p',x') - \hat{h}(p,x)\right| < \frac{2^5E_0}{\alpha_r d^2}E\delta^4 \tag{7.4}$$

$$\left|Z^{(r)}(p',x',\pi',\xi',\varepsilon)\right| < 2^5E\delta^4 \tag{7.5}$$

$$\left|\mathcal{R}^{(r)}(p',x',\pi',\xi',\varepsilon)\right| < \frac{2^6E}{r+1}\delta^4\left(\frac{\delta\overline{\delta}}{\delta_{*r}}\right)^{2r}\ . \tag{7.6}$$

The theorem has the following

Corollary 7.2: *The normalized part* $h_\Omega(\pi',\xi') + \hat{h}(p',x') + Z^{(r)}(p',x',\pi',\xi',\varepsilon)$ *of the Hamiltonian (7.2) admits* $\nu - \dim\mathcal{M}$ *independent prime integrals of the form*

$$I_\mu(\pi',\xi') = \sum_{l=1}^{\nu}\frac{\mu_l}{2}(\pi_l'^2 + \xi_l'^2)\ , \tag{7.7}$$

with $\mathcal{M} \perp \mu \in \mathbf{R}^\nu$. *Moreover, if* π',ξ' *and* π,ξ *are related by the canonical transformation* $T_{\chi^{(r)}}$, *for any such integral one has*

$$|I_\mu(\pi',\xi') - I_\mu(\pi,\xi)| < \frac{2^5E}{\alpha_r d^2}\delta^4\|I_\mu\|_R\ , \tag{7.8}$$

and its time derivative taking into account the full system (7.2) is bounded in the domain $\mathcal{D}_{R-2d,\varrho}$ *by*

$$\left|\dot{I}_\mu(p',x',\pi',\xi',\varepsilon)\right| < \frac{2^5e^2E}{d^2}\delta^4\left(\frac{\delta\overline{\delta}}{\delta_{*r}}\right)^{2r}\|I_\mu\|_R\ . \tag{7.9}$$

Now we can state the

Theorem 6.1: *Consider the Hamiltonian* $H(p,x,\pi,\xi,\varepsilon) = \sum_{s\geq 0} H_s(p,x,\pi,\xi,\varepsilon)$, *with* $H_0(p,x,\pi,\xi,\varepsilon) = h_\Omega(\pi,\xi) + \varepsilon^2 \hat{h}(p,x)$, $h_\Omega = i\sum_{l=1}^{\nu} \Omega_l \pi_l \xi_l$, H_s *of class* $\mathcal{P}_{2s+2,2s+2}$ *on the domain* $\mathcal{G}_R$ *defined by (2.12), and* ε *a real parameter. Assume that* $N_R(\varepsilon^2\hat{h}) \leq E_0$ *and* $N_R(H_s) \leq \gamma^{s-1}\mathcal{F}$ *for positive* γ, E_0 *and* $\mathcal{F}$. *With reference to the vector* Ω *of the frequencies of* h_Ω *and to the module* $\mathcal{M}_\Omega$ *related to it, let* $\mathcal{M} \supset \mathcal{M}_\Omega$ *be a resonance module, and* $\{\alpha_s\}_{s\geq 1}$ *a sequence of positive constants satisfying (6.10).*
Then, for any positive $d < R/2$ *and any integer* $r \geq 1$, *there exists a generating sequence* $\chi^{(r)} = \{\chi_l\}_{1\leq l\leq r}$, *with* χ_l *of class* $\mathcal{P}_{2l+2,4l}$ *on the domain* $\mathcal{G}_{R-d}$, *which brings the Hamiltonian into its normal form* $H^{(r)}$, *defined by (2.7), with respect to the module* $\mathcal{M}$, *and one has the estimates*

$$\begin{aligned} N_{R-d}(Z_l) &\leq \frac{\beta^{l-1}}{l}\mathcal{F} \ , \\ N_{R-d}(\chi_l) &\leq \frac{\beta^{l-1}}{l}\Phi \ , \end{aligned} \tag{6.11}$$

with

$$\Phi = \frac{\mathcal{F}}{\alpha_r} \ , \qquad \beta = \frac{12(\mathcal{F}+E_0)(r-1)}{\alpha_r d^2} + 2\gamma \ . \tag{6.12}$$

The proof of the theorem is deferred to the technical section 11.

7. Exponential estimates on the remainder

We can now use the canonical transformation of sect. 5, with the generating sequence of sect. 6 in order to get the normalized Hamiltonian in the form (6.1) with the remainder in the form (6.2). So, making a transformation up to a finite order r, we prove the following

Theorem 7.1: *Consider the Hamiltonian* $H(p,x,\pi,\xi,\varepsilon) = \sum_{s\geq 0} H_s(p,x,\pi,\xi,\varepsilon)$ *with* $H_0(p,x,\pi,\xi,\varepsilon) = h_\Omega(\pi,\xi) + \varepsilon^2\hat{h}(p,x)$, $h_\Omega(\pi,\xi) = \frac{1}{2}\sum_{l=1}^{\nu}\Omega_l(\pi_l^2+\xi_l^2)$, H_s *of class* $\mathcal{P}_{2s+2,2s+2}$ *on the domain* $\mathcal{G}_R$ *defined by (2.12), and* ε *a real parameter. Assume that* $N_R(\hat{h}) \leq E_0$ *and* $N_R(H_s) \leq \sigma^{s-1}E$ *for positive* σ, E_0 *and* E. *With reference to the vector* Ω *of the frequencies of* h_Ω *and to the module* $\mathcal{M}_\Omega$ *related to it, let* $\mathcal{M} \supset \mathcal{M}_\Omega$ *be a resonance module, and* $\{\alpha_s\}_{s>0}$ *a sequence of positive constants satisfying (6.10). Consider a domain* $\mathcal{D}_{R,\varrho} = \mathcal{G}_R \times \Delta_{\varrho R}$ *with* $\Delta_{\varrho R}$ *defined by (2.14) and with* $\varrho \geq \varepsilon/e^2$, module $\mathcal{M}_\Omega$ defined by (2.10), one chooses a module $\mathcal{M}$ with the condition $\mathcal{M} \supset \mathcal{M}_\Omega$, and splits Ψ_s into

$$\Psi_s = \overline{\Psi}_s + \tilde{\Psi}_s \ , \tag{6.7}$$

$\overline{\Psi}$ being the resonant part of Ψ_s with respect to $\mathcal{M}$, and $\tilde{\Psi}_s$ the nonresonant one. So, eq. (6.5) can be split into

$$Z_s = \overline{\Psi}_s \ , \qquad L_{h_\Omega}\chi_s = \tilde{\Psi}_s \ , \tag{6.8}$$

and the first of these equations determines the normalized part of the Hamiltonian, while the second one can be solved, and gives the generating sequence. In solving the latter equation, as is well known, small denominators appear of the form $k \cdot \Omega$ with $k \in \mathbf{Z}^\nu$, but, thanks to the polynomial dependence of Ψ_s on the variables π, ξ, one has the bound $|k| \leq 4s$. The details on the method of solution, as well as its relations with the classes $\mathcal{P}_{\Lambda,K}$, are deferred to the technical section 11.

This shows that the Hamiltonian can be formally put in normal form with respect to the resonance module $\mathcal{M}$ up to an arbitrary order r. In order to have a rigorous result we must produce estimates on the norms of the generating sequence $\{\chi_l\}_{1\leq l\leq r}$. Precisely, we look for constants β and Φ, possibly dependent on r, such that $N_{R-d}(\chi_l) \leq \frac{\beta^{l-1}}{l}\Phi$ for $1 \leq l \leq r$, as required by the theorems on canonical transformations. To this end, we make a convergence hypothesis on the Hamiltonian H by assuming that there exist positive real constants γ, E_0 and $\mathcal{F}$ such that

$$N_R(\varepsilon^2 \hat{h}) \leq E_0 \ , \qquad N_R(H_s) \leq \gamma^{s-1}\mathcal{F} \ , \quad s \geq 1 \ . \tag{6.9}$$

Moreover, we need a lower bound on the small denominators which appear in the generating function in solving eq. (6.5). To this end, let $\{\alpha_s\}_{s>0}$ be a nonincreasing sequence of positive real constants satisfying

$$|k \cdot \Omega| \geq \alpha_s \ \text{for}\ k \in \mathbf{Z}^\nu \setminus \mathcal{M} \ \text{and}\ \ |k| \leq 4s \ . \tag{6.10}$$

The condition $\mathcal{M} \supset \mathcal{M}_\Omega$ ensures that none of the α_s's needs to be zero. The apparently most natural choice would seem to be $\alpha_s = \min_k(|k \cdot \Omega|)$ for $k \in \mathbf{Z}^\nu \setminus \mathcal{M}$ and $|k| \leq 4s$. However, such choice can be handled in a numerical approach, but must be replaced by a more regular function of s if explicit analytic estimates are required (see for example ref. [10]). On the other hand, an explicit expression will be used only at the end of our perturbative treatment, while the monotonicity property will be enough at each step of the procedure. That's why we prefer to leave the sequence undefined until an explicit choice will be unavoidable.

$\sum_{s\geq 0}\varepsilon^2\hat{h}_s$. Assuming now that χ_s is of class $\mathcal{P}_{2s+2,4s}$, it is an easy matter to check that, for $s\geq 1$, h_s is of class $\mathcal{P}_{2s+2,4s}$ and $\varepsilon^2\hat{h}_s$ is of class $\mathcal{P}_{2s+4,4s+2}\subset\mathcal{P}_{2(s+1)+2,4(s+1)}$. This elementary remark is the key which allows to shift $\varepsilon^2\hat{h}_s$ to an higher order, thus allowing to consider the variables p,x essentially as parameters in our perturbation scheme. Moreover, using the properties (4.11), it is also easily checked that $H_{l,s-l}$ is of class $\mathcal{P}_{2s+2,4s-2l+2}\in\mathcal{P}_{2s+2,4s}$. Thus, eq. (6.3) gives the system

$$\begin{aligned}
Z_0 &= H_0\\
Z_1 &= h_1 + H_1\\
Z_s &= h_s + \varepsilon^2\hat{h}_{s-1} + \sum_{l=1}^{s} H_{l,s-l}\ , \qquad 2\leq s\leq r\\
H_s^{(r)} &= h_s + \varepsilon^2\hat{h}_{s-1} + \sum_{l=1}^{s} H_{l,s-l}\ , \qquad s>r\ .
\end{aligned} \tag{6.4}$$

Here, only the second and third equations need further discussion, because they actually allow to determine all the required quantities, i.e. $\chi_1,\ldots,\chi_r$ and $Z_1,\ldots,Z_r$. The fourth one just gives the explicit expression for the remainder, but contains no information about $Z^{(r)}$ and $\chi^{(r)}$; we shall use it in determining the size of the remainder, in sect. 12. So, we concentrate on the second and third equations.

By using the explicit expressions

$$h_1 = L_{\chi_1}h_\Omega\ ,\quad h_s = \sum_{l=1}^{s-1}\frac{l}{s}L_{\chi_l}h_{s-l} + L_{\chi_s}h_\Omega\ ,\quad 2\leq s\leq r$$

these equations immediately give

$$L_{h_\Omega}\chi_s + Z_s = \Psi_s\ , \tag{6.5}$$

where

$$\begin{aligned}
\Psi_1 &= H_1\ ,\\
\Psi_s &= \varepsilon^2\hat{h}_{s-1} + \sum_{l=1}^{s-1}\frac{l}{s}L_{\chi_l}h_{s-l} + \sum_{l=1}^{s}H_{l,s-l}\ ,\quad 2\leq s\leq r\ .
\end{aligned} \tag{6.6}$$

These expressions can also be used to check that the assumption that χ_s be of class $\mathcal{P}_{2s+2,4s}$ for $s\geq 1$ is consistent, provided that if Ψ_s is of class $\mathcal{P}_{2s+2,4s}$, the equation (6.5) above can be solved with χ_s and Z_s of the same class. Moreover, $H_s^{(r)}$ as defined by (6.4) is of class $\mathcal{P}_{2s+2,4s}$.

Eq. (6.5) is the standard one of perturbation theory, and can be solved in our case by considering the variables p,x as parameters. The solution of such equation is a well known topic. Given the set Ω of the frequencies of h_Ω, and the related resonance for $(p', x', \pi', \xi') \in \mathcal{D}_{R-d,\varrho}$.

In virtue of such theorem the construction of the transformed function $(f \circ T_\chi)$ can be done by an explicit algorithm, and does not involve any inversion. The proof is deferred to sect. 10.

6. Construction of the generating sequence

We come now to the heart of our perturbative scheme, i.e. the reduction of the Hamiltonian to normal form up to a finite order.

Starting with an Hamiltonian of the form (4.4) transformed to complex variables via the canonical transformation (2.8), i.e. with $h_\Omega(\pi,\xi)$ of the form (2.9) and H_s of class $\mathcal{P}_{2s+2,2s+2}$ on a domain $\mathcal{G}_R$, we look for a truncated generating sequence $\chi^{(r)} = \{\chi_l\}_{1\le l\le r}$ on a reduced domain $\mathcal{G}_{R-d}$, with $0 < d < R/2$, which puts the Hamiltonian into the normal form up to order r,

$$H^{(r)}(p,x,\pi,\xi,\varepsilon) = h_\Omega(\pi,\xi) + \hat{h}(p,x) + Z^{(r)}(p,x,\pi,\xi,\varepsilon) + \mathcal{R}^{(r)}(p,x,\pi,\xi,\varepsilon) ,$$

where

$$Z^{(r)}(p,x,\pi,\xi,\varepsilon) = \sum_{1\le s\le r} Z_s(p,x,\pi,\xi,\varepsilon) \tag{6.1}$$

is the normalized part, and

$$\mathcal{R}^{(r)}(p,x,\pi,\xi,\varepsilon) = \sum_{s>r} H_s^{(r)}(p,x,\pi,\xi,\varepsilon) \tag{6.2}$$

is the unnormalized remainder. The canonical transformation is the one generated by the operator T_χ of sect. 5, and the functions χ_s, Z_s and $H_s^{(r)}$ will turn out to be of class $\mathcal{P}_{2s+2,4s}$. Our aim is then to establish the equations needed in order to determine the generating sequence $\{\chi_l\}_{1\le l\le r}$ and the normal form $Z^{(r)}$, and to give estimates on the norms of χ_l.

Thus, we start with the formal deduction of the equations. Denoting $T_{\chi^{(r)}} H_s = \sum_{l\ge 0} H_{s,l}$, we write the equation $T_{\chi^{(r)}} H = H^{(r)}$ in the form

$$\sum_{s\ge 0}\sum_{l=0}^{s} H_{l,s-l} = \sum_{s=0}^{r} Z_r + \sum_{s>r} H_s^{(r)} . \tag{6.3}$$

We look now for a more explicit expression by isolating the terms of the same class. To this end we use $H_0 = h_\Omega + \varepsilon^2 \hat{h}$, and denote $T_{\chi^{(r)}} h_\Omega = \sum_{s\ge 0} h_s$ and $T_{\chi^{(r)}}(\varepsilon^2 \hat{h}) =$

Then, for any positive $d < R/2$ and for $0 < \delta\overline{\delta} \le \delta_*/\sqrt{2}$, with

$$\delta_* = \left(\frac{2e^2\Phi}{d^2} + \beta\right)^{-\frac{1}{2}} \tag{5.6}$$

the canonical transformation $(p, x, \pi, \xi) = T_\chi(p', x', \pi', \xi')$ defined by (5.4) analytically maps the domain $\mathcal{D}_{R-d,\varrho}$ into $\mathcal{D}_{R,\varrho}$ in such a way that $T_\chi(\mathcal{D}_{R-d,\varrho}) \supset \mathcal{D}_{R-2d,\varrho}$. Moreover, for $(p', x', \pi', \xi') \in \mathcal{D}_{R-d,\varrho}$ one has

$$\begin{aligned} |p_l - p'_l| &\le \frac{2\delta^4\Phi}{d} \ , \quad |x_l - x'_l| \le \frac{2\delta^4\Phi}{d} \ , \quad 1 \le l \le n \ , \\ |\pi_l - \pi'_l| &\le \frac{2\delta^3\Phi}{d} \ , \quad |\xi_l - \xi'_l| \le \frac{2\delta^3\Phi}{d} \ , \quad 1 \le l \le \nu \ . \end{aligned} \tag{5.7}$$

In virtue of such theorem, the canonical transformation (5.4) is not just formal, and moreover, thanks to (5.7), we can estimate the deformation of coordinates induced by T_χ. The proof is given in sect. 10.

Consider now the problem of applying the transformation to a function. More precisely, let $f(p, x, \pi, \xi)$ be a function of the coordinates (p, x, π, ξ), and denote by $f'(p', x', \pi', \xi') = (f \circ T_\chi)(p', x', \pi', \xi')$ the transformed function under the canonical transformation (5.3). By definition it is $f'(p', x', \pi', \xi') = f(T_\chi(p', x', \pi', \xi'))$, and this is an analytic function on the domain $\mathcal{D}_{R-d,\varrho}$ where the canonical transformation is defined and analytic. A standard result in the theory of Lie series, namely the exchange theorem,[12] states that f' is nothing but $T_\chi f$ as defined in (5.1). More precisely, we state the following

Theorem 5.2: *Let $\{\chi_l\}_{l\ge1}$, with $\chi_l(p, x, \pi, \xi, \varepsilon)$ of class $\mathcal{P}_{2l+2,4l}$ on $\mathcal{G}_R$ and ε a real parameter, be a generating sequence, and assume that there exist positive constants β and Φ such that $N_R(\chi_l) \le \frac{\beta^{l-1}}{l}\Phi$. Consider a domain $\mathcal{D}_{R,\varrho} = \mathcal{G}_R \times \Delta_{\varrho R}$, with canonical coordinates (p, x, π, ξ), where $\mathcal{G}_R$ has the form (2.12) and $\Delta_{\varrho R}$ the form (2.14). Assume $\varrho \ge \varepsilon/e^2$, and define δ and $\overline{\delta}$ as in (5.5). Let $f(p, x, \pi, \xi, \varepsilon)$ be a function of class $\mathcal{P}_{\Lambda,K}$ on the domain $\mathcal{G}_R$ for some $K \ge \Lambda \ge 0$, and consider the function $T_\chi f = \sum_{r\ge0} f_r$ defined as in (5.2).*
Then, for any positive $d < R$ and for $0 \le \delta\overline{\delta} \le \delta_*/\sqrt{2}$, with δ_* defined by (5.6), the function $T_\chi f$ is analytic in $\mathcal{D}_{R-d,\varrho}$, and coincides there with the transformed function of f by the canonical transformation (5.3), i.e.

$$(f \circ T_\chi)(p', x', \pi', \xi', \varepsilon) = (T_\chi f)(p', x', \pi', \xi', \varepsilon) \ . \tag{5.8}$$

5. Canonical transformations

We first recall how a canonical transformation can be defined by purely algebraic methods, and avoiding any inversion. At a formal level, as in ref. [11], we consider a generating sequence $\{\chi_l\}_{l\geq 1}$ with χ_l of class $\mathcal{P}_{2l+2,4l}$ on a domain $\mathcal{G}_R$. The choice that χ_l be of class $\mathcal{P}_{2l+2,4l}$ is adapted to our case: we do not look for a more general characterization. Let's also introduce the operator L_g by $L_g \cdot = \{g, \cdot\}$. Then, for any function f of class $\mathcal{P}_{\Lambda,K}$ we define the transformed function $T_\chi f$ as a series of the form

$$T_\chi f = \sum_{r\geq 0} f_r \ , \tag{5.1}$$

where the sequence $\{f_r\}_{r\geq 0}$ is recursively defined by

$$f_0 = f \ , \qquad f_r = \sum_{l=1}^{r} \frac{l}{r} L_{\chi_l} f_{r-l} \ . \tag{5.2}$$

The operator T_χ so defined turns out to be linear and invertible, and to preserve products and Poisson brackets. So, considering T_χ as acting on the canonical coordinates, we can define a formal canonical transformation T_χ on the domain $\mathcal{D}_R$, i.e.

$$(p, x, \pi, \xi) = T_\chi(p', x', \pi', \xi') \ . \tag{5.3}$$

by

$$\begin{aligned} p_l &= T_\chi p'_l \ , \quad x_l = T_\chi x'_l \ , \qquad 1 \leq l \leq n \ , \\ \pi_l &= T_\chi \pi'_l \ , \quad \xi_l = T_\chi \xi'_l \ , \qquad 1 \leq l \leq \nu \ . \end{aligned} \tag{5.4}$$

Such a transformation is near the identity, in the sense that the change of coordinates is of the order $O(\varepsilon)$. These formal properties can be directly checked, as was done in ref. [11], so the proof is omitted.

Coming now to rigorous results, we first consider the coordinate transformation (5.3). We have the

Theorem 5.1: *Let $\{\chi_l\}_{l\geq 1}$, with $\chi_l(p, x, \pi, \xi, \varepsilon)$ of class $\mathcal{P}_{2l+2,4l}$ on $\mathcal{G}_R$ and ε a real parameter, be a generating sequence, and assume that there exist positive constants β and Φ such that $N_R(\chi_l) \leq \frac{\beta^{l-1}}{l}\Phi$. Consider a domain $\mathcal{D}_{R,\varrho} = \mathcal{G}_R \times \Delta_{\varrho R}$, with canonical coordinates (p, x, π, ξ), where $\mathcal{G}_R$ has the form (2.12) and $\Delta_{\varrho R}$ the form (2.14). Assume $\varrho \geq \varepsilon/e^2$, and define*

$$\delta = \max(\varepsilon, \varrho) \ , \quad \overline{\delta} = \max(1, \delta) \ . \tag{5.5}$$

Then we introduce the norm for f

$$\|f\|_R = R^\Lambda \sum_{j,k} |f_{jk}|_R \ ; \tag{4.13}$$

the factor R^Λ, which is also suggested by dimensional reasons, will turn out to be very relevant in simplifying the estimates.

Coming now to a function f of class $\mathcal{P}_{\Lambda,K}$, and recalling the form (4.9) or (4.10) for f, we introduce the norm

$$N_R(f) = \sum_{\Lambda \le l+m \le K} \|f^{(l,m)}\|_R \ , \tag{4.14}$$

or, equivalently,

$$N_R(f) = \sum_{\Lambda \le l+|j|+|k| \le K} R^{|j|+|k|} |f^{(l)}_{jk}|_R \ .$$

The choice of the norms made above allows us to conveniently bound derivatives and Poisson brackets. The technical estimates are deferred to sect. 9.

We finally consider a function of the form

$$f(p,x,\pi,\xi,\varepsilon) = \sum_{s\ge 0} f_s(p,x,\pi,\xi,\varepsilon) \tag{4.15}$$

with f_s of class $\mathcal{P}_{\Lambda_s,K_s}$ on a domain $\mathcal{G}_R$, for given sequences $\{\Lambda_s\}_{s\ge 0}$ and $\{K_s\}_{s\ge 0}$ of nonnegative integers $\Lambda_s \le K_s$. In order to investigate the convergence of such a series, we consider $f_s(p,x,\pi,\xi,\varepsilon)$, for $s \ge 0$, as defined on the domain $\mathcal{D}_{R,\varrho}$ defined by (2.15), and look for values ε_* and ϱ_* such that the series (4.15) converges for $\varepsilon < \varepsilon_*$ and $\varrho < \varrho_*$. Such procedure obviously depends on the sequences $\{\Lambda_s\}_{s\ge 0}$ and $\{K_s\}_{s\ge 0}$, as well as on the norms $N_R(f_s)$. Precisely, by introducing the parameters

$$\delta = \max(\varepsilon, \varrho) \ , \quad \overline{\delta} = \max(1, \delta) \tag{4.16}$$

one easily sees, using the form (4.9) for f_s, that for $(p,x,\pi,\xi) \in \mathcal{D}_{R,\varrho}$ one has the bound

$$|f_s(p,x,\pi,\xi,\varepsilon)| \le \delta^{\Lambda_s} \overline{\delta}^{(K_s-\Lambda_s)} N_R(f_s) \ . \tag{4.17}$$

The problem of the convergence of (4.15) is thus reduced to that of the convergence of the series $\sum_{s\ge 0} \delta^{\Lambda_s} \overline{\delta}^{(K_s-\Lambda_s)} N_R(f_s)$. The actual application of such method will be done below, when needed.

the maximal degree, i.e. the second index, will allow to work at each step of the perturbative procedure with polynomials of finite order.

It is now an easy matter to check that for two functions f of class $\mathcal{P}_{\Lambda,K}$ and f' of class $\mathcal{P}_{\Lambda',K'}$ one has the following properties:

$$\begin{aligned}
&\text{i.} && \frac{\partial f}{\partial p_l}\ ,\quad \frac{\partial f}{\partial x_l} \quad \text{is of class } \mathcal{P}_{\Lambda,K}\ ;\\
&\text{ii.} && \frac{\partial f}{\partial \pi_l}\ ,\quad \frac{\partial f}{\partial \xi_l} \quad \text{is of class } \mathcal{P}_{\Lambda-1,K-1}\ ;\\
&\text{iii.} && f+f' \quad \text{is of class } \mathcal{P}_{\min(\Lambda,\Lambda'),\max(K,K')}\ ;\\
&\text{iv.} && f\cdot f' \quad \text{is of class } \mathcal{P}_{\Lambda+\Lambda',K+K'}\ ;\\
&\text{v.} && \{f,f'\}_{p,x} \quad \text{is of class } \mathcal{P}_{\Lambda+\Lambda',K+K'}\ ;\\
&\text{vi.} && \{f,f'\}_{\pi,\xi} \quad \text{is of class } \mathcal{P}_{\Lambda+\Lambda'-2,K+K'-2}\ ;\\
&\text{vii.} && \{f,f'\} \quad \text{is of class } \mathcal{P}_{\Lambda+\Lambda'-2,K+K'}\ .
\end{aligned} \tag{4.11}$$

Notice that all of these properties but the last one have a correspondent property in the framework of the spaces Π_Λ. It is just the property vii. that makes the characterization by the classes $\mathcal{P}_{\Lambda,K}$ more suitable for our perturbative scheme.

With such a formal algebraic framework the Hamiltonian (4.4) is characterized by the fact that $h_\Omega(\pi,\xi)$ and $\varepsilon^2\hat{h}(p,x)$ are of class $\mathcal{P}_{2,2}$, while H_s is of class $\mathcal{P}_{2s+2,2s+2}$ for $s \geq 1$.

To complete the formal algebraic framework we must consider an equation of the form $L_{h_\Omega} f = g$, where g is a known function of class $\mathcal{P}_{\Lambda,K}$ for some Λ and K. A classical result, recalled in sect. 11, is that such equation can be solved if g satisfies suitable conditions depending on the frequencies Ω, and that in such case the solution f is of class $\mathcal{P}_{\Lambda,K}$.

Within the above algebraic framework, the perturbation theory can be formally developed. In order to make it rigorous, we must introduce norms for functions and some technical inequalities.

Consider first a function $f \in \Pi_\Lambda$ of the form (4.7). As the coefficients $f_{jk}(p,x)$ are analytic functions of (p,x) in a domain $\mathcal{G}_R$, defined by (2.12), with positive R, we bound them by the usual supremum norm

$$|f_{jk}|_R = \sup_{(p,x)\in\mathcal{G}_R} |f_{jk}(p,x)| \ . \tag{4.12}$$

properties:

$$
\begin{aligned}
&\text{i.} && \frac{\partial f}{\partial p_l} \in \Pi_\Lambda \ , \quad \frac{\partial f}{\partial x_l} \in \Pi_\Lambda \ ; \\
&\text{ii.} && \frac{\partial f}{\partial \pi_l} \in \Pi_{\Lambda-1} \ , \quad \frac{\partial f}{\partial \xi_l} \in \Pi_{\Lambda-1} \ ; \\
&\text{iii.} && \text{if } \Lambda = \Lambda' \ , \quad \text{then } f + f' \in \Pi_\Lambda \ ; \\
&\text{iv.} && f \cdot f' \in \Pi_{\Lambda+\Lambda'} \ ; \\
&\text{v.} && \{f, f'\}_{p,x} \in \Pi_{\Lambda+\Lambda'} \ ; \\
&\text{vi.} && \{f, f'\}_{\pi,\xi} \in \Pi_{\Lambda+\Lambda'-2} \ .
\end{aligned}
\tag{4.8}
$$

Here, $\{\cdot,\cdot\}_{p,x}$ and $\{\cdot,\cdot\}_{\pi,\xi}$ denote the Poisson bracket restricted to the p, x and π, ξ variables respectively, so that $\{\cdot,\cdot\} = \{\cdot,\cdot\}_{p,x} + \{\cdot,\cdot\}_{\pi,\xi}$. The properties v. and vi. in particular show that, due to the inhomogeneity of the Poisson bracket between homogeneous functions, the spaces Π_Λ are not suitable for the development of our perturbation scheme, and that we need a weaker algebraic characterization, where sets of nonhomogeneous functions are considered. Moreover, in the latter characterization we also want to make reference to the order in ε; so, we introduce the classes of functions $\mathcal{P}_{\Lambda,K}$, with $K \geq \Lambda \geq 0$ defined as follows: a function $f(p, x, \pi, \xi, \varepsilon)$ is said to be of class $\mathcal{P}_{\Lambda,K}$ if it can be written in the form

$$
f(p, x, \pi, \xi, \varepsilon) = \sum_{\substack{l,m\geq 0 \\ \Lambda \leq l+m \leq K}} \varepsilon^l f^{(l,m)}(p, x, \pi, \xi) \tag{4.9}
$$

with $f^{(l,m)} \in \Pi_m$; equivalently one can write

$$
f(p, x, \pi, \xi, \varepsilon) = \sum_{\Lambda \leq l+|j|+|k| \leq K} \varepsilon^l f^{(l)}_{jk}(p, x) \pi^j \xi^k \tag{4.10}
$$

Notice in particular that if $\Lambda \leq \Lambda'$ and $K \geq K'$ then $\mathcal{P}_{\Lambda,K} \supset \mathcal{P}_{\Lambda',K'}$, and that the class $\mathcal{P}_{0,0}$ contains functions of the variables p, x only.

Let us give some motivations. As explained above, our aim is to develop our perturbation scheme by considering ε as a small parameter and π, ξ confined to a polydisk of radius ϱR, with ϱ of the same order of ε. So, it is natural to consider functions of a definite order with respect to π, ξ, ε. However, for reasons which will be evident below, we cannot have complete homogeneity in π, ξ, ε; that is why we allow the class $\mathcal{P}_{\Lambda,K}$ to contain functions which are non homogeneous polynomials in π, ξ, ε, at the same time keeping track of the minimal and maximal degree. The minimal degree, i.e. the first index, will work as a perturbation order in reordering the series arising from the perturbative algorithm, as was done above for the Hamiltonian, while using $f(p,x,\pi,\xi)$ of the form (4.2), can be written as

$$H_s(p,x,\pi,\xi,\varepsilon)=\sum_{l=0}^{L}\varepsilon^{s-l+2}f^{(l,s)}(p,x,\Omega^{1/2}\pi,\Omega^{-1/2}\xi)\ . \tag{4.6}$$

Such a reordering of the power series development of $f(p,x,\pi,\xi)$ requires some additional comment. In order to apply the methods of perturbation theory we must choose one or more perturbative parameters. On the other hand, we cannot simply identify such parameter with ε, because in the power series expansion of $f(p,x,\pi,\xi)$ one finds terms which are of the same order of $\varepsilon^2\hat{h}$. So, as was already done in the first part of this paper, we consider the variables π,ξ confined to a polydisk whose radius is of the same order of ε This means that, besides ε, we can use as a perturbative parameter also the size of the domain of the π,ξ variables. The reordering above then corresponds to considering $H_s(p,x,\pi,\xi,\varepsilon)$ as an homogeneous polynomial of degree $2s+2$ in ε,π,ξ, due to the fact that $f^{(l,s)}$ is an homogeneous polynomial of degree l in π and of degree s in ξ.

The algebraic framework we are going to build takes into account such remarks. Moreover, from now on we shall assume $L\le 2$, so that negative powers of ε do not appear in the expression (4.6) of H_s.

Starting with a formal viewpoint, we first introduce the spaces Π_Λ of the homogeneous polynomials of degree Λ in the canonical variables $(\pi,\xi)\in\mathbf{C}^{2\nu}$, whose coefficients are analytic bounded functions of $(p,x)\in\mathcal{G}_R$ (and independent of ε), with $\mathcal{G}_R$ defined by (2.12). A function $f\in\Pi_\Lambda$ can be represented as

$$f(p,x,\pi,\xi)=\sum_{|j|+|k|=\Lambda}f_{jk}(p,x)\pi^j\xi^k\ , \tag{4.7}$$

where $\pi^j\xi^k=\pi_1^{j_1}\ldots\pi_\nu^{j_\nu}\xi_1^{k_1}\ldots\xi_\nu^{k_\nu}$, j_l and k_l being nonnegative integers for $1\le l\le\nu$, and $|j|=|j_1|+\ldots+|j_\nu|$, and similarly for $|k|$.

It is an easy matter to check that, for $f\in\Pi_\Lambda$ and $f'\in\Pi_{\Lambda'}$, one has the following

Part B – The perturbation scheme.

4. The algebraic framework

In order to apply the apparatus of classical perturbation theory, along the lines of refs. [11] and [9], we need to build an algebraic framework compatible with the operations required by our perturbative algorithm.

To this end, we first use the fact that we are interested in a neighbourhood of the origin of the π, ξ variables, where we can perform a power series development of the perturbation $f(p, x, \pi, \xi)$ in the Hamiltonian (1.1), thus obtaining

$$f(p, x, \pi, \xi) = \sum_{s\geq 1} \sum_{l=0}^{L} \tilde{f}^{(l,s)}(p, x, \pi, \xi) \ , \tag{4.1}$$

where $\tilde{f}^{(l,s)}(p, x, \pi, \xi)$ is an homogeneous polynomial of degree l in π and of degree s in ξ, whose coefficients are analytic functions of (p, x) in the domain $\mathcal{G}$. Here, use has been made of the fact that $f(p, x, \pi, \xi)$ is assumed to vanish for $\xi = 0$ and to be a polynomial of finite order L in the π variables.

Next, we perform the canonical transformation (2.2), thus transforming $h_\Omega(\pi, \xi)$ to the form (2.3) and $f(p, x, \pi, \xi)$ as in (2.4), or, more explicitly, to the form

$$f'(p, x, \pi', \xi') = \sum_{s\geq 1} \sum_{l=0}^{L} \tilde{f}^{(l,s)}(p, x, (\lambda\Omega)^{1/2}\pi', (\lambda\Omega)^{-1/2}\xi') \ . \tag{4.2}$$

Divide now the whole Hamiltonian by λ, which corresponds to a rescaling of the time variable, and denote

$$\varepsilon = \lambda^{-1/2} \ ; \tag{4.3}$$

then the Hamiltonian takes the form, omitting the primes,

$$H(p, x, \pi, \xi, \varepsilon) = \sum_{s\geq 0} H_s(p, x, \pi, \xi, \varepsilon), \tag{4.4}$$

where

$$\begin{aligned} H_0(p, x, \pi, \xi, \varepsilon) &= h_\Omega(\pi, \xi) + \varepsilon^2 \hat{h}(p, x) \\ h_\Omega(\pi, \xi) &= \frac{1}{2} \sum_{l=1}^{\nu} \Omega_l \left(\pi_l^2 + \xi_l^2\right) \ , \end{aligned} \tag{4.5}$$

and H_s, for $s \geq 1$, is a nonhomogeneous polynomial of degree $s + L$ in π, ξ which, local estimates of theorem 2.1' the canonical transformation reduces to the identity (i.e. $\pi(t) - \pi'(t) \to 0$ and $\xi(t) - \xi'(t) \to 0$ as $t \to \pm\infty$). This means that we have proven the following

Corollary 3.1: *Consider the Hamiltonian system (1.3), describing the interaction of two diatomic molecules; assume that the interaction potential decays sufficiently rapidly as $|x_2 - x_1| \to \infty$, as discussed above, and let the two molecules have a collision, which satisfies the above assumptions i. and ii. Then there exists a constant $\hat{V}_0$ independent of λ, such that one has*

$$\limsup_{t\to\infty} |h_\Omega(\pi(t),\xi(t)) - h_\Omega(\pi(-t),\xi(-t))| \le \lambda^{-1} e^{-\frac{\lambda}{e\lambda_*}} \hat{V}_0 \ .$$

Let us now consider the problem of the relaxation times in a polyatomic gas, and discuss in particular the Hamiltonian system (1.3). In this case Ω is completely resonant, so that, as commented in sect. 2, one can take $\tau = 0$ and $\gamma = \Omega_1$, while on the other hand one must take $\mathcal{M} = \mathcal{M}_\Omega$ and μ parallel to Ω. Taking $\mu = \Omega_1$, so that $I_\mu = h_\Omega$ and $|\mu| = \nu\Omega_1 = 2\Omega_1$, one can use theorem 2.1' and estimate

$$\left|\dot{h}_\Omega(\pi', \xi')\right| \leq 2^{11} e^3 \Omega_1 \lambda^{-1} e^{-\frac{\lambda}{e\lambda_*}} \hat{E}(p, x) \ .$$

We now assume that the interaction potential between the two molecules decays with the distance $r = |x_2 - x_1|$ more rapidly than r^{-1}; this means that for large r, say for $r \geq r_0$, there exists a function $\hat{V}(r)$ such that $\hat{E}$ is bounded by $\hat{E}(p, x) \leq \hat{V}(r)$, and the integral $\int_{r_0}^{\infty} \hat{V}(r) \mathrm{d}r$ converges to a constant, $\hat{V}_0$ say.

We then assume that the two molecules have a collision, precisely that

i. for $t \to \pm\infty$ one has $r(t) \to \infty$;

ii. there exists a finite time T_0 and a positive constant w_0, such that $\dot{r}(t) \leq -w_0$ for all $t \leq -T_0$ and $\dot{r}(t) \geq w_0$ for all $t \geq T_0$ (this means essentially that the collision time is bounded). One can always take T_0 such that $r(\pm T_0) \geq r_0$.

It could be seen that if one makes the assumptions i. and ii. for a model of perfectly rigid molecules, i.e. for a model described by the Hamiltonian $H(p, x) = \hat{h}(p, x) = \frac{1}{2}(p_1^2 + p_2^2) + V(x_1, x_2, 0, 0)$, then the same assumptions are automatically satisfied also for the Hamiltonian (1.3), if λ is sufficiently large.

From the above assumptions one gets immediately

$$\begin{aligned} \limsup_{t\to\infty} & \left|h_\Omega(\pi'(t), \xi'(t)) - h_\Omega(\pi'(-t), \xi'(-t))\right| \\ & \leq 2^{11} e^3 \Omega_1 \lambda^{-1} e^{-\frac{\lambda}{e\lambda_*}} \int_{-\infty}^{+\infty} \hat{E}(p(t), x(t)) \mathrm{d}t \ , \end{aligned}$$

and the integral clearly converges: indeed one has

$$\int_{-\infty}^{+\infty} \hat{E}(p(t), x(t)) \mathrm{d}t \leq \int_{-\infty}^{-T_0} \hat{V}(r(t)) \mathrm{d}t + \int_{-T_0}^{+T_0} \hat{E}(p(t), x(t)) \mathrm{d}t + \int_{T_0}^{\infty} \hat{V}(r(t)) \mathrm{d}t \ ,$$

and both the first and the third integrals are bounded by $w_0^{-1} \int_{T_0}^{\infty} \hat{V}(r(t)) \mathrm{d}t \leq w_0^{-1} \hat{V}_0$, while the second one is clearly finite. Denoting by $\hat{V}_0$ the sum of the three integrals, multiplied by $2^{11} e^3 \Omega_1$, one gets finally

$$\limsup_{t\to\infty} \left|h_\Omega(\pi'(t), \xi'(t)) - h_\Omega(\pi'(-t), \xi'(-t))\right| \leq \lambda^{-1} e^{-\frac{\lambda}{e\lambda_*}} \hat{V}_0 \ .$$

Now, in the l.h.s. of this inequality one can clearly drop the primes: indeed, in the above assumptions, one has $\hat{E}(p(t), x(t)) \to 0$ for $t \to \pm\infty$, and, in virtue of the

The proofs of the corollaries are deferred to sect. 8.

3. Physical application

Before entering the perturbation scheme, let us briefly discuss the application to the physical systems discussed in the introduction, namely a constrained system with nonresonant frequencies, and a statistical model of a diatomic gas.

Consider first the case of a constrained system. In a practical application we are interested in real values of the p, x, π, ξ variables in a domain $\mathcal{G} \times B$, as discussed in the introduction. In particular we are interested in systems for which the energy surfaces $\hat{h}(p, x) = \mathcal{E}$ have a compact component for, say, $\mathcal{E} < \hat{E}_0$. Assume now that these energy surfaces are contained in the complex domain $\mathcal{G}_{\frac{1}{4}R}$, defined by (2.12), and take real initial data $(p_0, x_0, \pi_0, \xi_0) \in \mathcal{D}_{\frac{1}{4}R,\varrho}$, still with $\varrho = \lambda^{-1/2}$. Assume moreover that the harmonic frequencies are nonresonant, and that the usual diophantine condition

$$|k \cdot \omega| \geq \gamma \, |k|^{-\tau}$$

holds with real constants $\gamma > 0$ and $\tau > n - 1$, so that the normalized system admits ν approximate prime integrals $I_l = \frac{1}{2}(p_l^2 + x_l^2)$, $1 \leq l \leq \nu$. Then a sufficient condition to guarantee a small energy exchange among the subsystems described by $\hat{h}(p, x)$ and $h_\Omega(\pi, \xi)$ is that the initial datum satisfies the condition

$$\begin{aligned} &\hat{h}(p_0, x_0) \leq \hat{E}_0 - 2^8 \lambda^{-1} E \\ &I_l(\pi_0, \xi_0) \leq \left[\frac{1}{8} - \frac{1}{4} e^{-\tau} \left(\frac{\lambda_*}{\lambda} \right)^{\frac{1}{\tau+1}} \right] \lambda^{-1} R^2 \ , \quad 1 \leq l \leq \nu \end{aligned} \tag{3.1}$$

(notice that the second bound is certainly positive, since $\lambda > e^{\tau+1} \lambda_*$). Indeed, corollary 2.4 states that for $|t| < T$ one has $|\hat{h}(p(t), x(t))| < \hat{E}_0$ and $I_l(\pi(t), \xi(t)) < \frac{1}{8} \lambda^{-1} R^2$, so that $(p(t), x(t), \pi(t), \xi(t)) \in \mathcal{D}_{\frac{1}{4}R,\varrho}$. This in turn implies, by theorem 2.1, that the transformed point $(p'(t), x'(t), \pi'(t), \xi'(t))$ belongs to $\mathcal{D}_{\frac{1}{2}R,\varrho}$, so that the escape time T_0 from the latter domain actually exceeds T, and the energy sharing, according to the definition of T in corollary 2.4, takes an exponentially large time. In such case the constant λ_*, according to its definition (2.17) and the estimated values of σ and E in sect. 8, turns out to be proportional to some positive power of ν, and the exponent a in (1.4) turns out to be of order $1/\nu$.

(2.18), the time derivative of $I_\mu(\pi',\xi')$ is bounded in $\mathcal{D}_{\frac{1}{2}R,\varrho}$ by

$$|\dot I_\mu(p',x',\pi',\xi',\lambda)| < 2^{10}e^{\tau+3}\lambda^{-1}\,|\mu|\,E\exp\left[-\frac{\tau+1}{e}\left(\frac{\lambda}{\lambda_*}\right)^{\frac{1}{\tau+1}}\right] . \tag{2.21}$$

An easy consequence of theorem 2.1 is that one can estimate the change in time of the approximate integrals $I_\mu(\pi,\xi)$. In particular we are interested in orbits starting from initial data $(\pi_0,\xi_0)\in\mathcal{G}\times\mathbf{R}^\nu$ with (π_0,ξ_0) close enough to the origin of $\mathbf{R}^\nu$. Denoting by $(p(t),x(t),\pi(t),\xi(t))$ such an orbit, and by $(p'(t),x'(t),\pi'(t),\xi'(t))$ the transformed orbit under the canonical transformation $\mathcal{C}_\lambda$, we can prove the following

Corollary 2.3: *For an orbit of the Hamiltonian system (2.16) with the initial datum $(p_0,x_0)\in\mathcal{G}$ and real $(\pi_0,\xi_0)\in\Delta_{\frac{1}{4}\varrho R}$, with $\varrho=\lambda^{-1/2}$, one has the bound*

$$|I_\mu(\pi'(t),\xi'(t)) - I_\mu(\pi'(0),\xi'(0))| < 2^{-3}e^{-\tau}\,|\mu|\,R^2\frac{E}{E_0}\lambda^{-1}\left(\frac{\lambda_*}{\lambda}\right)^{\frac{1}{\tau+1}}\frac{|t|}{T} \tag{2.22}$$

for

$$|t| = \min(T_0,T)\ , \tag{2.23}$$

where T_0 is the (possibly infinite) escape time of $(p'(t),x'(t),\pi'(t),\xi'(t))$ from $\mathcal{D}_{\frac{1}{2}R,\varrho}$, and

$$T = 2^{-4}e^{-(\tau+3)}\,|\Omega|^{-1}\,\lambda^{-1}\exp\left[\frac{\tau+1}{e}\left(\frac{\lambda}{\lambda_*}\right)^{\frac{1}{\tau+1}}\right] . \tag{2.24}$$

Moreover, if $\mathcal{M}=\mathcal{M}_\Omega$, one has over the same time interval the bound

$$\begin{aligned} |h_\Omega(\pi'(t),\xi'(t)) - h_\Omega(\pi'(0),\xi'(0))| &< 2^6\lambda^{-2}E \\ |\hat h(p'(t),x'(t)) - \hat h(p'(0),x'(0))| &< 2^8\lambda^{-1}E\ . \end{aligned} \tag{2.25}$$

Analogous results hold for the same functions of the original variables. One has in this connection the following

Corollary 2.4: *For the same orbit of corollary 2.3 over the same time interval one has the bound*

$$|I_\mu(\pi(t),\xi(t)) - I_\mu(\pi(0),\xi(0))| < 2^3e^{-\tau}\,|\mu|\,R^2\frac{E}{E_0}\lambda^{-1}\left(\frac{\lambda_*}{\lambda}\right)^{\frac{1}{\tau+1}}\frac{|t|}{T}\ . \tag{2.26}$$

Moreover, if $\mathcal{M}=\mathcal{M}_\Omega$ one has

$$\begin{aligned} |h_\Omega(\pi(t),\xi(t)) - h_\Omega(\pi(0),\xi(0))| &< 2^6\lambda^{-2}E \\ |\hat h(p(t),x(t)) - \hat h(p(0),x(0))| &< 2^8\lambda^{-1}E\ . \end{aligned} \tag{2.27}$$

The proof is deferred to sect. 8, where an indication is also given on how to compute the constants E and σ. According to such indication, a possible choice is $\sigma = \frac{\nu}{\varrho_*}$ and $E = \frac{1-\sigma^3}{1-\sigma}\sigma \sup_{\mathcal{D}_{R,\varrho_*}} |f(p,x,\pi,\xi)|$.

In some applications, in particular in dealing with colliding molecules, the estimates (2.19) involving the "global" constant E turn out to be slightly rough, because they do not exploit the possibility that in some regions of phase space (say when the molecules are far apart) the coupling term $f(p,x,\pi,\xi)$ in the Hamiltonian (2.16) be negligible. One then needs a more detailed, so to say "local", version of the above estimates, obtained by replacing in the r.h.s. of (2.19) the constant E by a convenient function $\hat{E}(p,x)$. In fact, one can prove the following theorem, which will be used in discussing the collision of molecules in sect. 3.

Theorem 2.1': *In the same assumptions of theorem 2.1, there exists a function $\hat{E}(p,x)$, $(p,x) \in \mathcal{G}$, with $0 \le \hat{E}(p,x) \le E$, such that the same conclusions of that theorem hold, with $\hat{E}(p,x)$ in place of E everywhere in the inequalities (2.19). A possible choice of $\hat{E}$ is*

$$\hat{E}(p,x) = \frac{1-\sigma^3}{1-\sigma}\sigma \sup_{\mathcal{D}_{R,\varrho_*}(p,x)} |f(p',x',\pi,\xi)| \ ,$$

with $\mathcal{D}_{R,\varrho_*}(p,x) = D_R(p,x) \times \Delta_{\varrho_* R}$, and $D_R(p,x)$ defined by (2.13).

The proof of this local version of the theorem requires very minor changes with respect to the proof of theorem 2.1, and will be sketched in sect. 8. Essentially, the local formulation is possible because canonical transformations are themselves locally defined.

The theorem 2.1 has the following

Corollary 2.2: *The normalized part $\lambda h_\Omega(\pi',\xi') + \hat{h}(p',x') + Z(p',x',\pi',\xi',\lambda)$ of the Hamiltonian (2.18) admits $\nu - \dim \mathcal{M}$ independent prime integrals of the form $I_\mu(\pi',\xi') = \sum_{l=1}^{\nu} \frac{\mu_l}{2}({\pi'}_l^2 + {\xi'}_l^2)$, with $\mu \in \mathbf{R}^\nu$ and $\mu \perp \mathcal{M}$. If π',ξ' and π,ξ are related by the canonical transformation $\mathcal{C}_\lambda$, then one has the bound*

$$|I_\mu(\pi',\xi') - I_\mu(\pi,\xi)| < 2^2 e^{-\tau} \lambda^{-1} \left(\frac{\lambda_*}{\lambda}\right)^{\frac{1}{\tau+1}} |\mu| \, R^2 \frac{E}{E_0} \ , \qquad (2.20)$$

with $|\mu| = \sum_{l=1}^{\nu} |\mu_l|$. Moreover, if one takes into account the whole Hamiltonian and introducing the complex domain

$$\mathcal{D}_{R,\varrho} = \mathcal{G}_R \times \Delta_{\varrho R} \ . \tag{2.15}$$

The value of ϱ must be chosen in such a way that the Hamiltonian is convergent in $\mathcal{D}_{R,\varrho}$.

We can now state the main theorem of this paper, which is proven in Part B in a more general and mathematically more natural form.

Theorem 2.1: *Consider the Hamiltonian*

$$H(p,x,\pi,\xi,\lambda) = \lambda h_\Omega(\pi,\xi) + \hat{h}(p,x) + \frac{1}{\lambda} f_\lambda(p,x,\pi,\xi) \tag{2.16}$$

with $h_\Omega(\pi,\xi)$ of the form (2.6) and f_λ a polynomial in π of order $L \le 2$. Assume that H is analytic in the interior of a domain $\mathcal{D}_{R,\varrho_*}$ defined by (2.15), with positive ϱ_*, and bounded in $\mathcal{D}_{R,\varrho_*}$, and denote $E_0 = \sup_{(p,x)\in\mathcal{G}_R} |\hat{h}(p,x)|$. With reference to the vector Ω of the frequencies of h_Ω and to the module $\mathcal{M}_\Omega$ related to it, let $\mathcal{M} \supset \mathcal{M}_\Omega$ be a resonance module, and assume the nonresonance condition (2.11) with real constants $\gamma > 0$ and $\tau \ge 0$.
Then there exist positive constants E and σ, depending on ϱ_*, on ν and on the perturbation f_λ, such that for any $\lambda \ge e^{\tau+1}\lambda_*$, with

$$\lambda_* = \frac{2^8}{\gamma R^2}\left[2e^2 E + 3(4E+E_0) + \frac{1}{4}\gamma R^2 \sigma\right] \ , \tag{2.17}$$

and for $\varrho = \min(\lambda^{-1/2}, \varrho_*)$ there exists a real analytic, near to identity canonical transformation $\mathcal{C}_\lambda$ from $\mathcal{D}_{\frac{1}{2}R,\varrho}$ to $\mathcal{D}_{\frac{3}{4}R,\varrho}$, with $\mathcal{C}_\lambda(\mathcal{D}_{\frac{1}{2}R,\varrho}) \supset \mathcal{D}_{\frac{1}{4}R,\varrho}$, which puts the Hamiltonian into the form

$$\begin{aligned} H'(p',x',\pi',\xi',\lambda) = \lambda h_\Omega(\pi',\xi') + \hat{h}(p',x') \\ + Z(p',x',\pi',\xi',\lambda) + \mathcal{R}(p',x',\pi',\xi',\lambda) \ , \end{aligned} \tag{2.18}$$

where $Z(p',x',\pi',\xi',\lambda)$ is in normal form with respect to the resonance module $\mathcal{M}$. Moreover, for $(p',x',\pi',\xi') \in \mathcal{D}_{\frac{1}{2}R,\varrho}$ one has the bounds

$$\begin{aligned} |h_\Omega(\pi',\xi') - h_\Omega(\pi,\xi)| &< 2^5 \lambda^{-2} E \\ |\hat{h}(p',x') - \hat{h}(p,x)| &< 2^2 e^{-\tau} \lambda^{-1} \left(\frac{\lambda_*}{\lambda}\right)^{\frac{1}{\tau+1}} E \\ |Z(p',x',\pi',\xi',\lambda)| &< 2^5 \lambda^{-1} E \\ |\mathcal{R}(p',x',\pi',\xi',\lambda)| &< 2^6 e^{\tau+2} \lambda^{-1} \left(\frac{\lambda_*}{\lambda}\right)^{\frac{1}{\tau+1}} E \exp\left[-\frac{\tau+1}{e}\left(\frac{\lambda}{\lambda_*}\right)^{\frac{1}{\tau+1}}\right] \end{aligned} \tag{2.19}$$

As is typical in perturbation theory, we shall use the bound

$$|k \cdot \Omega| \geq \gamma \, |k|^{-\tau} \quad \text{for } k \in \mathbf{Z}^{\nu} \setminus \mathcal{M} \tag{2.11}$$

with real constants $\gamma > 0$ and $\tau \geq 0$. In particular we shall consider the two extreme cases of nonresonance, i.e. $k \cdot \Omega = 0$ implies $k = 0$, and of complete resonance with equal positive frequencies $\Omega_1 = \ldots = \Omega_\nu$. In the first case one has $\mathcal{M}_\Omega = \{0\}$, so that any $\mathcal{M}$ can be used; as is well known, even for $\mathcal{M} = \mathcal{M}_\Omega = \{0\}$, which gives the most stringent normal form, the condition (2.11) is satisfied, for $\tau > \nu - 1$, by a set of Ω's of large measure if γ is small enough. In the latter case $\mathcal{M}_\Omega$ is $(\nu-1)$–dimensional, and we are forced to take $\mathcal{M} = \mathcal{M}_\Omega$; concerning γ and τ, we can take $\gamma = \Omega_1$ and $\tau = 0$. Notice that in this latter case the expression $|k \cdot \Omega|$ either vanishes, if $k \in \mathcal{M}_\Omega$, or is a multiple of Ω_1, so that there are no small denominators at all. As will be shown in sect. 7, this elementary remark lies at the very heart of the ν–independence of the exponent a in (1.4).

In order to use the standard methods of perturbation theory, in particular the Cauchy's estimates for the derivatives of analytic functions, we must consider a complex domain $\mathcal{G}_R$, defined as the union of polydisks centered on every point of $\mathcal{G}$. More precisely, we define

$$\mathcal{G}_R = \bigcup_{(p,x) \in \mathcal{G}} D_R(p,x) \ , \tag{2.12}$$

where

$$D_R(p,x) = \left\{ (p', x') \in \mathbf{C}^{2n} : |p_l - p'_l| \leq R, |x_l - x'_l| \leq R, 1 \leq l \leq n \right\} \ . \tag{2.13}$$

Here, for the sake of simplicity, the (p,x) variables are assumed to be dimensionally homogeneous, as the variables (π, ξ) are (all of them being square roots of actions). The extension to the case of dimensionally nonhomogeneous variables is just a trivial technical fact. For what concerns the variables π, ξ, we make a power series expansion of the perturbation $f_\lambda(p, x, \pi, \xi)$ in the neighbourhood of the origin, and the coefficients of such expansion can be assumed to be analytic functions of p, x in the interior of the domain $\mathcal{G}_R$ and bounded in $\mathcal{G}_R$. The natural domain of definition of the Hamiltonian $H(p, x, \pi, \xi, \lambda)$ defined by (2.5), as well as of the normal form (2.7), can then be built up by considering the variables π, ξ confined, for a given dimensionless parameter $\varrho > 0$, in the polydisk $\Delta_{\varrho R}$ of radius ϱR around the origin of $\mathbf{C}^{2\nu}$, i.e.

$$\Delta_{\varrho R} = \left\{ (\pi, \xi) \in \mathbf{C}^{2\nu} : |\pi_l| \leq \varrho R, |\xi_l| \leq \varrho R, 1 \leq l \leq \nu \right\} \ , \tag{2.14}$$

in the form $f = \xi\tilde{f}$, and consequently the transformation (2.2) gives a factor $\lambda^{-1/2}$; moreover, we shall consider the new variables π', ξ' confined in a neighbourhood of the origin of size $\lambda^{-1/2}$, so that the perturbation turns out to be of size λ^{-1}.

According to the usual procedure in perturbation theory we try now to put the Hamiltonian (2.5) in normal form, i.e. we look for a near to identity canonical transformation $(p, x, \pi, \xi) = \mathcal{C}_\lambda(p', x', \pi', \xi')$, which puts the Hamiltonian in the form

$$\begin{aligned} H'(p', x', \pi', \xi', \lambda) &= \lambda h_\Omega(\pi', \xi') + \hat{h}(p', x') \\ &\quad + Z(p', x', \pi', \xi', \lambda) + \mathcal{R}(p', x', \pi', \xi', \lambda) \ , \end{aligned} \tag{2.7}$$

where Z is the normalized part, and $\mathcal{R}$ is a small unnormalized remainder (*noise*).

Since the variables p, x must be essentially considered as parameters, we define the normal form with respect to the variables π, ξ only. Precisely, we introduce the usual linear canonical transformation to complex variables $\tilde{\pi}, \tilde{\xi}$ defined by

$$\pi_l = \frac{1}{\sqrt{2}}\left(\tilde{\pi}_l + i\tilde{\xi}_l\right) \ , \quad \xi_l = \frac{i}{\sqrt{2}}\left(\tilde{\pi}_l - i\tilde{\xi}_l\right) \ , \quad (1 \le l \le \nu), \tag{2.8}$$

which puts the unperturbed Hamiltonian $h_\Omega(\pi, \xi)$ into the form

$$\tilde{h}_\Omega(\tilde{\pi}, \tilde{\xi}) = i\sum_{l=1}^{\nu} \Omega_l \tilde{\pi}_l \tilde{\xi}_l \ , \tag{2.9}$$

while $\hat{h}(p, x)$ is unchanged. Moreover, with reference to the vector Ω of the harmonic frequencies of $h_\Omega(\pi, \xi)$, we introduce the resonance module $\mathcal{M}_\Omega$ defined by

$$\mathcal{M}_\Omega = \{k \in \mathbf{Z}^\nu : k \cdot \Omega = 0\} \ . \tag{2.10}$$

Then, by considering a (possibly larger) module $\mathcal{M} \supset \mathcal{M}_\Omega$, Z is said to be in normal form with respect to $\mathcal{M}$ in case the power expansion of Z in the complex variables $\tilde{\pi}, \tilde{\xi}$ contains only monomials $\tilde{\pi}^j \tilde{\xi}^k$ such that $j - k \in \mathcal{M}$.

In fact, for our problem of giving a bound on the energy exchange between h_ω and $\hat{h}$, we are mainly interested in the case $\mathcal{M} = \mathcal{M}_\Omega$, because in this case, as is immediately checked, the Poisson bracket $\{h_\Omega, Z\}$ vanishes, and consequently $h_\Omega(\pi', \xi')$ turns out to be a constant of motion, up to the small noise due to $\mathcal{R}$. However, it is more natural in perturbation theory to take $\mathcal{M}$ possibly larger than $\mathcal{M}_\Omega$, so we will present our main result in this slightly more general case, and deduce from it some corollaries adapted to the case $\mathcal{M} = \mathcal{M}_\Omega$.

As is well known, in performing the normalization procedure with reference to the module $\mathcal{M}$, there appear small denominators of the form $k \cdot \Omega$, with $k \in \mathbf{Z}^\nu \setminus \mathcal{M}$.

Part A – Results.

2. Statement of the results

We start with the Hamiltonian (1.1) with h_ω given by (1.2), and assume both $\hat{h}(p,x)$ and $f(p,x,\pi,\xi)$ to be analytic functions of the canonical variables $(p,x,\pi,\xi) \in \mathcal{G}\times B$, where $\mathcal{G} \subset \mathbf{R}^{2n}$ is a bounded domain (it can be thought of as the natural domain of definition of $\hat{h}$, say a domain inside a compact energy surface), and $B \subset \mathbf{R}^{2\nu}$ is a suitable neighbourhood of the origin of $\mathbf{R}^{2\nu}$, to be specified later. Moreover, $f(p,x,\pi,\xi)$ is assumed to vanish for $\xi = 0$, and to be a polynomial of finite order L in π. In the following, we shall assume $L = 2$, which covers all the cases of physical interest illustrated in the introduction, and at the same time simplifies the task of building an adapted algebraic framework, as will be done in sect. 4.

Introduce now the dimensionless parameter λ by setting

$$\omega = \lambda\Omega \ , \tag{2.1}$$

where $\Omega = (\Omega_1, \ldots, \Omega_\nu)$ is a fixed set of frequencies (for example chosen of the same order of the inverse of a typical time scale of $\hat{h}$), and perform the usual canonical change of variables

$$\xi_l = (\lambda\Omega_l)^{-1/2}\xi'_l \ , \quad \pi_l = (\lambda\Omega_l)^{1/2}\pi'_l \ , \quad 1 \le l \le \nu \ , \tag{2.2}$$

which transforms h_ω into

$$h'_{\lambda\Omega}(\pi',\xi') = \frac{\lambda}{2}\sum_{l=1}^{\nu}\Omega_l\left(\pi'^2_l + \xi'^2_l\right) \ , \tag{2.3}$$

while $f(p,x,\pi,\xi)$ is transformed into

$$f'(p,x,\pi',\xi') = f(p,x,(\lambda\Omega)^{1/2}\pi',(\lambda\Omega)^{-1/2}\xi') \ . \tag{2.4}$$

Then the Hamiltonian takes the form, omitting primes,

$$H(p,x,\pi,\xi,\lambda) = \lambda h_\Omega(\pi,\xi) + \hat{h}(p,x) + \frac{1}{\lambda}f_\lambda(p,x,\pi,\xi) \ , \tag{2.5}$$

where

$$h_\Omega(\pi,\xi) = \frac{1}{2}\sum_{l=1}^{\nu}\Omega_l\left(\pi_l^2 + \xi_l^2\right) \ . \tag{2.6}$$

The factor $\frac{1}{\lambda}$ in front of f_λ recalls the fact that the perturbation is of order λ^{-1}, and is explained as follows: by hypothesis f vanishes for $\xi = 0$, so that it can be written of refs. [11], [9] and [10]. However, due to the peculiarities of the model, such a scheme must be adapted in several nontrivial points, and for this reason we give here all estimates in detail.

The paper is organized in three parts as follows. In Part A (sects. 2 and 3) the general results and the application to physical systems are discussed in synthetic form, without details on the proofs. In particular, sect. 3 contains a discussion on the relevance of our results in the thermodynamic limit. The rest of the paper is devoted to the development of the perturbative scheme; we try to make reading easier by separating the general scheme (Part B, sects. 4 to 8) from technical lemmas and proofs (Part C, sects. 9 to 12). For what concerns Part B, first the perturbation procedure is explained in sections 4 to 7 in a slightly more general framework than that of sect. 2. More precisely, in sect. 4 we give the Hamiltonian a suitable form and build up an algebraic framework adapted to our problem; in sect. 5 we illustrate an algebraic approach to canonical transformations and give the necessary estimates; in sect. 6 we produce the generating sequence of the canonical transformation which puts the Hamiltonian in suitable normal form; in sect. 7 the canonical transformation is used and the exponential estimates on the remainder are obtained. The theorems of sect. 7 can be considered as the main general results of the paper, and could be applied to several models. Then, in sect. 8 such results are applied to the models discussed above, and the results claimed in Part A are proven. The remaining part of the paper, namely Part C, is devoted to the detailed proofs of all theorems of sections 4 to 7.

Part of this work was done when two of the authors (G.B and A.G.) were guests of the Institute of Theoretical Physics of ETH, Zürich. We are grateful to Professor J. Fröhlich for his kind invitation.

during a collision is estimated by

$$\Delta E < \tau W_* e^{-\lambda/\lambda_*} \ , \tag{1.5}$$

where τ is the collision time. So, if one accepts that only the few body collisions are relevant in reaching the statistical equilibrium (as usual in discussing the Boltzmann equation), one has a *freezing* of the internal vibrational degrees of freedom of the molecules for times of the order e^{λ/λ_*}, as suggested by Boltzmann and Jeans; this is quite relevant for the dynamical foundations of classical statistical mechanics.

ii. In the problem of realization of constraints the energy exchange between the motion along the constraints and the transversal vibrations is negligible up to a time of order $e^{(\lambda/\lambda_*)^a}$.

From a technical point of view, this part II is substantially different and more complicated than part I. In particular, in order to have an effective control on the ν–dependence of the constants, in particular of a, the relevance of which was illustrated above, one needs a careful choice of the algebraic characterization of the problem and generically of all the ingredients of classical perturbation theory.

The most apparent technical difference is that here we work directly in cartesian coordinates π, ξ instead of using action–angle variables, I, φ say, as is customary in perturbation theory, and as was done in part I. The reason is that the action–angle variables have a singular point for $I = 0$, and this requires to work within domains excluding such points. This is not a serious difficulty if $\nu = 1$, and even in the general case $\nu > 1$, provided the frequencies ω are strongly nonresonant, because in such cases one can bound the motion away from $I = 0$. However, the difficulty becomes serious when resonance relations exist between the ω's, because in the latter case the point $I = 0$ cannot be excluded by dynamical considerations. On the other hand, this is obviously not an intrinsic difficulty, being just due to the choice of the coordinates. Taking instead into account the fact that we are interested in initial data for the variables π, ξ in a neighbourhood of the origin, the problem looks more similar to that of a system of harmonic oscillators, with a polynomial or analytic perturbation, at least if we can consider the variables p, x, in some sense, as parameters, as we shall do. In fact, we develop our perturbative scheme having in mind the formal results of Whittaker,[6] Cherry[7] and Birkhoff,[8] and just add rigorous estimates along the lines of refs. [9] and [10].

The perturbation scheme developed in this paper is strongly reminiscent of that the system we propose to keep in mind is a point mass confined to a closed curve in space, for example a circle of radius R in the x, y plane. Using cylindrical coordinates $r = R + \xi_1$, $\vartheta \equiv x$, $z = \xi_2$, one has

$$H = \frac{1}{2m}\left(\pi_1^2 + \pi_2^2 + \frac{p^2}{(r+\xi_1)^2}\right) + \frac{1}{2}\left(k_1\xi_1^2 + k_2\xi_2^2\right) + V(x,\xi_1,\xi_2)\ ,$$

where V is the external potential, while k_1, k_2 are large elastic constants, say $k_i = \lambda^2 K_i$, $i = 1, 2$, providing in the limit $\lambda \to \infty$ the required confinement. Here too one easily gets the above form (1.1), by putting $\hat{h} = \frac{p^2}{2mR^2} + V(x, 0, 0)$, $f = V(x, \xi_1, \xi_2) - V(x, 0, 0) + \frac{p^2}{2m(R+\xi)^2} - \frac{p^2}{2mR^2}$, and introducing as above a trivial rescaling of the π, ξ coordinates.

While in part I we were dealing, by a rather simple direct technique, with the particular case $\nu = 1$ (collision of one molecule with a wall; realization of just one constraint, as for example in a spherical pendulum), in the present part II we deal instead with the much more difficult general case $\nu > 1$.

Roughly speaking, denoting by W the rate of the energy transfer between the two subsystems, we look for Nekhoroshev's like exponential estimates[2] of the type

$$W < W_* e^{-(\lambda/\lambda_*)^a}\ , \tag{1.4}$$

with explicit expressions for the constants W_*, λ_* and a, which in general turn out to be dependent on ν, as typical of perturbation theory.

However, in the very relevant case of identical frequencies (collision of identical molecules), for the constant a, which is clearly the most relevant one in asymptotic estimates, we find $a = 1$ for any ν. This is exactly the value conjectured by Jeans already in the year 1903[3], later independently proposed, for example by Landau and Rapp[4], on the basis of heuristic considerations; for a numerical study oriented to the above problem i., also giving $a = 1$, see ref. [5]. On the other hand, the value obtained by just adapting the common perturbation techniques would be $1/\nu^2$ or $1/\nu$ (this fact led some authors to conclude that such a kind of results are irrelevant for statistical mechanics). One might point out, in this connection, the curious fact that the proof of the independence of the quantity a from the number of degrees of freedom is obtained by suitably exploiting resonance; indeed, according to a kind of folklore, resonance is often considered to be responsible for the pretended failure of the freezing of the energy exchanges in the thermodynamic limit.

From (1.4) we get two consequences:

i. In the problem of a collision of ν identical molecules, the energy exchange ΔE

1. Introduction

As in the first part of this paper[1] we consider a Hamiltonian system of the form

$$H(p,x,\pi,\xi) = h_\omega(\pi,\xi) + \hat{h}(p,x) + f(p,x,\pi,\xi) \ ; \tag{1.1}$$

here $h_\omega(\pi,\xi)$, with $(\pi,\xi) = (\pi_1,\ldots,\pi_\nu,\xi_1,\ldots,\xi_\nu) \in \mathbf{R}^{2\nu}$, is the Hamiltonian of a set of ν uncoupled harmonic oscillators of angular frequency $\omega = (\omega_1,\ldots,\omega_\nu)$, i.e.

$$h_\omega = \frac{1}{2}\sum_{l=1}^{\nu} \left(\pi_l^2 + \omega_l^2\xi_l^2\right) \ , \tag{1.2}$$

while $\hat{h}(p,x)$, with $(p,x) = (p_1,\ldots,p_n,x_1,\ldots,x_n) \in \mathcal{G} \subset \mathbf{R}^{2n}$, represents any dynamical system with n degrees of freedom, defined on a domain $\mathcal{G}$, and $f(p,x,\pi,\xi)$ is a coupling term which is assumed to vanish for $\xi = 0$. We are interested in the case of large ω's, say $\omega = \lambda\Omega$, with some fixed $\Omega = (\Omega_1,\ldots,\Omega_\nu) \in \mathbf{R}^\nu$ and large λ. Our aim is to study the rate of the energy exchange between the two subsystems described by the Hamiltonians $h_\omega(\pi,\xi)$ and $\hat{h}(p,x)$, due to the coupling term $f(p,x,\pi,\xi)$; notice that the Hamiltonian (1.1) is not in general a perturbation of an integrable system.

A Hamiltonian like (1.1) naturally appears in at least two typical problems in classical physics:

i. The relaxation times in statistical mechanics, in particular the efficiency of ν–body collisions in polyatomic gases for the energy exchanges among different degrees of freedom, possibly leading to equilibrium. The simplest model example one should have in mind is the collinear collision of two identical diatomic molecules, with short–range interaction forces, namely

$$H(p,x,\pi,\xi) = \sum_{l=1}^{2}\left(\frac{\pi_l^2}{2\mu} + \frac{\mu\omega^2\xi_l^2}{2} + \frac{p_l^2}{2M}\right) + V(x_1,x_2,\xi_1,\xi_2) \ , \tag{1.3}$$

where: x_1, x_2 are the coordinates of the centers of mass of the molecules; ξ_1,ξ_2 those of the internal degrees of freedom ($\xi = 0$ for molecules at rest); p_1,p_2,π_1,π_2 are the conjugate momenta; M and μ are the total and the reduced mass respectively of a molecule; finally, V is a short range interaction potential, as will be made precise below. One recovers the above form (1.1) by putting $\hat{h} = \frac{1}{2M}(p_1^2+p_2^2) + V(x_1,x_2,0,0)$, $f = V(x_1,x_2,\xi_1,\xi_2) - V(x_1,x_2,0,0)$; the trivial rescaling $\pi_l = \sqrt{\mu}\pi_l'$, $\xi_l = \xi_l'/\sqrt{\mu}$, $l = 1,2$, gives to h_ω the form (1.2).

ii. The realization of holonomic constraints in classical mechanics. This problem is discussed to some extent in part I, where some basic references are also given;

REALIZATION OF HOLONOMIC CONSTRAINTS AND FREEZING OF HIGH FREQUENCY DEGREES OF FREEDOM IN THE LIGHT OF CLASSICAL PERTURBATION THEORY. Part II

GIANCARLO BENETTIN
Dipartimento di Fisica dell'Università,
Via Marzolo 8, 35131 PADOVA, Italy

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

ANTONIO GIORGILLI
Dipartimento di Fisica dell'Università,
Via Celoria 16, 20133 MILANO, Italy

Abstract. As in Part I of this paper, we consider the problem of the energy exchanges between two subsystems, of which one is a system of ν harmonic oscillators, while the other one is any dynamical system of n degrees of freedom. Such a problem is of interest both for the realization of holonomic constraints of classical mechanics, and for the freezing of the internal degrees of freedom in molecular collisions. The results of Part I, which referred to the particular case $\nu = 1$, are here extended to the more difficult case $\nu > 1$. For the rate of energy transfer we find exponential estimates of Nekhoroshev's type, namely of the form $\exp(\lambda_*/\lambda)^{1/a}$, where λ is a positive real number giving the size of the involved frequencies, and λ_* and a are constants. For the particularly relevant constant a we find in general $a = 1/\nu$; however, in the particular case when the ν frequencies are equal (collision of identical molecules), we find $a = 1$ independently of ν, as conjectured by Jeans in the year 1903.