Zeitschrift: Commentarii Mathematici Helvetici
Herausgeber: Schweizerische Mathematische Gesellschaft
Band: 64 (1989)

Artikel: Convergence of Birkhoff normal forms for integrable systems.
Autor: Ito, Hidekazu
DOI: https://doi.org/10.5169/seals-48956

Comment. Math. Helvetici 64 (1989) 412–461 0010–2571/89/030412–50$01.50 + 0.20/0

Convergence of Birkhoff normal forms for integrable systems

HIDEKAZU ITO

1. Introduction

A basic tool for many differential equations is to transform them into a simpler form that is called *normal form*. In Hamiltonian systems, it is related to integrability of the system near an equilibrium point or a periodic motion. The purpose of this paper is to clarify this connection. Throughout this paper, we consider analytic or real analytic Hamiltonian systems and canonical transformations. We first consider analytic case and later treat real analytic system as its special case.

Let us consider a Hamiltonian system with n degrees of freedom

$$\frac{dx_k}{dt}=\frac{\partial H}{\partial y_k}, \qquad \frac{dy_k}{dt}=-\frac{\partial H}{\partial x_k} \qquad (k=1,\dots,n) \tag{1.1}$$

in a neighbourhood of an equilibrium point which we take at the origin $z=0\in\mathbf{C}^{2n}$, where $z=(x,y)$ with $x=(x_1,\dots,x_n)$, $y=(y_1,\dots,y_n)$. We assume that the Hamiltonian $H=H(x,y)$ is analytic in a neighbourhood of the origin with satisfying $H(0)=0$ and therefore its Taylor expansion begins with quadratic terms. The eigenvalues of the linearized system of (1.1) about the origin, which are determined by the quadratic terms of the H, occur in pairs $\pm\lambda_1,\dots,\pm\lambda_n$. The equilibrium point (the origin) is called *non-resonant* if the λ_k $(k=1,\dots,n)$ are rationally independent, that is, linearly independent over the field of rational numbers. This condition is equivalent to the condition

$$\sum_{k=1}^{n} m_k\lambda_k\neq 0 \quad \text{for any } (m_1,\dots,m_n)\in\mathbf{Z}^n\backslash\{0\},$$

which will be referred as the *non-resonance condition.*

In this paper, we consider the normalization of Hamiltonian systems near a non-resonant equilibrium point. Then the eigenvalues $\pm\lambda_1,\dots,\pm\lambda_n$ are all distinct and therefore we can find a linear canonical transformation which takes the Hamiltonian into the form

$$H = \sum_{k=1}^{n} \lambda_k x_k y_k + \cdots, \tag{1.2}$$

where the terms not written out explicitly denote a power series containing terms of order ≥ 3 only. For this fact, we refer to [11, §15]. Here a transformation (mapping) $z = \phi(\zeta)$ is called *canonical* (or *symplectic*) if the identity $\sum_{k=1}^{n} dx_k \wedge dy_k = \sum_{k=1}^{n} d\xi_k \wedge d\eta_k$ holds, where $\zeta = (\xi, \eta)$ with $\xi = (\xi_1, \ldots, \xi_n)$, $\eta = (\eta_1, \ldots, \eta_n)$. By a canonical transformation $z = \phi(\zeta)$, the Hamiltonian system (1.1) is transformed into a Hamiltonian system with Hamiltonian $H(\phi(\zeta))$.

In order to normalize higher order terms of the Hamiltonian, let us consider a canonical transformation ϕ of the form

$$\phi(\zeta) = \zeta + \text{terms of order } \geq 2. \tag{1.3}$$

G. D. Birkhoff [2] proved the following result.

THEOREM 1.1. *Let $H(z)\,(z = (x, y))$ be a power series of the form* (1.2). *Assume that $\lambda_1, \ldots, \lambda_n$ are rationally independent. Then there exists a canonical formal power series transformation $z = \phi(\zeta)\,(\zeta = (\xi, \eta))$ of the form* (1.3) *such that $H \circ \phi$ is a formal power series in n products $\xi_l \eta_l$ $(l = 1, \ldots, n)$.*

In the above, the transformation $z = \phi(\zeta)$ is not determined uniquely. However the function $H \circ \phi$ is uniquely determined independently of ϕ. The function $H \circ \phi$ is called *Birkhoff normal form* and the ϕ is called *Birkhoff transformation.*

If there exists a convergent Birkhoff transformation, the corresponding Hamiltonian system is solved explicitly for (ξ, η) coordinates. Indeed the system can be written as

$$\frac{d\xi_k}{dt} = \frac{\partial H}{\partial \omega_k} \xi_k, \qquad \frac{d\eta_k}{dt} = -\frac{\partial H}{\partial \omega_k} \eta_k \qquad (k = 1, \ldots, n),$$

where $\omega_k = \xi_k \eta_k$ $(k = 1, \ldots, n)$. Therefore we have $d\omega_k/dt = 0$, namely ω_k $(k = 1, \ldots, n)$ are integrals. Hence one can integrate the system in the form

$$\xi_k(t) = e^{tH_{\omega_k}} \xi_k(0), \qquad \eta_k(t) = e^{-tH_{\omega_k}} \eta_k(0) \qquad (k = 1, \ldots, n),$$

where the arguments of H_{ω_k} are the initial values $\omega_1(0), \ldots, \omega_n(0)$.

However the Birkhoff transformation ϕ is divergent in general (Siegel [9, 10]). Then, when does a convergent Birkhoff transformation ϕ exist? This problem was studied by H. Rüssmann [8] for two degrees of freedom case, and J. Vey [12] generalized it to general degrees of freedom case as follows:

THEOREM 1.2 (J. Vey [12]). *Let $G_k(z)$ $(k=1,\ldots,n; z=(x,y))$ be n Poisson commuting functions of the form*

$$G_k(z)=\sum_{l=1}^{n}\beta_{kl}x_ly_l+\cdots \qquad (k=1,\ldots,n); \qquad \det(\beta_{kl})\neq 0, \tag{1.4}$$

where $\beta_{kl}\in\mathbf{C}$ and the remainder part not written out explicitly denotes a convergent power series containing terms of order ≥ 3 only. Then there exists an analytic canonical transformation $z=\phi(\zeta)$ $(\zeta=(\xi,\eta))$ near the origin such that $\phi(0)=0$ and $G_k\circ\phi$ $(k=1,\ldots,n)$ are analytic functions of n variables $\xi_l\eta_l$ $(l=1,\ldots,n)$.

In the above, $G_1,\ldots,G_n$ are called *Poisson commuting* if the Poisson bracket

$$\{G_k, G_l\}:=\sum_{j=1}^{n}\left(\frac{\partial G_k}{\partial x_j}\frac{\partial G_l}{\partial y_j}-\frac{\partial G_k}{\partial y_j}\frac{\partial G_l}{\partial x_j}\right)$$

vanishes identically for any $k,l=1,\ldots,n$. The Poisson bracket is invariant under canonical transformations. Although the non-resonance condition is not assumed, it is "hidden" in this theorem. We will see this after the formulation of Theorem 1.3.

The system (1.1) is called *integrable* in a domain $\Omega\subset\mathbf{C}^{2n}$ (or $\mathbf{R}^{2n}$) if there exist n Poisson commuting integrals $G_1=H, G_2,\ldots,G_n$ which are functionally independent in Ω. Here the functional independence of $G_1,\ldots,G_n$ implies that n differentials $dG_1,\ldots,dG_n$ are linearly independent on an open dense subset of Ω. Clearly the system (1.1) is integrable near the origin in this sense if there exists a convergent Birkhoff transformation. On the other hand, Theorem 1.2 asserts that if the system with $H=G_1$ is integrable near the origin with integrals of the form (1.4), then there exists a convergent Birkhoff transformation. This theorem can be proved also for C^∞ function case (Eliasson [3]).

However it may happen that some integrals G_k begin with terms of order greater than two and their result cannot apply to that case. The aim of this paper is to show that, without any restriction such as (1.4), if the system is integrable near a non-resonant equilibrium point, then there exists an analytic (convergent) Birkhoff transformation. The result is stated as follows:

THEOREM 1.3. *Let the origin be a non-resonant equilibrium point of the system* (1.1) *and assume that the Hamiltonian* $H(z)$ $(z=(x, y))$ *is analytic in a neighbourhood of the origin. Assume that in addition to* $G_1 = H$ *the system* (1.1) *possesses* $n-1$ *analytic integrals* $G_2(z), \ldots, G_n(z)$ *near the origin such that* $G_1, G_2, \ldots, G_n$ *are functionally independent. Then there exists an analytic canonical transformation* $z=\phi(\zeta)$ $(\zeta=(\xi, \eta))$ *near the origin such that* $\phi(0)=0$ *and* $G_k \circ \phi$ $(k=1, \ldots, n)$ *are analytic functions of n variables* $\xi_l \eta_l$ $(l=1, \ldots, n)$.

In the above, the transformation ϕ is obtained as the composition of a linear canonical transformation taking the Hamiltonian into the form (1.2) and a nonlinear canonical transformation of the form (1.3) (Birkhoff transformation). We do not need to assume that $G_1, \ldots, G_n$ are Poisson commuting. However, as we shall see in the next section (Remark 2.6), the integrals near a non-resonant equilibrium point are necessarily Poisson commuting. This theorem implies that integrability near a non-resonant equilibrium point is equivalent to the existence of an analytic Birkhoff transformation.

One can prove Theorem 1.2 from Theorem 1.3. To see this, let $A = (\alpha_{kj}) \in GL(n, \mathbf{C})$ and consider

$$F_k = \sum_{j=1}^{n} \alpha_{kj} G_j \qquad (k=1, \ldots, n)$$

for the functions $G_1, \ldots, G_n$ in Theorem 1.2. These F_k have the same form as (1.4) with AB in place of $B=(\beta_{kj})$. Let $(\lambda_1, \ldots, \lambda_n)$ be the first row of AB. Then $F_1 = \sum_{k=1}^{n} \lambda_k x_k y_k$ and one can find a matrix A so that $\lambda_1, \ldots, \lambda_n$ are rationally independent. Therefore for the system with Hamiltonian F_1, the origin is a non-resonant equilibrium point and it follows from $\det AB \neq 0$ that $F_1, \ldots, F_n$ are functionally independent. Moreover since $G_1, \ldots, G_n$ are Poisson commuting, $F_1, \ldots, F_n$ are also Poisson commuting and they are integrals of the system. Consequently Theorem 1.3 is applicable and gives the assertion of Theorem 1.2. In this sense, Theorem 1.3 is a generalization of Theorem 1.2.

If the Hamiltonian is real analytic, we have the similar result. We state the result for the case when the eigenvalues of the linearized system are all purely imaginary. In this case the equilibrium point is called *elliptic*.

THEOREM 1.4. *Let the origin be a non-resonant elliptic equilibrium point of the system* (1.1) *and assume that the Hamiltonian* $H(z)$ $(z=(x, y))$ *is real analytic in a neighbourhood of the origin. Assume that in addition to* $G_1 = H$ *the system* (1.1) *possesses* $n-1$ *analytic integrals* $G_2(z), \ldots, G_n(z)$ *near the origin such that* $G_1, G_2, \ldots, G_n$ *are functionally independent. Then there exists a real analytic* *canonical transformation* $z = \phi(\zeta)$ $(\zeta = (\xi, \eta))$ *near the origin such that* $\phi(0) = 0$ *and* $G_k \circ \phi$ $(k = 1, \ldots, n)$ *are analytic functions of* n *variables* $(\xi_l^2 + \eta_l^2)/2$ $(l = 1, \ldots, n)$.

In the above, we do not need to assume that the additional integrals $G_2, \ldots, G_n$ are real analytic. However the new Hamiltonian $H \circ \phi$ is real analytic and the n functions $\xi_k^2 + \eta_k^2$ are integrals and solutions are given explicitly in (ξ, η) variables as periodic or quasiperiodic orbits on invariant tori $\xi_k^2 + \eta_k^2 =$ const. ≥ 0 $(k = 1, \ldots, n)$. As is well known, for integrable systems with real analytic (or C^∞) Poisson commuting integrals $G_1, \ldots, G_n$, Arnold–Liouville's theorem [1] asserts that if $dG_1, \ldots, dG_n$ are linearly independent on a compact and connected level set $E = \{G_k = \text{const.}\ (k = 1, \ldots, n)\}$, then there exists a canonical coordinate system (τ, θ) (which is called action-angle variables) in a neighbourhood of E such that the G_k are functions of $\tau = (\tau_1, \ldots, \tau_n)$ alone. This implies that the flow of the system becomes linear in θ variables and the system is solved explicitly in (τ, θ) variables. Theorem 1.4 also implies the existence of action-angle variables through a canonical transformation $\xi_k = \sqrt{2\tau_k} \cos \theta_k$, $\eta_k = \sqrt{2\tau_k} \sin \theta_k$. Our case is quite different from Arnold–Liouville's theorem in the sense that rank $(dG_1, \ldots, dG_n) = 0$ at the origin since $G_k \circ \phi$ does not contain linear terms.

This paper is organized as follows: In the next section we discuss the role of functional independence and reduce Theorem 1.3 to Theorem 2.4 formulated there. The Sections 3 to 6 are devoted to the proof of this theorem. Our proof is based on a rapidly convergent iteration process. In Section 3, we prove the existence of a formal canonical transformation ϕ which takes G_k $(k = 1, \ldots, n)$ into the normal form. The transformation ϕ will be given as a composition of infinite number of canonical transformations defined by the so called Lie series. In this step, by using the non-resonance condition the formal expansion of ϕ is determined by the requirement that $H \circ \phi$ is in normal form. However it turns out that automatically also $G_k \circ \phi$ $(k = 2, \ldots, n)$ will be formally in normal form. Using the system of n equations $G_k \circ \phi =$ normal form $(k = 1, \ldots, n)$, we can obtain estimates good enough for the convergence proof. This is the main point to avoid the small divisor difficulty and it is described more precisely in Section 3. Moreover we also reformulate the result as Theorem 3.6 a little more generally than Theorem 2.4, which will be useful in Section 8. The convergence proof of the formal transformation ϕ is given in Sections 4 to 6. In Sections 4 and 5, we consider one step of the iteration and prove several estimates. The final estimates are Propositions 5.3 and 5.4. We prove convergence of the iteration in Section 6. In Section 7, we consider the case when the Hamiltonian is real analytic. We obtain Theorem 1.4 from Theorem 1.3 by imposing a "reality condition" on the

Hamiltonian of the form (1.2). In Section 8, we prove analogous results for canonical mappings near a non-resonant fixed point. The Section 9 is an appendix where we present a proof of Lemma 2.1 on functional independence.

Acknowledgement. This work was done when I was staying at Forschungsinstitut für Mathematik, ETH Zürich in the academic year of 1986–87. I would like to thank the institute for its hospitality. I would like to thank Professor J. Moser for his kind hospitality and for many helpful discussions and advices during preparing this work. I also wish to thank H. Duistermaat and L. Chierchia for helpful discussions and to E. Zehnder and J. Pöschel for their useful suggestions.

2. Preliminaries

The aim of this section is to reduce Theorem 1.3 to Theorem 2.4 stated below as well as presenting preliminary facts for its proof.

Throughout this paper from this section, analytic functions defined near the origin are always assumed to be given as a convergent power series in appropriate polydisks.

We begin with a general discussion on functionally independent functions. Let $f_k(z) = f_k\ (z_1, \ldots, z_m)$ $(k = 1, \ldots, n)$ be analytic functions near the origin $z = 0$ which begin with terms of order s_k. We denote it by

$$f_k = f_k^0 + f_k^1 + \cdots, \qquad f_k^0(z) \not\equiv 0$$

where f_k^j is a homogeneous polynomial of degree $s_k + j$. We call f_k^0 the *lowest order part* of f_k. The functional independence of $f_1, \ldots, f_n$ does not necessarily imply the functional independence of their lowest order parts $f_1^0, \ldots, f_n^0$. However the following holds:

LEMMA 2.1. *Let $f_1, \ldots, f_n$ be functionally independent analytic (or real analytic) functions near the origin. Assume that $f_1^0, \ldots, f_{r-1}^0$ $(2 \leq r \leq n)$ are functionally independent and that $f_1^0, \ldots, f_r^0$ are functionally dependent. Then there exists a polynomial P of $f_1, \ldots, f_r$ with complex (resp. real) coefficients such that $f_1^0, \ldots, f_{r-1}^0, \hat{f}_r^0$ are functionally independent, where $\hat{f}_r = P(f_1, \ldots, f_r)$. Moreover $f_1, \ldots, f_{r-1}, \hat{f}_r, f_{r+1}, \ldots, f_n$ are functionally independent.*

This lemma is proved in Ziglin [13]. However for the sake of completeness we will give its proof in the appendix (Section 9). By using this lemma repeatedly we can construct functions $\hat{f}_1 = f_1, \hat{f}_2, \ldots, \hat{f}_n$ which are polynomials of $f_1, \ldots, f_n$ and whose lowest order parts $\hat{f}_1^0, \ldots, \hat{f}_n^0$ are functionally independent.

Remark 2.2. In Theorem 1.3, we are given n functionally independent integrals $G_1, \dots, G_n$, which can be assumed to vanish at the origin. Therefore, if we apply the above argument to $(f_1, \dots, f_n) = (G_1, \dots, G_n)$, we obtain n integrals $\hat{G}_1 = G_1, \hat{G}_2, \dots, \hat{G}_n$ whose lowest order parts are functionally independent.

Next we present some facts concerning operations by the Poisson bracket. Let $f = f(z)$ $(z = (x, y))$ be an analytic function near the origin $z = 0 \in \mathbf{C}^{2n}$. Here we consider $\mathbf{C}^{2n}$ as a symplectic manifold with the standard symplectic structure $\sum_{k=1}^n dx_k \wedge dy_k$ and (x, y) as its canonical coordinates. We say that the function f is in *normal form* if it is a function of n variables $x_1y_1, \dots, x_ny_n$. We introduce the following operator on the space of analytic functions near the origin:

$$P_N f(x, y) = \int_0^1 \cdots \int_0^1 f(e^{2\pi i\theta}x, e^{-2\pi i\theta}y)\, d\theta_1 \cdots d\theta_n, \tag{2.1}$$

where

$$e^{2\pi i\theta}x = (e^{2\pi i\theta_1}x_1, \dots, e^{2\pi i\theta_n}x_n), \qquad e^{-2\pi i\theta}y = (e^{-2\pi i\theta_1}y_1, \dots, e^{-2\pi i\theta_n}y_n).$$

We note that

$$P_N(x^\alpha y^\beta) = \begin{cases} x^\alpha y^\beta & \text{if } \alpha = \beta; \\ 0 & \text{if } \alpha \neq \beta, \end{cases}$$

where we used multi-index notation

$$x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}, \qquad y^\beta = y_1^{\beta_1} \cdots y_n^{\beta_n},$$
$$\alpha = (\alpha_1, \dots, \alpha_n), \qquad \beta = (\beta_1, \dots, \beta_n).$$

Therefore $P_N f$ is a power series consisting of all terms of the form $x^\alpha y^\alpha$ in the power series expansion of f. The following facts will be used later.

LEMMA 2.3. *Let $f(z)$ and $g(z)$ $(z = (x, y))$ be analytic functions near the origin $z = 0$ and assume that $\lambda_1, \dots, \lambda_n$ are rationally independent numbers. Then*
(i) *If $\{f, \sum_{k=1}^n \lambda_k x_k y_k\} = 0$, then f is in normal form.*
(ii) *If f and g are in normal form, then $\{f, g\} = 0$.*
(iii) *If g is in normal form, then $P_N\{f, g\} = 0$.*

Proof. We set

$$f = \sum_{|\alpha|+|\beta| \geq 0} c_{\alpha\beta} x^\alpha y^\beta \qquad (c_{\alpha\beta} \in \mathbf{C}).$$

Suppose that g is in normal form. Then we have

$$\{f, g\} = \sum_{k=1}^{n} g_{\omega_k}(f_{x_k}x_k - f_{y_k}y_k) = \sum_{|\alpha|+|\beta|\geq 1} \gamma_{\alpha\beta}(\omega)c_{\alpha\beta}x^{\alpha}y^{\beta}, \tag{2.2}$$

where

$$\gamma_{\alpha\beta}(\omega) = \sum_{k=1}^{n} g_{\omega_k}(\alpha_k - \beta_k); \qquad \omega_k = x_k y_k \qquad (k = 1, \ldots, n).$$

If $g = \sum_{k=1}^{n} \lambda_k x_k y_k$, then we have $\gamma_{\alpha\beta}(\omega) = \sum_{k=1}^{n} \lambda_k(\alpha_k - \beta_k)$, which vanishes if and only if $\alpha = \beta$ by the non-resonance condition. Hence the identity (2.2) implies the assertion (i). We can also prove (ii) and (iii) easily by using (2.2). □

Now let us consider the analytic functions $\hat{G}_k(z)$ $(k = 1, \ldots, n)$ described in Remark 2.2 and rewrite them by $G_k(z)$. Then G_k $(k = 1, \ldots, n)$ are analytic integrals of the Hamiltonian system (1.1) with $H = G_1$. We denote the G_k by

$$G_k = G_k^0 + G_k^1 + \cdots, \qquad G_k^0(z) \not\equiv 0, \tag{2.3}$$

where G_k^j is a homogeneous polynomial of degree $s_k + j$ in $z = (x, y)$. By the discussions of Section 1, under the non-resonance condition we can assume that the Hamiltonian has the form (1.2), i.e.,

$$G_1^0 = \sum_{k=1}^{n} \lambda_k x_k y_k. \tag{2.4}$$

Since G_k is an integral of (1.1), we have the identity $\{G_k, G_1\} = 0$, and the comparison of its lowest order terms gives

$$\{G_k^0, G_1^0\} = 0.$$

Then by Lemma 2.3(i) this identity implies that G_k^0 are polynomials of n variables $x_1y_1, \ldots, x_ny_n$. The functional independence of the lowest order parts $G_1^0, \ldots, G_n^0$ is equivalent to the condition

$$\det\left(\frac{\partial(G_1^0, \ldots, G_n^0)}{\partial(\omega_1, \ldots, \omega_n)}\right) \not\equiv 0 \qquad (\omega_k = x_k y_k; k = 1, \ldots, n). \tag{2.5}$$

We will prove the following theorem instead of Theorem 1.3.

THEOREM 2.4. *Let $G_k = G_k(z)$ $(k = 1, \ldots, n; z = (x, y))$ be analytic functions near the origin which satisfy the following conditions:*

(i) $\{G_k, G_1\} = 0$ $(k = 2, \ldots, n)$; (2.6)

(ii) *The lowest order part G_1^0 has the form* (2.4) *with rationally independent numbers $\lambda_1, \ldots, \lambda_n$;*

(iii) *The lowest order parts $G_1^0, \ldots, G_n^0$ are polynomials of $x_1y_1, \ldots, x_ny_n$ alone and satisfy the condition* (2.5).

Then there exists an analytic canonical transformation $z = \phi(\zeta)$ $(\zeta = (\xi, \eta))$ of the form (1.3) *near the origin such that $G_k \circ \phi$ $(k = 1, \ldots, n)$ are analytic functions of n variables $\xi_l\eta_l$ $(l = 1, \ldots, n)$.*

In the above, the transformation ϕ is not unique. However we shall see in the next section (Remark 3.4) that the normal form $G_k \circ \phi$ $(k = 1, \ldots, n)$ are uniquely determined independently of ϕ.

We say that the function G_k given in (2.3) is in *normal form up to terms of order $s_k + d$* if the polynomial $G_k^0 + G_k^1 + \cdots + G_k^d$ is in normal form. The following fact will play a basic role in the proof of Theorem 2.4.

PROPOSITION 2.5. *Let G_k be analytic functions near the origin satisfying the conditions* (i) *and* (ii) *of Theorem* 2.4. *If G_1 is in normal form up to terms of order $s_1 + d$, then G_k $(k = 2, \ldots, n)$ are in normal form up to terms of order $s_k + d$.*

Proof. The proof is easily done by induction. We already proved the case $d = 0$. Assume that G_1 is in normal form up to terms of order $s_1 + j$ $(0 \le j \le d)$ and that G_k $(k = 2, \ldots, n)$ are in normal form up to terms of order $s_k + j - 1$. The comparison of the homogeneous part of degree $s_k + j$ in (2.6) gives

$$\{G_k^j, G_1^0\} + \sum_{l=1}^{j} \{G_k^{j-l}, G_1^l\} = 0. \tag{2.7}$$

Since G_k^{j-l} and G_1^l $(l = 1, \ldots, j)$ are in normal form by the assumption, we have $\{G_k^{j-l}, G_1^l\} = 0$ by Lemma 2.3(ii). Therefore the identity (2.7) leads to

$$\{G_k^j, G_1^0\} = 0.$$

By Lemma 2.3(i) this identity implies that G_k^j is normal form. This completes the proof. □

Under the assumptions of Theorem 1.3 with the Hamiltonian of the form (1.2), the $\hat{G}_k$ $(k = 1, \ldots, n)$ in Remark 2.2 satisfy the assumptions of Theorem

2.4. Therefore one can conclude that there exists an analytic canonical transformation $z = \phi(\zeta)$ such that $G_1 \circ \phi$ is in normal form. Then by Proposition 2.5, the other integrals $G_k \circ \phi$ $(k = 2, \ldots, n)$ are automatically in normal form. Therefore Theorem 2.4 implies Theorem 1.3.

Proposition 2.5 implies also the following fact which was mentioned in Section 1.

Remark 2.6. The functions G_k $(k = 1, \ldots, n)$ which satisfy the conditions (i) and (ii) of Theorem 2.4 are Poisson commuting.

Indeed, under the non-resonance condition there exists an analytic canonical transformation which takes the Hamiltonian $G_1 = H$ into normal form up to terms of sufficiently high order. Then Proposition 2.5 implies that the other integrals G_k $(k = 2, \ldots, n)$ are also in normal form up to terms of sufficiently high order. Therefore we have

$$\{G_k, G_l\} = 0 \quad \text{for} \quad k, l = 1, \ldots, n.$$

up to terms of any order by Lemma 2.3(ii), and hence these are identities. □

In our proof of Theorem 2.4, the transformation ϕ is obtained by an iteration process and we consider transformations which are closer to the identity as iteration step proceeds. We complete this section by presenting a fact concerning such transformations.

Let $\mathfrak{S}_d$ be a set of all canonical formal power series transformations $z = \phi(\zeta)$ of the form

$$z = \zeta + O(\zeta^{d+1}). \tag{2.8}$$

Here and in what follows, the notation $O(\zeta^{d+1})$ denotes a (vector of) formal power series in $\zeta = (\xi, \eta)$ consisting of terms of order $\geq d + 1$ only. This set $\mathfrak{S}_d$ forms a group under compositions of transformations. We note the following fact.

LEMMA 2.7. *If $\phi \in \mathfrak{S}_d$, then it is written in the form*

$$z = \zeta + J\nabla W(\zeta) + O(\zeta^{2d+1}); \qquad J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \tag{2.9}$$

with a polynomial $W(\zeta)$ of the form

$$W(\zeta) = W^{d+2} + W^{d+3} + \cdots + W^{2d+1}, \tag{2.10}$$

where I is the $n \times n$ identity matrix and W^j is a homogeneous polynomial of degree j in ζ.

Proof. We note that the canonical character of ϕ means

$$\sum_{k=1}^{n} dx_k \wedge dy_k = \sum_{k=1}^{n} d\xi_k \wedge d\eta_k,$$

which leads to the fact that the 1-form

$$\sum_{k=1}^{n} (x_k - \xi_k)\, d\eta_k + \sum_{k=1}^{n} (\eta_k - y_k)\, dx_k$$

is closed. Therefore there exists a formal power series $W(\xi, \eta)$ such that this 1-form is equal to dW. Since the first n components of equation (2.8) can be solved for ξ as a formal power series of (x, η), we can express the W in (x, η) variables, and $x_k - \xi_k$ and $\eta_k - y_k$ are power series of x and η consisting of terms of order $\geq d+1$. Consequently we have the expression

$$x_k = \xi_k + W_{\eta_k}(x, \eta), \qquad y_k = \eta_k - W_{x_k}(x, \eta), \qquad (k = 1, \ldots, n) \tag{2.11}$$

with a formal power series $W = W(x, \eta)$ which begins with terms of order $\geq d+2$ in x, η. (This is exactly the so-called generating function.) Since $W(\xi, \eta) = O(\zeta^{d+2})$, it follows from (2.11) that this transformation has the form (2.9) with (2.10). □

3. Construction of the iteration of transformations

Our proof of Theorem 2.4 is based on a rapidly convergent iteration process [5]. The transformation ϕ is obtained as a composition of infinite number of transformations φ_ν $(\nu = 0, 1, \ldots)$, where φ_ν provides a better approximation to the normal form. In this section, we construct the iteration of transformations formally.

To describe one iteration step, assume that G_k $(k = 1, \ldots, n)$ are in normal form up to terms of order $s_k + d - 1$. We note that this assumption is already satisfied for $d = 1$. Our purpose is to find a canonical transformation which normalizes G_k up to terms of order $s_k + 2d - 1$. To this end, we consider a canonical transformation φ defined by the time 1 map (flow) of the Hamiltonian system

$$\frac{dz}{dt} = J\nabla W(z) \tag{3.1}$$

with a polynomial Hamiltonian

$$W = W^{d+2} + W^{d+3} + \cdots + W^{2d+1}, \tag{3.2}$$

where W^j is a homogeneous polynomial of degree j in $z = (x, y)$ $(= (z_1, \ldots, z_{2n}))$ and ∇W is the gradient of W, i.e., $2n$-dimensional vector with components W_{z_k} $(k = 1, \ldots, 2n)$. Let $z = \varphi^t(\zeta)$ be the solution of (3.1) satisfying an initial condition $z = \zeta$ at $t = 0$. In this and the next section, we consider the solution $\varphi^t(\zeta)$ formally. Then the transformation φ is written as

$$\varphi : \zeta \mapsto \varphi^1(\zeta) = \zeta + \int_0^1 J\nabla W(\varphi^t(\zeta))\, dt. \tag{3.3}$$

Throughout this paper, we say that a canonical transformation φ is *generated by the Hamiltonian flow with Hamiltonian* W if it is defined by the time 1 map (3.3) for the Hamiltonian system (3.1). Since the origin $\zeta = 0$ is an equilibrium solution of (3.1), we have the identity $\varphi^t(0) = 0$. Therefore the series expansion of $\varphi^t(\zeta)$ in the powers of ζ begins with linear terms and consequently by substituting it into (3.3), we can see that the solution $\varphi^t(\zeta)$ has the form (2.8), where the remainder part $O(\zeta^{d+1})$ depends on t. Then it follows from (3.3) that $\varphi \in \mathfrak{S}_d$ and has the form (2.9). Conversely, by Lemma 2.7 we have

Remark 3.1. Let $\varphi \in \mathfrak{S}_d$ and W be the polynomial of the form (3.2) determined by the expression (2.9) of the φ. Then up to terms of order $2d$, φ is equal to the canonical transformation generated by the Hamiltonian flow with Hamiltonian W.

Under the above assumption on G_k $(k = 1, \ldots, n)$, we can write the G_k in the form

$$\begin{aligned} &G_k(z) = g_k(z) + \hat{G}_k(z); \qquad g_k = P_N g_k, \\ &\hat{G}_k = O(z^{s_k+d}) \qquad (k = 1, \ldots, n), \end{aligned} \tag{3.4}$$

where $g_k(z)$ is a polynomial of degree less than $s_k + d$ in $z = (x, y)$ which is actually a polynomial of n products $x_1y_1, \dots, x_ny_n$. We denote g_k and $\hat{G}_k$ by

$$g_k = g_k^0 + \cdots + g_k^{d-1}, \qquad \hat{G}_k = \hat{G}_k^d + \hat{G}_k^{d+1} + \cdots,$$

where g_k^j and $\hat{G}_k^j$ are homogeneous polynomials of degree $s_k + j$ in $z = (x, y)$.

We first prove the following

LEMMA 3.2. *Let G_k $(k = 1, \dots, n)$ be power series of the form* (3.4), *i.e., in normal form up to terms of order $s_k + d - 1$. Let $z = \varphi(\zeta) \in \mathfrak{S}_d$ $(\zeta = (\xi, \eta))$ be a canonical transformation generated by the Hamiltonian flow with Hamiltonian W of the form* (3.2). *Then $G_k(\varphi(\zeta))$ $(k = 1, \dots, n)$ are in normal form up to terms of order $s_k + 2d - 1$ if and only if the polynomial W satisfies a system of n equations*

$$\{g_k(\zeta), W(\zeta)\} + \hat{G}_k(\zeta) = P_N^{2d-1}\hat{G}_k(\zeta) + O(\zeta^{s_k+2d}), \qquad (k = 1, \dots, n), \tag{3.5}$$

where $P_N^{2d-1}\hat{G}_k = P_N(\hat{G}_k^d + \cdots + \hat{G}_k^{2d-1})$. The function $G_k(\varphi(\zeta))$ is written as

$$\begin{aligned} &G_k(\varphi(\zeta)) = g_k'(\zeta) + \hat{G}_k'(\zeta); \qquad g_k' = g_k + P_N^{2d-1}\hat{G}_k, \\ &\hat{G}_k' = O(\zeta^{s_k+2d}), \qquad (k = 1, \dots, n). \end{aligned} \tag{3.6}$$

Proof. Since the transformation $\varphi \in \mathfrak{S}_d$ has the form (2.9), one obtains

$$G_k(\varphi(\zeta)) = g_k(\zeta) + \{g_k(\zeta), W(\zeta)\} + \hat{G}_k(\zeta) + O(\zeta^{s_k+2d}), \qquad (k = 1, \dots, n) \tag{3.7}$$

by substituting the form (2.9) into (3.4). Since $P_N\{g_k, W\} = 0$ by Lemma 2.3(iii), equation (3.7) implies that $G_k(\varphi(\zeta))$ $(k = 1, \dots, n)$ are in normal form up to terms of order $s_k + 2d - 1$ if and only if W satisfies the system (3.5), and the expression (3.6) follows easily. This proves the assertion. □

The comparison of the homogeneous parts of degree $s_k + j$ in the equations (3.5) gives

$$\sum_{i=1}^{n} \frac{\partial g_k^0}{\partial \omega_i} D_i W^{j+2} = F_k^j(\zeta); \qquad F_k^j(\zeta) = -(I - P_N)\hat{G}_k^j - \sum_{\nu=1}^{j-d} \{g_k^\nu, W^{j+2-\nu}\}$$

$$(j = d, \dots, 2d-1;\ k = 1, \dots, n) \tag{3.8}$$

where I is the identity operator and D_i is an operator defined by

$$D_i = \{\omega_i, \cdot\}, \qquad \omega_i = \xi_i \eta_i \qquad (i = 1, \ldots, n).$$

We now prove that under the conditions of (i) and (ii) of Theorem 2.4 the system (3.5) can be solved through the homogeneous equations (3.8).

PROPOSITION 3.3. *Let G_k $(k = 1, \ldots, n)$ be power series satisfying conditions* (i) *and* (ii) *of Theorem* 2.4. *Assume that G_k are in the form* (3.4), *i.e., normal form up to terms of order $s_k + d - 1$. Then there exists a unique canonical transformation $z = \varphi(\zeta) \in \mathfrak{S}_d$ such that* (i) *$G_k(\varphi(\zeta))$ $(k = 1, \ldots, n)$ are in the form* (3.6), *i.e., normal form up to terms of order $s_k + 2d - 1$, and* (ii) *φ is generated by the Hamiltonian flow with Hamiltonian W of the form* (3.2) *satisfying a condition*

$$P_N W = 0. \tag{3.9}$$

Proof. By Lemma 3.2, our purpose is to prove that there exists a unique polynomial W of the form (3.2) which satisfies the system of n equations (3.5) together with the condition (3.9). The system (3.5) for a function W is clearly overdetermined. However it suffices to solve the first equation ($k = 1$), and then the other equations are automatically solved as we shall see. The relations $\{G_k, G_1\} = 0$ can be viewed as the compatibility condition.

To see this, we set $W^{j+2} = \sum_{|\alpha|+|\beta|=j+2} c_{\alpha\beta}\xi^\alpha\eta^\beta$. Then equation (3.8) for $k = 1$ can be written as

$$\sum_{|\alpha|+|\beta|=j+2} \gamma c_{\alpha\beta}\xi^\alpha\eta^\beta = F_1^j(\zeta); \qquad \gamma = \sum_{k=1}^{n} \lambda_k(\beta_k - \alpha_k) \tag{3.10}$$

The right-hand side of this equation is determined by G_1 and $W^{d+2}, \ldots, W^{j+1}$ and therefore this equation gives a recursion formula to determine the function W of the form (3.2). Since γ does not vanish for $\alpha \neq \beta$ by the non-resonance condition, the coefficients $c_{\alpha\beta}$ for $\alpha \neq \beta$ are determined by this equation. On the other hand the condition (3.9) implies that $c_{\alpha\beta} = 0$ for $\alpha = \beta$. Here we note that $F_1^j(\zeta)$ does not contain any term in normal form. In this way we can determine W^{j+2} $(j = d, \ldots, 2d - 1)$ successively and hence a solution W of equation (3.5) for $k = 1$. The canonical transformation (3.3) determined by this W normalizes G_1 up to terms of order $s_1 + 2d - 1$. Therefore Proposition 2.5 implies that $G_k(\varphi(\zeta))$ is in normal form up to terms of order $s_k + 2d - 1$ for any $k = 1, \ldots, n$. Hence the W satisfies $n - 1$ other equations in (3.5) automatically, which together with (3.7), implies (3.6). This proves the assertion. □

Since the assumption of Proposition 3.3 is satisfied for $d=1$, we can take $d=2^{\nu}$ ($\nu=0, 1, \ldots$) successively. Let φ_ν be the canonical transformation described in Proposition 3.3 with $d=2^{\nu}$. Then we obtain a formal canonical transformation ϕ which takes G_k ($k=1, \ldots, n$) into the normal form as follows:

$$\phi = \lim_{\nu\to\infty} \phi_\nu; \qquad \phi_\nu = \varphi_0 \circ \varphi_1 \circ \cdots \circ \varphi_\nu, \qquad \varphi_\nu \in \mathfrak{S}_d \qquad (d=2^\nu). \tag{3.11}$$

Remark 3.4. In the above proof of Proposition 3.3, we did not use the existence of the integrals $G_2, \ldots, G_n$ to show that W is uniquely determined. Therefore the above argument proves Theorem 1.1 also. Moreover, by the same argument as in the proof we can prove that the normal form $G_k \circ \phi$ ($k=1, \ldots, n$) are uniquely determined independently of ϕ as follows:

Suppose that G_k ($k=1, \ldots, n$) are taken into normal form by formal canonical transformations ϕ and ψ. Then we have

$$G_k \circ \phi \circ (\phi^{-1} \circ \psi) = G_k \circ \psi,$$

namely the formal canonical transformation $\phi^{-1} \circ \psi \in \mathfrak{S}_1$ takes the normal form $G_k \circ \phi$ into the other normal form $G_k \circ \psi$. From Remark 3.1, this transformation $\phi^{-1} \circ \psi$ can be expressed as a composition of infinite number of transformations $\varphi_\nu \in \mathfrak{S}_d$ ($d=2^\nu$; $\nu=0, 1, \ldots$) which are generated by the Hamiltonian flow with Hamiltonian W of the form (3.2). The transformation φ_ν takes G_k which are in normal form up to terms of order s_k+d-1 into the normal form up to terms of order s_k+2d-1. If G_k are already in normal form, then $\hat{G}_k(\zeta) = P_N \hat{G}_k(\zeta)$ in the system (3.5) which the polynomial W satisfies. One can see from (3.10) with the definition of F_1^j in (3.8) that $W^{j+2} = P_N W^{j+2}$ and $F_1^j \equiv 0$ inductively for $j = d, \ldots, 2d-1$. Therefore the polynomial W is in normal form. Then the Hamiltonian system (3.1) is solved explicitly and its time 1 map is expressed in the form

$$x_k = \xi_k \exp(W_{\omega_k}), \qquad y_k = \eta_k \exp(-W_{\omega_k}) \qquad (k=1, \ldots, n).$$

Since the n products $\xi_k \eta_k$ are invariant under this transformation, the normal form $G_k \circ \phi$ cannot be changed by φ_ν. This implies that $G_k \circ \phi = G_k \circ \psi$. □

In the proof of Proposition 3.3, the small divisor γ appears in (3.10). However we can avoid the small divisor difficulty in the following way:

LEMMA 3.5. *Let G_k ($k=1, \ldots, n$) be power series in the form* (3.4) *and W be a polynomial of the form* (3.2) *which satisfy the system* (3.5). *Assume the* *condition* (iii) *of Theorem* 2.4 ($G_k^0 = g_k^0$). *Then* $D_k W^{j+2}$ *is expressed as follows:*

$$D_k W^{j+2} = \frac{q_k^j(\zeta)}{p(\zeta)} \qquad (k = 1, \ldots, n; j = d, \ldots, 2d-1), \tag{3.12}$$

where

$$q_k^j(\zeta) = \det \begin{array}{c} \begin{array}{ccccc} \phantom{g^0_{1_{\omega_1}}} & & (k) & & \phantom{g^0_{1_{\omega_n}}} \end{array} \\ \begin{pmatrix} g^0_{1_{\omega_1}} & \cdots & F_1^j & \cdots & g^0_{1_{\omega_n}} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ g^0_{n_{\omega_1}} & \cdots & F_n^j & \cdots & g^0_{n_{\omega_n}} \end{pmatrix} \end{array}, \qquad p(\zeta) = \det\left(\frac{\partial g_k^0}{\partial \omega_l}\right). \tag{3.13}$$

Here the numerator $q_k^j(\zeta)$ *is divisible by* $p(\zeta)$.

Proof. We consider equation (3.8) as a system of n linear equations for $D_i W^{j+2}$ ($i = 1, \ldots, n$). Then at points $\zeta = (\xi, \eta)$ on which $\det(\partial g_k^0 / \partial \omega_l) \neq 0$, one can solve (3.8) uniquely as (3.12) with (3.13). Under the condition (2.5), the inequality $p(\zeta) \neq 0$ is satisfied on an open dense subset Ω' of a neighbourhood Ω of the origin. Since $D_k W^{j+2}$ for the solution W of (3.5) satisfies (3.8), it satisfies (3.12) on Ω'. Then the $D_k W^{j+2}$ satisfies (3.12) on the whole neighbourhood Ω by the continuity. This completes the proof. □

The expression (3.12) will be used to estimate $D_k W^{j+2}$ in the following section without any restriction of the small divisors.

The conditions (i) and (ii) of Theorem 2.4 was used to show that there exists a formal canonical transformation ϕ defined by the limit (3.11) which takes G_k ($k = 1, \ldots, n$) into the normal form. The aim of the following sections is to prove the convergence of this limit and we will not use the conditions (i) and (ii) for this purpose. For the later application, it is useful to formulate our result without assuming these conditions as follows:

THEOREM 3.6. *Let* $G_k(z)$ ($k = 1, \ldots, n$) *be analytic functions near the origin satisfying the condition* (iii) *of Theorem* 2.4. *Assume that there exists a formal canonical transformation* $z = \phi(\zeta)$ *which is defined by* (3.11) *and satisfies the following conditions:*

(i) *For any* $\nu = 0, 1, \ldots$, $G_k^{(\nu+1)} = G_k^{(\nu)} \circ \varphi_\nu$ ($k = 1, \ldots, n$) *are in normal form up to terms of order* $s_k + 2^{\nu+1} - 1$, *where* $G_k^{(0)} = G_k$ *and* s_k *is the degree of the lowest order part of* G_k;

(ii) *The canonical transformation* φ_ν *is generated by the Hamiltonian flow with Hamiltonian* W *of the form* (3.2) *satisfying a condition* $P_N W = 0$.

Then the limit (3.11) *is convergent uniformly in a neighbourhood of the origin* $\zeta = 0$, *and hence it defines an analytic canonical transformation* ϕ *such that* $G_k \circ \phi$ $(k = 1, \ldots, n)$ *are in normal form.*

By Lemma 3.2, the above condition (i) implies that the polynomial W in the condition (ii) satisfies the system (3.5) with $d = 2^\nu$. Therefore from Lemma 3.5 we obtain the expression (3.12) with (3.13). We will prove this theorem in Sections 4 to 6. It will also give the proof of Theorem 2.4 since the assumptions of Theorem 3.6 are satisfied under those of Theorem 2.4.

4. Estimates of the norm of W and its derivatives

In this and the next sections, we consider one iteration step defined by φ_ν $(\nu = 0, 1, \ldots)$ in Theorem 3.6 and denote φ and G_k in place of φ_ν and $G_k^{(\nu)}$ respectively. In this section, we consider the polynomial W which defines φ by (3.3) and give the estimates of W and its derivatives with respect to some norm. To specify the polydisks where the functions are to be considered, we first prove the following.

LEMMA 4.1. *Let s be the degree of the polynomial $p(\zeta)$ given in* (3.13). *Then there exist positive constants δ_k $(k = 1, \ldots, n)$ such that $0 < \delta_k < 1$ and*

$$|p(\zeta)| \geq c_1 r^s \quad \textit{on} \quad \Delta_r = \{(\xi, \eta) \in \mathbf{C}^{2n} \mid |\xi_k| = |\eta_k| = \delta_k r \qquad (k = 1, \ldots, n)\}, \tag{4.1}$$

where $c_1 > 0$ is a constant which is independent of r.

Proof. We set

$$\xi_k = t^{\rho_k} u_k, \qquad \eta_k = t^{\rho_k} v_k, \qquad (k = 1, \ldots, n)$$

and denote this by $(\xi, \eta) = (t^\rho u, t^\rho v)$. Here $t > 0$ and $\rho_k > 0$ $(k = 1, \ldots, n)$ are real constants and we assume that $\rho_1, \ldots, \rho_n$ are rationally independent. The homogeneous polynomial $p(\zeta)$ in $\xi_1\eta_1, \ldots, \xi_n\eta_n$ can be written as $p(\zeta) = \sum_{|\alpha| = s/2} c_\alpha \xi^\alpha \eta^\alpha$. Then we have

$$p(t^\rho u, t^\rho v) = \sum_{|\alpha| = s/2} c_\alpha t^{2\langle \alpha, \rho \rangle} u^\alpha v^\alpha; \qquad \langle \alpha, \rho \rangle = \sum_{k=1}^{n} \alpha_k \rho_k.$$

Let

$$d = \min_{\alpha \in C} 2\langle \alpha, \rho \rangle, \qquad C = \left\{ \alpha \mid |\alpha| = \frac{s}{2}, c_\alpha \neq 0 \right\}.$$

Here we note that d is attained by only one α because of the non-resonantness of $\rho_1, \ldots, \rho_n$. Let α_* be the α which attains the minimum d. Then we can write $p(\zeta)$ as

$$p(\zeta) = c_{\alpha_*} t^d u^{\alpha_*} v^{\alpha_*} + t^d \hat{p}(uv, t), \qquad |\hat{p}(uv, t)| \to 0 \qquad (t \to 0).$$

Here $\hat{p}(uv, t)$ is a homogeneous polynomial of degree $s/2$ in $u_1 v_1, \ldots, u_n v_n$ whose coefficients tend to 0 as $t \to 0$. Therefore there exists a sufficiently small positive number t_0 which is independent of r such that

$$|\hat{p}(uv, t_0)| \leq \tfrac{1}{2} |c_{\alpha_*}| r^s \quad \text{on} \quad |u_k| = |v_k| = r \qquad (k = 1, \ldots, n).$$

Hence we have

$$|p(\zeta)| \geq \tfrac{1}{2} |c_{\alpha_*}| t_0^d r^s \quad \text{on} \quad \{(\xi, \eta) = (t_0^\rho u, t_0^\rho v) \mid |u_k| = |v_k| = r \ (k = 1, \ldots, n)\}.$$

If we define $\delta_k = t_0^{\rho_k}$ $(k = 1, \ldots, n)$ and $c_1 = \frac{1}{2} |c_{\alpha_*}| t_0^d$, we obtain (4.1). $\square$

We introduce the following norms. Let Ω_r be a neighbourhood of the origin (polydisk) defined by

$$\Omega_r = \{\zeta = (\xi, \eta) \in \mathbf{C}^{2n} \mid |\xi_k| < \delta_k r, |\eta_k| < \delta_k r \ (k = 1, \ldots, n)\},$$

where $\delta_1, \ldots, \delta_n$ are the positive constants given in Lemma 4.1. Let $A(\Omega_r)$ be a space of functions which are analytic in $\Omega_{r+\epsilon}$ with some $\epsilon > 0$. Then $f \in A(\Omega_r)$ can be written in the power series

$$f = f_0 + f_1 + f_2 + \cdots \tag{4.2}$$

which is absolutely convergent in $\bar{\Omega}_r$. Here f_j is a homogeneous polynomial of degree j in ζ. For $f \in A(\Omega_r)$, we define the following norms:

$$|f|_r := \max_{\zeta \in \bar{\Omega}_r} |f(\zeta)|, \qquad \|f\|_r := \sum_{j=0}^{\infty} |f_j|_r.$$

Here we note that $|f|_r \leq \|f\|_r$ and that $|f_j|_r = \|f_j\|_r$ for homogeneous polynomials

f_j. Moreover we introduce a space

$$A_m(\Omega_r)=\{f\in A(\Omega_r)\mid f_j(\zeta)\equiv 0 \quad \text{for} \quad j=0,1,\ldots,m-1\}.$$

and define for $f\in A_m(\Omega_r)$ a norm

$$\|f\|_{r,m}:=\frac{\|f\|_r}{r^m}.$$

Now assume that $G_k\in A(\Omega_r)$ $(k=1,\ldots,n)$ are written in the form (3.4) (i.e., normal form up to terms of order s_k+d-1 with $d=2^\nu$) and consider the polynomial W in the condition (ii) of Theorem 3.6. Then D_kW^{j+2} $(j=d,\ldots,2d-1)$ are given by (3.12) with (3.13). We note that for $D_kW^{j+2}\in A(\Omega_r)$ the maximum $\|D_kW^{j+2}\|_r$ is attained on Δ_r. (We can see this property by using, for example, the maximum principle for one complex variable repeatedly.) Therefore using the relation (3.12) one can estimate $\|D_kW^{j+2}\|_r$ in the form

$$\|D_kW^{j+2}\|_r\leq\frac{\|q_k^j(\zeta)\|_r}{\min\limits_{\zeta\in\Delta_r}|p(\zeta)|}. \tag{4.3}$$

By this formula, we can prove

LEMMA 4.2. *The solution of equation* (3.8) *satisfies*

$$\|D_kW^{j+2}\|_r\leq c_2\sum_{i=1}^n\|F_i^j\|_{r,s_i-2} \qquad (j=d,\ldots,2d-1), \tag{4.4}$$

where c_2 is a positive constant which is independent of r.

Proof. Expanding the $q_k^j(\zeta)$ in (3.13) according to the k-th column, one obtains

$$q_k^j(\zeta)=\sum_{i=1}^n(-1)^{i+k}F_i^j(\zeta)\det Q_{ik}(\zeta), \tag{4.5}$$

where $Q_{ik}(\zeta)$ denotes the matrix obtained from that of $q_k^j(\zeta)$ by deleting the i-th row and k-th column. The $\det Q_{ik}(\zeta)$ and $F_i^j(\zeta)$ are homogeneous polynomials in ζ of degree $s-s_i+2$ (or $\det Q_{ik}(\zeta)\equiv 0$) and degree s_i+j respectively. Therefore

$q_k^j(\zeta)$ is a homogeneous polynomial of degree $s+j+2$. It follows from (4.5) that

$$\|q_k^j(\zeta)\|_r \leq \sum_{i=1}^n \|F_i^j\|_r \|\det Q_{ik}\|_r = r^s \sum_{i=1}^n \|F_i^j\|_{r,s_i-2} \|\det Q_{ik}\|_{r,s-s_i+2}.$$

Since $\det Q_{ik}$ is a homogeneous polynomial of degree $s - s_i + 2$, we have

$$\|\det Q_{ik}\|_{r,s-s_i+2} \leq c_{ik},$$

where c_{ik} is a positive constant determined by the coefficients of the polynomial $\det Q_{ik}$. Setting

$$c_3 = \max_{i,k} c_{ik},$$

we have

$$\|q_k^j(\zeta)\|_r \leq c_3 r^s \sum_{i=1}^n \|F_i^j\|_{r,s_i-2}. \tag{4.6}$$

Then, setting $c_2 = c_3/c_1$ we obtain the estimate (4.4) by the formula (4.3) with (4.1) and (4.6). □

To give the estimate of $\|D_k W\|_r$, we introduce the following notation

$$|||\hat{g}_\omega|||_r := \sum_{i,j=1}^n \|\hat{g}_{i_{\omega_j}}\|_{r,s_i-2}; \qquad \hat{g}_i = g_i - g_i^0 \; (= g_i^2 + g_i^4 + \cdots + g_i^{d-2}). \tag{4.7}$$

Here we note that $\hat{g}_{i_{\omega_j}}$ begins with terms of order $\geq s_i$ and therefore $\|\hat{g}_\omega\|_r$ will be small if we take r sufficiently small. Moreover let us introduce the notation

$$|||\hat{G}|||_r := \sum_{k=1}^n \|\hat{G}_k\|_{r,s_k-2} = \sum_{k=1}^n r^{-s_k+2} \|\hat{G}_k\|_r. \tag{4.8}$$

Then we have

PROPOSITION 4.3. *Let $G_k \in A(\Omega_r)$ be of the form* (3.4) *and assume that*

$$c_2 \, |||\hat{g}_\omega|||_r < \tfrac{1}{2}. \tag{4.9}$$

Then the polynomial $W(\zeta)$ *in the condition* (ii) *of Theorem* 3.6 *satisfies*

$$\|D_kW\|_r \le 4c_2 \,|\!|\!|\hat{G}|\!|\!|_r \qquad (k=1,\ldots,n). \tag{4.10}$$

Proof. This is a consequence of the following estimate:

$$\|D_kW\|_r \le c_2\left\{\sum_{l=0}^{d/2-1}(c_2\,|\!|\!|\hat{g}_\omega|\!|\!|_r)^l\right\}\sum_{i=1}^{n}\sum_{j=d}^{2d-1}\|P_R\hat{G}_i^j\|_{r,s_i-2}, \qquad P_R = I - P_N. \tag{4.11}$$

From (2.1) it follows that $|P_N\hat{G}_i^j|_r \le |\hat{G}_i^j|_r$, and therefore

$$\sum_{j=d}^{2d-1}\|P_R\hat{G}_i^j\|_{r,s_i-2} \le 2\sum_{j=d}^{2d-1}\|\hat{G}_i^j\|_{r,s_i-2} \le 2\,\|\hat{G}_i\|_{r,s_i-2}.$$

Hence if the assumption (4.9) is satisfied, the estimate (4.11) implies (4.10).

The proof of (4.11) is based on an inductive argument. First we note that

$$|\!|\!|\hat{g}_\omega|\!|\!|_r = \sum_{i,j=1}^{n}\sum_{l=1}^{d/2-1}\|g_{i\omega_j}^{2l}\|_{r,s_i-2},$$

and that $\|g_{i\omega_j}^{2l}\|_{r,s_i-2} = O(r^{2l})$ and $\|P_R\hat{G}_i^j\|_{r,s_i-2} = O(r^{j+2})$ as $r\to 0$. Motivated by this fact, we denote

$$\|g_{i\omega_j}^{2l}\|_{r,s_i-2} = u_{ij}^{2l}, \qquad \|P_R\hat{G}_i^j\|_{r,s_i-2} = v_i^{j+2}$$

and call $2l$ and $j+2$ the degree of u_{ij}^{2l} and v_i^{j+2} respectively. Let us denote

$$U = c_2\sum_{l=0}^{d/2-1}(c_2\,|\!|\!|\hat{g}_\omega|\!|\!|_r)^l, \qquad V = \sum_{i=1}^{n}\sum_{j=d}^{2d-1}\|P_R\hat{G}_i^j\|_{r,s_i-2}.$$

We treat U and V as polynomials of u_{ij}^{2l} and v_i^{j+2} respectively. Then each monomial in U has the form

$$(\text{const.})\prod_{\nu=1}^{m}u_{i_\nu j_\nu}^{2l_\nu} \quad \text{with} \quad i_\nu, j_\nu \in \{1,\ldots,n\}, \qquad 1 \le l_\nu \le \frac{d}{2}-1,$$

where m is a positive integer $\le d/2-1$. Let us define the degree of this monomial by $\mu = \sum_{\nu=1}^{m} 2l_\nu$ and let U^μ denote the sum of monomials in U of degree μ. We call U^μ the homogeneous part of degree μ in U and can write

$$U = U^0 + U^2 + \cdots + U^{(d-2)^2/2}.$$

We define the degree of monomials in V and UV similarly and can write

$$V = V^{d+2} + V^{d+3} + \cdots + V^{2d+1}.$$

To prove (4.11), it suffices to show

$$\|D_k W^{j+2}\|_r \le \sum_{\mu+\lambda=j+2} U^\mu V^\lambda \qquad (j = d, d+1, \ldots, 2d-1). \tag{4.12}$$

We will prove this by induction. For $j = d, d+1$ it follows from (3.8) and (4.4) that

$$\|D_k W^{j+2}\|_r \le c_2 \sum_{i=1}^{n} \|P_R \hat{G}_i^j\|_{r,s_i-2} = U^0 V^{j+2} \qquad (j = d, d+1).$$

This implies that (4.12) is valid for $j = d, d+1$. Assume that (4.12) is valid for $j = d, d+1, \ldots, d+2\kappa-2, d+2\kappa-1$ $(\kappa \le (d-2)/2)$. Then from (3.8) and (4.4), for $j = d+2\kappa, d+2\kappa+1$ we have

$$\begin{aligned}
\|D_k W^{j+2}\|_r &\le c_2 \sum_{i=1}^{n} \left\| P_R \hat{G}_i^j + \sum_{l=1}^{\kappa} \{g_i^{2l}, W^{j+2-2l}\} \right\|_{r,s_i-2} \\
&\le c_2 \sum_{i=1}^{n} \left(\|P_R \hat{G}_i^j\|_{r,s_i-2} + \sum_{l=1}^{\kappa} \sum_{\nu=1}^{n} \|g_{i\omega_\nu}^{2l} D_\nu W^{j+2-2l}\|_{r,s_i-2} \right) \\
&\le c_2 \sum_{i=1}^{n} \left(v_i^{j+2} + \sum_{l=1}^{\kappa} \sum_{\nu=1}^{n} \|g_{i\omega_\nu}^{2l}\|_{r,s_i-2} \|D_\nu W^{j+2-2l}\|_r \right) \\
&= c_2 V^{j+2} + c_2 \sum_{i,\nu=1}^{n} \sum_{l=1}^{\kappa} u_{i\nu}^{2l} \Big(\sum_{\mu+\lambda=j+2-2l} U^\mu V^\lambda \Big).
\end{aligned} \tag{4.13}$$

Here the second term can be estimated from above by the homogeneous part of degree $j+2$ in $c_2 \,|||\hat{g}_\omega|||_r\, UV$. Moreover

$$|||\hat{g}_\omega|||_r\, UV = \left\{ \sum_{l=1}^{d/2} (c_2 \,|||\hat{g}_\omega|||_r)^l \right\} V,$$

and $(c_2 \,|||\hat{g}_\omega|||_r)^{d/2} V$ contains monomials of degree $\ge d + d + 2 = 2d + 2$ alone, which do not contribute to $\|D_k W^{j+2}\|_r$ since $j+2<2d+2$. Therefore the right-hand side of (4.13) is estimated from above by the homogeneous part of degree $j+2$ in

$$c_2 V + c_2 \sum_{l=1}^{d/2-1} (c_2 \, |||\hat{g}_\omega|||_r)^l V = UV.$$

This implies that (4.12) is valid for $j=d+2\kappa, d+2\kappa+1$. Hence we have completed the proof of Proposition 4.3. □

Now we derive the estimate of norm of W and its derivatives. We set

$$\delta = \min_k \delta_k, \qquad c_4 = 8\pi c_2 n. \tag{4.14}$$

LEMMA 4.4. *Let $0<\rho<r$. Under the assumptions of Proposition 4.3, we have*

(i) $\|W\|_r \le c_4 \, |||\hat{G}|||_r$,

(ii) $\|W_{\zeta_k}\|_\rho \le \dfrac{c_4}{\delta(r-\rho)} \, |||\hat{G}|||_r \qquad (k=1,\ldots,2n)$,

(iii) $\|W_{\zeta_k\zeta_l}\|_\rho \le \dfrac{4c_4}{\delta^2(r-\rho)^2} \, |||\hat{G}|||_r \qquad (k,l=1,\ldots,2n)$,

where $\zeta=(\xi,\eta)=(\zeta_1,\ldots,\zeta_{2n})$.

Proof. We introduce a transformation $\psi:(\tau,\theta)\mapsto(\xi,\eta)$ defined by

$$\xi_k = \tau_k e^{2\pi i\theta_k}, \qquad \eta_k = \tau_k e^{-2\pi i\theta_k} \qquad (k=1,\ldots,n).$$

Then one can easily prove that

$$\frac{\partial}{\partial\theta_k} = -2\pi i D_k \qquad (k=1,\ldots,n). \tag{4.15}$$

Let $\Omega'_r=\bar{\Omega}_r\setminus\{(\xi,\eta)\mid \prod_{k=1}^n \xi_k\eta_k=0\}$. For any point $(\xi,\eta)\in\Omega'_r$, one can find a point (τ,θ) such that $(\xi,\eta)=\psi(\tau,\theta)$. Then for any $\sigma=(\sigma_1,\ldots,\sigma_n)$ with $0\le\sigma_k\le 1$, the point $(\xi',\eta')=\psi(\tau,\theta+\sigma)$ belongs to Ω'_r. By using the relation (4.15), the mean value theorem gives

$$|W^{j+2}(\psi(\tau,\theta)) - W^{j+2}(\psi(\tau,\theta+\sigma))| \le 2\pi \sum_{k=1}^n |D_k W^{j+2}|_r. \tag{4.16}$$

Since $P_N W^{j+2} = 0$ by the condition (3.9), the integration of (4.16) from 0 to 1 with respect to $\sigma_1, \ldots, \sigma_n$ gives

$$|W^{j+2}(\xi, \eta)| \leq 2\pi \sum_{k=1}^{n} |D_k W^{j+2}|_r.$$

Here (ξ, η) is arbitrary in Ω_r' and $|W^{j+2}|$ does not attain its maximum $|W^{j+2}|_r$ at a point where $\Pi_{k=1}^n \xi_k \eta_k = 0$. Therefore this implies

$$\|W\|_r \leq 2\pi \sum_{k=1}^{n} \|D_k W\|_r.$$

Hence the estimate (i) follows from Proposition 4.3. The estimates (ii) and (iii) can be obtained by using Cauchy integral formula. Indeed, applying it to each homogeneous polynomial W^{j+2}, we have

$$\|W_{\zeta_k}\|_\rho \leq \delta^{-1}(r-\rho)^{-1} \|W\|_r \leq c_4 \delta^{-1}(r-\rho)^{-1} |||\hat{G}|||_r.$$

For the estimate of $\|W_{\zeta_k \zeta_l}\|_\rho$, we set $\rho' = (r+\rho)/2$. Then similarly we have

$$\|W_{\zeta_k}\|_{\rho'} \leq \delta^{-1}(r-\rho')^{-1} \|W\|_r, \qquad \|W_{\zeta_k \zeta_l}\|_\rho \leq \delta^{-1}(\rho'-\rho)^{-1} \|W_{\zeta_k}\|_{\rho'},$$

which lead to the estimate (iii). □

5. Estimates of remainder terms $\hat{G}_k'$

The aim of this section is to give the estimate of new remainder terms G_k' at each step of iteration process defined by φ_ν. As in the previous section, we denote φ and G_k in place of φ_ν and $G_k^{(\nu)}$ respectively, and assume that G_k $(k = 1, \ldots, n)$ are written in the form (3.4). First we set

$$c_5 = 2nc_4\delta^{-2} \tag{5.1}$$

and prove the following

LEMMA 5.1. *Let* $0 < \sigma < \rho < r$ *with* $\rho - \sigma = r - \rho$. *Assume that* $c_2 \, |||\hat{g}_\omega|||_r < \frac{1}{2}$ *and that*

$$c_5(r-\rho)^{-2} \, |||\hat{G}|||_r < 1. \tag{5.2}$$

Then for any $\zeta \in \Omega_\sigma$ *the solution* $z = \varphi^t(\zeta)$ *of the system* (3.1) *exists and is contained in* Ω_ρ *for* $|t| < 2$, *and hence* $\varphi = \varphi^1$ *is an analytic transformation from* Ω_σ *into* Ω_ρ.

Proof. By Lemma 4.4(ii) and (5.1), it follows that

$$\|W_{\zeta_k}\|_\rho < \frac{1}{2n}\,\delta(r-\rho) \qquad (k = 1, \ldots, 2n)$$

under the assumption (5.2). Therefore for any point $\zeta \in \Omega_\sigma$ we have

$$\max_k |W_{z_k}(z)| < \frac{1}{2n}\,\delta(r-\rho)$$

$$\text{on}\quad \Omega(\zeta) := \left\{ z = (z_1, \ldots, z_{2n}) \,\middle|\, \max_k |z_k - \zeta_k| < \delta(\rho - \sigma) \right\}.$$

Therefore by the fundamental theorem for differential equations, the solution $z = \varphi^t(\zeta)$ of (3.1) exists and is contained in the domain $\Omega(\zeta)$ for $|t| < 2n$, where $2n \geq 2$. Since $\Omega(\zeta) \subset \Omega_\rho$, this proves the assertion. □

In what follows, we use the notation

$$|f|_r := \max_k |f_k|_r \left(= \max_k \max_{\zeta \in \bar{\Omega}_r} |f_k(\zeta)| \right) \tag{5.3}$$

for a $2n$-dimensional vector function $f = (f_1, \ldots, f_{2n})$ with $f_k \in A(\Omega_r)$. Moreover we use the notation $\langle z, z' \rangle = \sum_k^{2n} z_k z'_k$ for $2n$-dimensional vectors $z, z' \in \mathbf{C}^{2n}$ whose components are z_k, z'_k $(k = 1, \ldots, 2n)$ respectively.

In addition to σ, ρ, let us introduce r', τ such that

$$r' < \tau < \sigma < \rho < r; \qquad r - \rho = \rho - \sigma = \sigma - \tau = \tau - r' = \tfrac{1}{4}(r - r'). \tag{5.4}$$

Our aim is to estimate $|||\hat{G}'|||_{r'} := \sum_{i=1}^n \|\hat{G}'_k\|_{r', s_k - 2}$ under the condition (5.2). First we define a $2n$-dimensional vector function $\hat{W}$ by

$$\hat{W}(\zeta) = \int_0^1 \nabla W(\varphi^t(\zeta))\, dt \tag{5.5}$$

so that the transformation φ is written in the form

$$\varphi(\zeta) = \zeta + J\hat{W}(\zeta). \tag{5.6}$$

To estimate $G_k(\varphi(\zeta))$, we write it in the form

$$G_k(\varphi(\zeta)) = g_k(\varphi(\zeta)) + \hat{G}_k(\varphi(\zeta)). \tag{5.7}$$

Then from (5.6) one obtains

$$\begin{aligned} g_k(\varphi(\zeta)) &= g_k(\zeta) + \langle \nabla g_k(\zeta), J\hat{W}(\zeta)\rangle + R_k^1(\zeta) \\ &= g_k(\zeta) + \{g_k(\zeta), W(\zeta)\} + R_k^1(\zeta) + R_k^2(\zeta), \end{aligned} \tag{5.8}$$

$$\hat{G}_k(\varphi(\zeta)) = \hat{G}_k(\zeta) + R_k^3(\zeta), \tag{5.9}$$

where the remainder terms R_k^1, R_k^2 and R_k^3 are estimated with respect to the norm $|\cdot|_\tau$ as follows:

$$\begin{aligned} |R_k^1(\zeta)|_\tau &\le \tfrac{1}{2} \sum_{i,j=1}^{2n} \left| \frac{\partial^2 g_k}{\partial \zeta_i\, \partial \zeta_j} \right|_\rho |\hat{W}|_\sigma^2 \le 2n^2 \max_{i,j} \left| \frac{\partial^2 g_k}{\partial \zeta_i\, \partial \zeta_j} \right|_\rho |\nabla W|_\rho^2, \\ |R_k^2(\zeta)|_\tau &= |\langle \nabla g_k, J(\hat{W}(\zeta) - \nabla W(\zeta))\rangle|_\tau \\ &\le 2n\, |\nabla g_k|_\tau \left| \int_0^1 (\nabla W(\varphi^t(\zeta)) - \nabla W(\zeta))\, dt \right|_\tau \\ &\le (2n)^2\, |\nabla g_k|_\tau \max_{i,j} |W_{\zeta_i\zeta_j}|_\rho\, |\nabla W|_\rho, \\ |R_k^3(\zeta)|_\tau &\le 2n\, |\nabla \hat{G}|_\rho\, |\nabla \hat{W}|_\sigma \le 2n\, |\nabla \hat{G}|_\rho\, |\nabla W|_\rho. \end{aligned}$$

From (5.7), (5.8) and (5.9), we can write $G_k(\varphi(\zeta))$ as

$$G_k(\varphi(\zeta)) = g_k(\zeta) + \{g_k(\zeta), W(\zeta)\} + \hat{G}_k(\zeta) + R_k^1 + R_k^2 + R_k^3$$

with

$$\{g_k(\zeta), W(\zeta)\} + \hat{G}_k(\zeta) = P_N^{2d-1}\hat{G}_k(\zeta) + R_k^4(\zeta). \tag{5.10}$$

Here $R_k^4(\zeta)$ contains terms of order $\ge s_k + 2d$ only since W satisfies equation (3.5). Hence the remainder part $\hat{G}_k'$ is given by

$$\hat{G}_k' = R_k^1 + R_k^2 + R_k^3 + R_k^4.$$

Now we first consider the estimate of $R_k^1 + R_k^2 + R_k^3$. By Lemma 4.4 and using the Cauchy integral formula, we can rewrite the estimates of $|R_k^1|_\tau$, $|R_k^2|_\tau$ and $|R_k^3|_\tau$

as follows:

$$|R_k^1|_\tau \le 2n^2 \frac{4}{\delta^2(r-\rho)^2} |g_k|_r \left\{ \frac{c_4}{\delta(r-\rho)} |||\hat{G}|||_r \right\}^2 = 2c_5^2(r-\rho)^{-4} |g_k|_r \, |||\hat{G}|||_r^2,$$

$$|R_k^2|_\tau \le (2n)^2 \frac{1}{\delta(r-\tau)} |g_k|_r \frac{4c_4}{\delta^2(r-\rho)^2} |||\hat{G}|||_r \cdot \frac{c_4}{\delta(r-\rho)} |||\hat{G}|||_r$$
$$\le \tfrac{4}{3} c_5^2 (r-\rho)^{-4} |g_k|_r \, |||\hat{G}|||_r^2,$$

$$|R_k^3|_\tau \le 2n \frac{1}{\delta(r-\rho)} \|\hat{G}_k\|_r \frac{c_4}{\delta(r-\rho)} |||\hat{G}|||_r = c_5(r-\rho)^{-2} \|\hat{G}_k\|_r \, |||\hat{G}|||_r$$

and hence we have

$$|R_k^1 + R_k^2 + R_k^3|_\tau \le \frac{c_5}{(r-\rho)^4} |||\hat{G}|||_r \left\{ \tfrac{10}{3} c_5 |g_k|_r \, |||\hat{G}|||_r + (r-\rho)^2 \|G_k\|_r \right\}. \tag{5.11}$$

To obtain the estimate with respect to the norm $\|\cdot\|_{r'}$, we use the following fact.

LEMMA 5.2. *For $f \in A(\Omega_\tau)$, we have*

$$\|f\|_{r'} \le \frac{|f|_\tau}{1 - \dfrac{r'}{\tau}}. \tag{5.12}$$

Proof. We may assume that the function f is expanded in homogeneous polynomials as in (4.2). Then for any $\zeta = (\xi, \eta) \in \bar{\Omega}_{r'}$ the function

$$\hat{f}(t) := f(t\zeta) = \sum_{j=0}^{\infty} t^j f_j(\zeta)$$

is analytic in $t \in \{|t| \le \tau/r'\}$. Therefore the Cauchy's estimate gives

$$|f_j(\zeta)| \le \left(\frac{\tau}{r'}\right)^{-j} \max_{|t| \le \tau/r'} |\hat{f}(t)| \le \left(\frac{\tau}{r'}\right)^{-j} |f|_\tau.$$

This implies

$$|f_j|_{r'} \le \left(\frac{r'}{\tau}\right)^j |f|_\tau.$$

By the definition of $\|f\|_{r'}$, the relation (5.12) follows easily. □

Here it follows from (5.4) that

$$1-\frac{\rho}{r}=\frac{r-\rho}{r}<\frac{r-\rho}{\tau}=1-\frac{r'}{\tau}$$

and therefore

$$\frac{1}{1-\frac{r'}{\tau}}<\frac{1}{1-\frac{\rho}{r}}.$$

By using the above lemma together with this relation, one obtains from (5.11) that

$$\sum_{k=1}^{n}\|R_k^1+R_k^2+R_k^3\|_{r',s_k-2}\le\frac{c_5c_6}{r^4\left(1-\frac{\rho}{r}\right)^5}|||\hat{G}|||_r^2 \times\left(\tfrac{10}{3}c_5\sum_{k=1}^{n}\|g_k\|_{r,s_k-2}+(r-\rho)^2\right), \tag{5.13}$$

where c_6 is a constant satisfying

$$\left(\frac{r}{r'}\right)^{s_k-2}\le c_6\quad\text{for}\quad k=1,\ldots,n. \tag{5.14}$$

To estimate $\sum_{k=1}^{n}\|g_k\|_{r,s_k-2}$ in terms of $|||\hat{g}_\omega|||_r$, we note that for each homogeneous polynomial g_k^{2l}

$$g_k^{2l}(\zeta)=\int_0^{\omega_j}g_{k_{\omega_j}}^{2l}\,d\omega_j\quad\text{for some}\quad\omega_j=\xi_j\eta_j.$$

Using this relation we have

$$|g_k^{2l}|_r\le r^2\delta_j^2\,|g_{k_{\omega_j}}^{2l}|_r<r^2\sum_{j=1}^{n}|g_{k_{\omega_j}}^{2l}|_r$$

and therefore

$$\|g_k\|_{r,s_k-2}\le r^2\sum_{j=1}^{n}\|g_{k_{\omega_j}}\|_{r,s_k-2},$$

which leads to

$$\sum_{k=1}^{n} \|g_k\|_{r,s_k-2} \leq r^2\left(\sum_{k,j=1}^{n} \|g^0_{k_{\omega_j}}\|_{r,s_k-2} + |||\hat{g}_\omega|||_r\right) \leq r^2\left(c_7 + \frac{1}{2c_2}\right) \tag{5.15}$$

by the assumption $c_2\,|||\hat{g}_\omega|||_r < \frac{1}{2}$. Here c_7 is a constant satisfying

$$\sum_{k,j=1}^{n} \|g^0_{k_{\omega_j}}\|_{r,s_k-2} \leq c_7.$$

Since $g^0_{k_{\omega_j}}$ is a homogeneous polynomial of degree $s_k - 2$, the constant c_7 can be taken independently of r.

Hence it follows from (5.13) and (5.15) that

$$\sum_{k=1}^{n} \|R^1_k + R^2_k + R^3_k\|_{r',s_k-2} \leq \frac{c_6 c_8}{r^2\left(1 - \frac{\rho}{r}\right)^5} |||\hat{G}|||_r^2 \tag{5.16}$$

where

$$c_8 = c_5\left\{\tfrac{10}{3} c_5\left(c_7 + \frac{1}{2c_2}\right) + 1\right\}. \tag{5.17}$$

Finally we will estimate $\|R^4_k\|_{r'}$. Since R^4_k contains terms of order $\geq s_k + 2d$ only, by applying Schwarz' lemma to each homogeneous polynomial in R^4_k one obtains

$$\|R^4_k\|_{r'} \leq \|R^4_k\|_r \left(\frac{r'}{r}\right)^{s_k+2d}.$$

From (5.10) we have

$$\|\{g_k, W\} + \hat{G}_k\|_r = \|P_N^{2d-1}\hat{G}_k\|_r + \|R^4_k\|_r$$

by the definition of the norm $\|\cdot\|_r$. This gives

$$\|R^4_k\|_r \leq \|\{g_k, W\} + \hat{G}_k\|_r$$

and therefore we have

$$\|R^4_k\|_{r',s_k-2} \leq c_6\,\|\{g_k, W\} + \hat{G}_k\|_{r,s_k-2}\left(\frac{r'}{r}\right)^{s_k+2d},$$

where c_6 is a constant satisfying (5.14). Moreover we have the estimate

$$\begin{aligned}\sum_{k=1}^{n} \|\{g_k, W\} + \hat{G}_k\|_{r,s_k-2} &\le \sum_{k,j=1}^{n} \|g_{k_{\omega_j}} D_j W\|_{r,s_k-2} + |||\hat{G}|||_r \\ &\le \left(c_7 + \frac{1}{2c_2}\right) 4c_2 \,|||\hat{G}|||_r + |||\hat{G}|||_r \\ &= (4c_2c_7 + 3)\,|||\hat{G}|||_r\end{aligned}$$

by (4.10) and the estimate used to obtain (5.15). Hence, noting that $s_k \ge 2$, one obtains

$$\sum_{k=1}^{n} \|R_k^4\|_{r',s_k-2} \le c_6(4c_2c_7+3)\,|||\hat{G}|||_r \left(\frac{r'}{r}\right)^{2d+2} \tag{5.18}$$

Then setting

$$c_9 = 4^5 \max(c_8, 4c_2c_7+3), \tag{5.19}$$

the estimates (5.16) and (5.18) give

$$\begin{aligned}|||\hat{G}'|||_{r'} &= \sum_{k=1}^{n} \|R_k^1 + R_k^2 + R_k^3 + R_k^4\|_{r',s_k-2} \\ &\le c_6c_9\,|||\hat{G}|||_r \left\{\frac{|||\hat{G}|||_r}{r^2\left(1-\frac{r'}{r}\right)^5} + \left(\frac{r'}{r}\right)^{2d+2}\right\}.\end{aligned}$$

This is the desired estimate of $|||\hat{G}'|||_{r'}$. For the iteration process, we require $|||\hat{G}|||_r$ to satisfy a stronger condition than (5.2). We summarize as follows:

PROPOSITION 5.3. *Let $r' < r$ and $G_k \in A(\Omega_r)$ $(k = 1, \ldots, n)$ be of the form (3.4) (i.e., in normal form up to terms of order $s_k + d - 1$). Assume that*

$$c_2\,|||\hat{g}_\omega|||_r < \tfrac{1}{2} \tag{5.20}$$

and

$$\frac{c_{10}}{r^2\left(1-\frac{r'}{r}\right)^5}\,|||\hat{G}|||_r < 1, \qquad c_{10} = c_6c_9. \tag{5.21}$$

Let $z=\varphi(\zeta)$ be the canonical transformation which satisfies the condition (ii) *of Theorem* 3.6 *and takes the G_k into* (3.6). *Then the φ is an analytic transformation from Ω_σ into Ω_ρ and $G_k(\varphi(\zeta))$ belongs to $A(\Omega_{r'})$ with the remainder part $\hat{G}'_k$ satisfying*

$$\||\hat{G}'\||_{r'} \leq c_{10} \||\hat{G}\||_r \left\{ \frac{\||\hat{G}\||_r}{r^2\left(1-\frac{r'}{r}\right)^5} + \left(\frac{r'}{r}\right)^{2d+2} \right\}, \tag{5.22}$$

where σ and ρ are defined by $\sigma=r-\frac{1}{2}(r-r')$, $\rho=r-\frac{1}{4}(r-r')$.

Since $c_6 \geq 1$ by (5.14), it follows from (5.17) and (5.19) that

$$c_{10} \geq c_9 \geq 4^5 c_8 \geq 4^5 c_5. \tag{5.23}$$

Therefore the condition (5.2) is satisfied under the condition (5.21), which implies the validity of the estimate (5.22).

The estimate of the form (5.22) contains the term $c_{10} \||\hat{G}\||_r (r'/r)^{2d+2}$ which depends linearly on $\||\hat{G}\||_r$. In this sense, it is a little different from the usual estimate in the rapidly convergent iteration method. This type of estimate is found in [7], to which we owe the idea of estimating $\|R_k^4\|_{r'}$ in the above as well as the convergent proof in the next section.

For the new normal form part $g'_k = g_k + P_N^{2d-1}\hat{G}_k$, we estimate

$$\||\hat{g}'_\omega\||_{r'} := \sum_{k,j=1}^{n} \|\hat{g}'_{k_{\omega_j}}\|_{r',s_k-2}; \qquad \hat{g}'_k = g'_k - g_k^0 = \hat{g}_k + P_N^{2d-1}\hat{G}_k. \tag{5.24}$$

PROPOSITION 5.4. *Under the assumptions of Proposition* 5.3, *we have*

$$\||\hat{g}'_\omega\||_{r'} \leq \||\hat{g}_\omega\||_r + \frac{c_6 n}{\delta^2} \frac{\||\hat{G}\||_r}{r'(r-r')}. \tag{5.25}$$

Proof. We note that $\partial/\partial\omega_i = \eta_i^{-1}\,\partial/\partial\xi_i$ and the maximum $|\partial(P_N\hat{G}_k^j)/\partial\omega_i|_{r'}$ is attained on $\Delta_{r'}$. Then using the Cauchy integral formula, we have

$$\begin{aligned} |\partial(P_N\hat{G}_k^j)/\partial\omega_i|_{r'} &\leq \frac{1}{\delta_i r'} \frac{|P_N\hat{G}_k^j|_r}{\delta_i(r-r')} \\ &\leq \frac{1}{\delta^2 r'(r-r')} |\hat{G}_k^j|_r \quad \text{for} \quad j=d,\ldots,2d-1, \end{aligned}$$

which leads to

$$\|\partial(P_N^{2d-1}\hat{G}_k)/\partial\omega_i\|_{r',s_k-2} \le \frac{c_6}{\delta^2 r'(r-r')}\|\hat{G}_k\|_{r,s_k-2}.$$

From (5.24) this gives (5.25). □

6. Convergence proof

The ν-th approximation step described in the assumption of Theorem 3.6 is defined by the transformation $\varphi = \varphi_\nu$ which takes $G_k = G_k^{(\nu)}$ in the form (3.4) into (3.6). More precisely we set in (3.4) and (3.6)

$$g_k = g_k^{(\nu)}, \qquad \hat{G}_k = \hat{G}_k^{(\nu)} \quad \text{and} \quad g_k' = g_k^{(\nu+1)}, \qquad \hat{G}_k' = \hat{G}_k^{(\nu+1)}.$$

Then Proposition 5.3 with $r = r_\nu$ and $r' = r_{\nu+1}$ implies that

$$\varphi_\nu : \Omega_{\sigma_\nu} \to \Omega_{\rho_\nu}; \qquad \sigma_\nu = r_\nu - \tfrac{1}{2}(r_\nu - r_{\nu+1}), \qquad \rho_\nu = r_\nu - \tfrac{1}{4}(r_\nu - r_{\nu+1})$$

and (5.22) gives the estimate of $|||\hat{G}^{(\nu+1)}|||_{r_{\nu+1}} = \sum_{k=1}^n \|\hat{G}_k^{(\nu+1)}\|_{r_{\nu+1},s_k-2}$. Our goal is to choose a sequence $\{r_\nu\}$ so that

$$|||\hat{G}^{(\nu)}|||_{r_\nu} \to 0 \quad \text{as} \quad \nu \to \infty \tag{6.1}$$

and that

$$\phi_\nu = \varphi_0 \circ \varphi_1 \circ \cdots \circ \varphi_\nu : \Omega_{\sigma_\nu} \to \Omega_{\rho_0} \tag{6.2}$$

converges in $\Omega_{r_0/2}$ to an analytic canonical transformation ϕ. For this purpose, we assume that $G_k = G_k^{(0)} \in A(\Omega_{r_0})$ $(r_0 > 0)$ and take a sequence $\{r_\nu\}$ defined by

$$r_\nu = \frac{r_0}{2}\left(1 + \frac{1}{\nu+1}\right) \qquad \nu = 0, 1, 2, \ldots.$$

Then since

$$1 - \frac{r_{\nu+1}}{r_\nu} = \frac{1}{(\nu+2)^2}, \tag{6.3}$$

we have

$$\frac{r_\nu}{r_{\nu+1}} \leq \tfrac{4}{3} \quad \text{for} \quad \nu = 0, 1, 2, \ldots, \tag{6.4}$$

and therefore the constant c_6 in (5.14) can be taken independently of r_ν.

To justify the iteration procedure with this sequence $\{r_\nu\}$, we have to prove that the conditions (5.20) and (5.21) hold at each step. We rewrite the condition (5.21) as

$$c_{10} \frac{|||\hat{G}^{(\nu)}|||_{r_\nu}}{r_\nu^2\left(1 - \frac{r_{\nu+1}}{r_\nu}\right)^5} < 1, \tag{6.5}$$

where $c_{10} = c_6 c_9$ is a positive constant which is independent of r_ν. The following lemma implies that if the condition (6.5) holds for any $\nu = 0, 1, \ldots,$ then the condition (5.20) is necessarily satisfied at each step of iteration procedure.

LEMMA 6.1. *If the condition* (6.5) *holds for* $\nu = 0, 1, \ldots, m$, *then*

$$c_2 \, |||\hat{g}_\omega^{(\nu)}|||_{r_\nu} < \tfrac{1}{2} \quad \text{for} \quad \nu = 0, 1, \ldots, m+1.$$

Proof. For the simplicity of notation, we set

$$d_\nu := c_2 \, |||\hat{g}_\omega^{(\nu)}|||_{r_\nu}.$$

By Proposition 5.4 we have

$$d_{\nu+1} \leq d_\nu + \frac{c_2 c_6 n}{\delta^2} \frac{|||\hat{G}^{(\nu)}|||_{r_\nu}}{r_{\nu+1}(r_\nu - r_{\nu+1})}.$$

Therefore the condition (6.5) implies that

$$d_{\nu+1} \leq d_\nu + \frac{c_2 n}{\delta^2 c_9} \frac{r_\nu}{r_{\nu+1}} \left(1 - \frac{r_{\nu+1}}{r_\nu}\right)^4.$$

Here we have

$$\frac{c_2 n}{\delta^2 c_9} \frac{r_\nu}{r_{\nu+1}} < 1$$

by (6.4) and the definitions of the constants (4.14), (5.1) and (5.23). Hence one has

$$d_{\nu+1} < d_\nu + \left(1 - \frac{r_{\nu+1}}{r_\nu}\right)^4 = d_\nu + \frac{1}{(\nu+2)^8} \tag{6.6}$$

by (6.3). Now we have $d_0 = 0$ by the definition of $|||\hat{g}_\omega^{(\nu)}|||_{r_\nu}$. Therefore (6.6) implies $d_1 < \frac{1}{2}$. If (6.5) holds for $\nu = 0, 1, \ldots, m$, then by using (6.6) inductively we have

$$d_{m+1} < \sum_{\nu=0}^{m} \frac{1}{(\nu+2)^8} < \frac{1}{2}.$$

This completes the proof. □

To prove (6.5) for all $\nu \geq 0$, we set

$$\epsilon_\nu = \frac{|||\hat{G}^{(\nu)}|||_{r_\nu}}{r_\nu^2\left(1 - \dfrac{r_{\nu+1}}{r_\nu}\right)^5} \tag{6.7}$$

and rewrite (5.22) in the form

$$\epsilon_{\nu+1} \leq c_{10}\left(\frac{r_\nu}{r_{\nu+1}}\right)^2\left(\frac{1 - r_{\nu+1}r_\nu^{-1}}{1 - r_{\nu+2}r_{\nu+1}^{-1}}\right)^5 \cdot \epsilon_\nu\left\{\epsilon_\nu + \left(\frac{r_{\nu+1}}{r_\nu}\right)^{2^{\nu+1}+2}\right\}.$$

Noting (6.5) and

$$\frac{1 - r_{\nu+1}r_\nu^{-1}}{1 - r_{\nu+2}r_{\nu+1}^{-1}} = \left(\frac{\nu+3}{\nu+2}\right)^2 \leq \left(\frac{3}{2}\right)^2$$

by (6.3), one obtains

$$\epsilon_{\nu+1} \leq c\epsilon_\nu(\epsilon_\nu + \lambda_\nu) \tag{6.8}$$

where

$$c = (\tfrac{4}{3})^2(\tfrac{3}{2})^{10}c_{10}$$

and

$$\lambda_\nu = \left(\frac{r_{\nu+1}}{r_\nu}\right)^{2^{\nu+1}+2} = \left(1 - \frac{1}{(\nu+2)^2}\right)^{2^{\nu+1}+2} \to 0 \qquad (\nu \to \infty). \tag{6.9}$$

We note that $c > 1$ since $c_{10} > 1$ because of (5.19) and $c_6 > 1$.

Our aim is to prove

$$c\epsilon_\nu < 1 \quad \text{for all} \quad \nu = 0, 1, \ldots, \tag{6.10}$$

which yields (6.5) for all $\nu = 0, 1, \ldots$ since $c > c_{10}$. We note from (6.9) that there exists a positive integer N such that

$$\lambda_\nu < (4c)^{-1} \quad \text{for} \quad \nu \geq N.$$

Since $\hat{G}_k^{(0)}$ begins with terms of order $\geq s_k + 1$, $r_0^{-2} \, |||\hat{G}^{(0)}|||_{r_0}$ can be so small that

$$c\epsilon_o < (4c)^{-1}(2c)^{-N}$$

by taking sufficiently small $r_0 > 0$. By using (6.8) with ϵ_ν, $\lambda_\nu < 1$ inductivity, one can prove that

$$c\epsilon_\nu < (4c)^{-1}(2c)^{-N+\nu} < 1 \quad \text{for} \quad \nu = 0, 1, \ldots, N. \tag{6.11}$$

For $\nu > N$ one finds, using (6.8) with ϵ_ν, $\lambda_\nu < (4c)^{-1}$ inductively,

$$c\epsilon_\nu \leq c\epsilon_N 2^{-\nu+N} < 1. \tag{6.12}$$

Hence the inequalities (6.11) and (6.12) imply (6.10), and the iteration procedure is justified. The assertion (6.1) follows from (6.7) and (6.12).

Finally we will prove the uniform convergence of $\phi_\nu = \varphi_0 \circ \varphi_1 \circ \cdots \circ \varphi_\nu$ in $\Omega_{r_0/2}$. First we note that

$$|\phi_{\nu+1}(\zeta) - \phi_\nu(\zeta)|_{\tau_{\nu+1}} \leq |d\phi_\nu|_{\tau_\nu} |\varphi_{\nu+1}(\zeta) - \zeta|_{\tau_{\nu+1}}; \qquad \tau_\nu = r_\nu - \tfrac{3}{4}(r_\nu - r_{\nu+1})$$

since $\phi_{\nu+1} = \phi_\nu \circ \varphi_{\nu+1}$. Here $|d\phi_\nu|_{\tau_\nu} := \sup_{\zeta \in \Omega_{\tau_\nu}} |d\phi_\nu(\zeta)|$, where $|d\phi_\nu(\zeta)|$ is the operator norm of the Jacobian $d\phi_\nu$ at the point ζ, and the other norms are supremum norm defined by (5.3) for $2n$-dimensional vector functions. Since

$d\phi_\nu = d\phi_{\nu-1} \circ d\varphi_\nu$, we have

$$|d\phi_\nu|_{\tau_\nu} \leq |d\phi_{\nu-1}|_{\tau_{\nu-1}} |d\varphi_\nu|_{\tau_\nu} \leq \prod_{j=0}^{\nu} |d\varphi_j|_{\tau_j}.$$

Here it follows from (5.6) that

$$d\varphi_\nu = I + J\nabla\hat{W}(\zeta). \tag{6.13}$$

Moreover from (5.5) and Lemma 4.4(ii) with (5.1), using Cauchy's integral formula one has

$$|\hat{W}_{k_{\zeta_l}}|_{\tau_\nu} \leq \frac{|\hat{W}_k|_{\sigma_\nu}}{\delta(\sigma_\nu - \tau_\nu)} \leq \frac{|W_{\zeta_k}|_{\rho_\nu}}{\delta(r_\nu - \rho_\nu)} \leq \frac{c_4}{\delta^2(r_\nu - \rho_\nu)^2} |||\hat{G}^{(\nu)}|||_{r_\nu}$$
$$\leq \frac{4^2 c_5}{2n} \frac{|||\hat{G}^{(\nu)}|||_{r_\nu}}{r_\nu^2\left(1 - \frac{r_{\nu+1}}{r_\nu}\right)^2} \qquad (k, l = 1, \ldots, 2n).$$

Then it follows from (5.23) and the relation $c > c_{10}$ that

$$|\hat{W}_{k_{\zeta_l}}|_{\tau_\nu} \leq \frac{c}{2n} \frac{|||\hat{G}^{(\nu)}|||_{r_\nu}}{r_\nu^2\left(1 - \frac{r_{\nu+1}}{r_\nu}\right)^5}.$$

By (6.13) this leads to

$$|d\varphi_\nu|_{\tau_\nu} < 1 + c\epsilon_\nu$$

and hence

$$|d\phi_\nu|_{\tau_\nu} < \prod_{\kappa=0}^{\nu} (1 + c\epsilon_\kappa) \leq \prod_{\kappa=0}^{\infty} (1 + c\epsilon_\kappa).$$

From (6.12) this infinite product converges. Setting

$$c_{11} = \prod_{\kappa=0}^{\infty} (1 + c\epsilon_\kappa),$$

one has

$$\begin{aligned}|\phi_{\nu+1}(\zeta)-\phi_\nu(\zeta)|_{\tau_{\nu+1}} &\le c_{11}\,|\varphi_{\nu+1}(\zeta)-\zeta|_{\tau_{\nu+1}}\le c_{11}\,|\nabla W|_{\rho_{\nu+1}}\\ &\le c_{11}\frac{c_4\,|||\hat{G}^{(\nu+1)}|||_{r_{\nu+1}}}{r_{\nu+1}-\rho_{\nu+1}}\le c_{11}c\epsilon_{\nu+1}\\ &\le c_{11}(4c)^{-1}2^{-\nu-1+N}\end{aligned}$$

for $\nu > N$. This implies the uniform convergence of ϕ_ν in $\Omega_{r_0/2}$. We have thus completed the proof of Theorem 3.6. □

7. The real analytic case

The aim of this section is to show that if the Hamiltonian is real analytic, the canonical transformation ϕ given in Theorem 1.3 can be taken as real analytic with replacing the normal form by "real" normal form. In order to see this, we impose a "reality condition" on the Hamiltonian of the form (1.2).

Let us assume that the original Hamiltonian is a real analytic function of $w=(u, v)$ and that the Hamiltonian $H=H(z)$ of the form (1.2) is obtained by a linear canonical transformation $w=Cz$. Then the original Hamiltonian is given by $H=H(C^{-1}w)$ and its reality gives the identity

$$H(z)=\bar{H}(Tz), \qquad T=\bar{C}^{-1}C. \tag{7.1}$$

Here and in what follows, for any power series (or vector of power series) f we denote by $\bar{f}$ the power series obtained from f by replacing the coefficients by their complex conjugates. The linear transformation T is canonical.

For any analytic function (power series) $f=f(z)$, we say that F satisfies the *reality condition* if the identity $f(z)=\bar{f}(Tz)$ holds, where $T=\bar{C}^{-1}C$ with a given symplectic matrix C. This means that $f(C^{-1}w)$ is a real analytic function of w. Let $\mathfrak{R}_F$ denote the set of all analytic functions satisfying the reality condition. Then we have

LEMMA 7.1. *If f, $W\in\mathfrak{R}_F$, then $\{f, W\}\in\mathfrak{R}_F$.*

Proof. The assertion follows from the identity

$$\{\bar{f}(Tz), \bar{W}(Tz)\}=\{\bar{f}, \bar{W}\}(Tz), \tag{7.2}$$

which can be easily seen by the canonical property of T, i.e., ${}^tTJT=J$. □

Let $\mathfrak{R}$ be the set of all canonical transformations φ satisfying a condition

$$\bar{\varphi}(T\zeta) = T\varphi(\zeta). \tag{7.3}$$

Then one can easily see that $\mathfrak{R}$ forms a group under compositions. We prove the following

LEMMA 7.2. *Let φ be a canonical transformation generated by the Hamiltonian flow with Hamiltonian W. If $W \in \mathfrak{R}_F$, then $\varphi \in \mathfrak{R}$ and $f \circ \varphi \in \mathfrak{R}_F$ for any function $f \in \mathfrak{R}_F$.*

Proof. We denote by $z = \varphi^t(\zeta)$ the Hamiltonian flow of equation (3.1) satisfying the initial condition $z = \zeta$ at $t = 0$. Let $f(z)$ be an analytic function. We note from Lemma 5.1 that $\varphi^t(\zeta)$ is defined for $|t| < 2$. Then we can express $f \circ \varphi^t(\zeta)$ in the form

$$f \circ \varphi^t(\zeta) = \exp(t ad_W f) := \sum_{m=0}^{\infty} \frac{t^m}{m!} ad_W^m f(\zeta) \tag{7.4}$$

where

$$ad_W^0 f = f, \qquad ad_W^m f = \{ad_W^{m-1} f, W\}.$$

Indeed we can easily prove that

$$\left(\frac{d}{dt}\right)^m f \circ \varphi^t(\zeta) = (ad_W^m f)(\varphi^t(\zeta)), \qquad (m = 1, 2, \ldots)$$

by using the above notation, and therefore the relation (7.4) represents the Taylor expansion of $f \circ \varphi^t(\zeta)$. Setting $t = 1$ in (7.4), we have

$$f \circ \varphi(\zeta) = \sum_{m=0}^{\infty} \frac{1}{m!} ad_W^m f(\zeta).$$

Assume that f, $W \in \mathfrak{R}_F$. Then one can see from Lemma 7.1 that $ad_W^m f(\zeta) \in \mathfrak{R}_F$ for any $m = 0, 1, \ldots$, inductively. Hence it follows that $f \circ \varphi(\zeta) \in \mathfrak{R}_F$.

Next we notice the identity

$$f \circ \varphi(\zeta) = f \circ \varphi \circ C^{-1}(w) = f \circ C^{-1} \circ (C \circ \varphi \circ C^{-1})(w), \qquad w = C\zeta \tag{7.5}$$

Here $f, f\circ\varphi \in \mathfrak{R}_F$, and therefore $f\circ C^{-1}(w)$ and $f\circ\varphi\circ C^{-1}(w)$ are real analytic functions of w. Since $f\circ C^{-1}(w)$ can be taken as an arbitrary real analytic function of w, the identity (7.5) implies that $C\circ\varphi\circ C^{-1}$ is real mapping, namely

$$\bar{C}\circ\bar{\varphi}\circ\bar{C}^{-1} = C\circ\varphi\circ C^{-1},$$

which is equivalent to (7.3). This completes the proof. □

Let us now specify the form of the linear transformation T. Notice that the linear transformation $z\mapsto Cz$ diagonalize the linearized vector field of the system (1.1) in such a way that $C^{-1}AC = \operatorname{diag}(\lambda_1, \ldots, \lambda_n, -\lambda_1, \ldots, -\lambda_n)$, where $A = J\operatorname{Hess} H(0)$. In the real case, the eigenvalues of A occur also in pairs $\lambda_k, \bar{\lambda}_k$. We can choose (see [11, §15]) the symplectic matrix C so that (i) $\bar{\lambda}_k \in \{\lambda_1, \ldots, \lambda_n\}$ unless λ_k is purely imaginary, and (ii) the transformation $z^* = Tz$ is given by

$$\begin{cases} x_k^* = -iy_k, \quad y_k^* = -ix_k & \text{if } \lambda_k \text{ is purely imaginary;} \\ x_k^* = x_{l_k}, \quad y_k^* = y_{l_k} & \text{otherwise,} \end{cases} \tag{7.6}$$

where l_k is the number $(1\le l_k\le n)$ such that $\lambda_{l_k} = \bar{\lambda}_k$.

We consider one iteration step described in Proposition 3.3.

PROPOSITION 7.3. *In addition to the assumption of Proposition* 3.3, *assume that* $G_1\in\mathfrak{R}_F$, *where the matrix T in the reality condition is given by* (7.6). *Let W be the polynomial whose Hamiltonian flow generates the canonical transformation φ in Proposition* 3.3. *Then* $W\in\mathfrak{R}_F$ *and* $\varphi\in\mathfrak{R}$.

Proof. Since the W satisfies the system (3.5), it satisfies the equation especially for $k=1$. In that equation, replacing the coefficients by their complex conjugates and changing the variable ζ into $T\zeta$, we obtain

$$\{\bar{g}_1, \bar{W}\}(T\zeta) + \overline{\hat{G}_1}(T\zeta) = \overline{P_N^{2d-1}\hat{G}_1}(T\zeta) + O(\zeta^{s_1+2d}).$$

Here we note that the linear transformation $\zeta\mapsto T\zeta$ given by (7.6) takes $\xi_k\eta_k$ into $-\xi_k\eta_k$ if λ_k is purely imaginary and into $\xi_{l_k}\eta_{l_k}$ otherwise. Therefore if $G_1\in\mathfrak{R}_F$, then not only $g_1, \hat{G}_1\in\mathfrak{R}_F$ but also $P_N^{2d-1}\hat{G}_1\in\mathfrak{R}_F$. Hence by using (7.2) the above identity leads to

$$\{g_1, \bar{W}(T\zeta)\} + \hat{G}_1 = P_N^{2d-1}\hat{G}_1 + O(\zeta^{s_1+2d}).$$

This implies that $\bar{W}(T\zeta)$ satisfies the first equation of (3.5) ($k=1$). Recall that in the proof of Proposition 3.3, the solution W was determined uniquely so that it satisfies the first equation of (3.5) and the condition $P_N W = 0$. Then since $P_N \bar{W}(T\zeta) = 0$, it follows that $\bar{W}(T\zeta) = W(\zeta)$, i.e., $W \in \mathfrak{R}_F$. It follows from Lemma 7.2 that $\varphi \in \mathfrak{R}$. This completes the proof. □

From Lemma 7.2 and Proposition 7.3, $G_1 \circ \varphi(\zeta) \in \mathfrak{R}_F$ if $G_1 \in \mathfrak{R}_F$ under the assumption of Proposition 3.3. Hence we can see inductively that if $G_1(z) = G_1^{(0)}(z) \in \mathfrak{R}_F$, then $G_1^{(\nu+1)} = G_1^{(\nu)} \circ \varphi_\nu \in \mathfrak{R}_F$ for any $\nu = 0, 1, \ldots$. Moreover since $\mathfrak{R}$ forms a group, $\phi_\nu = \varphi_0 \circ \cdots \circ \varphi_\nu \in \mathfrak{R}$ and consequently the limit ϕ defined by (3.11) belongs to $\mathfrak{R}$.

Finally we note that the condition $w = (u, v) = Cz \in \mathbf{R}^{2n}$ is equivalent to the condition $\bar{z} = Tz$, which is rewritten as

$$\bar{\phi}(\bar{\zeta}) = T\phi(\zeta) = \bar{\phi}(T\zeta),$$

where we have the last equality since $\phi \in \mathfrak{R}$. Hence the condition $w \in \mathbf{R}^{2n}$ is equivalent to the condition $\bar{\zeta} = T\zeta$ which is expressed by using the form (7.6) of T as follows:

$$\begin{cases} \eta_k = i\bar{\xi}_k & \text{if } \lambda_k \text{ is purely imaginary;} \\ \xi_{l_k} = \bar{\xi}_k, \quad \eta_{l_k} = \bar{\eta}_k & \text{otherwise.} \end{cases} \tag{7.7}$$

If the transformation $z = \phi(\zeta)$ is convergent, one can find a real analytic canonical transformation which takes the original (real analytic) Hamiltonian into the real Birkhoff normal form. It is defined by choosing real variable in place of $\zeta = (\xi, \eta)$ satisfying the condition (7.7). As an example, let us consider the case when all the eigenvalues $\pm\lambda_1, \ldots, \pm\lambda_n$ are purely imaginary. Then if we carry out a linear canonical transformation $\zeta = M(q, p)$ defined by

$$\xi_k = \frac{1}{\sqrt{2}}(q_k + ip_k), \qquad \eta_k = \frac{1}{\sqrt{2}}(p_k + iq_k) \qquad (k = 1, \ldots, n), \tag{7.8}$$

it follows from (7.7) that $w = (u, v) \in \mathbf{R}^{2n}$ if and only if $(q, p) \in \mathbf{R}^{2n}$. Therefore the canonical transformation $(u, v) = C \circ \phi \circ M(q, p)$ is real analytic. Since we have $\xi_k \eta_k = (i/2)(q_k^2 + p_k^2)$, this transformation takes the G_k ($k = 1, \ldots, n$) into analytic functions of n variables $(q_l^2 + p_l^2)/2$ ($l = 1, \ldots, n$). This proves Theorem 1.4. The quadratic part of the normal form of the Hamiltonian G_1 is given by

$$\tfrac{1}{2} \sum_{k=1}^{n} \alpha_k (q_k^2 + p_k^2), \qquad \alpha_k = i\lambda_k.$$

8. Canonical mapping case

In this section, as an application of Theorem 3.6 we prove analogous results to Theorems 1.3 and 1.4 for canonical mappings near a fixed point. As before, we first consider analytic case and later treat real analytic case by imposing the reality condition on the original mapping.

Let us consider an analytic canonical mapping (symplectic diffeomorphism) defined in a neighbourhood of the origin in $\mathbf{C}^{2n}$. We denote it by $z^* = f(z)$ with $z = (z, y)$ and $z^* = (x^*, y^*)$ and assume that $f(0) = 0$. Then it is written in the form

$$z^* = Az + \cdots, \tag{8.1}$$

where A is a $2n \times 2n$ nondegenerate matrix and the part not written out explicitly is a vector of convergent power series containing terms of order ≥ 2 only. Since A is the Jacobian matrix df at the origin, A is symplectic and consequently its eigenvalues occur in pairs $\lambda_k, \lambda_k^{-1}$ $(k = 1, \ldots, n)$ (see [1]). The fixed point (the origin $z = 0$) is called *non-resonant* if these λ_k $(k = 1, \ldots, n)$ satisfy a condition

$$\prod_{k=1}^{n} \lambda_k^{m_k} \neq 1 \quad \text{for any} \quad (m_1, \ldots, m_n) \in \mathbf{Z}^n \backslash \{0\}. \tag{8.2}$$

An analytic function $G(z)$ is called an *integral* of f if $G(z)$ is invariant under f, i.e., the identity $G(f(z)) = G(z)$ holds. Our result is stated as follows:

THEOREM 8.1. *Let $z^* = f(z)$ be an analytic canonical mapping defined in a neighbourhood of the origin $z = 0 \in \mathbf{C}^{2n}$. Assume that the origin is a non-resonant fixed point of f and that the f possesses n analytic functionally independent integrals $G_k(z)$ $(k = 1, \ldots, n)$. Then there exists an analytic canonical transformation $z = \phi(\zeta)$ $(\zeta = (\xi, \eta))$ near the origin such that $\phi(0) = 0$ and the resulting mapping $\zeta^* = \phi^{-1} \circ f \circ \phi(\zeta)$ has the form*

$$\xi_k^* = \xi_k \exp(H_{\omega_k}), \qquad \eta_k^* = \eta_k \exp(-H_{\omega_k}) \qquad (k = 1, \ldots, n), \tag{8.3}$$

where H is an analytic function of n variables $\omega_l = \xi_l \eta_l$ $(l = 1, \ldots, n)$. The integrals $G_k \circ \phi(\zeta)$ also become analytic functions of ω_l $(l = 1, \ldots, n)$.

A canonical mapping of the form (8.3) is called *Birkhoff normal form*. We note that it is generated by the integrable Hamiltonian flow with Hamiltonian H which is in normal form in the sense of functions.

Since the eigenvalues $\lambda_k, \lambda_k^{-1}$ $(k=1, \ldots, n)$ are all distinct under the non-resonance condition (8.2), we can find a linear canonical transformation $z \mapsto Cz$ such that the resulting mapping $z^* = C^{-1} \circ f \circ C(z)$ has the form

$$x_k^* = \lambda_k x_k + \cdots, \qquad y_k^* = \lambda_k^{-1} y_k + \cdots \qquad (k=1, \ldots, n). \tag{8.4}$$

In the following, we assume that f is in the form (8.4) and prove the existence of a normalizing transformation $z = \phi(\zeta)$ of the form (1.3) such that $\phi^{-1} \circ f \circ \phi$ is in normal form (8.3). This transformation ϕ is called *Birkhoff transformation.*

We introduce the following definition.

DEFINITION 8.2. (i) A canonical mapping $\zeta^* = f(\zeta)$ is called (*Birkhoff*) *normal form of order* d (≥ 1) if it has the form (8.3) with a polynomial H of degree $d+1$ in ξ, η which is actually a polynomial of n variables $\omega_k = \xi_k \eta_k$ $(k=1, \ldots, n)$.

(ii) A canonical mapping $\zeta^* = f(\zeta)$ is said to be in *normal up to terms of order* d if it is expressed as

$$f = f_d \circ \psi, \tag{8.5}$$

where f_d is the Birkhoff normal form of order d and $\psi \in \mathfrak{S}_d$.

We note that the mapping (8.4) is in normal form up to terms of order 1 and is expressed in the form (8.5) with

$$f_1 : x_k \mapsto \lambda_k x_k, \qquad y_k \mapsto \lambda_k^{-1} y_k \qquad (k=1, \ldots, n). \tag{8.6}$$

Here this mapping f_1 is expressed in the form (8.3) with

$$H = \sum_{k=1}^{n} \alpha_k x_k y_k, \qquad e^{\alpha_k} = \lambda_k.$$

Therefore using the following proposition inductively, one can prove the existence of the desired Birkhoff transformation formally. Here ϕ is obtained in the form (3.11), where $\varphi_\nu = \varphi$ in the following proposition with $d = 2^\nu$.

PROPOSITION 8.3. *Let $z^* = f(z)$ be a canonical mapping which is in normal form up to terms of order d. Assume that the origin is a non-resonant fixed point of f. Then there exists a unique canonical transformation $z = \varphi(\zeta) \in \mathfrak{S}_d$ such that* (i) *the transformation $\varphi^{-1} \circ f \circ \varphi$ is in normal form up to terms of order 2d, and* (ii) *φ is generated by the Hamiltonian flow with Hamiltonian W of the form* (3.2) *satisfying a condition $P_N W = 0$.*

Proof. To prove the existence of the desired polynomial W, we first prove the existence of a homogeneous polynomial $W = W^{d+2}$ such that the canonical transformation generated by the Hamiltonian flow with this Hamiltonian W takes f into normal form up to terms of order $d+1$. For the proof, we use the similar arguments to [6].

According to the Definition 8.2, we express f in the form (8.5), i.e., $f = f_d \circ \psi$, $\psi \in \mathfrak{S}_d$. Let $z = \varphi(\zeta)$ be the canonical transformation generated by the Hamiltonian flow with Hamiltonian $W = W^{d+2}$. Since $\varphi \in \mathfrak{S}_d$, $f' = \varphi^{-1} \circ f \circ \varphi$ can be also expressed as $f' = f_d \circ \psi'$ with $\psi' \in \mathfrak{S}_d$. Therefore the relation $\varphi \circ f' = f \circ \varphi$ gives

$$\varphi \circ f_d \circ \psi' = f_d \circ \psi \circ \varphi. \tag{8.7}$$

Let $U = U^{d+2}$ and $K = K^{d+2}$ denote the lowest order parts (homogeneous polynomials of degree $d+2$) of the polynomial W in the expression (2.9) for ψ and ψ' respectively. Then comparison of terms of order $d+1$ in (8.7) gives

$$J\nabla W(\Lambda\zeta) + \Lambda J\nabla K(\zeta) = \Lambda J\nabla U(\zeta) + \Lambda J\nabla W(\zeta), \tag{8.8}$$

where $\zeta \mapsto \Lambda\zeta$ denotes the linear transformation (8.6), namely the Λ is a $2n \times 2n$ diagonal matrix with components $\lambda_1, \ldots, \lambda_n, \lambda_1^{-1}, \ldots, \lambda_n^{-1}$. If we set $V(\zeta) = W(\Lambda\zeta)$, we have $\nabla V(\zeta) = \Lambda \nabla W(\Lambda\zeta)$. Since $J\Lambda^{-1} = \Lambda J$, equation (8.8) is equivalent to $\nabla V + \nabla K = \nabla U + \nabla W$. By the homogeneity of V, K, U and W, this implies that

$$W(\Lambda\zeta) + K(\zeta) = U(\zeta) + W(\zeta). \tag{8.9}$$

Setting $W = \sum_{|\alpha|+|\beta|=d+2} c_{\alpha\beta}\xi^\alpha\eta^\beta$, one obtains

$$\sum_{|\alpha|+|\beta|=d+2} \gamma c_{\alpha\beta}\xi^\alpha\eta^\beta = U(\zeta) - K(\zeta); \qquad \gamma = \prod_{k=1}^{n} \lambda_k^{\alpha_k-\beta_k} - 1.$$

Here, due to the non-resonance condition (8.2), $\gamma = 0$ only if $\alpha = \beta$. Therefore by imposing the condition $P_N W = 0$, we can determine the coefficients $c_{\alpha\beta}$ uniquely so that $K = P_N U$, i.e., $K = K^{d+2}$ is a polynomial of n products $\omega_k = \xi_k\eta_k$ $(k = 1, \ldots, n)$. Consequently we can write the transformation ψ' as a composition $\psi' = \psi_{d+1} \circ \psi''$, where

$$\psi_{d+1} : \xi_k \mapsto \xi_k \exp(K_{\omega_k}), \qquad \eta_k \mapsto \eta_k \exp(-K_{\omega_k}) \qquad (k = 1, \ldots, n),$$

and $\psi'' \in \mathfrak{S}_{d+1}$. Noting that $\xi_k\eta_k$ are invariant under ψ_{d+1}, the mapping $f_d \circ \psi_{d+1}$

is the Birkhoff normal form of order $d+1$. Hence $f' = \varphi^{-1} \circ f \circ \varphi$ is in normal form up to terms of order $d+1$. If we express f_d in the form (8.4) with $H = H_{d+1}$, a polynomial of degree $d+1$ in normal form, then $f_d \circ \psi_{d+1}$ is expressed in the same form with $H = H_{d+1} + K = H_{d+1} + P_N U$.

Let φ^{d+1} denote the canonical transformation φ which has been determined just above. Here we note that the above discussion is valid for any d and therefore we can define canonical transformations $\varphi^{d+2}, \ldots, \varphi^{2d}$ successively. Let us consider the composition $\hat{\varphi} = \varphi^{d+1} \circ \varphi^{d+2} \circ \cdots \circ \varphi^{2d}$. Then the mapping $\hat{\varphi}^{-1} \circ f \circ \hat{\varphi}$ is in normal form up to terms of order $2d$. Notice that each φ^{j+1} $(j = d, \ldots, 2d-1)$ has the form (2.9) with $W = W^{j+2}$. Therefore the composition $\hat{\varphi}$ is written in the form (2.9) with (2.10). On the other hand, if we consider a canonical transformation generated by the Hamiltonian flow with Hamiltonian $W = W^{d+2} + \cdots + W^{2d+1}$, the power series of this transformation is equal to $\hat{\varphi}$ up to terms of order $2d$. Therefore it also transforms f into the normal form up to terms of order $2d$. From the above discussions the polynomial W is uniquely determined by the condition $P_N W = 0$. Hence it is the desired unique canonical transformation. This completes the proof. □

Proof of Theorem 8.1. The proof of Theorem 8.1 is done by reducing it to Theorem 3.6. Let G_k $(k = 1, \ldots, n)$ be n integrals of f and we denote these G_k in the form (2.3) again. By the same arguments as in Section 2, we can assume that their lowest order parts G_k^0, which are polynomials of degree s_k, are functionally independent. Since G_k are integrals of $z^* = f(z)$, the identity

$$G_k(f(z)) = G_k(z) \tag{8.10}$$

holds. Here we can assume that f is in the form (8.4), and therefore comparison of the lowest order terms of this identity gives

$$G_k^0(\Lambda z) = G_k^0(z).$$

Due to the non-resonance condition (8.2), the G_k^0 are reduced to polynomials of n variables $x_k y_k$ $(k = 1, \ldots, n)$. Therefore they satisfy the condition (2.5). We note the following fact.

LEMMA 8.4. *Let $G_k = G_k(z)$ be integrals of a canonical transformation f. If f is in normal form up to terms of order d, then G_k is in normal form up to terms of order $s_k + d - 1$.*

Proof. By the transformation f which is in normal form up to terms of order d, $x_k y_k$ is transformed into the form $x_k y_k + O(z^{d+2})$. Assume that the integral G_k

is in normal form up to terms of order $s_k + j$ $(0 \le j \le d-2)$. Then the comparison of terms of order $s_k + j + 1$ in (8.10) gives

$$G_k^{j+1}(\Lambda z) = G_k^{j+1}(z).$$

It follows from the non-resonance condition that G_k^{j+1} is in normal form. Hence we complete the proof by induction. $\square$

Suppose that $G_k(z)$ $(k = 1, \ldots, n)$ are integrals of the canonical mapping $z^* = f(z)$ which is in normal form up to terms of order d. Let $z = \varphi(\zeta)$ be the canonical transformation described in Proposition 8.3. Then, since $G_k \circ \varphi$ are integrals of $\varphi^{-1} \circ f \circ \varphi$, one can see from Proposition 8.3 and Lemma 8.4 that the transformation φ takes G_k which is in normal form up to terms of order $s_k + d - 1$ into the normal form up to terms of order $s_k + 2d - 1$. If we set $\varphi = \varphi_\nu$ for $d = 2^\nu$. it satisfies the assumptions of Theorem 3.6. Consequently Theorem 8.1 follows from Theorem 3.6. $\square$

Next we consider the case when the original mapping (8.1) is real analytic. We consider the original mapping (8.1) in $w = (u, v)$ variables and suppose that the mapping (8.4), which we denote by f, is obtained by subjecting a linear canonical transformation $w = Cz$ to (8.1). Then the original mapping is given by $w^* = C \circ f \circ C^{-1}(w)$ and therefore its reality means $C \circ f \circ C^{-1} = \bar{C} \circ \bar{f} \circ \bar{C}^{-1}$, which is equivalent to the condition $f \in \mathfrak{R}$ (see Section 7).

The symplectic matrix C diagonalizes the linearized mapping of (8.1) in such a way that $C^{-1}AC = \operatorname{diag}(\lambda_1, \ldots, \lambda_n, \lambda_1^{-1}, \ldots, \lambda_n^{-1})$. We can choose the matrix C so that (i) $\bar{\lambda}_k \in \{\lambda_1, \ldots, \lambda_n\}$ unless $|\lambda_k| = 1$, and (ii) the transformation $z^* = Tz$ is given by

$$\begin{cases} x_k^* = -iy_k, & y_k^* = -ix_k \quad \text{if } |\lambda_k| = 1; \\ x_k^* = x_{l_k}, & y_k^* = y_{l_k} \quad \text{otherwise.} \end{cases} \tag{8.11}$$

where l_k is the number $(1 \le l_k \le n)$ such that $\lambda_{l_k} = \bar{\lambda}_k$.

We consider one iteration step described in Proposition 8.3. Let $z^* = f(z)$ be a canonical mapping satisfying the assumption of Proposition 8.3 and let $z = \varphi(\zeta) \in \mathfrak{S}_d$ be the canonical transformation described there. Then, since f and $\varphi^{-1} \circ f \circ \varphi$ are in normal form up to terms of order d and $2d$ respectively, we can write them in the form

$$f = f_d \circ \psi, \qquad \varphi^{-1} \circ f \circ \varphi = f_{2d} \circ \psi', \tag{8.12}$$

where f_j ($j = d, 2d$) is the Birkhoff normal form of order j and $\psi \in \mathfrak{S}_d$, $\psi' \in \mathfrak{S}_{2d}$. Suppose that f_j is expressed in the form (8.3) with $H = H_{j+1}$, where H_{j+1} ($j = d, 2d$) is a polynomial of degree $\leq j + 1$, which is in normal form in the sense of functions. Then we have

PROPOSITION 8.5. *Assume that $f \in \mathfrak{R}$ and $H_{d+1} \in \mathfrak{R}_F$ under the above assumptions and notations. Then $\varphi^{-1} \circ f \circ \varphi \in \mathfrak{R}$ and $H_{2d+1} \in \mathfrak{R}_F$.*

Proof. In the following, we use the same notations as in the proof of Proposition 8.3. Notice that $f_d \in \mathfrak{R}$ under the assumption. Therefore it follows that $\psi \in \mathfrak{R}$ with the notation (8.12). The $\psi \in \mathfrak{S}_d$ is written in the form

$$\psi(\zeta) = \zeta + J\nabla U(\zeta) + O(\zeta^{2d+1}); \qquad U = U^{d+2} + \cdots + U^{2d+1},$$

where U^j is a homogeneous polynomial of degree j in ζ. One can easily see that the condition $\psi \in \mathfrak{R}$ implies $U \in \mathfrak{R}_F$ by comparison of both sides of the relation $\bar{\psi}(T\zeta) = T\psi(\zeta)$. Recall that in the proof of Proposition 8.3, we have solved equation (8.9) for $W = W^{d+2}$ so that $K = P_N U$ with $U = U^{d+2}$, i.e.,

$$W(\Lambda\zeta) - W(\zeta) = P_R U(\zeta).$$

Since $U = U^{d+2} \in \mathfrak{R}_F$ and the linear transformation T is given by (8.11), we conclude $W(\zeta) \in \mathfrak{R}_F$ in the same way as in the proof of Proposition 7.3. From Lemma 7.2, this implies that $\varphi^{d+1} \in \mathfrak{R}$ and hence $(\varphi^{d+1})^{-1} \circ f \circ \varphi^{d+1} \in \mathfrak{R}$. Moreover we have $H_{d+2} := H_{d+1} + P_N U \in \mathfrak{R}_F$.

By repeating the above discussions, we can prove that $W = W^{d+2} + \cdots + W^{2d+1} \in \mathfrak{R}_F$ and $H_{2d+1} \in \mathfrak{R}_F$. Since the transformation φ in Proposition 8.3 is generated by Hamiltonian flow with Hamiltonian W, it follows from Lemma 7.2 that $\varphi \in \mathfrak{R}$. This completes the proof. $\square$

Notice that the mapping f given by (8.4) belongs to $\mathfrak{R}$ and $H_2 = \sum_{k=1}^{n} \alpha_k x_k y_k \in \mathfrak{R}_F$. Therefore by using the above proposition, we can see that if $f \in \mathfrak{R}$, the ϕ in Theorem 8.1 belongs to $\mathfrak{R}$ and the function H in (8.3) belongs to $\mathfrak{R}_F$. By the same way as in Section 7, we can prove that the condition $w \in \mathbf{R}^{2n}$ is equivalent to $\bar{\zeta} = T\zeta$ which is written in the form

$$\begin{cases} \eta_k = i\xi_k & \text{if } |\lambda_k| = 1; \\ \xi_{l_k} = \bar{\xi}_k, \quad \eta_{l_k} = \bar{\eta}_k & \text{otherwise.} \end{cases} \tag{8.13}$$

By choosing the real variable in place of $\zeta = (\xi, \eta)$ satisfying (8.13), we can find a real analytic canonical transformation which takes the original real analytic mapping into the real normal form. In particular, the fixed point (the origin) is called *elliptic* if $|\lambda_k| = 1$ for all $k = 1, \ldots, n$. In this case, if we carry out a linear canonical transformation (7.8), then we have $\omega_k = \xi_k \eta_k = (i/2)(q_k^2 + p_k^2)$. Therefore $H_{\omega_k} = -iH_{\tau_k}$ with $\tau_k = (q_k^2 + p_k^2)/2$. Hence we obtain the following result, where we use (ξ, η) in place of (q, p).

THEOREM 8.6. *Let $z^* = f(z)$ be a real analytic canonical mapping defined in a neighbourhood of the origin $z = 0 \in \mathbf{R}^{2n}$. Assume that the origin is a non-resonant elliptic fixed point of f and that the f possesses n analytic functionally independent integrals $G_k(z)$ $(k = 1, \ldots, n)$. Then there exists a real analytic canonical transformation $z = \phi(\zeta)$ $(\zeta = (\xi, \eta))$ near the origin such that $\phi(0) = 0$ and the resulting mapping $\zeta^* = \phi^{-1} \circ f \circ \phi(\zeta)$ has the form*

$$\xi_k^* = \xi_k \cos(H_{\tau_k}) + \eta_k \sin(H_{\tau_k}),$$
$$\eta_k^* = -\xi_k \sin(H_{\tau_k}) + \eta_k \cos(H_{\tau_k}) \qquad (k = 1, \ldots, n),$$

where H is a real analytic function of n variables $\tau_l = (\xi_l^2 + \eta_l^2)/2$ $(l = 1, \ldots, n)$. The integrals $G_k \circ \phi(\zeta)$ also become analytic functions of τ_l $(l = 1, \ldots, n)$.

9. Appendix

In this appendix, we give a proof of Lemma 2.1 following Ziglin [13].

First we present a fact obtained by using the knowledge of the field theory. Let $k[Z] = k[z_1, \ldots, z_m]$ be a polynomial ring over $k = \mathbf{C}$ or $\mathbf{R}$. For the proof of Lemma 2.1, the following fact plays a key role.

LEMMA 9.1. *If $f_1, \ldots, f_r \in k[Z]$ are algebraic independent, then they are functionally independent.*

The converse is obvious, and therefore this lemma implies the equivalence of algebraic independence and functional independence for polynomials.

Proof. In the following, for any $\alpha_1, \ldots, \alpha_j$ which are elements of some extension field of k, $k(\alpha_1, \ldots, \alpha_j)$ denotes a minimal field containing $\alpha_1, \ldots, \alpha_j$ and k. In particular, $K = k(z_1, \ldots, z_m)$ is the field of rational functions of the variables $z_1, \ldots, z_m$ over k. Since $k = \mathbf{C}$ or $\mathbf{R}$, we note that any extension of k is of characteristic 0. Under the assumption, $F = k(f_1, \ldots, f_r)$ is a separably generated and finitely generated extension of k of transcendence degree r. Here the separability follows from the fact that F is of characteristic 0. Then by a result on derivations over a field [4, Chap. X Theorem 10], there exist r derivations D_k $(k = 1, \ldots, r)$ on F such that

$$D_k f_j = \delta_{kj}. \tag{9.1}$$

Since K is also a separably generated and finitely generated extension of F (of transcendence degree $m - r$), there exist r derivations D_k^* $(k = 1, \ldots, r)$ on K such that $D_k^* = D_k$ on F. We note that the set of all derivations on K is a m-dimensional vector space over K which is spanned by $\partial/\partial z_1, \ldots, \partial/\partial z_m$. Therefore we can write D_k^* as

$$D_k^* = \sum_{l=1}^{m} p_{kl}(z) \frac{\partial}{\partial z_l}, \qquad p_{kl}(z) \in K$$

and hence

$$D_k^* f_j = \sum_{l=1}^{m} p_{kl}(z) \frac{\partial f_j}{\partial z_l}. \tag{9.2}$$

It follows from (9.1) and (9.2) that the rank of $r \times m$ Jacobian matrix $(\partial f_j/\partial z_l)$ is r. This proves the functional independence of $f_1, \ldots, f_r$. □

Proof of Lemma 2.1. Since $f_1, \ldots, f_r$ are functionally independent, we can find a non-vanishing (i.e., $\not\equiv 0$) minor $M = \det(\partial(f_1, \ldots, f_r)/\partial(z_{i_1}, \ldots, z_{i_r}))$ in the Jacobian matrix $\partial(f_1, \ldots, f_r)/\partial(z_1, \ldots, z_m)$. We set

$$\mu(f_1, \ldots, f_r) = d(M) + r - \sum_{k=1}^{r} d(f_k),$$

where and in what follows $d(f)$ denotes the degree of the lowest order part f^0 of a power series f. Notice that $M_0 := \det(\partial(f_1^0, \ldots, f_r^0)/\partial(z_{i_1}, \ldots, z_{i_r}))$ is either a homogeneous polynomial ($\not\equiv 0$) of degree $\sum_{k=1}^{r} d(f_k) - r$ or equal to 0 identically. Therefore $\mu(f_1, \ldots, f_r) \geq 0$ and $M_0 \not\equiv 0$ (which implies the functional independence of $f_1^0, \ldots, f_r^0$) if and only if $\mu(f_1, \ldots, f_r) = 0$.

Now assume that $f_1^0, \ldots, f_{r-1}^0$ are functionally independent and $f_1^0, \ldots, f_r^0$ are functionally dependent. Then by Lemma 9.1 $f_1^0, \ldots, f_r^0$ are also algebraic dependent, i.e., there exists a polynomial $P \in k[z_1, \ldots, z_r]$ $(k \in \mathbf{C}$ or $\mathbf{R})$ such that

$$P(f_1^0, \ldots, f_r^0) \equiv 0. \tag{9.3}$$

We choose the P such that its degree is minimal among the polynomials satisfying this identity. We set

$$F=P(f_1,\ldots,f_r)$$

and define $\hat{M}=\det(\partial(f_1,\ldots,f_{r-1},F)/\partial(z_{i_1},\ldots,z_{i_r}))$. Then we have

$$\hat{M}=\det\left(\frac{\partial(f_1,\ldots,f_{r-1},F)}{\partial(f_1,\ldots,f_{r-1},f_r)}\frac{\partial(f_1,\ldots,f_r)}{\partial(z_{i_1},\ldots,z_{i_r})}\right)=M\frac{\partial P}{\partial z_r}(f_1,\ldots,f_r).$$

Here if $\partial P/\partial z_r(z_1,\ldots,z_r)\equiv 0$, then the identity (9.3) implies the functional dependence of $f_1^0,\ldots,f_{r-1}^0$, which contradicts the assumption. Hence $\partial P/\partial z_r\not\equiv 0$ and it follows from the minimality of the degree of P that $(\partial P/\partial z_r)(f_1^0,\ldots,f_r^0)\not\equiv 0$. Therefore $\hat{M}\not\equiv 0$, that is, $f_1,\ldots,f_{r-1}$, F are functionally independent. Let $P(z)=\sum_\alpha c_\alpha z^\alpha$, $z^\alpha=z_1^{\alpha_1}\cdots z_r^{\alpha_r}$, and define a positive number ν as follows:

$$\nu:=\min\left\{\sum_{k=1}^r \alpha_k d(f_k)\mid c_\alpha\neq 0,\ \alpha_r\neq 0\right\}.$$

Then it follows that

$$d(\hat{M})=d(M)+d\left(\frac{\partial P}{\partial z_r}(f_1,\ldots,f_r)\right)=d(M)+\nu-d(f_r),$$

and that $\nu<d(F)$. Moreover the quantity $\mu(f_1,\ldots,f_{r-1},F)$ is calculated as follows:

$$\begin{aligned}\mu(f_1,\ldots,f_{r-1},F)&=d(\hat{M})+r-\sum_{k=1}^{r-1}d(f_k)-d(F)\\&=d(M)+\nu-d(f_r)+r-\sum_{k=1}^{r-1}d(f_k)-d(F)\\&=\mu(f_1,\ldots,f_r)+\nu-d(F)\\&<\mu(f_1,\ldots,f_r).\end{aligned}$$

Repeating this process, one can find a function $\hat{f}_r$ such that $f_1,\ldots,f_{r-1}$, $\hat{f}_r$ are functionally independent and $\mu(f_1,\ldots,f_{r-1},\hat{f}_r)=0$. This implies the functional independence of $f_1^0,\ldots,f_{r-1}^0,\hat{f}_r^0$. The last assertion of Lemma 2.1 can be easily seen. This completes the proof. □

REFERENCES

[1] V. I. ARNOLD, Mathematical Methods in Classical Mechanics, Springer (1978).
[2] G. D. BIRKHOFF, Dynamical Systems, AMS Colloquium Publ. IX (1927).
[3] H. ELIASSON, Hamiltonian systems with Poisson commuting integrals, Thesis at Univ. of Stockholm (1984).
[4] S. LANG, Algebra, Addison-Wesley (1971).
[5] J. MOSER, *Stable and random motions in dynamical systems,* Ann. Math. Studies *77* Princeton Univ. Press (1973).
[6] J. MOSER, Proof of a generalized form of a fixed point theorem due to G. D. Birkhoff, Lec. Notes in Math., Springer, *597* (1977), 464–494.
[7] J. MOSER, *Analytic surfaces in* $\mathbf{C}^2$ *and their local hull of holomorphy,* Annales Academiae Scientiarum Fenniae Ser. A. I., *10* (1985), 397–410.
[8] H. RÜSSMANN, *Über das Verhalten analytischer Hamiltonscher Differentialgleichungen in der Nähe einer Gleichgewichtslösung,* Math. Ann., *154* (1964), 285–300.
[9] C. L. SIEGEL, *On integrals of canonical systems,* Ann. Math., *42* (1941), 806–822.
[10] C. L. SIEGEL, *Über die Existenz einer Normalform analytischer Hamiltonscher Differentialgleichungen in der Nähe einer Gleichgewichslösung,* Math. Ann., *128* (1954), 144–170.
[11] C. L. SIEGEL and J. K. MOSER, Lectures on Celestial Mechanics, Springer (1971).
[12] J. VEY, *Sur certain systèmes dynamiques séparables,* Amer. J. Math., *100* (1978), 591–614.
[13] S. L. ZIGLIN, *Branching of solutions and nonexistence of first integrals in Hamiltonian mechanics I,* Functional. Anal. Appl. *16* (1983), 181–189.

Department of Mathematics
Tokyo Institute of Technology
Oh-okayama, Meguro-ku
Tokyo 152, Japan

Received November 30, 1987