Tôhoku Math. J.
49 (1997), 73–114

INTEGRABLE SYMPLECTIC MAPS AND THEIR BIRKHOFF NORMAL FORMS

Dedicated to Professor Junji Kato on his sixtieth birthday

HIDEKAZU ITO

(Received September 25, 1995, revised February 7, 1996)

Abstract. For a symplectic map that is analytic near a fixed point, the Birkhoff normal form is studied in connection with its integrability in the sense of Liouville. It is proved that, when the fixed point is non-resonant or simply resonant, there exists an analytic Birkhoff transformation if and only if the map is integrable.

1. Introduction. This paper is devoted to the study of normal forms for symplectic maps near a fixed point. We consider local real analytic diffeomorphisms having the origin $z=0$ as a fixed point. We assume that they are symplectic, that is, they preserve the standard symplectic structure $\omega_0=\sum_{k=1}^{n} dx_k \wedge dy_k$, where $z=(x_1, \dots, x_n, y_1, \dots, y_n)$ is the coordinate system of $\boldsymbol{R}^{2n}$. We denote by $\mathrm{Symp}(\boldsymbol{R}^{2n}, 0)$ the set of germs of all those local real analytic symplectic diffeomorphisms near the fixed point $z=0$.

The normal form we are concerned with is the so-called *Birkhoff normal form* and the normalizing transformation is called the *Birkhoff transformation*, named after G. D. Birkhoff, who first developed this theory (cf. [2], [3]). Let $f\in \mathrm{Symp}(\boldsymbol{R}^{2n}, 0)$ be a map with semisimple linear part $Df(0)$. Then, according to Moser [10], the map f is in the Birkhoff normal form if it commutes with its linear part $Df(0)$ in the real symplectic Jordan canonical form. If we ignore convergence, there always exists a Birkhoff transformation and f is in the Birkhoff normal form if and only if the nonlinear part $Df(0)^{-1}\circ f$ is equal to the time-one map of the Hamiltonian flow with a time-independent Hamiltonian H invariant under $Df(0)$. This implies that the function H is a formal power series integral of f. Therefore, when the fixed point is non-resonant or simply resonant (see Section 2 for definitions), one can easily see that the given map is integrable in the sense of Liouville [1] if there exists a convergent Birkhoff transformation together with a convergent interpolating Hamiltonian H.

The aim of this paper is to prove the converse of this assertion. We have already proved it in the non-resonance case [7]. In this paper we generalize it in the simple resonance cases. It corresponds to the result for normalization of Hamiltonian functions [8] and its proof has the same feature as that in [8]. However the result is not a direct

This research was supported by a grant from the Inamori Foundation.
1991 *Mathematics Subject Classification*. Primary 58F07.

consequence of [8], and in particular, we have a new problem to the effect that the convergence of the interpolating Hamiltonian H does not follow from the convergence of the Birkhoff transformation.

To overcome this difficulty, we interpolate the given map by the time-one map of the Hamiltonian flow with a generally *periodically time-dependent* Hamiltonian that is real analytic in both $z \in \boldsymbol{R}^{2n}$ and time $t \in \boldsymbol{R}$. Due to a recent result by Kuksin and Pöschel [9], this is possible under the condition that the linear part $Df(0)$ can be written as the time-one map of the Hamiltonian flow with a quadratic time-independent Hamiltonian. We see that this condition is satisfied under fairly general assumptions on those eigenvalues of $Df(0)$ which can be easily checked (Proposition 1). Moreover, we say that a map $f \in \mathrm{Symp}(\boldsymbol{R}^{2n}, 0)$ is in the Birkhoff normal form if it is the time-one map of the Hamiltonian flow with generally periodically (of period 1) time-dependent Hamiltonian in the Birkhoff normal form (as time-dependent functions), and we call a normalizing transformation a Birkhoff transformation. Then it turns out that if f is in the Birkhoff normal form in this sense, it is also in the Birkhoff normal form in Moser's sense and the nonlinear part $Df(0)^{-1} \circ f$ becomes the time-one map of the Hamiltonian flow with some time-independent real analytic Hamiltonian (Proposition 2). In particular, in the non-resonance case, the map f itself becomes the time-one map of a time-independent Hamiltonian flow. Our main result (Theorem 1) claims that, in the simple resonance case as well as in the non-resonance case, there exists a convergent Birkhoff transformation and a convergent interpolating Hamiltonian if and only if f is integrable. In the next section, we formulate it under more general setting where we allow cases in which the linear part $Df(0)$ is not semisimple, which has been studied in [4] under Moser's definition.

To prove Theorem 1, the crucial observation is that the integrability of the given map implies that of the interpolating time-dependent Hamiltonian system (Theorem 3) and hence the proof of Theorem 1 is reduced to that of the corresponding theorem for time-dependent Hamiltonian systems near an equilibrium (Theorem 2). We proved in [8] the corresponding theorem in the case of time-independent systems. Theorem 1 and Theorem 2 are analogs of this result for symplectic maps and for time-dependent Hamiltonian systems.

This paper is organized as follows. In the next section, we will give necessary definitions and state the results mentioned above. In Section 3, we will show that the proof of Theorem 1 is reduced to that of Theorem 2. For this purpose, we will prove an interpolation theorem mentioned above (Theorem 3). To prove Theorem 2, we will work with the complex Birkhoff normal form instead of the real one and also will extend the phase space so that the time-dependent Hamiltonian vector field can be considered as a time-independent one in the extended phase space of dimension $2n+2$. This extension of the phase space will be convenient to deal with the integrability of the given vector field. In Section 4, we will introduce this extension and will reformulate Theorem 2 as Theorem 4.4 concerning the normalization in the extended phase space.

Theorem 4.4 will be proved in Sections 5–10. In Section 5, we will give a power series expression for the normal form. By using it, we will prove Proposition 2 and show that the original symplectic map is integrable if there exists a convergent Birkhoff transformation in the non-resonance or simple resonance cases. In Section 6, we will prove the existence of a formal normalizing transformation (Theorem 6.3 and Corollary 6.4). In Section 7, we will describe the basic idea on the proof of its convergence and the proof will be carried out from Section 8 through Section 10. Finally, we will prove Theorem 2 in Section 11 by imposing reality condition on the original Hamiltonian.

2. Statement of the results. In this section, we state the main results. We first introduce symbols to be used throughout this paper: $\mathscr{A}(\boldsymbol{R}^{2n}, 0)$ denotes the set of germs of all functions of $2n$ variables, say $z \in \boldsymbol{C}^{2n}$, that are real analytic in a neighbourhood of the origin $z=0$. $\mathscr{A}(\boldsymbol{R}^{2n+1}, S^1)$ denotes the set of germs of all functions in $(z, t) \in \boldsymbol{C}^{2n} \times \boldsymbol{C}$ that are real analytic in a neighbourhood of the real t-axis $\{(z, t) \in \boldsymbol{C}^{2n+1} \mid z=0, t \in \boldsymbol{R}\}$ and are periodic in t with period 1. In addition to $\mathrm{Symp}(\boldsymbol{R}^{2n}, 0)$ introduced in Section 1, $\mathrm{Symp}(\boldsymbol{R}^{2n+1}, S^1)$ denotes the set of germs of all local real analytic symplectic diffeomorphisms $(\boldsymbol{R}^{2n}, 0) \to (\boldsymbol{R}^{2n}, 0)$ that depend on the parameter t (time) real analytically and periodically with period 1.

In the above, "real analytic" means that those analytic functions and maps are real-valued if the domains of definition are restricted to $\boldsymbol{R}^{2n}$ or $\boldsymbol{R}^{2n+1}$. For the sets of germs of analytic functions or maps without real-analyticity assumption, we denote them by the same symbols with $\boldsymbol{R}^{2n}$, $\boldsymbol{R}^{2n+1}$ replaced by $\boldsymbol{C}^{2n}$, $\boldsymbol{C}^{2n+1}$ respectively, that is, $\mathscr{A}(\boldsymbol{C}^{2n}, 0)$, and so on. We consider them as $\mathscr{A}(\boldsymbol{R}^{2n}, 0) \subset \mathscr{A}(\boldsymbol{C}^{2n}, 0)$, and so on.

Let X_H denote the Hamiltonian vector field with Hamiltonian $H \in \mathscr{A}(\boldsymbol{C}^{2n}, 0)$ or $H \in \mathscr{A}(\boldsymbol{C}^{2n+1}, S^1)$, which is given by the system of differential equations

$$\frac{dz}{dt} = JH_z; \quad J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}.$$

Here H_z is the vector of first derivatives of H with respect to z and I is the identity matrix of degree n. Also $\exp X_H$ denotes the time-one map of its flow, i.e., $\exp X_H : z(0) \mapsto z(1)$, where $z(t)$ is the solution of the above system. The Poisson bracket of two functions F, G is defined by

$$\{F, G\} := \langle F_z, JG_z \rangle \left(= \sum_{k=1}^{n} (F_{x_k} G_{y_k} - F_{y_k} G_{x_k}) \right),$$

where $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product, i.e., $\langle z, z' \rangle = \sum_{k=1}^{2n} z_k z'_k$ with z_k, z'_k coordinates of z, z' respectively.

Now let us introduce necessary definitions to state the results. First, the integrability in the sense of Liouville is defined as follows:

DEFINITION 1. (i) A symplectic map $f \in \mathrm{Symp}(\boldsymbol{C}^{2n}, 0)$ is said to be *integrable* if it possesses n analytic integrals $G_1(z), \dots, G_n(z) \in \mathscr{A}(\boldsymbol{C}^{2n}, 0)$ that are functionally independent and Poisson commuting.

(ii) A Hamiltonian vector field X_H with $H \in \mathscr{A}(\boldsymbol{C}^{2n+1}, S^1)$ is said to be *integrable* if it possesses n analytic integrals $G_1(z, t), \dots, G_n(z, t) \in \mathscr{A}(\boldsymbol{C}^{2n+1}, S^1)$ such that for any t fixed they are functionally independent and Poisson commuting functions of z.

REMARK. Since $\mathrm{Symp}(\boldsymbol{R}^{2n}, 0) \subset \mathrm{Symp}(\boldsymbol{C}^{2n}, 0)$ and $\mathscr{A}(\boldsymbol{R}^{2n+1}, S^1) \subset \mathscr{A}(\boldsymbol{C}^{2n+1}, S^1)$, this definition is valid also when $f \in \mathrm{Symp}(\boldsymbol{R}^{2n}, 0)$ or $H \in \mathscr{A}(\boldsymbol{R}^{2n+1}, S^1)$. In other words, we allow their integrals to be complex-valued for real variables.

In the definition (i) above, a function $G(z) \in \mathscr{A}(\boldsymbol{C}^{2n}, 0)$ is called an *integral of f* if it is invariant under f, i.e., $G \circ f = G$. The functional independence of $G_1(z), \dots, G_n(z)$ means that their gradient vectors with respect to z are linearly independent on an open and dense subset of a neighbourhood of the origin $z=0$, and those n functions are said to be *Poisson commuting* if $\{G_i, G_j\} \equiv 0$ for any $i, j = 1, \dots, n$. The meaning of the same terminology used in (ii) is obvious.

To define Birkhoff normal forms for maps, let us give preliminary considerations about interpolation of symplectic maps. Let $Df(0)$ denote the linear part of $f \in \mathrm{Symp}(\boldsymbol{R}^{2n}, 0)$ at the origin. It is symplectic and hence its eigenvalues occur in pairs $\lambda_k, \lambda_k^{-1}$ $(k=1, \dots, n)$ (cf. [1]). We assume that $Df(0) = \exp(JA)$ for some real symmetric matrix A, in other words, the linear map $z \mapsto Df(0)z$ is the time-one map

$$Df(0)z = \exp X_{H_2}(z) \qquad \text{with} \quad H_2(z) = \frac{1}{2} \langle Az, z \rangle . \tag{2.1}$$

Then the theorem due to Kuksin and Pöschel mentioned in Section 1 claims that $f = \exp X_H$ with some time-dependent Hamiltonian $H \in \mathscr{A}(\boldsymbol{R}^{2n+1}, S^1)$ of the form

$$H = H(z, t) = H_2(z) + \hat{H}(z, t) ; \qquad \hat{H}(z, t) = O(|z|^3) , \tag{2.2}$$

where $O(|z|^3) \in \mathscr{A}(\boldsymbol{R}^{2n+1}, S^1)$ denotes the terms of order ≥ 3 with respect to z.

Since the set of all quadratic forms on $\boldsymbol{R}^{2n}$ forms a semi-simple Lie algebra under the Poisson bracket $\{\cdot, \cdot\}$ (see [6], [11]), we have the Jordan decomposition

$$H_2 = S + N ; \quad \{S, N\} = 0 , \tag{2.3}$$

where S, N are quadratic forms such that the matrices JS_{zz} and JN_{zz} are semisimple and nilpotent, respectively. After a real linear symplectic coordinate transformation, we may assume that the quadratic form H_2 is in the normal form, which implies that the linear vector field X_{H_2} is taken into the real Jordan canonical form. In the semisimple (i.e., $N=0$) case, the normal form is given as follows:

$$H_2(z) = S(z) = \sum_{j=1}^{k} a_j p_j q_j + \sum_{j=1}^{l} \frac{b_j}{2} (u_j^2 + v_j^2) \tag{2.4}$$

$$+\sum_{j=1}^{m}\{c_j(x_{2j-1}y_{2j-1}+x_{2j}y_{2j})+d_j(x_{2j-1}y_{2j}-x_{2j}y_{2j-1})\},$$

where

$$z=(z_1,\dots,z_{2n})\,;\quad \begin{cases} z_j=p_j,\ z_{n+j}=q_j & (j=1,\dots,k)\,,\\ z_{k+j}=u_j,\ z_{n+k+j}=v_j & (j=1,\dots,l)\,,\\ z_{k+l+j}=x_j,\ z_{n+k+l+j}=y_j & (j=1,\dots,2m)\,. \end{cases}$$

In this case, the eigenvalues of the linear vector field X_{H_2} are $\pm\mu_j$ $(j=1,\dots,n)$ with

$$\begin{cases} \mu_j=a_j \quad (j=1,\dots,k)\,, & \mu_{k+j}=b_j i \quad (j=1,\dots,l)\,,\\ \mu_{k+l+2j-1}=c_j+d_j i\,, & \mu_{k+l+2j}=c_j-d_j i \quad (j=1,\dots,m)\,, \end{cases} \tag{2.5}$$

where

$$\begin{cases} a_j,\,b_j,\,c_j,\,d_j\in \boldsymbol{R}\,; \quad b_j\neq 0,\ c_j\neq 0,\ d_j\neq 0\,, \quad i=\sqrt{-1}\,,\\ k,\,l,\,m \text{ are nonnegative integers satisfying } k+l+2m=n\,. \end{cases}$$

For the original map $f\in\mathrm{Symp}(\boldsymbol{R}^{2n},0)$, this implies that, if $Df(0)$ is written in the form (2.1), negative eigenvalues of $Df(0)$ can only occur with even multiplicities (when $d_j=(2r+1)\pi$ for some j and $r\in\boldsymbol{Z}$). Also we note that $\lambda_j=1$ corresponds to $a_j=0$.

When can the linear part of the original map f be written in the form (2.1)? We have the following result about this question.

PROPOSITION 1. *The linear part of* $f\in\mathrm{Symp}(\boldsymbol{R}^{2n},0)$ *can be written in the form* (2.1) *if one of the following three conditions holds*:

(i) $Df(0)$ *has no negative eigenvalues*;

(ii) $Df(0)$ *is the square of another symplectic matrix on* $\boldsymbol{R}^{2n}$;

(iii) $Df(0)$ *is semisimple and does not have negative eigenvalues of odd multiplicities.*

For the proof of this proposition under conditions (i) or (ii), we refer to [9, Appendix]. Suppose that $f\in\mathrm{Symp}(\boldsymbol{R}^{2n},0)$ satisfies the condition (iii). Then the eigenvalues λ_j, λ_j^{-1} of $Df(0)$ are given by

$$\lambda_j=e^{\mu_j}\quad (j=1,\dots,n)\qquad \text{for}\quad \mu_j \text{ given by (2.5)}\,.$$

We note that the real Jordan canonical form of $Df(0)$ is equal to $\exp X_{H_2}$ with H_2 given by (2.4). Let C be a real symplectic matrix which takes $Df(0)$ into $\exp X_{H_2}$. Then we have

$$Df(0)=C\circ\exp X_{H_2}\circ C^{-1}=\exp X_{H_2\circ C^{-1}}\,,$$

which proves Proposition 1 under the condition (iii).

We omit writing down all normal forms, which was studied by Williamson [12]. See [1, Appendix] or [5] for the complete list. The semisimple part of the normal form is different from (2.3) when H_2 is not semisimple. On the other hand, if we consider the complex Jordan canonical form instead of the real one, the corresponding normal form of H_2 reduces to simpler forms. For example, the normal form of the semisimple part of H_2 is given by

$$S(z)=\sum_{k=1}^{n}\mu_k x_k y_k \qquad (\mu_k=\log\lambda_k) \tag{2.6}$$

for the non-semisimple case as well as for the semisimple case. Correspondingly, the vector field X_S and its time-one map $\exp X_S$ are in diagonal form.

Associated with the semisimple part S of the quadratic form H_2 in the normal form, we define the Birkhoff normal forms for functions and maps as follows:

DEFINITION 2. (i) A function $H\in\mathscr{A}(\boldsymbol{R}^{2n+1}, S^1)$ of the form (2.2) with (2.3) is said to be in the *Birkhoff normal form* if H_2 is in the normal form and if H is invariant under the flow of X_S, that is,

$$\{H, S\}+\frac{\partial H}{\partial t}\equiv 0\,. \tag{2.7}$$

More generally, an arbitrary function $G\in\mathscr{A}(\boldsymbol{R}^{2n+1}, S^1)$ is said to be in the *Birkhoff normal form* if the identity (2.7) holds with H replaced by G.

(ii) A map $f\in\mathrm{Symp}(\boldsymbol{R}^{2n}, 0)$ satisfying (2.1) with H_2 in the normal form is said to be in the *Birkhoff normal form* if $f=\exp X_H$ with $H\in\mathscr{A}(\boldsymbol{R}^{2n+1}, S^1)$ of the form (2.2) in the Birkhoff normal form.

(iii) A function $G\in\mathscr{A}(\boldsymbol{C}^{2n+1}, S^1)$ and a map $f\in\mathrm{Symp}(\boldsymbol{C}^{2n}, 0)$ satisfying (2.1) are said to be in the *complex Birkhoff normal form* if they satisfy the above definitions (i) and (ii) respectively, with H_2 replaced by the complex normal form (and hence S is given by (2.6)).

We observe that this definition gives rise to the same form as the Birkhoff normal form defined by Moser [10]. Namely, we have:

PROPOSITION 2. *Let $f=\exp X_H$ with $H\in\mathscr{A}(\boldsymbol{R}^{2n+1}, S^1)$ of the form* (2.2) *with* (2.3) *in the Birkhoff normal form. Let Λ be the semisimple part of $Df(0)$, i.e., $\Lambda=\exp X_S$. Then f commutes with Λ, i.e., $f\circ\Lambda=\Lambda\circ f$, and can be written in the form*

$$f=\Lambda\circ\exp X_{\hat{H}}\ (=\exp X_{\hat{H}}\circ\Lambda)\ \text{with}\ \hat{H}=\hat{H}(z)=H(z,0)-S(z)\,.$$

Here $\hat{H}(z)$ is a real analytic integral of f that is invariant under Λ.

This proposition will be proved in Section 5. To discuss the convergence of Birkhoff transformations, we give the definition of a non-resonant or simply resonant fixed point (and an equilibrium point). For a map $f=\exp X_H$ with H of the form (2.2), we consider its fixed point $z=0$ as the equilibrium point of X_{H_2} and recall that the eigenvalues λ_j of $Df(0)$ are given by $\lambda_j=e^{\mu_j}$. Let us consider the condition

$$\sum_{j=0}^{n} k_j \mu_j = 0 \qquad \text{for} \quad \mu_0 = 2\pi i \quad \text{and} \quad (k_0, k_1, \dots, k_n) \in \boldsymbol{Z}^{n+1}. \tag{2.8}$$

This is equivalent to the condition

$$\prod_{j=1}^{n} \lambda_j^{k_j} = 1 \qquad \text{for} \quad (k_1, \dots, k_n) \in \boldsymbol{Z}^n \qquad (\lambda_j = e^{\mu_j}). \tag{2.9}$$

Let $H \in \mathscr{A}(\boldsymbol{R}^{2n+1}, S^1)$ be of the form (2.2) and let $f = \exp X_H \in \mathrm{Symp}(\boldsymbol{R}^{2n}, 0)$.

DEFINITION 3. (i) The equilibrium point $z=0$ of X_H is said to be *non-resonant* if the condition (2.8) holds only for $k_0 = k_1 = \cdots = k_n = 0$.

(ii) The equilibrium point $z=0$ of X_H is said to be *simply resonant* if either of the following three conditions holds after changing the indices of $\mu_1, \dots, \mu_n$ if necessary:

(1) $\mu_1/\mu_0 \in \boldsymbol{Q} \setminus \{0\}$ and the condition (2.8) holds only for $k_2 = \cdots = k_n = 0$;

(2) $\mu_1/\mu_2 \in \boldsymbol{Q} \setminus \{0\}$ and the condition (2.8) holds only for $k_0 = k_3 = \cdots = k_n = 0$;

(3) $\mu_1 = 0$ and the condition (2.8) holds only for $k_0 = k_2 = \cdots = k_n = 0$.

(iii) The fixed point $z=0$ of f is said to be *non-resonant* or *simply resonant* if it is non-resonant or simply resonant respectively as an equilibrium point of X_H.

In the non-resonance case, the complex Birkhoff normal form becomes a power series with constant coefficients in n variables $x_1 y_1, \dots, x_n y_n$ alone. In the simple resonance case, it becomes a function of $n+1$ variables with coefficients that are periodic functions in t. We will present them in Section 5.

Our main results are as follows:

THEOREM 1. *Let $f \in \mathrm{Symp}(\boldsymbol{R}^{2n}, 0)$ satisfy the condition* (2.1) *and assume that its fixed point $z=0$ is non-resonant or simply resonant. Then, there exists a real analytic Birkhoff transformation $\varphi \in \mathrm{Symp}(\boldsymbol{R}^{2n}, 0)$ such that $f' := \varphi^{-1} \circ f \circ \varphi$ is in the Birkhoff normal form, if and only if f is integrable. Furthermore, for any integral $G \in \mathscr{A}(\boldsymbol{C}^{2n}, 0)$ of f, the function $G \circ \varphi$ is invariant under $\Lambda = \exp X_S$ as well as f'.*

This theorem is already proved in the non-resonance case (cf. [7]), where we did not need the interpolation idea because of the special form of the Birkhoff normal form (see the end of Section 3 for more details). In this paper, we will prove Theorem 1 as a corollary to the following:

THEOREM 2. *Let $H \in \mathscr{A}(\boldsymbol{R}^{2n+1}, S^1)$ be of the form* (2.2) *and assume that the equilibrium point $z=0$ of X_H is non-resonant or simply resonant. Then, there exists a real analytic transformation $z = \varphi(\zeta, t) \in \mathrm{Symp}(\boldsymbol{R}^{2n+1}, S^1)$ such that the Hamiltonian of the transformed system is in the Birkhoff normal form, if and only if the vector field X_H is integrable. Furthermore, for any integral $G(z, t) \in \mathscr{A}(\boldsymbol{C}^{2n+1}, S^1)$ of X_H, the function $G(\varphi(\zeta, t), t)$ is in the Birkhoff normal form.*

In the proof of these theorems, the proof of the "only if" part is straightforward.

As we will show in Section 3, we can represent an integrable symplectic map as the time-one map of the flow of an integrable time-dependent Hamiltonian system. Therefore, the proof of Theorem 1 is reduced to that of Theorem 2. We will prove Theorem 2 in the simple resonance case only, because the proof in the non-resonance case can be carried out along the same line and its technical complexity is smaller than that in the simple resonance case.

3. Reduction to the time-dependent Hamiltonian case. The aim of this section is to derive Theorem 1 from Theorem 2. To this end, we will prove the following:

THEOREM 3. *Let $f \in \mathrm{Symp}(\boldsymbol{R}^{2n}, 0)$ be an integrable map satisfying* (2.1). *Then it is the time-one map of an integrable vector field X_H with $H \in \mathscr{A}(\boldsymbol{R}^{2n+1}, S^1)$ of the form*

$$H(z, t) = \frac{1}{2}\langle Az, z\rangle + O(|z|^3). \tag{3.1}$$

PROOF. As we mentioned in Section 1, f can be written as the time-one map $f = \exp X_H$ with $H \in \mathscr{A}(\boldsymbol{R}^{2n+1}, S^1)$ of the form (3.1) even without assuming the integrability of f. This is the assertion of a theorem by Kuksin and Pöschel [9]. Our purpose is to prove that the vector field X_H is necessarily integrable provided that f is integrable.

Let $G_k = G_k(z) \in \mathscr{A}(\boldsymbol{C}^{2n}, 0)$ $(k = 1, \dots, n)$ be n analytic integrals of $f = \exp X_H$ which are functionally independent and Poisson commuting. We assume that those G_k as well as the map f are defined to be analytic in a neighbourhood U $(\subset \boldsymbol{C}^{2n})$ of the origin $z = 0$. In the following, let $\phi(t; z_0, t_0)$ be the solution of the vector field X_H through $(z_0, t_0) \in \boldsymbol{C}^{2n} \times \boldsymbol{C}$. Further, let $U_\delta \subset \boldsymbol{C}^{2n}$ be the δ-neighbourhood of the origin, and set $V_K := \{t \in \boldsymbol{C} \mid |t| < K\}$ for any constant $K > 0$. Since the origin $z = 0$ is an equilibrium of X_H, for an arbitrary number $K > 0$ we can choose a small number $\delta > 0$ so that if $z_0 \in U_\delta$ and $t_0 \in V_K$, then $\phi(t; z_0, t_0) \in U \cap f^{-1}(U)$ for all $t \in V_K$.

Let us define functions $\hat{G}_k(z, t)$ on $U_\delta \times V_K$ by

$$\hat{G}_k(z, t) := G_k(\phi(0; z, t)) \qquad \text{for} \quad (z, t) \in U_\delta \times V_K. \tag{3.2}$$

Let us take $(z_0, t_0) \in U_\delta \times V_K$ arbitrarily and fix it. Then, by the uniqueness of solutions, we have

$$\hat{G}_k(\phi(t; z_0, t_0), t) = \hat{G}_k(z_0, t_0) \qquad \text{for} \quad t \in V_K.$$

This implies that $\hat{G}_k$ are invariant under the flow of X_H. Furthermore, by the invariance of G_k under f we have

$$\hat{G}_k(z, t+1) = G_k(\phi(0; z, t+1)) = G_k(f \circ \phi(0; z, t+1)) = G_k(\phi(1; z, t+1)).$$

Here we note $\phi(1; z, t+1) = \phi(0; z, t)$ in view of the periodicity of the vector field X_H with respect to t. Hence we have proved the relation

$$\hat{G}_k(z, t+1)=\hat{G}_k(z, t) \qquad \text{for} \quad (z, t)\in U_\delta \times V_{K-1}\,.$$

By the analytic dependence of solutions on the initial conditions, $\hat{G}_k$ are analytic in $(z, t)\in U\times V_K$ and periodic in t with period 1. By the analytic continuation using the relation above, $\hat{G}_k$ can be extended to a neighbourhood of the whole real t axis $\{(z, t)\in \boldsymbol{C}^{2n+1}\,|\, z=0,\, t\in \boldsymbol{R}\}$. Clearly they are integrals of X_H and hence we have proved that the existence of the integrals $G_k(z)$ of f leads to that of the integrals $\hat{G}_k$ of X_H which are analytic in a neighbourhood of the real t axis in the complex (z, t) space.

For each t fixed, we define a symplectic map Φ as

$$\Phi: z \longmapsto \phi(0; z, t)\,.$$

By the definition (3.2), we have

$$\frac{\partial \hat{G}_k}{\partial z}={}^t\Phi_z \frac{\partial G_k}{\partial z}\,, \tag{3.3}$$

where Φ_z is the Jacobian matrix of Φ with respect to z and ${}^t\Phi_z$ denotes its transpose. Hence the functional independence of $G_1, \dots, G_n$ implies that of $\hat{G}_1, \dots, \hat{G}_n$. Furthermore, from (3.3) as well as the symplectic character of Φ, it follows that

$$\{\hat{G}_i, \hat{G}_j\}=\{G_i, G_j\}$$

and therefore $\hat{G}_1, \dots, \hat{G}_n$ are Poisson commuting. Hence the vector field X_H is integrable. □

Using Theorem 3, we now deduce Theorem 1 from Theorem 2.

Proof of Theorem 1. Assume that $f\in \mathrm{Symp}(\boldsymbol{R}^{2n}, 0)$ is integrable. Then it follows from Theorem 3 that $f=\exp X_H$ with an integrable Hamiltonian $H\in \mathscr{A}(\boldsymbol{R}^{2n+1}, S^1)$. Let $z=z(t)$ be the solution of X_H through $z(0)=z_0$ at $t=0$. Then f is given by

$$f: z(0) \longmapsto z(1)\,.$$

Since the Hamiltonian H satisfies the assumption of Theorem 2 under that of Theorem 1, it follows from Theorem 2 that there exists a tansformation $z=\varphi(\zeta, t)\in \mathrm{Symp}(\boldsymbol{R}^{2n+1}, S^1)$ which takes the vector field X_H into the new vector field $X_{H'}$ with $H'=H'(\zeta, t)$ in the Birkhoff normal form. Let $\zeta(t)$ be the solution of $X_{H'}$ through $\zeta(0)=\varphi_0^{-1}(z_0)$ at $t=0$, where $\varphi_0=\varphi(\,\cdot\,, 0)$. Noting that $z(t)=\varphi(\zeta(t), t)$ and the periodicity of φ in t, one can see that f is transformed by φ_0 into

$$\varphi_0^{-1}\circ f\circ \varphi_0: \zeta(0) \longmapsto \zeta(1)\,,$$

which is equal to the time-one map $\exp X_{H'}$. Hence $\varphi_0^{-1}\circ f\circ \varphi_0$ is the Birkhoff normal form.

Conversely, assume that f is in the Birkhoff normal form, that is, $f=\exp X_H$ with $H\in \mathscr{A}(\boldsymbol{R}^{2n+1}, S^1)$ in the Birkhoff normal form. Then X_H is integrable by Theorem 2 and therefore there exist n integrals $G_k(z, t)$ of X_H ($k=1, \dots, n$) which are functionally independent and Poisson commuting. By Definition 1 and the periodicity of $G_k(z, t)$ in t, the functions $G_k(z, 0)$ are integrals of f and they are functionally independent and Poisson commuting. Hence f is integrable.

To prove the final assertion, let $f \in \mathrm{Symp}(\boldsymbol{R}^{2n}, 0)$ be in the Birkhoff normal form and $G \in \mathscr{A}(\boldsymbol{C}^{2n}, 0)$ its integral. According to Definition 2, we write $f = \exp X_H$ with $H \in \mathscr{A}(\boldsymbol{R}^{2n+1}, S^1)$ in the Birkhoff normal form. By the preceding arguments, $\hat{G}(z, t) := G(\phi(0; z, t))$ is an integral of X_H. Then it follows from Theorem 2 that $\hat{G}$ is in the Birkhoff normal form and hence we have $\{\hat{G}, S\} + \hat{G}_t \equiv 0$, that is,

$$\frac{d}{dt} \hat{G}(\exp tX_S(z), t) = 0 ,$$

where $\exp tX_S$ denotes the flow of the linear vector field X_S. Consequently we have $\hat{G}(\Lambda z, 1) = \hat{G}(z, 0)$ for $\Lambda = \exp X_S$, or equivalently $\hat{G}(\Lambda z, 0) = \hat{G}(z, 0)$ because of the periodicity of $\hat{G}$ in t. This implies that $G(\Lambda z) = G(z)$, which proves the final assertion of Theorem 1. $\square$

REMARK. In the previous paper [7], we proved Theorem 1 in the non-resonance case without the interpolation idea. For an integrable symplectic map $f \in \mathrm{Symp}(\boldsymbol{C}^{2n}, 0)$ near a non-resonant fixed point, we proved the existence of a convergent Birkhoff transformation $\varphi \in \mathrm{Symp}(\boldsymbol{C}^{2n}, 0)$ so that the transformed map $f' := \varphi^{-1} \circ f \circ f: (x, y) \mapsto (x', y')$ is written in the form

$$x'_k = x_k \exp H_{\tau_k}, \quad y'_k = y_k \exp(-H_{\tau_k}) \qquad \text{with} \quad \tau_k = x_k y_k \quad (k = 1, \dots, n), \tag{3.4}$$

where H is a power series in n variables τ_k with constant coefficients. When H is convergent, we called maps of the form (3.4) to be in the Birkhoff normal form. This is clearly the time-one map of the time-independent vector field X_H and therefore is in the complex Birkhoff normal form in the sense of Definition 2 (iii). The convergence of φ implies that of the Hamiltonian H in the expression (3.4) for f' (we omitted its proof in [7]). Indeed, the n power series $\exp H_{\tau_k}$ are convergent since f' is convergent. This implies the convergence of H_{τ_k} and hence that of H itself. The proof of Theorem 1 follows from the arguments above by imposing the reality condition on the original map f (see [7] and also Section 11 of this paper).

4. Reformulation of Theorem 2 in the extended phase space. For time-independent Hamiltonian systems, we have proved the result corresponding to Theorem 2 [8]. In order to proceed in the same way as in [8], we will mainly work with the complex Birkhoff normal form and also consider the time-dependent Hamiltonian as the time-independent one by extending the phase space to the $2n+2$ dimensional one.

The aim of this section is to reformulate Theorem 2 as a theorem about normalization in the extended complex phase space. Again, let $H \in \mathscr{A}(\boldsymbol{C}^{2n+1}, S^1)$ be a time-dependent Hamiltonian of the form

$$(4.1)\qquad H(z,t)=H_2(z)+O(|z|^3)\,;\quad H_2(z)=S(z)+N(z)\qquad\text{with}\quad S(z)=\sum_{k=1}^{n}\mu_k x_k y_k\,,$$

where S and N respectively are the semisimple and nilpotent parts of H_2, satisfying the relation $\{S,N\}\equiv 0$. Setting $t=x_{n+1}$, we define a function $\bar{H}$ as

$$(4.2)\qquad \bar{H}(z,x_{n+1},y_{n+1}):=H(z,x_{n+1})+y_{n+1}\,.$$

This is a function of $2n+2$ variables $z=(x,y)$, x_{n+1}, y_{n+1} and the corresponding Hamiltonian vector field $X_{\bar{H}}$ is given by

$$(4.3)\qquad \dot{z}=JH_z\,,\quad \dot{x}_{n+1}=1\,,\quad \dot{y}_{n+1}=-H_{x_{n+1}}\,,$$

where the dot ($\cdot$) denotes the differentiation with respect to t, and J is the $2n\times 2n$ symplectic matrix given in Section 2. We call the (x,y,x_{n+1},y_{n+1})-space *the extended phase space* and its symplectic structure is given by $\sum_{k=1}^{n+1}dx_k\wedge dy_k$. The function $\bar{H}$ is called *the extended Hamiltonian* (*function*) *of* H, and $X_{\bar{H}}$ is called *the extended Hamiltonian vector field* of X_H. The Poisson bracket of any two functions F, G on the extended phase space is defined by

$$[F,G]:=\{F,G\}+F_{x_{n+1}}G_{y_{n+1}}-F_{y_{n+1}}G_{x_{n+1}}\,,$$

where $\{F,G\}=\langle F_z, JG_z\rangle$. Also we define

$$\bar{S}(z,y_{n+1}):=S(z)+y_{n+1}\,;\quad S(z)=\sum_{k=1}^{n}\mu_k x_k y_k\,,$$

and introduce the following:

DEFINITION 4.1. A function $G=G(z,x_{n+1},y_{n+1})$ is said to be in the *normal form* if the following identity holds:

$$(4.4)\qquad [G,\bar{S}]\equiv 0\,.$$

Then we have:

LEMMA 4.2. *Let $G=G(z,t)\in\mathscr{A}(C^{2n+1},S^1)$ and $h(y_{n+1})$ an arbitrary function of y_{n+1}. Then*:

(i) *A function of the form $G(z,x_{n+1})+h(y_{n+1})$ is in the normal form if and only if $G(z,t)$ is in the complex Birkhoff normal form. In particular, the extended Hamiltonian $\bar{H}$ is in the normal form if and only if $H=H(z,t)$ is in the complex Birkhoff normal form.*

(ii) *$G(z,x_{n+1})$ is an integral of $X_{\bar{H}}$ if and only if $G(z,t)$ is an integral of X_H.*

PROOF. By the definition of the Poisson bracket $[\cdot,\cdot]$, we have

$$[G+h,\bar{S}]=\{G,S\}+G_{x_{n+1}}\,,\quad [G,\bar{H}]=\{G,H\}+G_{x_{n+1}}\,.$$

By the definition of the complex Birkhoff normal form, the first identity implies the assertion (i) and the second one implies (ii). □

The assertion (ii) above implies that the extended vector field $X_{\bar{H}}$ is integrable if X_H is integrable.

To find a normalizing transformation, we restrict ourselves to symplectic transformations of the form

$$\phi:(\zeta, \xi_{n+1}, \eta_{n+1}) \longmapsto (z, x_{n+1}, y_{n+1})\,; \quad \zeta=(\xi, \eta)\,, \quad z=(x, y) \in \boldsymbol{C}^n \times \boldsymbol{C}^n \tag{4.5}$$

with

$$z=\varphi(\zeta, \xi_{n+1})\,, \quad x_{n+1}=\xi_{n+1}+m \ (m \in \boldsymbol{Z})\,, \quad y_{n+1}=\eta_{n+1}+\psi(\zeta, \xi_{n+1})\,, \tag{4.6}$$

where all components of the vector function φ as well as ψ belong to $\mathscr{A}(\boldsymbol{C}^{2n+1}, S^1)$. We denote by $\mathscr{S}$ the set of all such symplectic transformations. Then we have:

LEMMA 4.3. (i) *$\mathscr{S}$ forms a group under composition of transformations.*

(ii) *For any function $H \in \mathscr{A}(\boldsymbol{C}^{2n+1}, S^1)$, the time-one map $\exp X_{\bar{H}}$ in the extended phase space belongs to $\mathscr{S}$.*

(iii) *Let ϕ be a transformation of the form* (4.5) *with* (4.6) *satisfying* $\det(\partial x/\partial\xi) \neq 0$. *Then ϕ is symplectic (i.e., $\phi \in \mathscr{S}$) if and only if $\varphi \in \mathrm{Symp}(\boldsymbol{C}^{2n+1}, S^1)$ and*

$$\psi(\zeta, \xi_{n+1})=\frac{\partial v(x, \eta, x_{n+1})}{\partial x_{n+1}} \qquad \text{for some } v \in \mathscr{A}(\boldsymbol{C}^{2n+1}, S^1)\,,$$

where $v=v(x, \eta, t)$ is the generating function of $\varphi \in \mathrm{Symp}(\boldsymbol{C}^{2n+1}, S^1)$.

(iv) *Let $H \in \mathscr{A}(\boldsymbol{C}^{2n+1}, S^1)$ be a function of the form* (4.1) *and $\bar{H}$ its extended Hamiltonian. Further let $\phi \in \mathscr{S}$ satisfy the condition* $\det(\partial x/\partial\xi) \neq 0$. *Then the vector field X_H is transformed by $\varphi \in \mathrm{Symp}(\boldsymbol{C}^{2n+1}, S^1)$ into the Hamiltonian vector field $X_{H'}$ with*

$$H'(\zeta, t)=H(\varphi(\zeta, t), t)+\psi(\zeta, t)\,, \tag{4.7}$$

and $\bar{H} \circ \phi$ is the extended Hamiltonian of H'.

REMARK. The assertion (iii) implies that the transformation $\phi \in \mathscr{S}$ is determined by φ uniquely under the condition $\det(\partial x/\partial\xi) \neq 0$. In other words, $\varphi \in \mathrm{Symp}(\boldsymbol{C}^{2n+1}, S^1)$ can be extended uniquely to a symplectic transformation $\phi \in \mathscr{S}$.

PROOF. First we prove that $\varphi \in \mathscr{A}(\boldsymbol{R}^{2n+1}, S^1)$ if ϕ is symplectic. We note that the symplectic property of ϕ means the identity

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

By the special form of ϕ, this implies that $\sum_{k=1}^{n} dx_k \wedge dy_k = \sum_{k=1}^{n} d\xi_k \wedge d\eta_k$ for any ξ_{n+1} fixed, which proves $\varphi \in \mathrm{Symp}(\boldsymbol{C}^{2n+1}, S^1)$. From this fact and the periodicity of φ and ψ in t, one can easily prove the assertion (i). Also the assertion (ii) can be easily proved.

To prove (iii), note that under the assumption $\det(\partial x/\partial\xi)\neq 0$ the symplectic transformation φ is expressed in terms of a generating function $v=v(x,\eta,t)\in\mathscr{A}(\boldsymbol{C}^{2n+1},S^1)$ as follows:

$$\xi=\frac{\partial v}{\partial\eta},\quad y=\frac{\partial v}{\partial x}.$$

Also, because of the form (4.6), ϕ is expressed as

$$\xi=\frac{\partial\hat{v}}{\partial\eta},\quad y=\frac{\partial\hat{v}}{\partial x},\quad \xi_{n+1}=\frac{\partial\hat{v}}{\partial\eta_{n+1}},\quad y_{n+1}=\frac{\partial\hat{v}}{\partial x_{n+1}}\tag{4.8}$$

in terms of the generating function

$$\hat{v}(x,\eta,x_{n+1},\eta_{n+1})=v(x,\eta,x_{n+1})+(x_{n+1}-m)\eta_{n+1}.$$

Clearly we have $y_{n+1}=\eta_{n+1}+\partial v/\partial x_{n+1}$. The converse assertion of (iii) can also be easily proved by constructing the generating function $\hat{v}$ from v. Hence (iii) is proved. For the proof of (iv), it is easy to see that $\bar{H}\circ\phi$ is written as

$$\bar{H}\circ\phi(\zeta,\xi_{n+1},\eta_{n+1})=H'(\zeta,\xi_{n+1})+\eta_{n+1}$$

with H' given by (4.7). We note that the vector field $X_{\bar{H}\circ\phi}$ is given by (4.3) with H replaced by H'. Therefore $\zeta=(\xi,\eta)$ satisfies the Hamiltonian system with the Hamiltonian H'. Since z, x_{n+1} are independent of η_{n+1}, this implies the assertion (iv). □

We consider the complex Birkhoff normalization of time-dependent Hamiltonians. In view of Lemma 4.2 and Lemma 4.3, the equivalence between the integrability and the convergence of Birkhoff transformation is formulated as the following theorem on the normalization in the extended phase space.

THEOREM 4.4. *Let $H\in\mathscr{A}(\boldsymbol{C}^{2n+1},S^1)$ be a function of the form* (4.1) *such that the origin $z=0$ is a non-resonant or simply resonant equilibrium of X_H. Let $\bar{H}$ be the extended Hamiltonian of H. Then, there exists an analytic symplectic transformation $\phi\in\mathscr{S}$ such that the new extended Hamiltonian $\bar{H}\circ\phi$ is in the normal form, if and only if X_H is integrable.*

The aim of Sections 5–10 is to prove this theorem.

5. Power series expansion of the normal form. We note that analytic functions $f(z,t)\in\mathscr{A}(\boldsymbol{C}^{2n+1},S^1)$ can be expanded as absolutely convergent power series in $z=(x,y)\in\boldsymbol{C}^{2n}$ at the origin whose coefficients are absolutely convergent Fourier series in $t=x_{n+1}\in\boldsymbol{C}$ with period 1. For convenience of notation, instead of $\mathscr{A}(\boldsymbol{R}^{2n+1},S^1)$ we denote by $\mathscr{P}$ the vector space over $\boldsymbol{C}$ of all such convergent series in z and t. Further, we denote by $\bar{\mathscr{P}}$ the direct sum of $\mathscr{P}$ and the one-dimensional vector space spanning by y_{n+1}. Namely $\bar{\mathscr{P}}$ is the vector space consisting of all those series $g=g(x,y,x_{n+1},y_{n+1})$

of the form

$$g=f(x, y, x_{n+1})+cy_{n+1}\,;\quad f\in\mathscr{P}\,,\quad c\in\boldsymbol{C}\,.$$

In this section, we express the normal form as such a series. It will lead to the proof of Proposition 2 and that of the "only if" part of Theorem 4.4.

Let us write $f\in\mathscr{P}$ as

$$f=\sum_{\alpha,\beta} c_{\alpha\beta}(t)x^{\alpha}y^{\beta}\,;\quad c_{\alpha\beta}(t)=\sum_{k\in\boldsymbol{Z}} d^{k}_{\alpha\beta}e^{2\pi ikt}\qquad (d^{k}_{\alpha\beta}\in\boldsymbol{C})\,,\tag{5.1}$$

where $x^{\alpha}=x_1^{\alpha_1}\cdots x_n^{\alpha_n}$, $y^{\beta}=y_1^{\beta_1}\cdots y_n^{\beta_n}$, $\alpha=(\alpha_1, \dots, \alpha_n)$, $\beta=(\beta_1, \dots, \beta_n)$ with nonnegative integers α_i, β_i.

PROPOSITION 5.1. *A normal form $f\in\mathscr{P}$ is written as*

$$f=f(z, t)=\sum_{\langle\mu,\alpha-\beta\rangle\in 2\pi i\boldsymbol{Z}} c_{\alpha\beta}e^{-\langle\mu,\alpha-\beta\rangle t}x^{\alpha}y^{\beta}\qquad (c_{\alpha\beta}\in\boldsymbol{C})\,.\tag{5.2}$$

PROOF. By the expression (5.1), we have

$$\begin{aligned}[f, \bar{S}]&=\sum_{\alpha,\beta}(\langle\mu, \alpha-\beta\rangle c_{\alpha\beta}(t)+c'_{\alpha\beta}(t))x^{\alpha}y^{\beta}\\ &=\sum_{\alpha,\beta}\left(\sum_{k\in\boldsymbol{Z}} d^{k}_{\alpha\beta}(\langle\mu, \alpha-\beta\rangle+2\pi ik)e^{2\pi ikt}\right)x^{\alpha}y^{\beta}\,,\end{aligned}\tag{5.3}$$

where $'=d/dt$ and $\mu=(\mu_1, \dots, \mu_n)$. Therefore the identity $[f, \bar{S}]\equiv 0$ holds if and only if the Fourier coefficients of $c_{\alpha\beta}(t)$ satisfy the condition

$$d^{k}_{\alpha\beta}=0\qquad\text{if}\quad \langle\mu, \alpha-\beta\rangle+2\pi ik\neq 0\,.$$

This implies that $c_{\alpha\beta}(t)\equiv 0$ if $\langle\mu, \alpha-\beta\rangle\notin 2\pi i\boldsymbol{Z}$ and further that a normal form $f\in\mathscr{P}$ can be written as

$$f=\sum_{\langle\mu,\alpha-\beta\rangle\in 2\pi i\boldsymbol{Z}} d^{k}_{\alpha\beta}e^{-\langle\mu,\alpha-\beta\rangle t}x^{\alpha}y^{\beta}\,,$$

where k is the integer determined by $\langle\mu, \alpha-\beta\rangle+2\pi ik=0$. By setting $d^{k}_{\alpha\beta}=c_{\alpha\beta}\in\boldsymbol{C}$, this can be written as (5.2). □

Using Proposition 5.1, one can prove Proposition 2 stated in Section 2.

PROOF OF PROPOSITION 2. Let $f=\exp X_H$ with $H\in\mathscr{A}(\boldsymbol{C}^{2n+1}, S^1)$ in the complex Birkhoff normal form. Then, by Lemma 4.2, $\bar{H}$ and $\bar{S}$ commute in the extended phase space and consequently their time-one maps commute. By the special form of $\bar{H}$ and $\bar{S}$, this implies that $f=\exp X_H$ and $\Lambda=\exp X_S$ commute in the original phase space.

Let $\varphi^t\colon \zeta\mapsto z=\varphi^t(\zeta)$ be the flow of the vector field X_S with $S(z)=\sum_{k=1}^{n}\mu_k x_k y_k$. It is linear symplectic but not periodic in t and therefore does not belong to $\mathrm{Symp}(\boldsymbol{C}^{2n+1}, S^1)$. It is given by

(5.4) $$\varphi^t(\zeta)=\operatorname{diag}(e^{\mu_1 t}\xi_1,\ldots,e^{\mu_n t}\xi_n,e^{-\mu_1 t}\eta_1,\ldots,e^{-\mu_n t}\eta_n)$$

and can be represented in terms of a generating function:

$$y_k=W_{x_k}\,,\quad \xi_k=W_{\eta_k}\qquad\text{for}\quad W=W(x,\eta,t)=\sum_{k=1}^{n}e^{-\mu_k t}x_k\eta_k\,.$$

By this transformation, the vector field X_H is taken into another vector field X_K with the Hamiltonian

$$K(\zeta,t):=H(\varphi^t(\zeta),t)+W_t(x,\eta,t)=H(\zeta,0)-S(\zeta)\,.$$

Here the second equality follows from (5.4) and Proposition 5.1. The function K is independent of t and is invariant under the linear map $\exp X_S:\zeta\mapsto\Lambda\zeta$. Moreover we have

$$\exp X_H=\varphi^1\circ\exp X_K\circ\{\varphi^0\}^{-1}\,.$$

Since $\varphi^0=\mathrm{id}$ and $\varphi^1=\exp X_S=\Lambda$, this leads to

$$f=\exp X_H=\Lambda\circ\exp X_K\,.$$

In view of this expression, the commuting relation $f\circ\Lambda=\Lambda\circ f$ is equivalent to $\exp X_K\circ\Lambda=\Lambda\circ\exp X_K$. Furthermore, since K is an integral of the vector field X_K, it is invariant under $\exp X_K$ as well as under Λ. Using this invariance we have

$$K(f(z))=K(\Lambda\circ\exp X_K(z))=K(z)\,.$$

Finally, let $f\in\mathrm{Symp}(\boldsymbol{R}^{2n},0)$ be in the (real) Birkhoff normal form, i.e., $f(z)=\exp X_H(z)$ with H in the Birkhoff normal form. Then one can find a linear (complex) symplectic transformation $z=C\zeta$, such that $H(C\zeta)$ is in the complex Birkhoff normal form. The arguments above imply that $C^{-1}\circ f\circ C=\exp X_{H\circ C}=\Lambda\circ\exp X_K$, which gives the desired expression for f with $\bar{H}(z)=K(C^{-1}z)$ as well as its invariance under f and $C\Lambda C^{-1}$. This completes the proof of Proposition 2.

Next, let us investigate normal forms when the equilibrium point $z=0$ is non-resonant or simply resonant. It will lead to the proof of the "only if" part of Theorem 4.4.

In the non-resonance case, in view of Definition 3 (i) and Proposition 5.1, a normal form $f\in\mathscr{P}$ is a power series in n products $x_1y_1,\ldots,x_ny_n$ with constant coefficients. Therefore, if $\bar{H}$ is in the normal form, $x_1y_1,\ldots,x_ny_n$ are n integrals of X_H which are functionally independent and Poisson commuting.

In the simple resonance case, we divide our discussions into three cases, namely, cases (1)–(3) of Definition 3 (ii).

In the case (1), we may put

$$\mu_1=\frac{2\pi iq}{p}\quad(p\in\boldsymbol{N},\ q\in\boldsymbol{Z}\setminus\{0\},\ p\text{ and }|q|\text{ are relatively prime})$$

and a normal form $f\in\mathscr{P}$ can be written as

$$f=\sum_{\substack{\alpha_1-\beta_1\in p\boldsymbol{Z}\\ \alpha_2,\dots,\alpha_n\in\boldsymbol{Z}_+}} c_{\alpha\beta}e^{-\mu_1(\alpha_1-\beta_1)t}x_1^{\alpha_1}y_1^{\beta_1}(x_2y_2)^{\alpha_2}\cdots(x_ny_n)^{\alpha_n},$$

where $\boldsymbol{Z}_+$ is the set of nonnegative integers. Therefore the normal form $f\in\mathscr{P}$ is a power series with constant coefficients in $n+2$ variables $x_1y_1,\dots,x_ny_n$ and $e^{-2\pi iqt}x_1^p$, $e^{2\pi iqt}y_1^p$. For technical reasons, we set

$$\begin{cases}\tau_0=y_{n+1}, & \tau_1=y_{n+1}+\mu_1x_1y_1, \quad \tau_k=x_ky_k \qquad (k=2,\dots,n),\\ \tau_{n+1}=e^{-2\pi iqt}x_1^p, & \tau_{n+2}=e^{2\pi iqt}y_1^p \qquad (t=x_{n+1}).\end{cases}\tag{5.5}$$

Here we note that the following relation holds:

$$\tau_{n+1}\tau_{n+2}=\left(\frac{\tau_1-\tau_0}{\mu_1}\right)^p.\tag{5.6}$$

In the case (2), we may put

$$\frac{\mu_1}{\mu_2}=\pm\frac{q}{p}\quad (p,q\in N \text{ are relatively prime})$$

and the normal form can be written as

$$f=\sum_{\mu_1(\alpha_1-\beta_1)+\mu_2(\alpha_2-\beta_2)=0} c_{\alpha\beta}x_1^{\alpha_1}y_1^{\beta_1}x_2^{\alpha_2}y_2^{\beta_2}(x_3y_3)^{\alpha_3}\cdots(x_ny_n)^{\alpha_n}.$$

This is a power series with constant coefficients in $n+2$ variables $x_1y_1,\dots,x_ny_n$, $x_1^pv^q$, $y_1^pu^q$, where

$$(u,v)=\begin{cases}(x_2,y_2) & \text{if } p\mu_1=q\mu_2,\\ (y_2,x_2) & \text{if } p\mu_1=-q\mu_2.\end{cases}$$

In this case, we set

$$\begin{cases}\tau_0=y_{n+1}, \quad \tau_1=\mu_1x_1y_1-\mu_2x_2y_2, \quad \tau_2=\mu_1x_1y_1+\mu_2x_2y_2,\\ \tau_k=x_ky_k \quad (k=3,\dots,n), \quad \tau_{n+1}=x_1^pv^q, \quad \tau_{n+2}=y_1^pu^q,\end{cases}\tag{5.7}$$

and the following relation holds:

$$\tau_{n+1}\tau_{n+2}=\left(\frac{\tau_1+\tau_2}{2\mu_1}\right)^p\left(\frac{\tau_2-\tau_1}{2\mu_2}\right)^q.$$

In the case (3), the normal form can be written as

$$f=\sum_{\alpha_1,\alpha_2,\dots,\alpha_n,\beta_1\in\boldsymbol{Z}_+} c_{\alpha\beta}x_1^{\alpha_1}y_1^{\beta_1}(x_2y_2)^{\alpha_2}\cdots(x_ny_n)^{\alpha_n},$$

which is a power series with constant coefficients in $n+1$ variables $x_1, y_1, x_2y_2,\dots,x_ny_n$.

In this case, we set $\tau_0, \tau_1, \dots, \tau_{n+1}$ as follows:

$$\tau_0 = y_{n+1}, \quad \tau_1 = x_1, \quad \tau_k = x_k y_k \quad (k=2, \dots, n), \quad \tau_{n+1} = y_1. \tag{5.8}$$

In the cases (2) and (3) above, the normal forms are independent of t and the same as those of time-independent Hamiltonian functions in the simple resonance cases [8]. The proof of Theorem 4.4 in these cases can be carried out along the same line as in the case (1) with more repetition of the arguments in [8]. Hence we omit it and will concentrate on the case (1) from now on.

Then, the extended Hamiltonian $\bar{H}$ is in the normal form if and only if it is a power series with constant coefficients in $n+3$ variables $\tau_0, \tau_1, \dots, \tau_{n+2}$ with the relation (5.6). Recall that the identity $[\bar{H}, \bar{S}] \equiv 0$ holds for the normal form $\bar{H}$. This also implies that $\bar{S}$ is an integral of $X_{\bar{H}}$ and therefore $\bar{H} - \bar{S} = H - S$ is an integral of $X_{\bar{H}}$ which is independent of y_{n+1}. Then, in view of Lemma 4.2, $H-S, \tau_2, \dots, \tau_n$ are n integrals of X_H which are functionally independent and Poisson commuting functions of z in general for each t fixed. Here we note that they are functionally dependent only if $H-S$ is a function of $\tau_2, \dots, \tau_n$ only. However, in this case also, X_H is integrable. Thus we have proved that, in the non-resonance and simple resonance cases, the vector field X_H with H in the Birkhoff normal form is integrable. This proves the "only if" part of Theorem 4.4.

Finally, for later use, let us consider the Poisson bracket of any two functions in the normal form. It vanishes identically in the non-resonance case, but does not vanish in general in the simple resonance case. To compute it, let $f = f(z, x_{n+1}, y_{n+1})$ be a function in the normal form in the simple resonance case. Using the relation (5.6), we can write it uniquely in the form

$$f(z, x_{n+1}, y_{n+1}) = f_1(\tau, \tau_{n+1}) + f_2(\tau, \tau_{n+2}); \quad \tau = (\tau_0, \tau_1, \dots, \tau_n), \tag{5.9}$$

where $f_i(\tau, \tau_{n+i})$ are power series in $\tau_0, \dots, \tau_n$ and τ_{n+i} $(i=1, 2)$. If f_2 does not vanish identically, then by eliminating τ_{n+2} from $f_2(\tau, \tau_{n+2})$ using the relation (5.6), one can consider f as a Laurent series in $n+2$ variables $\tau_0, \tau_1, \dots, \tau_{n+1}$. Then, we obtain the following formula.

LEMMA 5.2. *Let f, g be in the normal form and consider them as Laurent series in $\tau_0, \tau_1, \dots, \tau_{n+1}$ given by* (5.5). *Then*

$$[f, g] = \left(\frac{\partial f}{\partial \tau_0} \frac{\partial g}{\partial \tau_{n+1}} - \frac{\partial f}{\partial \tau_{n+1}} \frac{\partial g}{\partial \tau_0} \right) [\tau_0, \tau_{n+1}] \tag{5.10}$$

with

$$[\tau_0, \tau_{n+1}] = p\mu_1 \tau_{n+1}. \tag{5.11}$$

PROOF. Since f and g are functions in $\tau_0, \tau_1, \dots, \tau_{n+1}$, we have

$$[f, g] = \sum_{i,j=0}^{n+1} \frac{\partial f}{\partial \tau_i} \frac{\partial g}{\partial \tau_j} [\tau_i, \tau_j] .$$

Here $[\tau_i, \tau_j] = 0$ except for $(i, j) = (0, n+1)$ or $(n+1, 0)$. Hence we obtain the desired formula (5.10) with (5.11).

6. Construction of a formal normalizing transformation. We will obtain a normalizing transformation ϕ described in Theorem 4.4 as the limit of an iteration process. The aim of this section is to construct this process and prove the existence of ϕ as a formal transformation.

We begin by proving some lemmas. For the vector space $\mathscr{P}$ and $\bar{\mathscr{P}}$ introduced in Section 5, let us consider the linear maps

$$\operatorname{ad}\bar{S} : \bar{\mathscr{P}} \ni f \longmapsto [f, \bar{S}] \in \mathscr{P} , \quad \operatorname{ad} N : \bar{\mathscr{P}} \ni f \longmapsto [f, N] \in \mathscr{P} , \quad \operatorname{ad}\bar{H}_2 = \operatorname{ad}\bar{S} + \operatorname{ad} N ,$$

where S, N and H_2 are quadratic forms defined by (4.1). By Definition 4.1, a series $f \in \bar{\mathscr{P}}$ is in the normal form if and only if $f \in \operatorname{Ker}\operatorname{ad}\bar{S}$. Let $\mathscr{P}_m$ denote the vector space over $\boldsymbol{C}$ of all homogeneous polynomials in $z = (x, y)$ of degree m with coefficients being Fourier series of t with period 1. Also we define the vector space $\bar{\mathscr{P}}_m$ in the same way as defining $\bar{\mathscr{P}}$ from $\mathscr{P}$. Let us define linear maps $\operatorname{ad}_m \bar{S}$, $\operatorname{ad}_m N$ and $\operatorname{ad}_m \bar{H}_2$ by restriction

$$\operatorname{ad}_m \bar{S} := \operatorname{ad}\bar{S} \,|\, \bar{\mathscr{P}}_m , \quad \operatorname{ad}_m N := \operatorname{ad} N \,|\, \bar{\mathscr{P}}_m , \quad \operatorname{ad}_m \bar{H}_2 := \operatorname{ad}\bar{H}_2 \,|\, \bar{\mathscr{P}}_m .$$

The following lemma will play an important role in the construction of the normalizing transformation.

LEMMA 6.1.
(i) $\bar{\mathscr{P}}_m = \operatorname{Ker}\operatorname{ad}_m \bar{S} \oplus \operatorname{Im}\operatorname{ad}_m \bar{S}$, $\bar{\mathscr{P}} = \operatorname{Ker}\operatorname{ad}\bar{S} \oplus \operatorname{Im}\operatorname{ad}\bar{S}$.
(ii) *The restriction of* $\operatorname{ad}_m \bar{H}_2$ *to* $\operatorname{Im}\operatorname{ad}_m \bar{S}$ *is an invertible map from* $\operatorname{Im}\operatorname{ad}_m \bar{S}$ *onto itself.*
(iii) *If* $f, g \in \operatorname{Ker}\operatorname{ad}\bar{S}$, *then* $[f, g] \in \operatorname{Ker}\operatorname{ad}\bar{S}$.
(iv) *If* $f \in \operatorname{Ker}\operatorname{ad}\bar{S}$ *and* $g \in \operatorname{Im}\operatorname{ad}\bar{S}$, *then* $[f, g] \in \operatorname{Im}\operatorname{ad}\bar{S}$.

PROOF. If we write $f \in \mathscr{P}$ in the form (5.1), we see from (5.3) that

$$\begin{cases} f \in \operatorname{Ker}\operatorname{ad}_m \bar{S} & \text{if and only if} \quad d_{\alpha\beta}^k = 0 \quad \text{for } \langle \mu, \alpha - \beta \rangle + 2\pi i k \neq 0 , \\ f \in \operatorname{Im}\operatorname{ad}_m \bar{S} & \text{if and only if} \quad d_{\alpha\beta}^k = 0 \quad \text{for } \langle \mu, \alpha - \beta \rangle + 2\pi i k = 0 . \end{cases}$$

This implies the assertion (i). To prove the assertion (ii), let $f \in \mathscr{P}_m$ and $g \in \operatorname{Im}\operatorname{ad}_m \bar{S}$ satisfy the equation

$$[f, \bar{S}] = g \qquad \text{with} \quad g = \sum_{|\alpha| + |\beta| = m} \left(\sum_{k \in \boldsymbol{Z}} e_{\alpha\beta}^k e^{2\pi i k t} \right) x^\alpha y^\beta \quad (e_{\alpha\beta}^k \in \boldsymbol{C}) .$$

From (5.3), this equation is solved uniquely for a formal series f as

$$f=\sum_{|\alpha|+|\beta|=m}\left(\sum_{k\in Z'}\frac{e_{\alpha\beta}^{k}}{\langle\mu,\alpha-\beta\rangle+2\pi ik}e^{2\pi ikt}\right)x^{\alpha}y^{\beta},$$

where Z' is the set of integers satisfying the condition $\langle\mu,\alpha-\beta\rangle+2\pi ik\neq 0$. Since the absolute value $|\langle\mu,\alpha-\beta\rangle+2\pi ik|$ is bounded away from zero (for α, β fixed), the coefficients of f are absolutely convergent Fourier series of t and therefore f is holomorphic and belongs to $\operatorname{Im}\operatorname{ad}_m\bar{S}$. This implies that $\operatorname{ad}_m\bar{S}$ is an invertible map from $\operatorname{Im}\operatorname{ad}_m\bar{S}$ onto itself. Furthermore, we note that $\operatorname{ad}_m\bar{S}$ and $\operatorname{ad}_m N$ commute because $[\bar{S},N]=\{S,N\}=0$. Hence we have

$$\operatorname{ad}_m N(\operatorname{Im}\operatorname{ad}_m\bar{S})=\operatorname{ad}_m\bar{S}(\operatorname{Im}\operatorname{ad}_m N)\subset\operatorname{Im}\operatorname{ad}_m\bar{S},$$

which implies that $\operatorname{ad}_m H_2$ is a map from $\operatorname{Im}\operatorname{ad}_m\bar{S}$ to itself. Since $\operatorname{ad}_m N$ is nilpotent, there exists a positive integer $m\in N$ such that $(\operatorname{ad}_m N)^m\neq 0$ and $(\operatorname{ad}_m N)^n=0$ for any integer $n>m$. This implies that

$$\begin{aligned}(\operatorname{ad}_m H_2)^{-1}&=(\operatorname{ad}_m\bar{S}+\operatorname{ad}_m N)^{-1}\\&=\{I-(\operatorname{ad}_m\bar{S})^{-1}(\operatorname{ad}_m N)+\cdots+(-(\operatorname{ad}_m\bar{S})^{-1}(\operatorname{ad}_m N))^m\}(\operatorname{ad}_m\bar{S})^{-1},\end{aligned}$$

which gives an expression for the inverse map of $\operatorname{ad}_m H_2\,|\operatorname{Im}\operatorname{ad}_m\bar{S}$. Thus we have proved the assertion (ii). Furthermore, one can prove (iii) using the Jacobi identity. Also, the assumption of (iv) implies $[f,\bar{S}]=0$ and $g=[h,\bar{S}]$ for some $h\in\mathscr{P}$, and therefore we have $[f,g]=[[f,h],\bar{S}]$, proving the assertion (iv). □

Associated with the decomposition (i) of Lemma 6.1, we introduce projection operators P_N and P_R as follows:

$$P_N:\bar{\mathscr{P}}\to\operatorname{Ker}\operatorname{ad}\bar{S},\quad P_R:\bar{\mathscr{P}}\to\operatorname{Im}\operatorname{ad}\bar{S}.$$

Then any series $f\in\bar{\mathscr{P}}$ is represented by

$$f=P_Nf+P_Rf.$$

The relation $f=P_Nf$ implies that f is in the normal form. We call P_Nf and P_Rf the normal form part of f and the remainder part of f respectively.

Since we define the normalizing transformation by an iteration process, we have to consider normal forms up to finite order. We consider series belonging to $\mathscr{P}$ modulo constants and write the power series expansion of $f=f(z,t)\in\mathscr{P}$ with respect to z as follows:

$$f=f^0+f^1+\cdots;\quad f^0\not\equiv\text{const.},\tag{6.1}$$

where f^d $(d=0,1,\dots)$ are homogeneous polynomials in $z=(x,y)$ of degree $s+d$ with coefficients being Fourier series in t with period 1. Here s is the degree of the polynomial f^0. Here and in what follows, we often use t instead of x_{n+1} or ξ_{n+1} for the convenience of notation. We call f^0 the *lowest order part* of f. For the Hamiltonian $\bar{H}$, we consider

$\bar{H}_2=y_{n+1}+H_2(z)$ as its lowest order part and denote it by $\bar{H}^0$.

DEFINITION 6.2. (i) A function $f\in\mathscr{P}$ is said to be in the *normal form up to order* $s+d$ $(s=\deg f^0)$ if $f^0+\cdots+f^d$ is in the normal form.

(ii) The extended Hamiltonian $\bar{H}=y_{n+1}+H(z, x_{n+1})$ is said to be in the *normal form up to order* s_0+d $(s_0=2)$ if $H\in\mathscr{P}$ is in the normal form up to order s_0+d. The lowest order part of $\bar{H}$ is defined as $\bar{H}^0=\bar{H}_2$ and we write $\bar{H}$ in the form (6.1), that is,

$$\bar{H}=\bar{H}^0+\bar{H}^1+\cdots; \quad \bar{H}^0=\bar{H}_2, \quad \bar{H}^d=H^d. \tag{6.2}$$

Now in order to define the iteration process, we consider symplectic transformations which are time-one maps of the Hamiltonian vector fields with Hamiltonians of the form

$$\bar{W}(z, x_{n+1}, y_{n+1}):=W(z, x_{n+1})+y_{n+1} \qquad \text{with} \quad W\in\mathscr{A}(\boldsymbol{R}^{2n+1}, S^1). \tag{6.3}$$

By Lemma 4.3 (ii), the time-one map $\phi=\exp X_{\bar{W}}$ belongs to $\mathscr{S}$. Therefore any composite of those time-one maps belongs to $\mathscr{S}$. The iteration procedure is described as follows:

THEOREM 6.3. *Let $H=H(z, t)\in\mathscr{A}(\boldsymbol{C}^{2n+1}, S^1)$ be a function of the form* (4.1) *and assume that H is holomorphic in a domain $\Omega=U\times R$, where $U\subset\boldsymbol{C}^{2n}$ is a neighbourhood of the origin $z=0$ and $R\subset\boldsymbol{C}$ is a strip domain of the form*

$$R=\{t\in\boldsymbol{C}\mid |\operatorname{Im} t|<r\}, \qquad r>0 \text{ a constant}.$$

Assume that its extended Hamiltonian $\bar{H}$ is in the normal form up to order s_0+d-1 $(s_0=2,\ d\geq 0)$. Then there exists a unique function $W(z, t)\in\mathscr{A}(\boldsymbol{C}^{2n+1}, S^1)$ such that

(i) *W has the form*

$$W=W^{d+2}+\cdots+W^{2d+1}; \quad P_N W=0, \tag{6.4}$$

where $W^l=W^l(z, t)\in\mathscr{A}(\boldsymbol{C}^{2n+1}, S^1)$ are homogeneous polynomials of degree l in z whose coefficients are holomorphic Fourier series of $t\in R$ with period 1;

(ii) *for the transformation $\phi=\exp X_{\bar{W}}$ with $\bar{W}$ of the form* (6.3), *$\bar{H}\circ\phi$ is in the normal form up to order s_0+2d-1.*

Clearly this implies the following:

COROLLARY 6.4. *For the extended Hamiltonian $\bar{H}$ of $H\in\mathscr{A}(\boldsymbol{C}^{2n+1}, S^1)$ in the form* (4.1), *there exists a unique sequence of symplectic transformations ϕ_ν $(\nu=0, 1, \ldots)$ such that*

(i) *$\phi_\nu=\exp X_{\bar{W}}$, $\bar{W}$ being a function described in Theorem* 6.3 *with $d=2^\nu$;*

(ii) *for the transformation $\phi^{(\nu)}:=\phi_0\circ\cdots\circ\phi_\nu$, $\bar{H}\circ\phi^{(\nu)}$ is in the normal form up to order $s_0+2^{\nu+1}-1$ $(s_0=2)$.*

REMARKS. (i) In the above, we need neither the assumption of integrability nor that of non-resonance or simple resonance.

(ii) Obviously $\phi:=\lim_{\nu\to\infty}\phi^{(\nu)}$ is a formal symplectic transformation such that

$\bar{H} \circ \phi$ is in the normal form.

PROOF OF THEOREM 6.3. By assumption, we can write $\bar{H}$ in the form

$$\bar{H} = h + \hat{H}\,; \quad h = P_N h\,, \quad \hat{H} = O(|z|^{s_0+d})\,.$$

Let $\bar{W}$ be a polynomial of the form (6.3) with (6.4). Then the time-one map $\phi := \exp X_{\bar{W}}$: $(\zeta, \xi_{n+1}, \eta_{n+1}) \mapsto (z, x_{n+1}, y_{n+1})$ can be written as (see [7, p. 423])

$$\phi: \begin{cases} z = \zeta + JW_\zeta(\zeta, \xi_{n+1}) + O(|\zeta|^{2d+1})\,, \\ x_{n+1} = \xi_{n+1} + 1\,, \\ y_{n+1} = \eta_{n+1} - W_{\xi_{n+1}}(\zeta, \xi_{n+1}) + O(|\zeta|^{2d+1})\,. \end{cases}$$

This implies that

$$\bar{H} \circ \phi(\zeta, \xi_{n+1}, \eta_{n+1}) = h + [h, W] + \hat{H} + O(|\zeta|^{s_0+2d})\,, \tag{6.5}$$

where the argument of h is $(\zeta, \xi_{n+1}, \eta_{n+1})$ and that of W, $\hat{H}$ is (ζ, ξ_{n+1}). Since h is in the normal form, it follows from Lemma 6.1 (iv) that $P_N[h, W] = 0$ under the assumption $P_N W = 0$. Therefore $\bar{H} \circ \phi$ is in the normal form up to order $s_0 + 2d - 1$ if and only if W satisfies the equation

$$[h, W] = -P_R \hat{H} + O(|\zeta|^{s_0+2d})\,. \tag{6.6}$$

Writing $\bar{H}$ in the form (6.2) and comparing the homogeneous parts of degree $s_0 + l$ $(d \le l \le 2d-1)$, we have

$$[h^0, W^{l+2}] = -P_R \hat{H}^l - \sum_{\nu=1}^{l-d} [h^\nu, W^{l+2-\nu}] \qquad (l = d, \dots, 2d-1)\,, \tag{6.7}$$

where $h^\nu = \bar{H}^\nu$, $\hat{H}^\nu = \bar{H}^\nu$. By Lemma 6.1 (ii) and (iv), this equation can be solved uniquely for $W^{l+2} \in \operatorname{Im} \operatorname{ad}_{l+2} \bar{S}$, provided that $P_N W^{d+2} = \cdots = P_N W^{l+1} = 0$. It follows from the proof of Lemma 6.1 (ii) that the coefficients of the terms $\xi^\alpha \eta^\beta$ in W^{l+2} are holomorphic Fourier series of $t \in R$. By induction this implies the unique existence of a polynomial W of the form (6.4) satisfying (6.6). □

7. Idea of the convergence proof. The essential part of the proof of Theorem 4.4 is the proof of convergence for the transformation ϕ obtained by Theorem 6.3. We now formulate it in the following theorem:

THEOREM 7.1. *Let $H \in \mathscr{A}(C^{2n+1}, S^1)$ be a function of the form* (4.1) *such that $z = 0$ is a non-resonant or simply resonant equilibrium point of X_H. Assume that the system X_H is integrable. Then the sequence $\{\phi^{(\nu)}\}$ described in Corollary* 6.4 *converges uniformly to an analytic symplectic transformation $\phi \in \mathscr{S}$.*

The aim of this section is to describe our idea of proving this theorem. A key point of the proof is that $\phi = \lim_{\nu \to \infty} \phi^{(\nu)}$ takes n additional integrals of $X_{\bar{H}}$ as well as $\bar{H}$ into the normal form. This fact is a consequence of the following:

LEMMA 7.2. *Let $G=G(z,t)\in\mathscr{A}(C^{2n+1},S^1)$ be an integral of X_H and assume that $\bar{H}$ is in the normal form up to order s_0+d, where $s_0=\deg H_2=2$ and $d\geq 0$. Then $s:=\deg G^0\geq 1$ and G is in the normal form up to order $s+d$.*

PROOF. Since $G(z,x_{n+1})$ is an integral of $X_{\bar{H}}$ by Lemma 4.2, we have the identity $[G,\bar{H}]=0$. Writing G and $\bar{H}$ in the form (6.1) and (6.2), the comparison of the homogeneous parts of degree $s+l$ $(0\leq l\leq d)$ in this identity gives

$$\sum_{i+j=l}[G^i,\bar{H}^j]=0\,. \tag{7.1}$$

For $l=0$, this reads $[G^0,\bar{H}^0]=0$, which implies that G^0 is in the normal form. If $\deg G^0=0$, $G^0(z,x_{n+1})$ is a function of x_{n+1} only and the relation $[G^0,\bar{H}^0]=0$ means $\partial G^0/\partial x_{n+1}\equiv 0$. Then G^0 is a constant, which contradicts (6.1). Therefore we have $\deg G^0\geq 1$. From the identity (7.1) for $l=1,2,\dots,d$, we can prove inductively that $G^1,G^2,\dots,G^d$ are in the normal form by using Lemma 6.1 (see the proof of Proposition 3.2 of [8]). □

Suppose that the vector field X_H is integrable with n functionally independent integrals $G_k(z,t)\in\mathscr{A}(C^{2n+1},S^1)$ $(k=1,\dots,n)$, which we write in the expansion in terms of homogeneous polynomials (6.1). For technical reasons, we need the functional independence of the lowest order parts $G_0^0,\dots,G_n^0$, which does not follow directly from that of $G_1,\dots,G_n$ in general. However, we have:

LEMMA 7.3. *Let $H\in\mathscr{A}(C^{2n+1},S^1)$ be a function of the form (4.1) such that $z=0$ is a non-resonant or simply resonant equilibrium point of X_H. Let $G_1(z,t),\dots,G_n(z,t)\in\mathscr{A}(C^{2n+1},S^1)$ be n integrals of X_H that are functionally independent functions of z for any t fixed. Then there exist n integrals of X_H in the form*

$$G_k'(z,t):=P_k(G_1(z,t),\dots,G_n(z,t))\in\mathscr{A}(C^{2n+1},S^1)\qquad(k=1,\dots,n)\,, \tag{7.2}$$

where P_k are polynomials of $G_1,\dots,G_n$ with complex constant coefficients, such that the lowest order parts of $G_1',\dots,G_n'$ are functionally independent polynomials of z for any t fixed. Furthermore, if $G_1,\dots,G_n$ are Poisson commuting, those $G_1',\dots,G_n'$ are also Poisson commuting.

PROOF. By assumption, $G_1(z,0),\dots,G_n(z,0)$ are functionally independent functions of z. Then, using Ziglin's lemma [13, Lemma 2.1], we can determine the polynomials P_k with complex coefficients so that $P_k(G_1(z,0),\dots,G_n(z,0))$ $(k=1,\dots,n)$ are functionally independent (see [7, Appendix] for the proof). Let us define the functions G_k' by (7.2), which are integrals of X_H. Since the lowest order part of H is in the normal form under the assumption, it follows from Lemma 7.2 that the lowest order parts of G_k' are in the normal form and their degrees are greater than or equal to 1. Recall that the normal forms in the simple resonance case are power series with constant coefficients in the variables $\tau_0, \dots, \tau_{n+1}, \tau_{n+2}$. Therefore, the coefficients of any term $x^\alpha y^\beta$ in the lowest order part of G_k' are either identically zero or nonzero everywhere. This implies that the lowest order parts of $G_k'(z, t)$ are determined independently of t and are functionally independent functions of z for any t fixed. The same assertion holds also in the non-resonance case. The final assertion is obvious. $\square$

Now, writing G_k for G_k' obtained by Lemma 7.3 under the assumption of Theorem 7.1, we may assume that $G_1, \dots, G_n \in \mathscr{A}(C^{2n+1}, S^1)$ are Poisson commuting integrals of X_H such that their lowest order parts $G_1^0, \dots, G_n^0$ are functionally independent polynomials of z for each t fixed. For the convenience of notation, we set $\bar{H} = G_0$ and assume that it is in the normal form up to order $s_0 + d - 1$ ($s_0 = 2$). Then it follows from Lemma 7.2 that those $n+1$ functions G_k are in the normal form up to order $s_k + d - 1$, where s_k are the degrees of the lowest order parts G_k^0. Let us write them in the form

$$G_k(z, x_{n+1}, y_{n+1}) = g_k(z, x_{n+1}, y_{n+1}) + \hat{G}_k(z, x_{n+1}), \qquad (k = 0, 1, \dots, n) \tag{7.3}$$

with g_k being in the normal form and $\hat{G}_k(z, x_{n+1}) = O(|z|^{s_k + d})$. Although we write $n+1$ functions $G_0, \dots, G_n$ in this manner, they (and g_k) are actually independent of y_{n+1} except for G_0 (and g_0). Let ϕ be the symplectic transformation described in Theorem 6.3. Similarly as in the proof of Theorem 6.3, we have

$$G_k \circ \phi(\zeta, \xi_{n+1}, \eta_{n+1}) = g_k + [g_k, W] + \hat{G}_k + O(|\zeta|^{s_k + 2d}) \qquad (k = 0, 1, \dots, n), \tag{7.4}$$

where the arguments of functions g_k, W, $\hat{G}_k$ are $(\zeta, \xi_{n+1}, \eta_{n+1})$. Those functions as well as the terms $O(|\zeta|^{s_k + 2d})$ are independent of η_{n+1} except for g_0. By Theorem 6.3 and Lemma 7.2, $G_k \circ \phi$ are in the normal form up to order $s_k + 2d - 1$ and therefore W satisfies $n+1$ equations:

$$[g_k, W] = -P_R \hat{G}_k + O(|\zeta|^{s_k + 2d}) \qquad (k = 0, 1, \dots, n). \tag{7.5}$$

Here we note that $G_k^0 = g_k^0$ ($k = 0, 1, \dots, n$). Since $g_0^0 = \eta_{n+1} + H_2(\zeta)$ and $g_1^0, \dots, g_n^0$ are independent of η_{n+1}, the functional independence of $g_1^0, \dots, g_n^0$ is equivalent to that of $g_0^0, g_1^0, \dots, g_n^0$ as functions of ζ and η_{n+1} for each $t = \xi_{n+1}$ fixed. Recall that $g_0^0, g_1^0, \dots, g_n^0$ can be considered as rational functions of $\tau_0, \tau_1, \dots, \tau_{n+1}$ by eliminating τ_{n+2} using the relation (5.6). Here and in what follows, τ_k are those given by (5.5) with x_i, y_i replaced by ξ_i, η_i. Then their functional independence implies that

$$\operatorname{rank}\left(\frac{\partial(g_0^0, \dots, g_n^0)}{\partial(\tau_0, \dots, \tau_{n+1})} \right) = n+1 \tag{7.6}$$

on an open and dense subset of C^{2n+1} for each t fixed. Furthermore, recall that G_i and G_j are Poisson commuting for all $i, j = 0, 1, \dots, n$. Then comparing the lowest order parts of the identity $[G_i, G_j] \equiv 0$, we see that the Poisson bracket $[g_i^0, g_j^0]$ vanishes identically. By Lemma 5.2, this leads to

$$\frac{\partial g_i^0}{\partial \tau_0}\frac{\partial g_j^0}{\partial \tau_{n+1}}-\frac{\partial g_i^0}{\partial \tau_{n+1}}\frac{\partial g_j^0}{\partial \tau_0}=0 \qquad \text{for all} \quad i,j=0,1,\dots,n\,.$$

This implies that, for the vector $g^0=(g_0^0, g_1^0, \dots, g_n^0)$, $\partial g^0/\partial \tau_0$ and $\partial g^0/\partial \tau_{n+1}$ are linearly dependent. Notice that at least one of these two vectors does not vanish because of the condition (7.6). Then, since $\partial g_0^0/\partial \tau_0=\partial g_0^0/\partial \tau_{n+1}=0$, we may assume that either $\partial g_n^0/\partial \tau_0\neq 0$ or $\partial g_n^0/\partial \tau_{n+1}\neq 0$ after changing the indices of $G_1, \dots, G_n$ if necessary. Corresponding to these two cases, the functional independence (7.6) implies that either of the conditions

$$\det\frac{\partial(g_0^0,\dots,g_n^0)}{\partial(\tau_0,\dots,\tau_n)}\neq 0 \quad \text{and} \quad \frac{\partial g_n^0}{\partial \tau_0}\neq 0 \tag{7.7}$$

or

$$\det\frac{\partial(g_0^0,\dots,g_n^0)}{\partial(\tau_1,\dots,\tau_{n+1})}\neq 0 \quad \text{and} \quad \frac{\partial g_n^0}{\partial \tau_{n+1}}\neq 0 \tag{7.8}$$

holds for any t fixed.

Setting $\xi_{n+1}=t$, let us rewrite the equations (7.5) as

$$\sum_{j=0}^{n+1}\frac{\partial g_i}{\partial \tau_j}[\tau_j, W(\zeta,t)]=-P_R\hat{G}_i(\zeta,t)+O(|\zeta|^{s_i+2d}) \qquad (i=0,1,\dots,n)\,. \tag{7.9$_i$}$$

Here, in order to take derivatives $\partial g_i/\partial \tau_j$, g_i are considered as Laurent polynomials in $\tau_0, \tau_1, \dots, \tau_{n+1}$, where τ_j are given by (5.5) with x_i, y_i replaced by ξ_i, η_i.

The equation (7.9) can be considered as a system of $n+1$ equations for $n+2$ quantities $[\tau_0, W], \dots, [\tau_{n+1}, W]$. However, we can reduce (7.9) to n equations for $[\tau_1, W], \dots, [\tau_n, W]$, depending on which of the conditions (7.7) and (7.8) holds. Those equations will be given in Lemma 7.4 below. To state it, we introduce a Laurent polynomial $a_{ij}(\tau)$ as follows:

$$a_{ij}(\tau):=\begin{cases} P(\tau)\left(\dfrac{\partial g_n}{\partial \tau_0}\dfrac{\partial g_i}{\partial \tau_j}-\dfrac{\partial g_n}{\partial \tau_j}\dfrac{\partial g_i}{\partial \tau_0}\right) & \text{in the case (7.7)}\,; \\[2ex] P(\tau)\left(\dfrac{\partial g_n}{\partial \tau_{n+1}}\dfrac{\partial g_i}{\partial \tau_j}-\dfrac{\partial g_n}{\partial \tau_j}\dfrac{\partial g_i}{\partial \tau_{n+1}}\right) & \text{in the case (7.8)}\,, \end{cases} \tag{7.10}$$
$$(i=0,\dots,n-1, j=0,1,\dots,n+1)$$

where

$$P(\tau)=\mu_1^{-2}(\tau_1-\tau_0)^2\tau_{n+1}\,. \tag{7.11}$$

Here we note that $a_{ij}(\tau)$ can be considered as polynomials of $\tau_0, \dots, \tau_{n+2}$. Indeed, after writing $g_0, \dots, g_n$ as the sum of polynomials of the form (5.9), their differentiations with respect to $\tau_0, \ldots, \tau_{n+1}$ are calculated according to the following formula:

$$(7.12) \qquad \begin{cases} \dfrac{\partial}{\partial \tau_i} = D_{\tau_i} + \dfrac{\partial \tau_{n+2}}{\partial \tau_i} D_{\tau_{n+2}} \qquad (i=0, 1), \\ \dfrac{\partial}{\partial \tau_i} = D_{\tau_i} \qquad (i=2, \ldots, n), \quad \dfrac{\partial}{\partial \tau_{n+1}} = D_{\tau_{n+1}} + \dfrac{\partial \tau_{n+2}}{\partial \tau_{n+1}} D_{\tau_{n+2}} \end{cases}$$

with

$$(7.13) \qquad \frac{\partial \tau_{n+2}}{\partial \tau_0} = -\frac{p\mu_1 \tau_{n+1}}{\tau_1 - \tau_0}, \quad \frac{\partial \tau_{n+2}}{\partial \tau_1} = \frac{p\mu_1 \tau_{n+1}}{\tau_1 - \tau_0}, \quad \frac{\partial \tau_{n+2}}{\partial \tau_{n+1}} = -\frac{\tau_{n+2}}{\tau_{n+1}}.$$

Here D_{τ_i} ($i=0, \ldots, n+2$) denotes the derivative with respect to τ_i when $\tau_0, \ldots, \tau_{n+2}$ are considered as independent variables. Then it follows from (7.11) that $a_{ij}(\tau)$ are polynomials in $\tau_0, \ldots, \tau_{n+2}$.

Also, if we consider $a_{ij}(\tau)$ as polynomials in $\zeta = (\xi, \eta)$, they can be written as

$$a_{ij}(\tau) = a_{ij}^0 + a_{ij}^1 + \cdots,$$

where a_{ij}^d are homogeneous polynomials of degree $m_i + d$ in ζ with coefficients dependent on $\xi_{n+1} = t$. Here

$$(7.14) \qquad m_i = \begin{cases} s_n + s_i + p & \text{in the case (7.7)}; \\ s_n + s_i + 2 & \text{in the case (7.8)}, \end{cases}$$

and we note

$$a_{ij}^0(\tau) = \begin{cases} P(\tau)\left(\dfrac{\partial g_n^0}{\partial \tau_0}\dfrac{\partial g_i^0}{\partial \tau_j} - \dfrac{\partial g_n^0}{\partial \tau_j}\dfrac{\partial g_i^0}{\partial \tau_0}\right) & \text{in the case (7.7)}; \\ P(\tau)\left(\dfrac{\partial g_n^0}{\partial \tau_{n+1}}\dfrac{\partial g_i^0}{\partial \tau_j} - \dfrac{\partial g_n^0}{\partial \tau_j}\dfrac{\partial g_i^0}{\partial \tau_{n+1}}\right) & \text{in the case (7.8)}. \end{cases}$$

Now we describe the reduction of the equation (7.9) together with summary about the simultaneous normalization of $n+1$ functions $G_0, \ldots, G_n$.

LEMMA 7.4. *Under the assumptions of Theorem* 7.1, *let* $G_0 = \bar{H}$ *and let* $G_k(z, t)$ ($k=1, \ldots, n$) *be Poisson commuting integrals of* X_H *such that their lowest order parts* $G_1^0, \ldots, G_n^0$ *are functionally independent functions of* z *for each* t *fixed. Suppose that* G_0 *is in the normal form up to order* $s_0 + d - 1$. *Then, for the transformation* $\phi = \exp X_{\bar{W}}$ *described in Theorem* 6.3, *the functions* $G_k \circ \phi$ ($k=0, 1, \ldots, n$) *are in the normal form up to order* $s_k + 2d - 1$. *Furthermore, each homogeneous polynomial* W^{l+2} ($l=d, \ldots, 2d-1$) *satisfies the following system of equations*:

$$(7.15) \qquad \sum_{j=1}^{n} a_{ij}^0(\tau)[\tau_j, W^{l+2}] = F_i^l(\zeta, t) \qquad (i=0, 1, \ldots, n-1),$$

where

$$F_i^l(\zeta, t) := \begin{cases} \displaystyle\sum_{\nu=0}^{l-d} P(\tau)\left(\frac{\partial g_i^\nu}{\partial \tau_0} P_R \hat{G}_n^{l-\nu} - \frac{\partial g_n^\nu}{\partial \tau_0} P_R \hat{G}_i^{l-\nu}\right) - \sum_{j=1}^{n} \sum_{\nu=1}^{l-d} a_{ij}^\nu [\tau_j, W^{l+2-\nu}] \\ \qquad \text{in the case (7.7)}, \\ \displaystyle\sum_{\nu=0}^{l-d} P(\tau)\left(\frac{\partial g_i^\nu}{\partial \tau_{n+1}} P_R \hat{G}_n^{l-\nu} - \frac{\partial g_n^\nu}{\partial \tau_{n+1}} P_R \hat{G}_i^{l-\nu}\right) - \sum_{j=1}^{n} \sum_{\nu=1}^{l-d} a_{ij}^\nu [\tau_j, W^{l+2-\nu}] \\ \qquad \text{in the case (7.8)}. \end{cases}$$

Moreover, in both cases (7.7) *and* (7.8), *we have*

$$p(\tau) := \det(a_{ij}^0(\tau))_{i=0,\dots,n-1;\, j=1,\dots,n} \not\equiv 0 \tag{7.16}$$

and $D_k W^{l+2} := [\tau_k, W^{l+2}]$ *are expressed in the form*

$$D_k W^{l+2}(\zeta, t) = \frac{q_k^l(\zeta, t)}{p(\tau)} \qquad (k=1,\dots,n;\, l=d,\dots,2d-1) \tag{7.17}$$

with

$$q_k^l(\zeta, t) = \det \begin{array}{c} \qquad\qquad\quad (k-1) \\ \begin{pmatrix} a_{01}^0 & \cdots & F_0^l & \cdots & a_{0n}^0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n-1,1}^0 & \cdots & F_{n-1}^l & \cdots & a_{n-1,n}^0 \end{pmatrix} \end{array}. \tag{7.18}$$

REMARK. The formula (7.17) shows that the numerator $q_k^l(\zeta, t)$ is divisible by $p(\tau)$. It will play a key role in getting the estimate for W.

PROOF. We already proved the first assertion. To prove the second assertion in the case (7.7), we multiply $(7.9)_i$ and $(7.9)_n$ by $P(\tau)\partial g_n/\partial\tau_0$ and $P(\tau)\partial g_i/\partial\tau_0$ respectively and take their difference. Then we have n equations

$$\begin{aligned} \sum_{j=1}^{n+1} a_{ij}(\tau)[\tau_j, W] &= P(\tau)\left(\frac{\partial g_i}{\partial \tau_0} P_R \hat{G}_n - \frac{\partial g_n}{\partial \tau_0} P_R \hat{G}_i\right) \\ &\quad + P(\tau)\left(\frac{\partial g_n}{\partial \tau_0} O(|\zeta|^{s_i+2d}) - \frac{\partial g_i}{\partial \tau_0} O(|\zeta|^{s_n+2d})\right) \quad (i=0, 1, \dots, n-1). \end{aligned} \tag{7.19$_i$}$$

Here, using Lemma 5.2 and the identity $[G_n, G_i] \equiv 0$, we have

$$\frac{p\mu_1}{(\tau_1-\tau_0)^2} a_{i,n+1}(\tau) = [g_n, g_i] = -([g_n, \hat{G}_i] + [\hat{G}_n, g_i] + [\hat{G}_n, \hat{G}_i]) = O(|\zeta|^{s_n+s_i+d-2}).$$

Therefore, $a_{i,n+1}(\tau)[\tau_{n+1}, W] = O(|\zeta|^{m_i+2d+2})$ with $m_i = s_n + s_i + p$ and the second part of the right-hand side of $(7.19)_i$ has the same order estimate. Hence the system of equations $(7.19)_0$–$(7.19)_{n-1}$ can be written as

$$\sum_{j=1}^{n} a_{ij}(\tau)[\tau_j, W] = P(\tau)\left(\frac{\partial g_i}{\partial \tau_0} P_R \hat{G}_n - \frac{\partial g_n}{\partial \tau_0} P_R \hat{G}_i\right) + O(|\zeta|^{m_i+2d+2}) \tag{7.20}$$

$$(i=0, 1, \dots, n-1).$$

Comparing the homogeneous parts of degree m_i+l+2 in these equations, we obtain the system of equations (7.15) which is satisfied by the function W^{l+2}. Since the condition (7.7) implies (7.16), we have $p(\tau)\neq 0$ on an open and dense subset Ω of $C^{2n}\times R$, where $R\subset C$ is the neibourhood of the real t-axis given in Theorem 6.3. Therefore, we have the expression (7.17) with (7.18) for every $(\zeta, t)\in\Omega$. By the analyticity of the functions $D_k W^{l+2}$, q_k^l and p, the same expression holds for every $(\zeta, t)\in C^{2n}\times R$. This completes the proof in the case (7.7). The proof in the case (7.8) is the same as above and we obtain the equations (7.20) with $\partial g_i/\partial\tau_0$ replaced by $\partial g_i/\partial\tau_{n+1}$ $(i=0, 1, \dots, n)$. This leads to the assertions in the case (7.8). □

8. Estimate for W. In order to estimate the transformation $\phi=\exp X_{\bar{W}}$ described in Theorem 6.3, the basic task is to estimate the function W. In this section, we will estimate W (not $\bar{W}$) with respect to an appropriate norm. The idea is to make use of the formula (7.17) to estimate $D_k W=[\tau_k, W]$, which will lead to the estimate for W itself.

First of all, we will determine the domains in which all functions are to be considered. To this end, for appropriate positive constants $\delta_1, \dots, \delta_{2n+1}$ to be determined in Lemma 8.1 below, we introduce the following complex domains of C^{2n+1}:

$$\begin{cases} \Omega_r := \{(\zeta, t)\in C^{2n}\times C \mid |\zeta_i|<\delta_i r \ (i=1, \dots, 2n), |\operatorname{Im} t|<\delta_{2n+1} r\}, \\ \Delta_r := \{(\zeta, t)\in C^{2n}\times C \mid |\zeta_i|=\delta_i r \ (i=1, \dots, 2n), |\operatorname{Im} t|\leq\delta_{2n+1} r\}. \end{cases} \tag{8.1}$$

We note that for any holomorphic function in a neibourhood of Ω_r, the maximum of its absolute value on $\bar{\Omega}_r$ can be attained at a point on Δ_r. This can be easily seen by the repeated use of the maximum principle for holomorphic functions of one variable. From this fact, we can prove the following lemma.

LEMMA 8.1. *Let s be the degree of the polynomial $p(\tau)$ given by* (7.16), *where $p(\tau)$ is considered as a polynomial of $\zeta_1, \dots, \zeta_{2n}$ with coefficients being periodic functions of t. Then there exist constants δ_i $(i=1, \dots, 2n+1)$ such that $0<\delta_i<1$ and*

$$|p(\tau)|\geq c_1 r^s \quad on \quad \Delta_r, \tag{8.2}$$

where r is a small positive constant given arbitrarily and $c_1>0$ is a constant which is independent of r.

PROOF. This lemma corresponds to Lemma 4.1 of [7], in which the constants δ_k can be chosen in such a way that $\delta_i=\delta_{i+n}$ for $i=1, \dots, n$ because $p(\tau)$ is a polynomial of $\xi_i\eta_i$ only. This is not the case here and moreover, $p(\tau)$ depends on t. However, one can easily see that, choosing the constant $\delta_{2n+1}>0$ sufficiently small, its proof works also for the case here with trivial modifications (see the proof of Lemma 4.1 of [7]). □

Using the constants $\delta_1, \dots, \delta_{2n+1}$ given above, we define the complex domains Ω_r and Δ_r by (8.1). Let $A(\Omega_r)$ be the space of functions in (ζ, t) that are holomorphic in $\Omega_{r+\varepsilon}$ with some $\varepsilon>0$ and are periodic in t with period 1. Then a function $f\in A(\Omega_r)$ can be written as

$$f=f_0+f_1+f_2+\cdots,$$

where f_j are homogeneous polynomials of degree j in $\zeta_1, \dots, \zeta_{2n}$ with coefficients being Fourier series in t that are holomorphic in $|\mathrm{Im}\, t|<\delta_{2n+1}(r+\varepsilon)$. In view of Cauchy's estimate, the series $f\in A(\Omega_r)$ has a majorant which is a convergent power series in $\zeta_1, \dots, \zeta_{2n}$ with positive constant coefficients. This is because f is holomorphic in $\Omega_{r+\varepsilon}\supset\bar{\Omega}_r$. Therefore, we can define

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

Furthermore, we introduce the space

$$A_m(\Omega_r):=\{f\in A(\Omega_r)\,|\,f_j(\zeta,t)\equiv 0 \ \text{ for } j=0,1,\dots,m-1\}$$

and define

$$\|f\|_{r,m}:=\frac{\|f\|_r}{r^m} \qquad \text{for} \quad f\in A_m(\Omega_r).$$

Now let $G_k=G_k(\zeta,t)$ $(k=0,1,\dots,n)$ be functions that are in the normal form up to order s_k+d-1. We write those G_k in the form (7.3) with (z, x_{n+1}, y_{n+1}) replaced by $(\zeta, \xi_{n+1}, \eta_{n+1})$ and assume that $G_0-\eta_{n+1}, G_1, \dots, G_n$ belong to $A(\Omega_r)$, where we set $\xi_{n+1}=t$. Then it follows from Theorem 6.3 that $W^{l+2}\in A(\Omega_r)$ and hence $D_kW^{l+2}\in A(\Omega_r)$. Let us consider the maximum $\|D_kW^{l+2}\|_r=|D_kW^{l+2}|_r$. Since it can be attained at a point on $\Delta_r\subset\bar{\Omega}_r$, the following estimate follows from the identity (7.17) and the inequality (8.2):

$$\|D_kW^{l+2}\|_r\le\frac{\|q_k^l(\zeta,t)\|_r}{\min_{(\zeta,t)\in\Delta_r}|p(\tau)|}\le\frac{1}{c_1}\|q_k^l(\zeta,t)\|_{r,s}. \tag{8.3}$$

Using this formula, we will derive the estimate for $\|D_kW\|_r$ for $W=W^{d+2}+\cdots+W^{2d+1}$ First we prove:

LEMMA 8.2. *The function W satisfies*

$$\|D_kW^{l+2}\|_r\le c_2\sum_{i=0}^{n-1}\|F_i^l\|_{r,m_i} \qquad (k=1,\dots,n;\, l=d,\dots,2d-1),$$

where c_2 is a positive constant that is independent of r, and m_i are the constants given by (7.14).

PROOF. We note that the degree of the polynomial F_i^l is m_i+l+2 (in both cases (7.7) and (7.8)). The proof is based on the estimate (8.3) and the expansion of the determinant $q_k^l(\zeta, t)$ with respect to the $(k-1)$-th column consisting of $F_i^l(\zeta, t)$. It is the same as that of Lemma 4.2 of [7], so it is omitted. □

To proceed further, it is convenient to introduce some notation. Let $\tilde{f}$ be the power series in $\zeta_1, \dots, \zeta_{2n}$ with constant coefficients which is defined by

$$\tilde{f} := \sum_{\alpha,\beta} \max_{t\in R} |c_{\alpha\beta}(t)| \xi^\alpha \eta^\beta \qquad \text{for} \quad f = \sum_{\alpha,\beta} c_{\alpha\beta}(t) \xi^\alpha \eta^\beta ,$$

where

$$R = \{ t \in \boldsymbol{C} \,|\, 0 \le \operatorname{Re} t \le 1, |\operatorname{Im} t| \le \delta_{2n+1} r \} .$$

Let f be one of the following three vector functions:

$$\begin{aligned} & f = g := (g_0 - \eta_{n+1}, g_1, \dots, g_n) , \\ (8.4) \qquad & f = \hat{g} := (\hat{g}_0, \hat{g}_1, \dots, \hat{g}_n) \quad \text{with} \quad \hat{g}_i = g_i - g_i^0 , \\ & f = g^0 := (g_0^0 - \eta_{n+1}, g_1^0, \dots, g_n^0) . \end{aligned}$$

For this vector function $f = (f_0, f_1, \dots, f_n)$, let us define a constant $|||\partial f|||_r$ by

$$(8.5) \qquad |||\partial f|||_r := e^{2\pi|q|r} \sum_{i,j=0}^{n} \|\widetilde{D_{\tau_j} f_i}\|_{r, s_i - 2} + p e^{4\pi|q|r} \sum_{i=0}^{n} \sum_{j=n+1}^{n+2} \|\widetilde{D_{\tau_j} f_i}\|_{r, s_i - p} .$$

Here each element f_i is in the normal form and belongs to $A_{s_i}(\Omega_r)$ for η_{n+1} fixed, and therefore the quantity $|||\partial f|||_r$ is well defined. One can easily see that

$$(8.6) \qquad |||\partial g^0|||_r \le c_3 \qquad \text{for some constant} \quad c_3 > 0 .$$

Finally, we introduce the quantity

$$(8.7) \qquad |||\hat{G}|||_r := \sum_{i=0}^{n} \|\widetilde{\hat{G}_i}\|_{r, s_i - 2} .$$

Then, introducing a constant

$$(8.8) \qquad c_4 := \max(c_2, 1 + 2c_2 c_3) ,$$

we can estimate $\|D_k W\|_r$ as follows.

LEMMA 8.3. *Assume that*

$$(8.9) \qquad c_4 |||\partial \hat{g}|||_r < \frac{1}{2} .$$

Then the function $W(\zeta, t)$ *satisfies*

$$\|D_k W\|_r \leq 2c_4 |||\hat{G}|||_r \qquad (k=1, \dots, n). \tag{8.10}$$

PROOF. We can actually prove a better estimte than (8.10) with $\tilde{\hat{G}}_i$ replaced by $\hat{G}_i$ in the definition (8.7) of $|||\hat{G}|||$. More precisely, we can prove

$$\|D_k W\|_r \leq UV; \quad U := c_4 \sum_{l=0}^{d-1} (c_4 |||\partial \hat{g}|||_r)^l, \quad V := \sum_{i=0}^{n} \sum_{l=d}^{2d-1} \|P_R \hat{G}_i^l\|_{r, s_i - 2}, \tag{8.11}$$

from which (8.10) follows in view of (8.9) and the relation $\|P_R \hat{G}_i^j\|_{r, s_i-2} \leq \|\tilde{\hat{G}}_i^j\|_{r, s_i-2}$. We can prove (8.11) by an inductive argument in the same way as in the proof of Lemma 5.3 of [8], so we only give its outline below. We set

$$u_{i,j}^l := \begin{cases} e^{2\pi|q|r} \|D_{\tau_j} \tilde{g}_i^l\|_{r, s_i-2} & \text{for } j=0, 1, \dots, n, \\ p e^{4\pi|q|r} \|D_{\tau_j} \tilde{g}_i^l\|_{r, s_i-p} & \text{for } j=n+1, n+2 \end{cases} \qquad (i=0, 1, \dots, n)$$

and

$$v_i^{l+2} := \|P_R \hat{G}_i^l\|_{r, s_i-2}.$$

Then we can write $|||\partial \hat{g}|||_r = \sum_{l=1}^{d-1} \sum_{i=0}^{n} \sum_{j=0}^{n+2} u_{i,j}^l$ and consider U as the sum of 'monomials'

$$w := (\text{const.}) \prod_{\nu=1}^{m} u_{i_\nu, j_\nu}^{l_\nu},$$

where i_ν, j_ν, l_ν and m run over all integers satisfying $0 \leq i_\nu \leq n$, $0 \leq j_\nu \leq n+2$, $1 \leq l_\nu \leq d-1$ and $1 \leq m \leq d-1$. We define the degree of w as $\lambda := \sum_{\nu=1}^{m} l_\nu$. Also V is the sum of monomials v_i^{l+2} ($l=d, \dots, 2d-1$) whose degree we define as $l+2$, and we define the degree of the monomial $w v_i^{l+2}$ as $\lambda + l + 2$. For our purpose, it suffices to prove

$$\|D_k W^{l+2}\|_r \leq \text{the sum of monomials of degree } l+2 \text{ in } UV \quad (l=d, \dots, 2d-1). \tag{8.12}$$

In the following, we only consider the case (7.7) since the proof is the same in the case (7.8). In Lemma 8.2, we note that the function $F_i^l(\zeta, t)$ can be written as $F_i^l(\zeta, t) = A_i^l + B_i^l$, where

$$A_i^l = \sum_{\nu=0}^{l-d} P(\tau) \left(\frac{\partial g_i^\nu}{\partial \tau_0} P_R \hat{G}_n^{l-\nu} - \frac{\partial g_n^\nu}{\partial \tau_0} P_R \hat{G}_i^{l-\nu} \right), \quad B_i^l = -\sum_{j=1}^{n} \sum_{\nu=1}^{l-d} a_{ij}^\nu [\tau_j, W^{l+2-\nu}].$$

Since $|\tau_{n+1}|_r \leq e^{2\pi|q|r} r^p$, we can deduce from (7.11)–(7.14) that

$$\|A_i^l\|_{r, m_i} \leq \sum_{\nu=0}^{l-d} \{(u_{i,0}^\nu + u_{i,n+2}^\nu) v_n^{l+2-\nu} + (u_{n,0}^\nu + u_{n,n+2}^\nu) v_i^{l+2-\nu}\}.$$

In view of (8.6), this implies that

$$\sum_{i=0}^{n-1} \|A_i^l\|_{r, m_i} \leq \text{the sum of monomials of degree } l+2 \text{ in } (c_3 + |||\partial \hat{g}|||_r) V. \tag{8.13}$$

In the case $l=d$, we have $B_i^l=0$. Hence, by Lemma 8.2 and (8.8), the assertion (8.13) implies that (8.12) holds for $l=d$, since V contains monomials of degree $\geq d+2$ only.

To estimate B_i^l, using (7.11)–(7.14) we can similarly prove

$$e^{2\pi|q|r}\|a_{ij}^{\nu}\|_{r,m_i}\leq \sum_{\lambda+\mu=\nu}\{u_{i,j}^{\lambda}(u_{n,0}^{\mu}+u_{n,n+2}^{\mu})+u_{n,j}^{\lambda}(u_{i,0}^{\mu}+u_{i,n+2}^{\mu})\}$$

$$+\varepsilon_j\sum_{\lambda+\mu=\nu}(u_{n,0}^{\lambda}u_{i,n+2}^{\mu}+u_{n,n+2}^{\lambda}u_{i,0}^{\mu}),$$

where $\varepsilon_1=1$ and $\varepsilon_2=\cdots=\varepsilon_n=0$. This implies that

$$\sum_{i=0}^{n-1}\sum_{j=1}^{n}\|a_{ij}^{\nu}\|_{r,m_i}\leq \text{the sum of monomials of degree } \nu \text{ in } 2c_3|||\partial\hat{g}|||_r+|||\partial\hat{g}|||_r^2 .$$

Suppose that (8.12) holds for $l=d,\ldots,k-1$ ($\leq 2d-2$). Then the estimate above implies that

$$\sum_{i=0}^{n-1}\|B_i^k\|_{r,m_i}\leq \text{the sum of monomials of degree } k+2 \text{ in } (2c_3|||\partial\hat{g}|||_r+|||\partial\hat{g}|||_r^2)UV. \tag{8.14}$$

Here we note that

$$(2c_3|||\partial\hat{g}|||_r+|||\partial\hat{g}|||_r^2)U=2c_3(c_4|||\partial\hat{g}|||_r)+(c_4^{-1}+2c_3)\sum_{j=2}^{d+1}(c_4|||\partial\hat{g}|||_r)^j$$

and that V contains monomials of degree $\geq d+2$ only. From Lemma 8.2, we have

$$\|D_kW^{k+2}\|_r\leq c_2\sum_{i=0}^{n-1}(\|A_i^k\|_{r,m_i}+\|B_i^k\|_{r,m_i})$$

and therefore, in view of (8.8), the estimates (8.13) and (8.14) yield the inequality (8.12) with $l=k$. This completes the induction to prove (8.11). □

Now we derive estimates for W and its derivatives by using Lemma 8.3. To this end, we note that for a series $f\in\mathscr{P}$ its normal form part P_Nf is given by

$$P_Nf(\xi,\eta,t)=\int_0^1\cdots\int_0^1 f(e^{2\pi i\theta}\xi, e^{-2\pi i\theta}\eta, t+p\theta_1)d\theta_1\cdots d\theta_n . \tag{8.16}$$

Here $e^{2\pi i\theta}\xi$ and $e^{-2\pi i\theta}\eta$ are n-dimensional vectors defined by

$$\begin{aligned} e^{2\pi i\theta}\xi &:=(e^{2\pi iq\theta_1}\xi_1, e^{2\pi i\theta_2}\xi_2,\ldots,e^{2\pi i\theta_n}\xi_n),\\ e^{-2\pi i\theta}\eta &:=(e^{-2\pi iq\theta_1}\eta_1, e^{-2\pi i\theta_2}\eta_2,\ldots,e^{-2\pi i\theta_n}\eta_n),\end{aligned} \tag{8.17}$$

(the k (≥ 2)-th component of $e^{2\pi i\theta}\xi$ (resp. $e^{-2\pi i\theta}\eta$) is $e^{2\pi i\theta_k}\xi_k$ (resp. $e^{-2\pi i\theta_k}\eta_k$)) where $i=\sqrt{-1}$, $\theta_j\in\boldsymbol{R}$ and q, p are integers such that $\mu_1=2\pi iq/p$ and $p>0$. We set

$$\delta = \min_{1 \le i \le 2n+1} \delta_i, \quad c_5 = 2c_4 \max(2\pi, p).$$

Then we have:

LEMMA 8.4. *Let* $0 < \rho < r$. *Under the assumption of Lemma* 8.3, *the following holds*:

(i) $\|W\|_r \le c_5 |||\hat{G}|||_r$,

(ii) $\|W_{\zeta_i}\|_\rho \le \dfrac{c_5}{\delta(r-\rho)} |||\hat{G}|||_r \qquad (i=1, \dots, 2n+1, \zeta_{2n+1}=t)$,

(iii) $\|W_{\zeta_i \zeta_j}\|_\rho \le \dfrac{4c_5}{\delta^2(r-\rho)^2} |||\hat{G}|||_r \qquad (i, j=1, \dots, 2n+1, \zeta_{2n+1}=t)$.

PROOF. Let $(\zeta, t) \in \Omega_r$ and $\zeta = (\xi, \eta)$ be taken arbitrarily and fixed. Using the notation (8.17), we define the function

$$\hat{W}^{l+2}(\theta) := W^{l+2}(e^{2\pi i\theta}\xi, e^{-2\pi i\theta}\eta, t + p\theta_1), \qquad \theta = (\theta_1, \dots, \theta_n) \in [0, 1]^n.$$

Here we note that $(e^{2\pi i\theta}\xi, e^{-2\pi i\theta}\eta, t + p\theta_1) \in \Omega_r$ and $\hat{W}^{l+2}(0) = W^{l+2}(\zeta, t)$. Then by the mean value theorem we have

$$|\hat{W}^{l+2}(\theta) - \hat{W}^{l+2}(0)| \le \sum_{k=1}^{n} \max_{\theta \in [0,1]^n} \left| \frac{\partial \hat{W}^{l+2}}{\partial \theta_k} \right| \le 2\pi \sum_{k=2}^{n} |[W^{l+2}, \tau_k]|_r + p|[W^{l+2}, \tau_1]|_r.$$

In view of (8.16) and by the condition $P_N W^{l+2} = 0$, this lead to

$$|W^{l+2}(\zeta, t)| = |\hat{W}^{l+2}(0)| \le \max(2\pi, p) \sum_{k=1}^{n} |D_k W^{l+2}|_r$$

by integration from 0 to 1 with respect to $\theta_1, \dots, \theta_n$. By Lemma 8.3, this implies the desired estimate (i). The estimates (ii) and (iii) follow easily by using Cauchy's integral formula (see the proof of Lemma 4.4 of [7]). □

9. Estimates for one iteration step. We continue to consider the iteration step described in Lemma 7.4. In this section, we will give the estimates for $G_k \circ \phi$. They will be summarized in Lemmas 9.3–9.4 below.

We begin with estimates for domains transformed by $\phi = \exp X_{\bar{W}}$. Let $\phi^t(\zeta, \xi_{n+1}, \eta_{n+1})$ be the solution of the vector field $X_{\bar{W}}$ through $(\zeta, \xi_{n+1}, \eta_{n+1})$ at $t=0$. Let D_r be the domain in C^{2n+2} defined by

$$D_r := \Omega_r \times \{\eta_{n+1} \in C \mid |\eta_{n+1}| < r\}.$$

We introduce a constant

$$c_6 = 2nc_5\delta^{-2}$$

and prove:

LEMMA 9.1. *Let $G_0-\eta_{n+1}, G_1, \dots, G_n$ belong to $A(\Omega_r)$ and let $0<\sigma<\rho<r$ with $\rho-\sigma=r-\rho$. Assume that $c_4|||\partial\hat{g}|||_r<1/2$ and*

$$c_6(r-\rho)^{-2}|||\hat{G}|||_r<1\,.$$

Then, for any $(\zeta, \xi_{n+1}, \eta_{n+1})\in D_\sigma$ the solution $\phi^t(\zeta, \xi_{n+1}, \eta_{n+1})$ exists and is contained in D_ρ for all $t\in\boldsymbol{R}$ satisfying $|t|<2$. Moreover, the time-one map $\phi=\exp X_{\bar{W}}$ is a holomorphic transformation from D_σ into D_ρ.

PROOF. By setting $\phi^t(\zeta, \xi_{n+1}, \eta_{n+1})=(z(t), x_{n+1}(t), y_{n+1}(t))$, we have $\dot{x}_{n+1}=1$ and hence $x_{n+1}(t)=t+\xi_{n+1}$. This implies that $(z(t), y_{n+1}(t))$ is the solution of the system

$$\dot{z}=JW_z(z, t+\xi_{n+1})\,, \quad \dot{y}_{n+1}=-W_t(z, t+\xi_{n+1})\,. \tag{9.1}$$

By assumption and Lemma 8.4, we have an estimate

$$\|W_{z_k}\|\leq\frac{c_5}{\delta(r-\rho)}|||\hat{G}|||_r<\frac{1}{2n}\delta(r-\rho) \qquad (k=1, \dots, 2n+1, z_{2n+1}=t)\,. \tag{9.2}$$

Let $(\zeta, \xi_{n+1}, \eta_{n+1})\in D_\sigma$ and let D be its neighbourhood defined by

$$|z_k-\zeta_k|<\delta(r-\rho) \quad (k=1, \dots, 2n)\,, \quad |\mathrm{Im}(x_{n+1}-\xi_{n+1})|<\delta(\rho-\sigma)\,,$$
$$|y_{n+1}-\eta_{n+1}|<\rho-\sigma\,.$$

Then, in view of the estimate (9.2) the fundamental theorem for differential equations implies that the solution $(z(t), y_{n+1}(t))$ of (9.1) is a holomorphic function of t and (ζ, η) as far as $|t|\leq\delta(\rho-\sigma)/(2n)^{-1}\delta(r-\rho)=2n$. Hence, for the solution of the original vector field $X_{\bar{W}}$ for $t\in\boldsymbol{R}$, we have proved the desired assertion. □

Next we will estimate the difference between $\phi=\phi^1=\exp X_{\bar{W}}$ and its linearization at $(\zeta, \xi_{n+1}, \eta_{n+1})$. We set

$$(\zeta', \xi'_{n+1}, \eta'_{n+1})=\phi(\zeta, \xi_{n+1}, \eta_{n+1})$$

and write $\varphi^t(\zeta, \xi_{n+1})$ for the z-coordinate of $\phi^t(\zeta, \xi_{n+1}, \eta_{n+1})$, that is, φ^t for the flow of X_W. Then, considering the integral equation corresponding to $X_{\bar{W}}$, we have

$$\left\{\begin{aligned}
&\zeta'=\zeta+\int_0^1 JW_z(\varphi^t(\zeta, \xi_{n+1}), \xi_{n+1}+t)dt=\zeta+JW_\zeta(\zeta, \xi_{n+1})+R^1(\zeta, \xi_{n+1})\,,\\
&\xi'_{n+1}=\xi_{n+1}+1\,,\\
&\eta'_{n+1}=\eta_{n+1}-\int_0^1 W_{\xi_{n+1}}(\varphi^t(\zeta, \xi_{n+1}), \xi_{n+1}+t)dt\\
&\qquad=\eta_{n+1}-W_{\xi_{n+1}}(\zeta, \xi_{n+1})+R^2(\zeta, \xi_{n+1})\,,
\end{aligned}\right. \tag{9.3}$$

where $R^1(\zeta, \xi_{n+1})$ and $R^2(\zeta, \xi_{n+1})$ are a $2n$-dimensional vector function and a scalar function given respectively by

$$R^1(\zeta, \xi_{n+1}) = \int_0^1 J\{W_z(\varphi^t(\zeta, \xi_{n+1}), \xi_{n+1}+t) - W_z(\zeta, \xi_{n+1})\}dt\,,$$

$$R^2(\zeta, \xi_{n+1}) = -\int_0^1 \{W_{\xi_{n+1}}(\varphi^t(\zeta, \xi_{n+1}), \xi_{n+1}+t) - W_{\xi_{n+1}}(\zeta, \xi_{n+1})\}dt\,.$$

In the following, our estimates will be for functions on Ω_r and not for functions on D_r. From this viewpoint, although the function g_0 depends on η_{n+1}, we define

$$|g_0|_r := |g_0 - \eta_{n+1}|_r\,, \quad |\tilde{g}_0|_r := |\tilde{g}_0 - \eta_{n+1}|_r\,,$$

and the norm $|\cdot|_r$ will be used only for functions on Ω_r. Also, for a vector function $f=(f_1, \dots, f_m)$ with $f_k \in A(\Omega_r)$, we define the norms

$$|f|_r := \max_{1\le k\le m} |f_k|_r\,, \quad \|f\|_r := \max_{1\le k\le m} \|f_k\|_r\,.$$

Furthermore, for later use we introduce constants τ, κ, r' in addition to r, ρ in such a way that

$$r' < \kappa < \tau < \sigma < \rho < r \qquad \text{with}$$

$$r-\rho = \rho-\sigma = \sigma-\tau = \tau-\kappa = \kappa-r'\left(=\frac{1}{5}(r-r')\right).$$

Then, in view of Lemma 8.4 and Lemma 9.1, R^1 and R^2 are estimated as follows:

$$|R^1|_\tau \le 2|W_z|_\rho \le \frac{c_6\delta}{n(r-\rho)} |||\hat{G}|||_r\,, \quad |R^2|_\tau \le 2|W_{x_{n+1}}|_\rho \le \frac{c_6\delta}{n(r-\rho)} |||\hat{G}|||_r\,. \tag{9.4}$$

We now turn to the estimate for $G_k \circ \phi$. By the periodicity of G_k with respect to t with period 1, we can write

$$G_k \circ \phi(\zeta, \xi_{n+1}, \eta_{n+1}) = g_k(\zeta', \xi_{n+1}, \eta'_{n+1}) + \hat{G}_k(\zeta', \xi_{n+1}) \qquad (k=0, 1, \dots, n)\,.$$

From (9.3), we can estimate $\zeta'-\zeta$ as follows:

$$|\zeta'-\zeta|_\tau \le |W_z|_\rho \le \frac{c_6\delta}{2n(r-\rho)} |||\hat{G}|||_r\,.$$

Let δ_{k0} be the constant defined by

$$\delta_{k0} = \begin{cases} 1 & k=0\,, \\ 0 & k\neq 0\,. \end{cases}$$

Then, using Taylor's theorem and Cauchy's estimate, we have

$$g_k(\zeta', \xi_{n+1}, \eta'_{n+1}) = g_k(\zeta, \xi_{n+1}, \eta_{n+1}) + \left\langle \frac{\partial g_k}{\partial \zeta}, \zeta'-\zeta \right\rangle + \delta_{k0}(\eta'_{n+1}-\eta_{n+1}) + R_k^3(\zeta, \xi_{n+1})\,,$$

with

$$(9.5)\quad |R_k^3|_\tau \le \frac{1}{2}\sum_{i,j=1}^{2n}\left|\frac{\partial^2 g_k}{\partial\zeta_i\partial\zeta_j}\right|_\rho |\zeta'-\zeta|_\tau^2 \le 2n^2\cdot\frac{4|g_k|_r}{\delta^2(r-\rho)^2}|W_z|_\rho^2 \le 2c_6^2\cdot\frac{|g_k|_r |||\hat{G}|||_r^2}{(r-\rho)^4}.$$

Therefore, it follows from (9.3) that

$$(9.6)\quad g_k(\zeta', \xi_{n+1}, \eta'_{n+1}) = g_k + [g_k, W] + \left\langle\frac{\partial g_k}{\partial\zeta}, R^1\right\rangle + \delta_{k0}R^2 + R^3 \quad (k=0, 1, \dots, n),$$

where the arguments for the functions on the right-hand sides are $(\zeta, \xi_{n+1}, \eta_{n+1})$. Similarly we have

$$(9.7)\quad \hat{G}_k(\zeta', \xi_{n+1}) = \hat{G}_k(\zeta, \xi_{n+1}) + R_k^4(\zeta, \xi_{n+1}) \qquad (k=0, 1, \dots, n)$$

with

$$(9.8)\quad |R_k^4|_\tau \le \sum_{j=1}^{2n}\left|\frac{\partial\hat{G}_k}{\partial\zeta_j}\right|_\rho |\zeta'-\zeta|_\tau \le 2n\left|\frac{\partial\hat{G}_k}{\partial\zeta}\right|_\rho |W_z|_\rho \le c_6\cdot\frac{|\hat{G}_k|_r |||\hat{G}|||_r}{(r-\rho)^2}.$$

Let $G'_k = G_k\circ\phi$, and let g'_k and $\hat{G}'_k$ be its normal form part and its remainder part respectively. Then, by (9.6) and (9.7) together with (7.5) it can be written as

$$G'_k(\zeta, \xi_{n+1}, \eta_{n+1}) = g'_k(\zeta, \xi_{n+1}, \eta_{n+1}) + \hat{G}'_k(\zeta, \xi_{n+1})$$

with

$$(9.9)\quad \begin{cases} g'_k(\zeta, \xi_{n+1}, \eta_{n+1}) = g_k(\zeta, \xi_{n+1}, \eta_{n+1}) + P_N^{2d-1}\hat{G}_k(\zeta, \xi_{n+1}), \\ \hat{G}'_k(\zeta, \xi_{n+1}) = \left\langle\dfrac{\partial g_k}{\partial\zeta}, R^1\right\rangle + \delta_{k0}R^2 + R_k^3 + R_k^4 + R_k^5, \end{cases}$$

where

$$R_k^5 = [g_k, W] + \hat{G}_k(\zeta, \xi_{n+1}) - P_N^{2d-1}\hat{G}_k(\zeta, \xi_{n+1})$$

and the arguments for the functions on the right-hand sides of (9.9) are $(\zeta, \xi_{n+1}, \eta_{n+1})$.

In order to estimate $|||\hat{G}'|||_{r'}$, we need to estimate the majorants $\tilde{\hat{G}}'_k$. For this purpose, we will use the following:

LEMMA 9.2. *For $f\in A_m(\Omega_\tau)$, we have*

$$\text{(i)}\quad \|\tilde{f}\|_\kappa \le \frac{|f|_\tau}{\left(1-\dfrac{\rho}{r}\right)^{2n}}, \qquad \text{(ii)}\quad \|\tilde{f}\|_{r'} \le \frac{|f|_\tau}{\left(1-\dfrac{\rho}{r}\right)^{2n}}\left(\frac{r'}{\kappa}\right)^m.$$

PROOF. The assertion (i) corresponds to Lemma 6.2 of [8], where $\|\tilde{f}\|_\kappa$ is estimated from above by $|f|_\tau\{1-(\rho/r)\}^{-2n-1}$. Since $\|\tilde{f}\|_\kappa = |\tilde{f}|_\kappa$, we deduce the estimate (i) by Cauchy's estimate and by the relation $1-(\rho/r) < 1-(\kappa/\tau)$. The estimate (ii) is a direct consequence of Schwarz' lemma for functions beginning with terms of order m. $\square$

Recall that $\hat{G}_k'(\zeta, \xi_{n+1})$ consists of terms of order $\geq s_k+2d$ as power series in ζ. Then we have the following estimates from (9.4) and Lemma 9.2 (ii).

$$\left\| \left\langle \frac{\partial \tilde{g}_k}{\partial \zeta}, \widetilde{R^1} \right\rangle \right\|_{r'} \leq 2n \left| \frac{\partial \tilde{g}_k}{\partial \zeta} \right|_{\tau} \frac{|R^1|_{\tau}}{\left(1-\dfrac{\rho}{r}\right)^{2n}} \left(\frac{r'}{\kappa}\right)^{s_k+2d} \leq \frac{2}{3} \cdot \frac{c_6 |\tilde{g}_k|_r |||\hat{G}|||_r}{r^2\left(1-\dfrac{\rho}{r}\right)^{2n+2}} \left(\frac{r'}{\kappa}\right)^{s_k+2d},$$

$$\|\delta_{k0}\widetilde{R^2}\|_{r'} \leq \frac{\delta_{k0}|R^2|_{\tau}}{\left(1-\dfrac{\rho}{r}\right)^{2n}} \left(\frac{r'}{\kappa}\right)^{s_k+2d} \leq \delta_{k0} \cdot \frac{c_6\delta}{n} \cdot \frac{|||\hat{G}|||_r}{r\left(1-\dfrac{\rho}{r}\right)^{2n+1}} \left(\frac{r'}{\kappa}\right)^{s_k+2d}.$$

By introducing the symbol

$$|||g|||_r := \sum_{k=0}^{n} \|\tilde{g}_k\|_{r,s_k-2}$$

and a constant c_7 satisfying

$$\left(\frac{r}{r'}\right)^{s_k-2} \leq c_7 \qquad (k=0, 1, \ldots, n), \tag{9.10}$$

the estimates above lead to

$$\sum_{k=0}^{n} \left\| \left\langle \frac{\partial \tilde{g}_k}{\partial \zeta}, \widetilde{R^1} \right\rangle \right\|_{r', s_k-2} \leq \frac{2}{3} \cdot \frac{c_6 c_7 |||g|||_r |||\hat{G}|||_r}{r^2\left(1-\dfrac{\rho}{r}\right)^{2n+2}} \left(\frac{r'}{\kappa}\right)^{s_k+2d},$$

$$\sum_{k=0}^{n} \|\delta_{k0}\widetilde{R^2}\|_{r', s_k-2} \leq \frac{c_6\delta}{n} \cdot \frac{|||\hat{G}|||_r}{r\left(1-\dfrac{\rho}{r}\right)^{2n+1}} \left(\frac{r'}{\kappa}\right)^{s_0+2d}.$$

Also, applying Lemma 9.2 (i), we derive from (9.5) and (9.8) that

$$\sum_{k=0}^{n} \|\widetilde{R_k^3}\|_{r', s_k-2} \leq \frac{2c_6^2 c_7 |||g|||_r |||\hat{G}|||_r^2}{r^4\left(1-\dfrac{\rho}{r}\right)^{2n+4}}, \quad \sum_{k=0}^{n} \|\widetilde{R_k^4}\|_{r', s_k-2} \leq \frac{c_6 c_7 |||\hat{G}|||_r^2}{r^2\left(1-\dfrac{\rho}{r}\right)^{2n+2}}.$$

As for the estimate for $\|\widetilde{R_k^5}\|_{\kappa}$, we note that

$$\|\widetilde{R_k^5}\|_{\kappa} \leq \|\widetilde{[g_k, W]}\|_{\kappa} + \|\widetilde{\hat{G}_k}\|_{\kappa}$$

with

$$\|\widetilde{[g_k, W]}\|_{\kappa} \leq \left\| \sum_{j=1}^{n} \left(\frac{\partial \tilde{g}_k}{\partial \xi_j} \frac{\partial \tilde{W}}{\partial \eta_j} + \frac{\partial \tilde{g}_k}{\partial \eta_j} \frac{\partial \tilde{W}}{\partial \xi_j} \right) + \delta_{k0} \frac{\partial \tilde{W}}{\partial \xi_{n+1}} \right\|_{\kappa}.$$

Then, recalling that R_k^5 consists of terms of order $\geq s_k + 2d$, we are led to the following estimate by means of Cauchy's estimate and Lemma 8.4:

$$\sum_{k=0}^{n} \|\widetilde{R_k^5}\|_{r', s_k - 2} \leq \left\{ \frac{c_6 c_7 |||g|||_r}{r^2\left(1-\dfrac{\rho}{r}\right)^{2n+2}} + \frac{c_6 \delta}{2nr\left(1-\dfrac{\rho}{r}\right)^{2n+1}} + c_7 \right\} |||\hat{G}|||_r \left(\frac{r'}{\kappa}\right)^{2d+1} .$$

In the above, we note that there exists a constant $A > 0$, depending on g but independent of r, such that

$$|||g|||_r \leq A r^2$$

since $\tilde{g}_0 - \eta_{n+1}$ and $\tilde{g}_k$ begin with terms of order 2 and s_k respectively. Furthermore, we note that

$$1 - \frac{\rho}{r} = \frac{1}{5}\left(1 - \frac{r'}{r}\right) < 1 , \quad \frac{r'}{\kappa} = \frac{5r'}{4r' + r} .$$

Then, from (9.9) and all the estimates obtained above, one can easily prove the estimate for $|||\hat{G}'|||_{r'}$ described in the following lemma.

LEMMA 9.3. *In addition to the assumption of Lemma* 7.4, *assume that* $G_0 - \eta_{n+1}$, $G_1, \ldots, G_n$ *belong to* $A(\Omega_r)$ *and satisfy*

$$\text{(9.11)} \qquad |||g|||_r \leq A r^2 , \quad c_4 |||\partial \hat{g}|||_r < \frac{1}{2} \quad \text{and} \quad c_6 (r - \rho)^{-2} |||\hat{G}|||_r < 1 .$$

Let $0 < r' < r$. *Then* $\phi = \exp X_{\bar{W}}$ *is a holomorphic transformation from* D_σ *into* D_ρ, *where* $\sigma = r - (2/5)(r - r')$ *and* $\rho = r - (1/5)(r - r')$. *Furthermore, the remainder parts* $\hat{G}_k'$ *of* $G_k' := G_k \circ \phi$ *satisfy the following estimate*:

$$\text{(9.12)} \qquad |||\hat{G}'|||_{r'} \leq c_8 (A+1) |||\hat{G}|||_r \left\{ \frac{|||\hat{G}|||_r}{r^2\left(1-\dfrac{r'}{r}\right)^{2n+4}} + \frac{\left(\dfrac{5r'}{4r'+r}\right)^{2d+1}}{r\left(1-\dfrac{r'}{r}\right)^{2n+2}} \right\} ,$$

where $c_8 = c_8(r/r', n)$ *is a positive constant that increases with* r/r' *and* n.

REMARK. The constant c_8 can be given explicitly as $c_8 = 2c_6^2 c_7 \cdot 5^{2n+4}$.

As for the new normal form parts, we have to estimate $|||g'|||_{r'}$ and $|||\partial \hat{g}'|||_{r'}$. Similarly to what we had above, we can prove:

LEMMA 9.4. *Under the assumption of Lemma* 9.3, *we have*

$$\text{(i)} \qquad |||g'|||_{r'} \leq \left(\frac{r'}{r}\right)^2 \left\{ |||g|||_r + |||\hat{G}|||_r \left(\frac{r'}{r}\right)^d \right\} ,$$

(ii) $$|||\partial \hat{g}'|||_{r'} \le |||\partial \hat{g}|||_r + \frac{c_7 c_9}{\delta^{p+1}} \frac{|||\hat{G}|||_r}{rr'\left(1-\dfrac{r'}{r}\right)},$$

where

$$c_9 = c_9(r') = \left(\frac{2p}{\pi|q|} + n - 1\right) e^{2\pi|q|r'} + 2e^{6\pi|q|r'}. \tag{9.13}$$

PROOF. Since the new normal form parts g'_k are given by (9.9), we have

$$\|\widetilde{g'_k}\|_{r'} \le \|\tilde{g}_k\|_{r'} + \|P_N^{2d-1}\widetilde{\hat{G}_k}\|_{r'} \le \|\tilde{g}_k\|_{r'} + \|\widetilde{\hat{G}_k}\|_{r'}.$$

Application of Schwarz' lemma to $\tilde{g}_k$ and $\widetilde{\hat{G}_k}$ leads to the estimate (i). To get the estimate for $|||\partial \hat{g}'|||_{r'}$, we need to estimate $\|D_{\tau_j}(P_N^{2d-1}\widetilde{\hat{G}_k})\|_{r', s_k-2}$ $(j=1, \dots, n)$ and $\|D_{\tau_j}(P_N^{2d-1}\widetilde{\hat{G}_k})\|_{r', s_k-p}$ $(j=n+1, n+2)$ according to the definition (8.5) of $|||\partial f|||_r$. We recall that any power series f in the normal form can be written in the form (5.9). Then we have

$$D_{\tau_1} f = -D_{\tau_0} f = \frac{1}{\mu_1}\left(\frac{1}{x_1}\frac{\partial f_1}{\partial y_1} + \frac{1}{y_1}\frac{\partial f_2}{\partial x_1}\right), \quad D_{\tau_j} f = \frac{1}{y_j}\frac{\partial f}{\partial x_j} \quad (j=2, \dots, n).$$

$$D_{\tau_{n+1}} f = \frac{1}{p\tau_{n+1}}\left(x_1\frac{\partial f_1}{\partial x_1} - y_1\frac{\partial f_1}{\partial y_1}\right), \quad D_{\tau_{n+2}} f = \frac{1}{p\tau_{n+2}}\left(y_1\frac{\partial f_2}{\partial y_1} - x_1\frac{\partial f_2}{\partial x_1}\right).$$

Assume that $f = \tilde{f}$. Then $\|f\|_r = \|f_1\|_r + \|f_2\|_r$. Also recall that for any holomorphic function in $\Omega_{r'}$ the maximum of its absolute value is attained at a point on $\Delta_{r'}$. Then using Cauchy's integral formula, we deduce from the above formulas that

$$\|D_{\tau_0} f\|_{r'} + \|D_{\tau_1} f\|_{r'} \le \frac{p}{\pi|q|\delta^2} \cdot \frac{\|f\|_r}{r'(r-r')}, \quad \|D_{\tau_j} f\|_{r'} \le \frac{\|f\|_r}{\delta^2 r'(r-r')} \quad (j=2, \dots, n),$$

$$\|D_{\tau_{n+1}} f\|_{r'} + \|D_{\tau_{n+2}} f\|_{r'} \le \frac{2e^{2\pi|q|r'}}{p\delta^{p+1}(r')^{p-1}} \cdot \frac{\|f\|_r}{r-r'}.$$

Let $P_N^{2d-1}\widetilde{\hat{G}}$ be the n-dimensional vector whose components are $P_N^{2d-1}\widetilde{\hat{G}_i}$ $(i=1, \dots, n)$. Then, using the above formulas and definition (8.5), one can prove

$$|||\partial(P_N^{2d-1}\widetilde{\hat{G}})|||_{r'} \le c_7\left(\frac{2pe^{2\pi|q|r'}}{\pi|q|\delta^2} + \frac{n-1}{\delta^2}e^{2\pi|q|r'} + \frac{2e^{6\pi|q|r'}}{\delta^{p+1}}\right)\frac{|||\hat{G}|||_r}{r'(r-r')},$$

which leads to the estimate (ii) with (9.13). □

10. The convergence proof. We will finally prove the convergence of the iteration process described in Corollary 6.4. It will complete the proof of Theorem 7.1 and hence that of Theorem 4.4.

Let us consider the estimates obtained in the preceding section. To consider the ν-th iteration step ($\nu=0, 1, \dots$), we set

$$g_k=g_k^{(\nu)}, \quad g'_k=g_k^{(\nu+1)}, \quad \hat{G}_k=\hat{G}_k^{(\nu)}, \quad \hat{G}'_k=\hat{G}_k^{(\nu+1)}, \quad \phi=\phi_\nu$$

and

$$r=r_\nu, \quad r'=r_{\nu+1}, \quad d=2^\nu.$$

The ν-th iteration step consists of taking $G_k^{(\nu)}:=g_k^{(\nu)}+\hat{G}_k^{(\nu)}$ into $G_k^{(\nu+1)}:=g_k^{(\nu+1)}+\hat{G}_k^{(\nu+1)}$ by the transformation ϕ_ν. Also we replace the symbols $|||\hat{G}|||_r$, $|||g|||_r$, $|||\partial\hat{g}|||_r$ by $|||\hat{G}^{(\nu)}|||_{r_\nu}$, $|||g^{(\nu)}|||_{r_\nu}$, $|||\partial\hat{g}^{(\nu)}|||_{r_\nu}$ respectively. Then, by Lemma 9.3 we have

$$\phi_\nu\colon \Omega_{\sigma_\nu}\to\Omega_{\rho_\nu}; \quad \sigma_\nu=r_\nu-\frac{2}{5}(r_\nu-r_{\nu+1}), \quad \rho_\nu=r_\nu-\frac{1}{5}(r_\nu-r_{\nu+1})$$

and (9.12) gives the estimate for $|||\hat{G}^{(\nu+1)}|||_{r_{\nu+1}}$ in terms of $|||\hat{G}^{(\nu)}|||_{r_\nu}$.

Our purpose is to prove that, with an appropriate choice of the sequence $\{r_\nu\}$, the sequence of transformations ϕ_ν is well-defined and their composite $\phi^{(\nu)}:=\phi_0\circ\phi_1\circ\cdots\circ\phi_\nu$ is uniformly convergent. To this end, we define the sequence $\{r_\nu\}$ by

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

Then we have

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

and hence

$$\frac{r_\nu}{r_{\nu+1}}\le\frac{4}{3} \qquad \text{for} \quad \nu=0, 1, 2, \dots. \tag{10.1}$$

This implies that the constant c_7 in (9.10) can be taken independently of the iteration step. Also, since $r_\nu\le r_0$, the constant $c_9=c_9(r_{\nu+1})$ defined by (9.13) can be taken independently of the iteration step and is assumed to satisfy the following inequality by choosing r_0 suitably small:

$$c_9(r_\nu)<2\left(\frac{2p}{\pi|q|}+n-1\right) \qquad (\nu=1, 2, \dots). \tag{10.2}$$

Let A_0 be a constant satisfying the condition

$$|||g^0|||_{r_\nu}\le A_0 r_\nu^2,$$

where A_0 can be taken independently of r_ν. Moreover we set

$$c_{10}=c_8\gamma\left(A_0+\frac{3}{2}\right) \qquad \text{with} \quad \gamma=\max(1, \delta^{3-p}).$$

The iteration procedure with this sequence $\{r_\nu\}$ will be justified by proving the three conditions given by (9.11) at each step. Instead of those conditions, we will prove that $|||\hat{G}^{(\nu)}|||_{r_\nu}$ satisfies the condition

$$c_{10}\varepsilon_\nu < 1\,; \quad \varepsilon_\nu = \frac{|||\hat{G}^{(\nu)}|||_{r_\nu}}{r_\nu^2\left(1-\dfrac{r_{\nu+1}}{r_\nu}\right)^{2n+4}} \tag{10.3}$$

for all $\nu \geq 0$. By the definitions of constants c_6 and c_{10}, this condition is clearly stronger than the third one in (9.11). Moreover, it also implies the first and the second conditions of (9.11), as is shown by the following lemma.

LEMMA 10.1. *Let $r_0 > 0$ be small enough to satisfy* (10.2). *Assume that the condition* (10.3) *holds for* $\nu = 0, 1, \ldots, m$. *Then*

(i) $$|||g^{(\nu)}|||_{r_\nu} < \left(A_0 + \frac{1}{2}\right) r_\nu^2 \qquad \textit{for} \quad \nu = 0, 1, \ldots, m+1\,;$$

(ii) $$c_4 |||\partial \hat{g}^{(\nu)}|||_{r_\nu} < \frac{1}{2} \qquad \textit{for} \quad \nu = 0, 1, \ldots, m+1\,.$$

In the above, the assertion (i) implies that the constant $A = A(g)$ in the first condition of (9.11) can be chosen as $A = A_0 + (1/2)$ independently of the iteration step. This lemma is a direct consequence of Lemma 9.4. The proof is the same as that of Lemma 7.1 of [8] and is omitted. Furthermore, by Lemma 9.3 one can prove that

$$\varepsilon_{\nu+1} \leq c\varepsilon_\nu(\varepsilon_\nu + \lambda_\nu)\,,$$

where

$$c = c_{10}\left(\frac{4}{3}\right)^2 \left(\frac{3}{2}\right)^{2(2n+4)}, \quad \lambda_\nu = (\nu+2)^{2(2n+2)}\left(1 - \frac{1}{5(\nu+2)^2 - 4}\right)^{2^{\nu+1}+1}$$

This implies that $c\varepsilon_\nu < 1$ for all $\nu \geq 0$ if $r_0 > 0$ is chosen sufficiently small, and in particular that there exists an integer $N > 0$ such that

$$c\varepsilon_\nu < (4c)^{-1} 2^{-\nu+N} \qquad \text{for} \quad \nu \geq N\,.$$

For the proof, we refer to [8]. Thus we have justified the iteration procedure and see that $|||\hat{G}^{(\nu)}|||_{r_\nu} \to 0$ as $\nu \to \infty$. Also it is easy to see that $\phi^{(\nu)} = \phi_0 \circ \phi_1 \circ \cdots \phi_\nu$ converges uniformly to a symplectic transformation which is analytic in the domain $\Omega_{r_0/2}$ (see [7], [8]). This completes the proof of Theorem 7.1 and hence that of Theorem 4.4.

11. Proof of Theorem 2. We will finally prove Theorem 2 using Theorem 4.4. We will give the proof only in the semisimple case. In the non-semisimple case, we leave its proof to the reader (see [8]).

First we see that the real quadratic form (2.4) is taken into $H_2(\zeta) = S(z) = \sum_{k=1}^n \mu_k \xi_k \eta_k$

by the following (complex) linear symplectic transformation:

$$z=C\zeta\,; \quad z={}^t(p,u,x,q,v,y)\,, \quad \zeta={}^t(\xi,\eta)\in \boldsymbol{C}^n\times\boldsymbol{C}^n$$

with

$$\begin{cases} p_j=\xi_j\,, \quad q_j=\eta_j \qquad (j=1,\dots,k)\,, \\ u_j=\dfrac{1}{\sqrt{2}}(\xi_{k+j}+i\eta_{k+j})\,, \quad v_j=\dfrac{1}{\sqrt{2}}(\eta_{k+j}+i\xi_{k+j}) \qquad (j=1,\dots,l)\,, \\ x_{2j-1}=\dfrac{1}{\sqrt{2}}(\xi_{r+2j-1}+\xi_{r+2j})\,, \quad y_{2j-1}=\dfrac{1}{\sqrt{2}}(\eta_{r+2j-1}+\eta_{r+2j})\,, \\ x_{2j}=\dfrac{i}{\sqrt{2}}(-\xi_{r+2j-1}+\xi_{r+2j})\,, \quad y_{2j}=\dfrac{i}{\sqrt{2}}(\eta_{r+2j-1}-\eta_{r+2j}) \\ \qquad\qquad\qquad\qquad\qquad\qquad (r=k+l;\, j=1,\dots,m)\,. \end{cases}$$

Then $z\in \boldsymbol{R}^{2n}$ if and only if

$$\begin{cases} \xi_j\,,\ \eta_j\in \boldsymbol{R} \qquad (j=1,\dots,k)\,, \\ \xi_{k+j}=-i\bar{\eta}_{k+j} \qquad (j=1,\dots,l)\,, \\ \xi_{r+2j}=\bar{\xi}_{r+2j-1}\,, \quad \eta_{r+2j}=\bar{\eta}_{r+2j-1} \qquad (j=1,\dots,m)\,. \end{cases}$$

If a function $f(\zeta,t)\in\mathscr{A}(\boldsymbol{C}^{2n+1},S^1)$ is written as $f(\zeta,t)=\sum_{\alpha,\beta}c_{\alpha\beta}(t)\xi^\alpha\eta^\beta$, we say that f satisfies the reality condition if the following relation holds:

$$(11.1) \qquad f(\zeta,t)=\bar{f}(T\zeta,t) \qquad \text{for} \quad T=\bar{C}^{-1}C\,,$$

where $\bar{f}$ is the power series in ζ defined by

$$\bar{f}(\zeta,t)=\sum_{\alpha,\beta}\bar{c}_{\alpha,\beta}(t)\xi^\alpha\eta^\beta\,,$$

with $\bar{c}_{\alpha\beta}(t)$ being Fourier series of t obtained from $c_{\alpha,\beta}(t)$ with coefficients replaced by their complex conjugates. The condition (11.1) is equivalent to the requirement that $f(C^{-1}z,t)$ is a real analytic function of z.

Let $H(\zeta,t)\in\mathscr{A}(\boldsymbol{C}^{2n+1},S^1)$ be the Hamiltonian in Theorem 4.4 which is obtained from a real analytic function as in Theorem 2 by the transformation $z=C\zeta$. Then $H(\zeta,t)$ clearly satisfies the reality condition. Also we say that a symplectic transformation $\phi\in\mathscr{S}$, which is of the form (4.5) and (4.6), satisfies the reality condition if the following relation holds:

$$\bar{\varphi}(T\zeta,\xi_{n+1})=T\varphi(\zeta,\xi_{n+1})\,,$$

where the meaning of $\bar{\varphi}$ is the same as that of $\bar{f}$. Then, in the same way as in Section 7 of [7] one can prove that the function $W(\zeta,t)$ and the transformation $\phi=\exp X_{\bar{W}}$ in

Theorem 6.3 satisfy the reality condition. This implies that $\phi=\lim_{\nu\to\infty}\phi^{(\nu)}$ obtained by Theorem 7.1, as well as the normal form $\bar{H}\circ\phi-\eta_{n+1}$, satisfies the reality condition. (Here the meaning of the "bar" of $\bar{H}$ is: $\bar{H}=H+y_{n+1}$.) Since ϕ has the form (4.5) with (4.6), we obtain a time-dependent real analytic symplectic transformation $C\circ\varphi\circ C^{-1}$, which is the desired real analytic transformation in Theorem 2. This completes the proof of Theorem 2.

References

[1] V. I. Arnol'd, Mathematical Methods of Classical Mechanics, 2nd ed., Graduate Texts in Math. 60, Berlin–Heidelberg–New York, Springer-Verlag, 1989.

[2] G. D. Birkhoff, Surface transformations and their dynamical applications, Acta Math. 43 (1920), 1–119.

[3] G. D. Birkhoff, Dynamical Systems, Amer. Math. Soc. Colloqium Publ. 1927, revised edition 1966.

[4] T. J. Bridges and R. H. Cushman, Unipotent normal forms for symplectic maps, Phisica D 65 (1993), 211–241.

[5] D. M. Galin, Versal deformations of linear hamiltonian systems, Amer. Math. Soc. Transl. Ser 2 118 (1982), 1–12.

[6] J. E. Humphreys, Introduction to Lie algebras and representation theory, 2nd printing. Graduate Texts in Math. 9, Berlin-Heidelberg-New York, Springer-Verlag, 1972.

[7] H. Ito, Convergence of Birkhoff normal forms for integrable systems, Comment. Math. Helv. 64 (1989), 412–461.

[8] H. Ito, Integrability of Hamiltonian systems and Birkhoff normal forms in the simple resonance case, Math. Ann. 292 (1992), 411–444.

[9] S. B. Kuksin, and J. Pöschel, On the inclusion of analytic symplectic maps in analytic Hamiltonian flows and its applications, in Seminar on Dynamical Systems (S. Kuksin, V. Lazutkin, J. Pöschel, eds.), Birkhäuser, (1994), 96–116.

[10] J. Moser, Lectures on Hamiltonian systems, Mem. Amer. Math. Soc. 81, (1968).

[11] J. C. van der Meer, The Hamiltonian Hopf Bifurcation, Lecture Notes in Math. 1160, Berlin-Heidelberg-New York, Springer-Verlag, (1985).

[12] J. Williamson, On the normal forms of linear canonical transformations in dynamics, Amer. J. Math. 59 (1937), 599–617.

[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