TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 331, Number 1, May 1992

NORMAL FORM AND LINEARIZATION FOR QUASIPERIODIC SYSTEMS

SHUI-NEE CHOW, KENING LU AND YUN-QIU SHEN

ABSTRACT. In this paper, we consider the following system of differential equations:

$$\dot{\theta} = \omega + \Theta(\theta, z), \qquad \dot{z} = Az + f(\theta, z),$$

where $\theta \in C^m$, $\omega = (\omega_1, \dots, \omega_m) \in R^m$, $z \in C^n$, A is a diagonalizable matrix, f and Θ are analytic functions in both variables and 2π-periodic in each component of the vector θ, $\Theta = O(|z|)$ and $f = O(|z|^2)$ as $z \to 0$. We study the normal form of this system of the equations and prove that this system can be transformed to a system of linear equations

$$\dot{\theta} = \omega, \qquad \dot{z} = Az$$

by an analytic transformation provided that the eigenvalues of A and the frequency ω satisfy certain small-divisor conditions.

1. INTRODUCTION

Poincaré normal form theory plays an important role in the study of existence, stability, approximation and bifurcation of solutions of differential equations. This theory is well known for differential equations in the neighborhood of an equilibrium point or a periodic motion and may be found in Arnold [1], Chow and Hale [4], Guckenheimer and Holmes [5], and Meyer [6]. For differential equations in the neighborhood of invariant tori, the reader may find the normal form theory, for example, in the recent works of Braaksma and Broer [3] and B. L. J. Braaksma, H. W. Broer and G. B. Huitema [10].

In this paper, we consider the following system of differential equations:

$$\dot{\theta} = \omega + \Theta(\theta, z), \tag{1.1}$$

$$\dot{z} = Az + f(\theta, z), \tag{1.2}$$

where $\theta \in C^m$, $\omega = (\omega_1, \dots, \omega_m) \in R^m$, $z \in C^n$, A is a diagonalizable matrix, f and Θ are analytic functions in both variables and 2π-periodic in each component of the vector θ, $\Theta = O(|z|)$ and $f = O(|z|^2)$ as $z \to 0$. Without loss of generality, we can assume $A = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$.

The idea of normal form theory is to find a transformation which changes the system of equations (1.1) and (1.2) into the "simplest" one. A special case occurs when the system of equations (1.1) and (1.2) can be changed into a linear system.

Received by the editors February 26, 1990.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 58F27, 34C27, 34C50, 34C29.
The first author partially supported by NSF grant DMS8401719 and DARPA.

If the following small-divisor conditions are satisfied

$$|i\langle \omega, k\rangle + \langle \alpha, \lambda\rangle - \varepsilon\lambda_j| \geq \frac{C_0}{(|k| + |\alpha|)^\mu}, \quad \text{for } j = 1, \dots, n,$$

where k is an integer vector, α is a nonnegative integer vector with $1 + \varepsilon \leq |\alpha|$, $\varepsilon = 0$ or 1, $\langle\ ,\ \rangle$ is the scalar product, $|k| = |k_1| + \cdots + |k_m|$, and $|\alpha| = \alpha_1 + \cdots + \alpha_n$, then we will prove that the system of equations (1.1) and (1.2) can be transformed to a system of linear equations

$$\dot\theta = \omega, \qquad \dot z = Az$$

by an analytic transformation.

For differential equations with no angle variables, the above result is the well-known Siegel linearization theorem. For the general case, the result was announced by Belaga [2] without proof.

In [1] Arnold gives a proof of Siegel's theorem for the analytic map $Lx = Ax + f(x)$. In [9] Zehnder gives a very nice, simple proof of Siegel's Theorem using a different approach. However, the method used by Zehnder, in fact, requires that $f(x)$ be a high order term (much higher than 2). But this can be carried out by normal form theory.

The purpose of this paper is to present a normal form theory for the system of equations (1.1) and (1.2) and a proof of the linearization theorem for the equations (1.1) and (1.2). The proof of the linearization is based on the normal form theory we present and the method in [9].

The organization of this paper is as follows. In §2, we give the basic lemmas; in §3, we present a normal form theorem; and in §4, we give a proof of the linearization theorem.

Acknowledgement. We would like to thank H. Scott Dumas for his helpful suggestions.

2. Basic lemmas

Set $D_r = \{(\theta, z) \in C^m \times C^n \,|\, |\mathrm{Im}\,\theta_j| < r \text{ for } j = 1, \dots, m \text{ and } |z| < r\}$ for $r \leq 1$. Let $f(\theta, z)$ be bounded analytic and 2π-periodic in $\theta_1, \dots, \theta_m$ in D_r. We define

$$|f|_r = \sup_{(\theta, z)\in D_r} |f(\theta, z)|. \tag{2.1}$$

Using the Fourier-Taylor expansion, we have

$$f = \sum_{\alpha\in N_n^+} \sum_{k\in N_m} f_{\alpha,k} e^{i\langle\theta, k\rangle} z^\alpha, \tag{2.2}$$

where

$$N_n^+ = \{\alpha = (\alpha_1, \dots, \alpha_n) | \alpha_j \geq 0 \text{ is an integer for } j = 1, \dots, n\},$$
$$N_m = \{k = (k_1, \dots, k_m) | k_j \text{ is an integer for } j = 1, \dots, m\},$$
$$\langle\theta, k\rangle = \theta_1 k_1 + \cdots + \theta_m k_m.$$

Define

$$\hat f(\theta, z) = \sum_{\alpha\in N_n^+} \sum_{k\in N_m} \frac{f_{\alpha,k}}{C_{\alpha,k}} e^{i\langle\theta, k\rangle} z^\alpha, \tag{2.3}$$

where $C_{\alpha,k}$ are constants satisfying

$$|C_{\alpha,k}| \geq \frac{C_0}{(|k|+|\alpha|)^{\mu}} \tag{2.4}$$

for some positive constants C_0 and μ.

The conditions (2.4) are small-divisor conditions. Various small-divisor problems can be found in [1, 7 and 9]. One of the basic ways to solve these problems is to prove that the series $\hat{f}$ is convergent and to get estimates on $\hat{f}$ in terms of f and the domain. For our case, we discuss the Fourier-Taylor series (2.3).

Lemma 2.1. *Let f be bounded analytic and 2π-periodic in each component of θ in D_r. Then $|f_{\alpha,k}| \leq |f|_\rho \rho^{-|\alpha|}\varepsilon^{-\rho|k|}$ for any $0 < \rho \leq r$, where $|\alpha| = \alpha_1 + \cdots + \alpha_n$, $|k| = |k_1| + \cdots + |k_m|$.*

Proof. Let $f_\alpha = \sum_{k\in N_m} f_{\alpha,k} e^{i\langle\theta,k\rangle}$. Then

$$f_{\alpha,k} = \frac{1}{(2\pi)^m}\int_0^{2\pi}\cdots\int_0^{2\pi} f_\alpha e^{-i\langle\theta,k\rangle}\,d\theta_1\cdots d\theta_m.$$

Since f_α is analytic and 2π-periodic in $\theta_1, \dots, \theta_m$, by Cauchy's theorem, the path of integration in the above integral can be shifted to $\theta_j = x_j \pm i\rho$, $0 < \rho < r$, $0 \leq x_j \leq 2\pi$ $(j = 1, \dots, m)$, and choosing the sign equal to the sign of $-k_j$, if $k_j \neq 0$ and arbitrarily if $k_j = 0$, we have

$$|f_{\alpha,k}| \leq \sup_{|\operatorname{Im}\theta|\leq\rho} |f_\alpha(\theta)| e^{-\rho|k|}. \tag{2.5}$$

It is clear that

$$f_\alpha(\theta) = \frac{1}{|\alpha|!}\left.\frac{\partial^\alpha f(\theta, z)}{\partial z^\alpha}\right|_{z=0}.$$

Using a Cauchy estimate, we have

$$|f_\alpha(\theta)| \leq |f|_\rho \rho^{-|\alpha|}. \tag{2.6}$$

The conclusion follows from (2.5) and (2.6) when $0 < \rho < r$ and also holds when $\rho = r$ by the boundedness of f.

Lemma 2.2. (i) *$\hat{f}$ is analytic, 2π-periodic in each component of θ in $D_{r-\sigma}$.*

(ii) *$|\hat{f}|_{r-\sigma} \leq C_1|f|_r\sigma^{-(m+n+\mu+1)}$, where $C_1 = C_1(C_0, m, n, \mu) > 0$ is constant, and $0 < \sigma < r$.*

Proof. For each $(\theta, z) \in D_{r-\sigma}$, using Lemma 2.1 and the small-divisor conditions (2.4), we obtain

$$\begin{aligned}
|\hat{f}| &\leq \sum_{\alpha\in N_n^+}\sum_{k\in N_m} \frac{|f_{\alpha,k}|}{|C_{\alpha,k}|}|e^{i\langle\theta,k\rangle}||z|^\alpha \\
&\leq \sum_{\alpha\in N_n^+}\sum_{k\in N_m} |f_{\alpha,k}|\frac{1}{C_0}(|\alpha|+|k|)^\mu e^{(r-\sigma)|k|}(r-\sigma)^{|\alpha|} \\
&\leq \sum_{\alpha\in N_n^+}\sum_{k\in N_m} |f|_r r^{-|\alpha|}e^{-r|k|}\frac{1}{C_0}(|\alpha|+|k|)^\mu e^{(r-\sigma)|k|}(r-\sigma)^{|\alpha|} \\
&\leq \frac{1}{C_0}|f|_r \sum_{\alpha\in N_n^+}\sum_{k\in N_m}\left(\frac{r-\sigma}{r}\right)^{|\alpha|}(|\alpha|+|k|)^\mu e^{-\sigma|k|}.
\end{aligned}$$

Since

$$\left(\frac{r-\sigma}{r}\right)^{|\alpha|} = e^{|\alpha|[\ln(r-\sigma)-\ln r]} \le e^{-\sigma|\alpha|},$$

the last term is less than or equal to

$$\frac{1}{C_0}|f|_r \sum_{\alpha\in N_n^+} \sum_{k\in N_m} (|\alpha|+|k|)^{\mu} e^{-(|k|+|\alpha|)\sigma}.$$

Since the number of integer vectors (α, k) satisfying $|\alpha|+|k| = j$ is not bigger than j^{n+m}, we have

$$\begin{aligned}
&\frac{1}{C_0}|f_r| \sum_{\alpha\in N_n^+} \sum_{k\in N_m} (|\alpha|+|k|)^{\mu} e^{-(|k|+|\alpha|)\sigma} \\
&\le \frac{1}{C_0}|f|_r \sum_{j\ge 0} j^{m+n+\mu} e^{-j\sigma} \\
&\le \frac{1}{C_0}|f|_r \sum_{j\ge 0} (j\sigma)^{m+n+\mu} e^{-j\sigma} \sigma^{-(m+n+\mu)}.
\end{aligned}$$

It is easy to see that $(j\sigma)^{m+n+\mu}e^{-\frac{1}{2}j\sigma}$ has an upper bound

$$M = (2(m+n+\mu))^{m+n+\mu}.$$

Hence the above term is less than

$$\begin{aligned}
\frac{1}{C_0}|f|_r M \sum_{j\ge 0} e^{-\frac{1}{2}j\sigma}\sigma^{-(m+n+\mu)} &= \frac{1}{C_0}|f|_r M \frac{1}{1-e^{-\frac{1}{2}\sigma}} \sigma^{-(m+n+\mu)} \\
\le \frac{1}{C_0}|f|_r M \frac{1}{\frac{1}{4}\sigma} \sigma^{-(m+n+\mu)} &= \frac{4}{C_0} M |f|_r \sigma^{-(m+n+\mu+1)}.
\end{aligned}$$

Let $C_1 = 4M/C_0$, we have

$$|\hat{f}|_{r-\sigma} \le C_1 |f|_r \sigma^{-(m+n+\mu+1)}.$$

This completes the proof.

3. Normal form

Consider the following system of differential equations

$$\dot{\theta} = \omega + \Theta(\theta, z), \tag{3.1}$$

$$\dot{z} = Az + f(\theta, z), \tag{3.2}$$

where $\theta \in C^m$, $z \in C^n$, $\omega = (\omega_1, \dots, \omega_m) \in R^m$, $A = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, and Θ and f are analytic and 2π-periodic in each component of θ in D_r.

Assume $\Theta = O(|z|)$ and $f = O(|z|^2)$ as $z \to 0$. We have the following normal form theorem.

Theorem 3.1. *Assume the following small-divisor conditions hold for a fixed integer $M > 0$,*

$$|i\langle\omega, k\rangle + \langle\lambda, \alpha\rangle - \varepsilon\lambda_j| \ge C_0/|k|^{\mu}, \tag{3.3}$$

where $k \in N_m$, $k \neq 0$, $\alpha \in N_n^+$, $1+\varepsilon \leq |\alpha| \leq M$, $\varepsilon = 0$ *or* 1, *and* C_0 *and* μ *are positive constants. Then the equations* (3.1) *and* (3.2) *can be changed to*

$$\dot{\beta} = \omega + \sum_{1\leq|\alpha|\leq M} a_\alpha y^\alpha + o(|y|^M), \tag{3.4}$$

$$\dot{y} = Ay + \sum_{2\leq|\alpha|\leq M} b_\alpha y^\alpha + o(|y|^M) \tag{3.5}$$

by a transformation

$$\theta = \beta + \Phi(\beta, y), \qquad z = y + \Psi(\beta, y), \tag{3.6}$$

where $a_\alpha = (a_\alpha^1, \ldots, a_\alpha^m)$ *and* $b_\alpha = (b_\alpha^1, \ldots, b_\alpha^n)$ *are constant vectors satisfying* $a_\alpha^j = 0$ *if* $\langle\lambda, \alpha\rangle \neq 0$ *and* $b_\alpha^j = 0$ *if* $\langle\lambda, \alpha\rangle - \lambda_j \neq 0$. $\Phi = O(|y|)$, $\Psi = O(|y|^2)$ *are analytic and* 2π*-periodic in each component of* β *in* D_{r_M}, *where* r_M *is a positive constant.*

Remark. If we do not have the θ variable, then Theorem 3.1 is the normal form theorem around a fixed point. If $M = 1$, the equation (3.2) is already of the normal form since $f = O(|z|^2)$.

Proof. We prove this theorem by induction on M. Suppose the transformation (3.6) transforms the system of equations (3.1) and (3.2) to

$$\dot{\beta} = \omega + h(\beta, y), \tag{3.7}$$

$$\dot{y} = Ay + g(\beta, y). \tag{3.8}$$

Let $M = 1$. Assume Φ and Ψ are first order and second order homogeneous polynomials in y respectively. The transformation changes the system of equations (3.1) and (3.2) to the system of equations (3.7) and (3.8) is equivalent to (Φ, Ψ) satisfying the following equations

$$h + D_\beta\Phi\omega + D_y\Phi Ay + D_y\Phi g + D_\beta\Phi h - \Theta \circ (I + (\Phi, \Psi)) = 0, \tag{3.9}$$

$$g + D_\beta\Psi\omega + D_y\Psi Ay - A\Psi + D_\beta\Psi h + D_y\Psi g - f \circ (I + (\Phi, \Psi)) = 0. \tag{3.10}$$

Now we solve the equations (3.9) and (3.10). Since $M = 1$, by observation, we take $\Psi = 0$. Then the equations (3.9) and (3.10) can be reduced to

$$h + D_\beta\Phi\omega + D_y\Phi Ay + D_y\Phi g + D_\beta\Phi h - \Theta \circ (I + (\Phi, 0)) = 0. \tag{3.11}$$

Write $h = h^1 + h^+$, where h^1 is first order in y and h^+ is a higher order term in y. Using a Taylor expansion for Θ, we have

$$\Theta(\beta + \Phi, y) = \Theta(\beta, y) + D_\beta\Theta(\beta, y)\Phi + R(\Phi).$$

Writing $\Theta(\beta, y) = \Theta^1 + \Theta^+$ as for h and comparing the orders of y in (3.11), we have

$$D_\beta\Phi\omega + D_y\Phi Ay + h^1 - \Theta^1 = 0, \tag{3.12}$$

$$h^+ + D_y\Phi g + D_\beta\Phi h - \Theta^+ - D_\beta\Theta\Phi - R(\Phi) = 0. \tag{3.13}$$

Note the left hand of (3.12) is a first order homogeneous polynomial in y and that (3.13) is $O(|y|^2)$.

Let us first solve (3.12) by finding the Φ which makes h^1 as simple as possible. Using Fourier expansions for Φ, h^1, Θ^1, we have

$$\Phi^j = \sum_{|\alpha|=1}\sum_{k\in N_m} \Phi^j_{\alpha,k} e^{i\langle k,\beta\rangle} y^\alpha,$$

$$h^{1,j} = \sum_{|\alpha|=1}\sum_{k\in N_m} h^{1,j}_{\alpha,k} e^{i\langle k,\beta\rangle} y^\alpha,$$

$$\Theta^{1,j} = \sum_{|\alpha|=1}\sum_{k\in N_m} \Theta^{1,j}_{\alpha,k} e^{i\langle k,\beta\rangle} y^\alpha,$$

where $\Phi = (\Phi^1, \dots, \Phi^m)$, $h^1 = (h^{1,1}, \dots, h^{1,m})$ and $\Theta^1 = (\Theta^{1,1}, \dots, \Theta^{1,m})$. Putting these into (3.12) and comparing the coefficients of y^α, we have

$$\Theta^{1,j}_{\alpha,k} - h^{1,j}_{\alpha,k} = (i\langle k, \omega\rangle + \langle \lambda, \alpha\rangle)\Phi^j_{\alpha,k}. \tag{3.14}$$

Because of the small-divisor conditions (3.3), the following choice is the simplest for h^1.

$$h^{1,j}_{\alpha,k} = \begin{cases} 0 & \text{for } |k| \neq 0 \text{ or } \langle\lambda, \alpha\rangle \neq 0, \\ \Theta^{1,j}_{\alpha,0} & \text{for } |k| = 0 \text{ and } \langle\lambda, \alpha\rangle = 0. \end{cases}$$

Take

$$\Phi^j_{\alpha,k} = \begin{cases} \dfrac{\Theta^{1,j}_{\alpha,k}}{i\langle k, \omega\rangle + \langle\lambda, \alpha\rangle} & \text{for } k \neq 0 \text{ or } \langle\lambda, \alpha\rangle \neq 0, \\ 0 & \text{for } |k| = 0 \text{ and } \langle\lambda, \alpha\rangle = 0. \end{cases}$$

Using the small-divisor conditions (3.3) and Lemma 2.2, we find that

$$\Phi^j = \sum_{|\alpha|=1}\sum_{k\in N_m} \Phi^j_{\alpha,k} e^{i\langle k,\beta\rangle} y^\alpha$$

is analytic and 2π-periodic in $\beta_1, \dots, \beta_m$ in $D_{r-\delta}$, $0 < \delta < r$. It is clear that $\Phi = O(|y|)$. Now we determine h^+. Write (3.13) as follows

$$(I + D_\beta\Phi)h^+ = -D_y\Phi g - D_\beta\Phi h^1 + \Theta^+ + D_\beta\Theta\Phi + R(\Phi).$$

Since $D_\beta\Phi = O(|y|)$, we can choose sufficiently small r_1, $0 < r_1 < r - \delta$ such that $(I + D_\beta\Phi)$ is invertible and $(\beta, y) + (\Phi, \Psi) \in D_r$ for $(\beta, y) \in D_{r_1}$. Hence

$$h^+ = (I + D_\beta\Phi)^{-1}(-D_y\Phi g - D_\beta\Phi h^1 + \Theta^+ + D_\beta\Theta\Phi + R(\Phi)).$$

Therefore the transformation $\theta = \beta + \Phi(\beta, y)$, $z = y$ changes the system of equations (3.1) and (3.2) to

$$\dot\beta = \omega + \sum_{|\alpha|=1} a_\alpha y^\alpha + O(|y|^2), \quad \dot y = Ay + O(|y|^2),$$

where $a_\alpha = (a^1_\alpha, \dots, a^m_\alpha)$,

$$a^j_\alpha = \begin{cases} 0, & \langle\lambda, \alpha\rangle \neq 0, \\ \Theta^{1,j}_{\alpha,0}, & \langle\lambda, \alpha\rangle = 0. \end{cases}$$

Assume Theorem 3.1 holds for $M = l \geq 1$. We will show that Theorem 3.1 is valid for $M = l + 1$.

By the induction hypothesis, there exists a transformation $\theta = \beta + \Phi^l(\beta, y)$, $z = y + \Psi^l(\beta, y)$ which satisfies the requirements of Theorem 3.1 and changes the system of equations (3.1) and (3.2) to

$$\dot{\beta} = \omega + \sum_{1\le|\alpha|\le l} a_\alpha y^\alpha + \widetilde{\Theta}(\beta, y), \tag{3.15}$$

$$\dot{y} = Ay + \sum_{2\le|\alpha|\le l} b_\alpha y^\alpha + \tilde{f}(\beta, y), \tag{3.16}$$

where $\widetilde{\Theta} = O(|y|^{l+1})$ and $\tilde{f} = O(|y|^{l+1})$ are analytic in both variables and 2π-periodic in each component of the vector β in D_{r_l}, where $r_l > 0$ is a constant.

Consider the transformation $\beta = \eta + \phi(\eta, x)$, $y = x + \psi(\eta, x)$, where ϕ and ψ are $(l+1)$th order homogeneous polynomials in x and 2π-periodic in each component of the vector η. This transformation changes the system of equations (3.15) and (3.16) to the following system

$$\dot{\eta} = \omega + \sum_{1\le|\alpha|\le l} a_\alpha x^\alpha + \tilde{h}(\eta, x), \tag{3.17}$$

$$\dot{x} = Ax + \sum_{2\le|\alpha|\le l} b_\alpha x^\alpha + \tilde{g}(\eta, x), \tag{3.18}$$

where $\tilde{h} = O(|x|^{l+1})$ and $\tilde{g} = O(|x|^{l+1})$ are to be determined later. As before, equivalently, $\beta = \eta + \phi$ and $y = x + \psi$ satisfy the following equations.

$$\begin{aligned} &\sum_{1\le|\alpha|\le l} a_\alpha (x + \psi(\eta, x))^\alpha + \widetilde{\Theta}(\eta + \phi, x + \psi) \\ &\quad = \sum_{1\le|\alpha|\le l} a_\alpha x^\alpha + D_\eta\phi\omega + D_x\phi A x + \tilde{h} \\ &\qquad + D_\eta\phi\left(\sum_{1\le|\alpha|\le l} a_\alpha x^\alpha + \tilde{h}\right) + D_x\phi\left(\sum_{2\le|\alpha|\le l} b_\alpha x^\alpha + \tilde{g}\right) \end{aligned} \tag{3.19}$$

and

$$\begin{aligned} &A\psi + \sum_{2\le|\alpha|\le l} b_\alpha (x + \psi(\eta, x))^\alpha + \tilde{f} \\ &\quad = \sum_{2\le|\alpha|\le l} b_\alpha x^\alpha + \tilde{g} + D_\eta\psi\omega + D_x\psi A x \\ &\qquad + D_\eta\psi\left(\sum_{1\le|\alpha|\le l} a_\alpha x^\alpha + \tilde{h}\right) + D_x\psi\left(\sum_{2\le|\alpha|\le l} b_\alpha x^\alpha + \tilde{g}\right). \end{aligned} \tag{3.20}$$

Since the functions we consider are analytic, we can classify the system of equations (3.19) and (3.20) into two systems according to the order of x. One contains only $(l+1)$th order terms with respect to x, the other contains higher order terms.

By an elementary calculation, we have

$$\sum_{1\le|\alpha|\le l} a_\alpha (x+\psi(\eta\,,\,x))^\alpha = \sum_{1\le|\alpha|\le l} a_\alpha x^\alpha + R_1(\psi)\,, \tag{3.21}$$

$$\sum_{2\le|\alpha|\le l} b_\alpha (x+\psi(\eta\,,\,x))^\alpha = \sum_{2\le|\alpha|\le l} b_\alpha x^\alpha + R_2(\psi)\,, \tag{3.22}$$

where $R_1(\psi) = O(|x|^{l+2})$ and $R_2(\psi) = O(|x|^{l+2})$. Using a Taylor expansion for $\widetilde{\Theta}(\eta+\phi\,,\,x+\psi)$ and $\tilde{f}(\eta+\phi\,,\,x+\psi)$, we have

$$\widetilde{\Theta}(\eta+\phi\,,\,x+\psi) = \widetilde{\Theta}(\eta\,,\,x) + R_3(\phi\,,\,\psi)\,, \tag{3.23}$$

$$\tilde{f}(\eta+\phi\,,\,x+\psi) = \tilde{f}(\eta\,,\,x) + R_4(\phi\,,\,\psi)\,, \tag{3.24}$$

where $R_3(\phi\,,\,\psi) = O(|x|^{l+2})$ and $R_4(\phi\,,\,\psi) = O(|x|^{l+2})$.

We write

$$\widetilde{\Theta}(\eta\,,\,x) = \widetilde{\Theta}^{l+1} + \widetilde{\Theta}^{+}\,, \quad \tilde{f}(\eta\,,\,x) = \tilde{f}^{l+1} + \tilde{f}^{+}\,,$$
$$\tilde{h}(\eta\,,\,x) = \tilde{h}^{l+1} + \tilde{h}^{+}\,, \quad \tilde{g}(\eta\,,\,x) = \tilde{g}^{l+1} + \tilde{g}^{+}\,,$$

where $\widetilde{\Theta}^{l+1}$, $\tilde{f}^{l+1}$, $\tilde{h}^{l+1}$ and $\tilde{g}^{l+1}$ are $(l+1)$th order homogeneous polynomials in x and $\widetilde{\Theta}^{+} = O(|x|^{l+2})$, $\tilde{f}^{+} = O(|x|^{l+2})$, $\tilde{h}^{+} = O(|x|^{l+2})$ and $\tilde{g}^{+} = O(|x|^{l+2})$. Hence we can write the system of equations (3.19) and (3.20) as the following two systems

$$\widetilde{\Theta}^{l+1} - \tilde{h}^{l+1} = D_\eta\phi\omega + D_x\phi Ax\,, \tag{3.25}$$

$$\tilde{f}^{l+1} - \tilde{g}^{l+1} = D_\eta\psi\omega + D_x\psi Ax - A\psi \tag{3.26}$$

and

$$\begin{aligned} \tilde{h}^{+} + D_\eta\phi\tilde{h}^{+} + D_x\phi\tilde{g}^{+} = \widetilde{\Theta}^{+} + R_3 - D_\eta\phi\left(\sum_{1\le|\alpha|\le l} a_\alpha x^\alpha + \tilde{h}^{l+1}\right) \\ - D_x\phi\left(\sum_{2\le|\alpha|\le l} a_\alpha x^\alpha + \tilde{g}^{l+1}\right) - R_1(\psi)\,, \end{aligned} \tag{3.28}$$

$$\begin{aligned} \tilde{g}^{+} + D_\eta\psi\tilde{h}^{+} + D_x\psi\tilde{g}^{+} = \tilde{f}^{+} + R_4 - D_\eta\psi\left(\sum_{1\le|\alpha|\le l} a_\alpha x^\alpha + \tilde{h}^{l+1}\right) \\ - D_x\psi\left(\sum_{2\le|\alpha|\le l} b_\alpha x^\alpha + \tilde{g}^{l+1}\right) - R_2(\psi). \end{aligned} \tag{3.28}$$

Using Fourier expansions for the functions $\widetilde{\Theta}^{l+1} = (\widetilde{\Theta}^{l+1\,,\,1}\,,\,\ldots\,,\,\widetilde{\Theta}^{l+1\,,\,m})$, $\tilde{f}^{l+1} = (\tilde{f}^{l+1\,,\,1}\,,\,\ldots\,,\,\tilde{f}^{l+1\,,\,n})$, $\tilde{h}^{l+1} = (\tilde{h}^{l+1\,,\,1}\,,\,\ldots\,,\,\tilde{h}^{l+1\,,\,m})$, $\tilde{g}^{l+1} = (\tilde{g}^{l+1\,,\,1}\,,\,\ldots\,,$

$\tilde{g}^{l+1,n})$, $\phi = (\phi^1, \dots, \phi^m)$ and $\psi = (\psi^1, \dots, \psi^n)$, we have

$$\tilde{\Theta}^{l+1,j} = \sum_{|\alpha|=l+1} \sum_{k \in N_m} \tilde{\Theta}^{l+1,j}_{\alpha,k} e^{i\langle \eta, k\rangle} x^\alpha$$

$$\tilde{f}^{l+1,j} = \sum_{|\alpha|=l+1} \sum_{k \in N_m} \tilde{f}^{l+1,j}_{\alpha,k} e^{i\langle \eta, k\rangle} x^\alpha,$$

$$\tilde{h}^{l+1,j} = \sum_{|\alpha|=l+1} \sum_{k \in N_m} \tilde{h}^{l+1,j}_{\alpha,k} e^{i\langle \eta, k\rangle} x^\alpha,$$

$$\tilde{g}^{l+1,j} = \sum_{|\alpha|=l+1} \sum_{k \in N_m} \tilde{g}^{l+1,j}_{\alpha,k} e^{i\langle \eta, k\rangle} x^\alpha,$$

$$\phi^j = \sum_{|\alpha|=l+1} \sum_{k \in N_m} \phi^j_{\alpha,k} e^{i\langle \eta, k\rangle} x^\alpha,$$

$$\psi^j = \sum_{|\alpha|=l+1} \sum_{k \in N_m} \psi^j_{\alpha,k} e^{i\langle \eta, k\rangle} x^\alpha.$$

Putting the above functions into the system of equations (3.25) and (3.26) and comparing the coefficients of $e^{i\langle \eta, k\rangle} x^\alpha$, we have

$$\tilde{\Theta}^{l+1,j}_{\alpha,k} - \tilde{h}^{l+1,j}_{\alpha,k} = (i\langle \omega, k\rangle + \langle \lambda, \alpha\rangle)\phi^j_{\alpha,k},$$

$$\tilde{f}^{l+1,j}_{\alpha,k} - \tilde{g}^{l+1,j}_{\alpha,k} = (i\langle \omega, k\rangle + \langle \lambda, \alpha\rangle - \lambda_j)\psi^j_{\alpha,k}.$$

Our purpose is to find ϕ, ψ such that $\tilde{h}^{l+1}$ and $\tilde{g}^{l+1}$ have the simplest form. By using the small-divisor conditions (3.3), the best choices for $\tilde{h}^{l+1}$, $\tilde{g}^{l+1}$ are the following

$$\tilde{h}^{l+1,j}_{\alpha,k} = \begin{cases} 0 & \text{for } |k| \neq 0 \text{ or } \langle \lambda, \alpha\rangle \neq 0, \\ \tilde{\Theta}^{l+1,j}_{\alpha,0} & \text{for } |k| = 0 \text{ and } \langle \lambda, \alpha\rangle = 0. \end{cases}$$

$$\tilde{g}^{l+1,j}_{\alpha,k} = \begin{cases} 0 & \text{for } |k| \neq 0 \text{ or } \langle \lambda, \alpha\rangle - \lambda_j \neq 0, \\ \tilde{f}^{l+1,j}_{\alpha,0} & \text{for } |k| = 0 \text{ and } \langle \lambda, \alpha\rangle - \lambda_j = 0. \end{cases}$$

Define

$$\phi^j_{\alpha,k} = \begin{cases} \dfrac{\tilde{\Theta}^{l+1,j}_{\alpha,0}}{i\langle k, \omega\rangle + \langle \lambda, \alpha\rangle} & \text{for } |k| \neq 0 \text{ or } \langle \lambda, \alpha\rangle \neq 0, \\ 0 & \text{for } |k| = 0 \text{ and } \langle \lambda, \alpha\rangle = 0. \end{cases}$$

and

$$\psi^j_{\alpha,k} = \begin{cases} \dfrac{\tilde{f}^{l+1,j}}{i\langle k, \omega\rangle + \langle \lambda, \alpha\rangle - \lambda_j} & \text{for } |k| \neq 0 \text{ or } \langle \lambda, \alpha\rangle - \lambda_j \neq 0, \\ 0 & \text{for } |k| = 0 \text{ and } \langle \lambda, \alpha\rangle - \lambda_j = 0. \end{cases}$$

By using the small-divisor conditions (3.3) and Lemma 2.2, we obtain that

$$\phi^j = \sum_{|\alpha|=l+1} \sum_{k \in N_m} \phi^j_{\alpha,k} e^{i\langle \eta, k\rangle} x^\alpha$$

and

$$\psi^j = \sum_{|\alpha|=l+1} \sum_{k \in N_m} \psi^j_{\alpha,k} e^{i\langle \eta, k\rangle} x^\alpha$$

are analytic and 2π-periodic in η in $D_{r_l-\delta}$, where $0 < \delta < r_{l+1}$. Moreover $\tilde{h}^{l+1}$, $\tilde{g}^{l+1}$, ϕ and ψ are solutions of the system of equations (3.25) and (3.26).

We choose r_{l+1} sufficiently small such that $0 < r_{l+1} < r_l - \delta$,

$$\begin{bmatrix} \mathrm{Id} + D_\eta \phi & D_x \phi \\ D_\eta \psi & \mathrm{Id} + D_x \psi \end{bmatrix}$$

has an inverse and $(\eta, x) + (\phi, \psi) \in D_{r_l}$ for $(\eta, x) \in D_{r_{l+1}}$. Hence we can solve the system of equations (3.27) and (3.28) for h^+ and q^+. Therefore the transformation $\beta = \eta + \phi(\eta, x)$, $g = x + \psi(\eta, x)$ changes the system of equations (3.15) and (3.16) to

$$\dot{\eta} = \omega + \sum_{1 \le |\alpha| \le l} a_\alpha x^\alpha + \sum_{|\alpha| = l+1} a_\alpha x^\alpha + h^+(\eta, x),$$

$$\dot{x} = Ax + \sum_{2 \le |\alpha| \le l} b_\alpha x^\alpha + \sum_{|\alpha| = l+1} b_\alpha x^\alpha + g^+(\eta, x),$$

where for $|\alpha| < l + 1$, a_α and b_α are the same as those in the system of equations (3.15) and (3.16), which is given by the induction hypotheses; and for $\alpha = l + 1$, $a_\alpha = (a_\alpha^1, \dots, a_\alpha^m)$ and $b_\alpha = (b_\alpha^1, \dots, b_\alpha^n)$ are constant vectors given by

$$a_\alpha^j = \begin{cases} 0 & \text{for } \langle \lambda, \alpha \rangle \ne 0, \\ \tilde{\Theta}_{\alpha,0}^{l+1,j} & \text{for } \langle \lambda, \alpha \rangle = 0, \end{cases}$$

$$b_\alpha^j = \begin{cases} 0 & \text{for } \langle \lambda, \alpha \rangle - \lambda_j \ne 0, \\ \tilde{f}_{\alpha,0}^{l+1,j} & \text{for } \langle \lambda, \alpha \rangle - \lambda_j = 0, \end{cases}$$

and $h^+ = O(|x|^{l+2})$ and $g^+ = O(|x|^{l+2})$.

Take

(3.29)

$$\theta = \eta + \Phi(\eta, x) = (I + \Phi^l) \circ (I + (\phi, \psi)) = \eta + \phi + \Phi^l(\eta + \phi, x + \psi),$$

$$z = x + \Psi(\eta, x) = (I + \Psi^l) \circ (I + (\phi, \psi)) = y + \psi + \Psi^l(\eta + \phi, x + \psi).$$

Then the transformation (3.29) changes the system of equations (3.1) and (3.2) to the system of equations (3.4) and (3.5) in which we recognize (β, y) as (η, x). Hence the theorem holds for $M = l + 1$. This completes the proof of Theorem 3.1.

Corollary 3.2. *Let* $M \ge 2$ *be fixed. If the small-divisor conditions*

$$|i\langle \omega, k \rangle + \langle \lambda, \alpha \rangle - \varepsilon \lambda_j| \ge \frac{C_0}{(|k| + |\alpha|)^\mu}$$

holds for all $k \in N_m$ *and* $1 + \varepsilon \le |\alpha| \le M$, $\varepsilon = 0$ *or* 1, *then the equations* (3.1) *and* (3.2) *can be transformed to*

$$\dot{\beta} = \omega + o(|y|^M), \qquad \dot{y} = Ay + o(|y|^M)$$

by an analytic transformation.

4. Linearization

Consider the following system of differential equations

(4.1) $$\dot{\theta} = \omega + \Theta(\theta, z),$$

(4.2) $$\dot{z} = Az + f(\theta, z),$$

where $\theta \in C^m$, $z \in C^n$, $\omega = (\omega_1, \ldots, \omega_m) \in R^m$, $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, and Θ and f are analytic and 2π-periodic in each component of the vector θ in $D_{\bar{r}}$, where $\bar{r} > 0$ is a constant.

We assume $\Theta = O(|z|)$ and $f = O(|z|^2)$ as $z \to 0$. We have the following theorem.

Theorem 4.1. *If the frequencies* $\omega = (\omega_1, \ldots, \omega_m)$ *and eigenvalues* $\lambda = (\lambda_1, \ldots, \lambda_n)$ *satisfy the small-divisor conditions*

$$|i\langle \omega, k\rangle + \langle \lambda, \alpha\rangle - \varepsilon\lambda_j| \geq \frac{C_0}{(|k| + |\alpha|)^{\mu}} \quad \text{for } j = 1, \ldots, n, \tag{4.3}$$

and for $k \in N_m$, $\alpha \in N_n^+$, $|\alpha| \geq 1 + \varepsilon$, *where* $\varepsilon = 0$ *or* 1, *then the system of equations* (4.1) *and* (4.2) *can be transformed to*

$$\dot{\beta} = \omega, \tag{4.4}$$

$$\dot{y} = Ay, \tag{4.5}$$

by a transformation

$$\theta = \beta + \Phi(\beta, y), \tag{4.6}$$

$$z = y + \Psi(\beta, y), \tag{4.7}$$

where Φ *and* Ψ *are analytic and* 2π*-periodic in each component of the vector* β *in* $D_{\frac{1}{2}r}$ *for some small* r, $0 < r < \bar{r}$, *and* $\Phi = O(|y|)$, $\Psi = O(|y|^2)$ *as* $y \to 0$.

First we describe the idea of the proof. This idea is essentially due to Rüssmann [8] and Zehnder [9]. In fact, Zehnder uses it to prove Siegel's theorem for maps. Then we prove the theorem precisely.

Without losing generality, we can assume that Θ and f are high order with respect to z. In fact, by using Corollary 3.2 of the normal form theorem, we can transform the system of equations (4.1) and (4.2) to a system whose nonlinear terms have the desired order.

It is not hard to see that (4.6) and (4.7) change the system of equations (4.1) and (4.2) to the linear system of equations (4.4) and (4.5) if and only if the transformation satisfies the following equation

$$F(\Phi, \Psi) = D(\Phi, \Psi)(\omega, A) - (0, A\Psi) - (\Theta, f) \circ (I + (\Phi, \Psi)) = 0. \tag{4.8}$$

We want to apply Newton's method to solve the equation (4.8). Suppose $F(\Phi, \Psi)$ is small (this is reasonable because if we take $(\Phi_0, \Psi_0) = (0, 0)$ as an initial iteration, then $F(0, 0) = (\Theta(\beta, y), f(\beta, y))$ is small as long as y is small). Our goal is to find a better next approximation $(\Phi + u, \Psi + v)$ such that $F(\Phi + u, \Psi + v)$ is smaller, where (u, v) is a small error.

Using a Taylor expansion, we have

$$F(\Phi + u, \Psi + v) = F(\Phi, \Psi) + F'(\Phi, \Psi)(u, v) + R(\Phi, \Psi; u, v), \tag{4.9}$$

where R is a high order term in terms of (u, v), and

$$\begin{aligned} &F'(\Phi, \Psi)(u, v) \\ &\quad = D(u, v)(\omega, A) - (0, Av) - D(\Theta, f) \circ (I + (\Phi, \Psi)) \cdot (u, v). \end{aligned} \tag{4.10}$$

If we can solve the following linear equation for (u, v)

$$F(\Phi, \Psi) + F'(\Phi, \Psi)(u, v) = 0, \tag{4.11}$$

then we obtain a better approximation $(\Phi + u, \Psi + v)$.

Unfortunately, F' has, in general, no inverse nor even a right inverse. So we cannot solve the linear equation (4.11).

However, $F'(\Phi, \Psi)$ has an approximate inverse. Namely, $F'(\Phi, \Psi)$ can be written as

$$F'(\Phi, \Psi) = L(u, v) + H(u, v),$$

where L is a linear operator with a right inverse, and $H(u, v)$ is "smaller" than (u, v).

If we can construct such an approximate inverse L^{-1}, then we can solve the linear equation $F(\Phi, \Psi)+L(u, v) = 0$ for (u, v) and $(\Phi+u, \Psi+v)$ is a better approximation since $F(\Phi+u, \Psi+v) = H(u, v)+R(\Phi, \Psi; u, v) =$ "smaller" + "smaller." Let

$$T(u, v) = D(u, v)(\omega, A) - (0, Av). \tag{4.12}$$

Then

$$F'(\Phi, \Psi)(u, v) = T(u, v) - D(\Theta, f) \circ (I + (\Phi, \Psi))(u, v),$$

where $D(\Theta, f)$ is the derivative of (Θ, f). Note T has a right inverse (we will see this later).

The first attempt is to pick T^{-1} as the approximate inverse of $F'(\Phi, \Psi)$. However $D(\Theta, f)\circ(I+(\Phi, \Psi))(u, v)$ is not smaller than (u, v). This implies that we need to choose a different one. Hence, we separate $D(\Theta, f)\circ(I+(\Phi, \Psi))$ into two parts. One of them will be added to T so that the summation with T has a right inverse and other part is a smaller term. This strategy can be carried out as follows.

Differentiating the function $(\beta, y) \to F(\Phi(\beta, y), \Psi(\beta, y))$, we have

$$\begin{aligned} dF(\Phi, \Psi)(p, q) &= d[D(\Phi, \Psi)(\omega, A)](p, q) - d(0, A\Psi)(p, q) \\ &\quad - D(\Theta, f) \circ (I + (\Phi, \Psi))(\mathrm{Id} + d(\Phi, \Psi))(p, q), \end{aligned} \tag{4.13}$$

where (p, q) is analytic and 2π-periodic in $\beta_1, \dots, \beta_m$ from D_r to $C^m \times C^n$.

Comparing (4.13) with (4.10), taking $(u, v) = (\mathrm{Id} + d(\Phi, \Psi))(p, q)$ and subtracting (4.13) from (4.10), we have

$$\begin{aligned} &F'(\Phi, \Psi)(u, v) - dF(\Phi, \Psi)(\mathrm{Id} + d(\Phi, \Psi))^{-1}(u, v) \\ &= (\mathrm{Id} + d(\Phi, \Psi)) \Big\{ D[(\mathrm{Id} + d(\Phi, \Psi))^{-1}(u, v)](\omega, A) \\ &\qquad - \begin{pmatrix} 0 & 0 \\ 0 & A \end{pmatrix} (\mathrm{Id} + d(\Phi, \Psi))^{-1}(u, v) \Big\} \\ &\equiv L(u, v). \end{aligned} \tag{4.14}$$

Hence $-D(\Theta, f) \circ (I + (\Phi, \Psi))$ is decomposed into

$$-D(\Theta, f) \circ (I + (\Phi, \Psi)) = [-T + L] + [dF(\Phi, \Psi)(\mathrm{Id} + d(\Phi, \Psi))^{-1}(u, v].$$

L has a right inverse (to be proved later), and $dF(\Phi, \Psi)(\mathrm{Id} + d(\Phi, \Psi))^{-1}(u, v)$ is smaller than (u, v) since $F(\Phi, \Psi)$ is small, by a Cauchy estimate, $dF(\Phi, \Psi)$ is small. We write (4.9) as

$$\begin{aligned} F(\Phi + u, \Psi + v) &= F(\Phi, \Psi) + L(u, v) \\ &\quad + dF(\Phi, \Psi)(\mathrm{Id} + d(\Phi, \Psi))^{-1}(u, v) + R(\Phi, \Psi; u, v). \end{aligned} \tag{4.15}$$

We should also mention here that each iteration is defined on a smaller domain than that of the previous iteration. Fortunately, we will see that the domain shrinks to a fixed domain rather than one point after infinitely many iterations.

Before we prove Theorem 4.1, we discuss the operators T and L. From (4.12) and (4.14) we can see that T and L have the following relation

$$L(u, v) = (\mathrm{Id} + d(\Phi, \Psi))T[(\mathrm{Id} + d(\Phi, \Psi))^{-1}(u, v]$$

provided that $(\mathrm{Id} + d(\Phi, \Psi))$ has an inverse. Hence L has a right inverse if and only if T has a right inverse.

Define a function space $H_{0,r}$ consisting of all analytic functions $g(\beta, y) = (g^1(\beta, y), g^2(\beta, y)) : D_r \to C^m \times C^n$ which satisfy the following conditions

(i) g is analytic and 2π-periodic in $\beta_1, \dots, \beta_m$;
(ii) $g^1(\beta, y) = O(|y|)$, $g^2(\beta, y) = O(|y|^2)$;
(iii) $\sup_{(\beta, y)\in D_r} |g(\beta, y)| < \infty$;

with norm $|g|_r = \sup_{(\beta, y)\in D_r} |g(\beta, y)|$ and a function space $H_{1,r} = \{g \in H_{0,r} | |dg|_r < \infty\}$ with norm $\|g\|_r = |g|_r + |dg|_r$. It is clear that $H_{0,r}$ and $H_{1,r}$ are Banach spaces.

For the operator T we have the following lemma.

Lemma 4.2. *Assume that the small-divisor conditions* (4.3) *are satisfied. For each* $g \in H_{0,\rho}$ $(0 < \rho \le r)$, *the equation*

$$T(p, q) = g \tag{4.16}$$

has a unique solution $(p, q) \in H_{1,\rho-\delta}$ *for all* $0 < \delta < \rho$. *Moreover,*

$$\|(p, q)\|_{p-\delta} \le C_2 |g|_\rho \delta^{-(m+n+\mu+2)}, \quad \text{for } \delta, 0 < \delta < \rho,$$

where $C_2 = C_2(C_0, m, n, \mu)$ *is a constant depending on* C_0, m, n, μ.

Note the above inequality implies that T has an "inverse" which is bounded from $H_{0,\rho}$ to $H_{1,\rho-\delta}$. We denote this operator: $g \to (p, q)$ by T^{-1}.

Proof. Let $p = (p^1, \dots, p^m)$, $q = (q^1, \dots, q^n)$, $g = (g^1, g^2)$, $g^1 = (g^{1,1}, \dots, g^{1,m})$, and $g^2 = (g^{2,1}, \dots, g^{2,n})$. Writing these in the form of Fourier-Taylor series; namely

$$p^j(\beta, y) = \sum_{\substack{\alpha\in N_n^+ \\ |\alpha|\ge 1}} \sum_{k\in N_m} p^j_{\alpha,k} e^{i\langle\beta, k\rangle} y^\alpha,$$

$$q^j(\beta, y) = \sum_{\substack{\alpha\in N_n^+ \\ |\alpha|\ge 2}} \sum_{k\in N_m} q^j_{\alpha,k} e^{i\langle\beta, k\rangle} y^\alpha,$$

$$g^{1,j}(\beta, y) = \sum_{\substack{\alpha\in N_n^+ \\ |\alpha|\ge 1}} \sum_{k\in N_m} g^{1,j}_{\alpha,k} e^{i\langle\beta, k\rangle} y^\alpha,$$

$$g^{2,j}(\beta, y) = \sum_{\substack{\alpha\in N_n^+ \\ |\alpha|\ge 2}} \sum_{k\in N_m} g^{2,j}_{\alpha,k} e^{i\langle\beta, k\rangle} y^\alpha,$$

then placing these into equation (4.16) and comparing all the coefficients of $e^{i\langle\beta,k\rangle}y^{\alpha}$, we have

$$p^j_{\alpha,k}=\frac{g^{1,j}_{\alpha,k}}{i\langle\omega,k\rangle+\langle\lambda,\alpha\rangle},\quad q^j_{\alpha,k}=\frac{g^{2,j}_{\alpha,k}}{i\langle\omega,k\rangle+\langle\lambda,\alpha\rangle-\lambda_j}.$$

Using Lemma 2.2, we conclude that $(p,q)\in H_{0,\rho-\delta}$ for any $0<\delta<\rho$, and $|(p,q)|_{\rho-\delta}\le C_1|g|_\rho\delta^{-(m+n+\mu+1)}$. Using Cauchy estimates, we have

$$|d(p,q)|_{\rho-\delta}\le|(p,q)|_{\rho-\frac{1}{2}\delta}\left(\frac{\delta}{2}\right)^{-1}\le 2^{m+n+\mu+2}C_1|g|_\rho\delta^{-(m+n+\mu+2)}.$$

Hence $\|(p,q)\|_{\rho-\delta}\le C_2|g|_\rho\delta^{-(m+n+\mu+2)}$, where $C_2=2^{m+n+\mu+2}C_1$. This completes the proof.

Now we prove Theorem 4.1. Define the iteration scheme as follows

$$(\Phi_0,\Psi_0)=(0,0), \tag{4.17}$$

$$(\Phi_{j+1},\Psi_{j+1})=(\Phi_j,\Psi_j)+(u_j,v_j), \tag{4.18}$$

$$\begin{aligned}(u_j,v_j)&=-L^{-1}F(\Phi_j,\Psi_j)\\&=-(\mathrm{Id}+d(\Phi_j,\Psi_j))T^{-1}(\mathrm{Id}+d(\Phi_j,\Psi_j))^{-1}F(\Phi_j,\Psi_j)\end{aligned} \tag{4.19}$$

for $j=0,1,\dots$, where (Φ_j,Ψ_j) and (u_j,v_j) are defined on D_{r_j} and $D_{r_{j+1}}$ respectively, and $r_j=\frac{1}{2}r(1+2^{-(j+1)})$, $0<r<1$, $j=0,1,\dots$, where r is to be determined later. We will show that (Φ_j,Ψ_j) is well-defined and converges to $(\Phi,\Psi)\in H_{1,\frac{1}{2}r}$ and that $F(\Phi,\Psi)=0$.

We have

$$r_j-r_{j+1}=2^{-(j+3)}r. \tag{4.20}$$

Define a sequence $\{\varepsilon_j\}$ by induction, $\varepsilon_{j+1}=C^{j+1}\varepsilon_j^2$, $\varepsilon_0=1/2C^2$, where $C>1$ is a constant to be chosen later. It is not hard to check that

(i) $\varepsilon_j=C^{-(j+2)}(\frac{1}{2})^{2^j}$;
(ii) $\varepsilon_j\to0$ as $j\to\infty$;
(iii) $\varepsilon_{j+1}\le\frac{1}{2}\varepsilon_j\le\varepsilon_j-\varepsilon_{j+1}$.

As we mentioned before, by using the normal form Theorem 3.1, we can take $(\Theta,f)=o(|z|^{5(m+n+\mu+3)})$. This will decrease the value of $\overline{r}$. However, this will have no effect in our proof since we will choose $r>0$ to be sufficient small. Let $C=8C_3(\frac{16}{r})^{(m+n+\mu+3)}$, where $C_3=\max\{C_2,1\}$ and $\varepsilon_0=1/2C^2$. We choose r, $0<r<1$, such that $\|(\Theta,f)\|_r\le\varepsilon_0^2$ and $\varepsilon_0\le\frac{1}{8}r$. Then we claim that (Φ_j,Ψ_j) and (u_j,v_j) have the following properties:

(A_j) (Φ_j,Ψ_j) is well defined and $\|(\Phi_j,\Psi_j)\|_{r_j}\le\varepsilon_0-\varepsilon_j$;
(B_j) $|F(\Phi_j,\Psi_j)|_{r_j}\le\varepsilon_j^2$;
(C_j) $\|(u_j,v_j)\|_{r_j}\le\varepsilon_{j+1}$.

We prove this claim by induction on j. It is sufficient to show the following statements.

(a) (A_0) and (B_0) are true;
(b) (A_j) and (B_j) imply (C_j);
(c) (A_j) and (C_j) imply (A_{j+1});
(d) (A_j), (B_j), (C_j) and (A_{j+1}) imply (B_{j+1}).

Let us look at (a). A_0 is trivial since $(\Phi_0, \Psi_0) = (0, 0)$. Because of

$$(\Theta, f) = o(|z|^{5(m+n+\mu+3)}),$$

(B_0) immediately follows by the choice of r.

Proof of (b). Take $\delta = \frac{1}{2}(r_j - r_{j+1})$. By using (A_j), we have

$$|d(\Phi_j, \Psi_j)|_{r_j} \leq \tfrac{1}{2}.$$

Hence

$$|(\mathrm{Id} + d(\Phi_j, \Psi_j))^{-1}|_{r_j} \leq 2.$$

Using Lemma 4.2, (4.19), the choice of δ and (4.20), we have

$$\begin{aligned}(u_j, v_j)|_{r_j-\delta} &= |(\mathrm{Id} + d(\Phi_j, \Psi_j))T^{-1}(\mathrm{Id} + d(\Phi_j, \Psi_j))^{-1}F(\Phi_j, \Psi_j)|_{r_j-\delta} \\ &\leq 4C_2|F(\Phi_j, \Psi_j)|_{r_j}\delta^{-(m+n+\mu+2)} \leq 4C_2\delta^{-(m+n+\mu+2)}\varepsilon_j^2 \\ &= 4C_2\left(\frac{2^{j+4}}{r}\right)^{m+n+\mu+2}\varepsilon_j^2.\end{aligned}$$

By using Cauchy estimates, we have

$$|d(u_j, v_j)|_{r_{j+1}} \leq |(u_j, v_j)|_{r_j-\delta}\delta^{-1} \leq 4C_2\left(\frac{2^{j+4}}{r}\right)^{m+n+\mu+3}\varepsilon_j^2$$

Hence (C_j) follows by the choice of C

$$\begin{aligned}\|(u_j, v_j)\|_{r_{j+1}} &\leq 8C_2\left(\frac{2^{j+4}}{r}\right)^{m+n+\mu+3} \\ &\leq 8C_2\left(\left(\frac{16}{r}\right)^{m+n+\mu+3}\right)^{j+1}\varepsilon_j^2 \qquad (4.21) \\ &\leq C^{j+1}\varepsilon_j^2 = \varepsilon_{j+1}.\end{aligned}$$

Proof of (c). By the choice of r and (A_j), we have

$$|I + (\Phi_j, \Psi_j)|_{r_j} < r \quad \text{and} \quad |d(\Phi_j, \Psi_j)| < \tfrac{1}{2}.$$

Hence $F(\Phi_j, \Psi_j)$ is well defined, $(\mathrm{Id} + d(\Phi_j, \Psi_j))^{-1}$ exists, and

$$(|(\mathrm{Id} + d(\Phi_j, \Psi_j))^{-1}| \leq 2.$$

Therefore (Φ_{j+1}, Ψ_{j+1}) is well defined. By (A_j) and (C_j), we have (A_{j+1}).

$$\begin{aligned}\|(\Phi_{j+1}, \Psi_{j+1})\|_{j+1} &\leq \|(\Phi_j, \Psi_j)\|_{r_j} + \|(u_j, v_j)\|_{r_{j+1}} \\ &\leq \varepsilon_0 - \varepsilon_j + \varepsilon_{j+1} \leq \varepsilon_0 - \varepsilon_{j+1}.\end{aligned}$$

Finally we show (d). By (4.15), (4.18) and (4.19), we have

$$F(\Phi_{j+1}, \Psi_{j+1}) = dF(\Phi_j, \Psi_j)(\mathrm{Id} + d(\Phi_j, \Psi_j))^{-1}(u_j, v_j) + R(\Phi_j, \Psi_j; u_j, v_j),$$

where

$$R(\Phi_j, \Psi_j; u_j, v_j) = \int_0^1 (1-t)\frac{d^2}{dt^2}(\Theta, f) \circ (I + (\Phi_j, \Psi_j) + t(u_j, v_j))\, dt.$$

Using Cauchy estimates, (4.20), (B_j) and (C_j), we have

$$\begin{aligned}|dF(\Phi_j, \Psi_j)(\mathrm{Id} + d(\Phi_j, \Psi_j))^{-1}(u_j, v_j)|_{r_{j+1}} &\le |F(\Phi_j, \Psi_j)|_{r_j}(r_j - r_{j+1})^{-1} 2\varepsilon_{j+1} \\ &\le \left(\frac{2^{j+3}}{r}\right) 2\varepsilon_j^2 \varepsilon_{j+1} \le \frac{1}{2} C^{j+1} \varepsilon_j^2 \varepsilon_{j+1} = \frac{1}{2}\varepsilon_{j+1}^2.\end{aligned}$$

It follows from the choice of r, (A_j), and (C_j) that $|I + (\Phi_j, \Psi_j) + t(u_j, v_j)|_{r_j} < r$. Hence $R(\Phi_j, \Psi_j; u_j, v_j)$ is well defined. By a Cauchy estimate, we have $|d^2(\Theta, f)|_{r_{j+1}} \le (\frac{8}{r})^2 |(\Theta, f)|_r \le 1$. Therefore, we have

$$|R(\Phi_j, \Psi_j; u_j, v_j)|_{r_{j+1}} \le \tfrac{1}{2}|d^2(\Theta, f)|_{r_{j+1}} |(u_j, v_j)|_{r_{j+1}}^2 \le \tfrac{1}{2}\varepsilon_{j+1}^2.$$

Thus $|F(\Phi_{j+1}, \Psi_{j+1})|_{r_{j+1}} \le \varepsilon_{j+1}^2$. This completes the proof of the claim.

Proof of Theorem 4.1. By the claim (C_j), we have $(\Phi, \Psi) = \lim_{j\to\infty}(\Phi_j, \Psi_j)$ exists in $H_{1, r/2}$. From (B_j) we conclude $\lim_{j\to\infty} F(\Phi_j, \Psi_j) = 0$, i.e., $F(\Phi, \Psi) = 0$. It follows from (A_j) that $I + (\Phi, \Psi)$ has an analytic inverse. This completes the proof.

References

1. V. I. Arnold, *Geometrical methods in the theory of ordinary differential equations*, Springer-Verlag, New York, 1983.
2. E. G. Belaga, *On the reducibility of a system of differential equations in the neighborhood of conditionally periodic motion*, Dokl. Akad. Nauk SSSR **8** (1962), 255–258.
3. B. L. J. Braaksma and H. W. Broer, *On a quasi-periodic Hopf bifurcation*, Ann. Inst. Henri Poincaré **4** (1987), 115–168.
4. S.-N. Chow and J. K. Hale, *Methods of bifurcation theory*, Springer-Verlag, New York, 1982.
5. J. Guckenheimer and P. J. Holmes, *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, Springer-Verlag, New York, 1983.
6. K. R. Meyer, *The implicit function theorem and analytic differential equations*, Lecture Notes in Math., vol. 468, Springer-Verlag, New York, 1975, pp. 191–208.
7. J. Moser, *Convergent series expansions for quasi-periodic motions*, Math. Ann. **196** (1967), 136–176.
8. H. Rüssmann, *Kleine Nenner.* II: *Bemerkungen zur Newtonschen Methode*, Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II, (1972), 1–10.
9. E. Zehnder, *A simple proof of a generalization of a theorem by C. L. Siegel*, Lecture Notes in Math., vol. 597, Springer-Verlag, 1977, pp. 855–866.
10. B. L. J. Braaksma, H. W. Broer and G. B. Huitema, *Toward a quasi-periodic bifurcation theory*, Mem. Amer. Math. Soc., no. 421, 1990, pp. 83–175.

Center for Dynamical Systems and Nonlinear Studies, School of Mathematics, Georgia Institute of Technology, Atlanta, Georgia 30332

Institute for Mathematics and Its Applications, University of Minnesota, Minneapolis, Minnesota 55455

Current address: Department of Mathematics, Brigham Young University, Provo, Utah 84602

Department of Mathematics, Western Washington University, Bellingham, Washington 98225