Illinois Journal of Mathematics
Volume 56, Number 1, Spring 2012, Pages 85–94
S 0019-2082

EXISTENCE OF DIVERGENT BIRKHOFF NORMAL FORMS OF HAMILTONIAN FUNCTIONS

XIANGHONG GONG

ABSTRACT. By the work of Siegel it is well known that as a rule the Birkhoff normal form of a real analytic Hamiltonian system whose eigenvalues satisfies suitable non-resonance condition cannot be realized by convergent symplectic transformations. We show the existence of divergent Birkhoff normal forms for suitable Hamiltonian systems. Our calculation shows how the small divisors appear in the normal forms, from which the divergence is derived by using Siegel's methods of small divisors.

1. Introduction

We consider the standard symplectic space $\mathbf{R}^4$, equipped with the symplectic 2-form $\omega = dx_1 \wedge dy_1 + dx_2 \wedge dy_2$. Let $h(x,y)$ be a real analytic function, defined near $0 \in \mathbf{R}^4$, that has the form

$$h(x,y) = \lambda_1 x_1 y_1 + \lambda_2 x_2 y_2 + E(x,y), \tag{1.1}$$

where $E(x,y) = \sum_{\alpha,\beta} E_{\alpha,\beta} x^\alpha y^\beta$ is a convergent power series in x, y satisfying $E_{\alpha\beta} = 0$ for $|\alpha| + |\beta| < 3$. For brevity, we denote the latter condition by $E(x,y) = O(3)$. We say that λ_1 and λ_2 are *non-resonant*, if $\lambda \cdot \alpha = \lambda_1\alpha_1 + \lambda_2\alpha_2 \neq 0$ for all multi-indices of integers $\alpha = (\alpha_1, \alpha_2) \neq 0$. Under the non-resonance condition on λ, there is a formal *symplectic* real map φ of $\mathbf{R}^4$, i.e. $\varphi^*\omega = \omega$, such that $\varphi(0) = 0$ and $h \circ \varphi^{-1}(x,y)$ is a real formal power series in x_1y_1, x_2y_2. The formal power series $\hat{h} = h \circ \varphi^{-1}$ is called a Birkhoff normal form of h (e.g., see [14], p. 209). Note that $\hat{h}(x,y)$ is not unique. However, it depends on the choice of coordinates in a simple way; namely, the only other Birkhoff normal forms are obtained from $\hat{h}(x,y)$ by a permutation of x_1, x_2, y_1, y_2 that preserves the symplectic 2-form. Therefore, the normal

Received November 23, 2011; received in final form July 2, 2012.
Supported in part by NSF grant DMS-0305474.
2010 *Mathematics Subject Classification.* 37J40.

forms of the h are either all convergent or all divergent. Throughout the paper, we refer $\hat{h}$, whose quadratic part is the same as h, as the Birkhoff normal form of h.

The following existence theorem is our main result.

THEOREM 1.1. *There exist some non-resonant* λ_1, λ_2 *and a real analytic function* $h(x,y) = \lambda_1 x_1 y_1 + \lambda_2 x_2 y_2 + O(3)$ *such that the Birkhoff normal form of* h *is divergent.*

In [12], Siegel showed that the Birkhoff normal form cannot be realized by convergent symplectic transformations in general. In fact, Siegel [13] showed that for some real analytic function $h(x,y) = \lambda_1(x_1^2 + y_1^2) + \lambda_2(x_2^2 + y_2^2) + O(3)$ having *any* given non-resonant λ_1, λ_2 and *generic* higher order terms, there exists no convergent symplectic mapping transforming $h(x,y)$ into its normal form. Note that for this type of quadratic part of h, the normal forms are formal power series in $x_1^2 + y_1^2$ and $x_2^2 + y_2^2$.

Despite Siegel's divergence results and many other results, the existence of a divergent Birkhoff normal form arising from a real analytic function is new (for instance, see [2]). The divergence of Birkhoff normal form implies, of course, the divergence of all normalizing transformations of the given function. The importance of the existence of a divergent Birkhoff normal form was demonstrated by Pérez-Marco [9].

For the Birkhoff normal form theory, the reader is referred to the works of Moser [7], Rüssmann [10], [11], Brjuno [1], Vey [16], Ito [6], Stolovitch [15], Giorgilli [3], and the author [4], [5] besides the above mentioned references. Papers by Brjuno [1] and Pérez-Marco [9] contain extensive references also.

The proof of Theorem 1.1 is based on Siegel's method of small divisors. One would expect that the present approach will have applications for other small-divisor problems. We will however focus on the Hamiltonian functions to demonstrate how the small-divisors enter the normal forms. We will state our result and its proof for the higher dimensions at the end of the paper.

2. Proof of the theorem

The proof consists of 3 steps. The first step is to recall how the Birkhoff normal form is derived. Here we do not claim any originality, and we present the details for the sake of the reader. These details are crucial in our construction of divergent normal forms. In this step, one sees how the small divisors enter the formal map φ that normalizes the function h. The second step is to show how the small divisors enter the normal form $\hat{h}$. Here the computation is crucial for our proof. Once we find small divisors in the coefficients of the normal form $\hat{h}$, the proof of the divergence of $\hat{h}$ follows from Siegel's arguments [12].

Consider a real analytic (real-valued) function

$$h(x,y) = \lambda_1 x_1 y_1 + \lambda_2 x_2 y_2 + \sum_{|\alpha|+|\beta|\geq 3} h_{\alpha\beta} x^\alpha y^\beta,$$

where λ_1, λ_2 are non-resonant.

Let $S(x,y) = O(3)$ be a real analytic function defined near $0 \in \mathbf{R}^2 \times \mathbf{R}^2$. Here we denote $S(x,y) = O(d)$ if $S_{\alpha\beta} = 0$ for all $|\alpha| + |\beta| < d$. Let $(\hat{x}, \hat{y}) = \varphi(x,y)$ be the symplectic map defined implicitly by

$$\hat{x}_j = x_j - S_{\hat{y}_j}(x, \hat{y}), \qquad \hat{y}_j = y_j + S_{x_j}(x, \hat{y}), \quad j = 1, 2. \tag{2.1}$$

We want to show that there exists a unique formal power series

$$S(x,y) = \sum_{|\alpha|+|\beta|\geq 3} S_{\alpha\beta} x^\alpha y^\beta, \quad S_{\alpha\alpha} = 0$$

such that $\hat{h}(\hat{x}, \hat{y}) = h \circ \varphi^{-1}(\hat{x}, \hat{y})$ is a formal power series in $\hat{x}_1\hat{y}_1, \hat{x}_2\hat{y}_2$. Write

$$\varphi(x,y) = \big(x + u(x,y), y + v(x,y)\big), \quad u = (u_1, u_2),\ v = (v_1, v_2).$$

From (2.1), we see that $u(x,y) = O(2)$, $v(x,y) = O(2)$ and

$$u_i(x,y) = -S_{\hat{y}_i}\big(x, y + v(x,y)\big), \qquad v_i(x,y) = S_{x_i}\big(x, y + v(x,y)\big).$$

Let $\delta_1 = (1,0)$ and $\delta_2 = (0,1)$. By comparing coefficients of $x^\alpha y^\beta$ and using $u(x,y) = O(2)$ and $v(x,y) = O(2)$ in the above two identities, we get

$$\begin{aligned} u_{j,\alpha\beta} &= -(\beta_j + 1) S_{\alpha\beta+\delta_j} + U_{j,\alpha\beta}(S), \\ v_{j,\alpha\beta} &= (\alpha_j + 1) S_{\alpha+\delta_j\beta} + V_{j,\alpha\beta}(S), \end{aligned} \tag{2.2}$$

where $U_{j,\alpha\beta}(S), V_{j,\alpha\beta}(S)$ are polynomials in $S_{\alpha'\beta'}$ with $|\alpha'| + |\beta'| \leq |\alpha| + |\beta|$. We need to solve the equation $h(x,y) = \hat{h}(x + u(x,y), y + v(x,y))$ under the normalizing condition that $\hat{h}_{\alpha\beta} = 0$ for $\alpha \neq \beta$. Comparing the coefficients we obtain

$$\begin{aligned} h_{\alpha\beta} &= \varepsilon_{\alpha\beta}\hat{h}_{\alpha\alpha} + \sum \lambda_j (v_{j,\alpha-\delta_j\beta} + u_{j,\alpha\beta-\delta_j}) + E_{\alpha\beta}(\hat{h}, u, v) \\ &= \varepsilon_{\alpha\beta}\hat{h}_{\alpha\alpha} + \lambda \cdot (\alpha - \beta) S_{\alpha\beta} + F_{\alpha\beta}(\hat{h}, S) \quad \text{(by (2.2))}. \end{aligned}$$

Here $\varepsilon_{\alpha\beta} = 0$ for $\alpha \neq \beta$ and $\varepsilon_{\alpha\alpha} = 1$. Also $E_{\alpha\beta}(\hat{h}, u, v)$ is a polynomial in $\hat{h}_{\alpha'\alpha'}, u_{\alpha''\beta''}, v_{\alpha''\beta''}$ with $\max\{2|\alpha'|, |\alpha''| + |\beta''| + 1\} < |\alpha| + |\beta|$, and $F_{\alpha\beta}(\hat{h}, S)$ is a polynomial in $\hat{h}_{\alpha'\alpha'}, S_{\alpha'\beta'}$ with $\max\{2|\alpha'|, |\alpha'| + |\beta'|\} < |\alpha| + |\beta|$. Therefore, we get

$$\begin{aligned} S_{\alpha\beta} &= \frac{1}{\lambda \cdot (\alpha - \beta)} \big\{ h_{\alpha\beta} - F_{\alpha\beta}(\hat{h}, S) \big\}, \quad \alpha \neq \beta, \\ \hat{h}_{\alpha\alpha} &= h_{\alpha\alpha} - F_{\alpha\alpha}(\hat{h}, S). \end{aligned}$$

It will be convenient to denote

$$\mu_{\alpha\beta}=\frac{1}{(\alpha-\beta)\cdot\lambda},\quad \alpha\neq\beta.$$

Recursively, we substitute $S_{\alpha'\beta'},\hat{h}_{\alpha'\alpha'}$ into $F_{\alpha\beta}(\hat{h},S)$ to obtain

$$\begin{aligned}S_{\alpha\beta}&=\frac{1}{\lambda\cdot(\alpha-\beta)}\{h_{\alpha\beta}+D_{\alpha\beta}(h,\mu)\},\quad \alpha\neq\beta,\\ \hat{h}_{\alpha\alpha}&=h_{\alpha\alpha}+D_{\alpha\alpha}(h,\mu),\end{aligned}$$

where $D_{\alpha\beta}(\mu,h)$ is a polynomial in $\mu_{\alpha'\beta'}$ and $h_{\alpha''\beta''}$ with $|\alpha'|+|\beta'|<|\alpha|+|\beta|,\alpha'\neq\beta'$ and $|\alpha''|+|\beta''|<|\alpha|+|\beta|$.

We remark that $S_{\alpha\beta}$ is uniquely determined by $h_{\alpha\beta}$ and $h_{\alpha'\beta'}$ with $|\alpha'|+|\beta'|<|\alpha|+|\beta|$. Also $D_{\alpha\beta}=S_{\alpha\beta}=0$ if $h_{\alpha'\beta'}=0$ for all $\alpha'\neq\beta'$ satisfying $2<|\alpha'|+|\beta'|<|\alpha|+|\beta|$. Therefore, if $d>2$ is fixed and $h_{\alpha'\beta'}=0$ for all $|\alpha'|+|\beta'|<d$ with $\alpha'\neq\beta'$, then

$$S_{\alpha\beta}=\frac{h_{\alpha\beta}}{\lambda\cdot(\alpha-\beta)},\quad \alpha\neq\beta,\ |\alpha|+|\beta|=d. \tag{2.3}$$

To see small divisors in $\hat{h}$, we need to calculate $D_{\alpha\alpha}(\mu,h)$ more explicitly. To this end, we apply a preliminary change of coordinates by truncating the above $S(x,y)$. We fix $d>2$. Let φ_1 be the symplectic mapping defined by

$$\hat{x}_j=x_j-S^*_{\hat{y}_j}(x,\hat{y}),\qquad \hat{y}_j=y_j+S^*_{x_j}(x,\hat{y}),$$

where $S^*(x,y)=\sum_{3\leq|\alpha|+|\beta|<d}S_{\alpha\beta}x^\alpha y^\beta$ with $S_{\alpha\beta}$ being determined above. Note that when $d=3$, we have taken φ_1 to be the identity. Then $f=h\circ\varphi_1^{-1}$ has the form

$$\begin{aligned}f(x,y)&=\sum_{|\alpha|+|\beta|\geq 2}f_{\alpha\beta}x^\alpha y^\beta=\hat{h}(x,y)+O(d),\\ f_{\alpha\beta}&=h_{\alpha\beta}+Q_{\alpha\beta}(h,\mu),\quad |\alpha|+|\beta|\geq d,\end{aligned}\tag{2.4}$$

where $Q_{\alpha\beta}(h,\mu)$ is a polynomial in $\mu_{\alpha'\beta'}$ and $h_{\alpha''\beta''}$ with $|\alpha'|+|\beta'|<d,\alpha'\neq\beta'$ and $|\alpha''|+|\beta''|<|\alpha|+|\beta|$.

Define the projection

$$\mathcal{N}\sum_{\alpha\beta}h_{\alpha\beta}x^\alpha y^\beta=\sum_\alpha h_{\alpha\alpha}x^\alpha y^\alpha.$$

LEMMA 2.1. *Let $f(x,y)=h\circ\varphi_1^{-1}(x,y)=\hat{h}(x,y)+O(d)$ be as above. Let φ_2 be the the unique mapping defined by (2.1), where $S(x,y)$ is replaced by $K(x,y)=O(d)$ and $K_{\alpha\alpha}=0$ for all α, such that $f\circ\varphi_2^{-1}=\hat{h}$. Let $T=[K]_d$*

denote the sum of all monomials in K *of order* $d>2$. *Then*

$$\begin{aligned}\hat{h}(x,y)-\mathcal{N}f(x,y) &= \mathcal{N}\left\{\sum_{j,k=1}^{2}\lambda_j\big(y_jT_{y_jy_k}(x,y)T_{x_k}(x,y)\right.\\ &\qquad \left.-x_jT_{x_jy_k}(x,y)T_{x_k}(x,y)\big)+\sum_{j=1}^{2}\lambda_jT_{x_j}(x,y)T_{y_j}(x,y)\right\}\\ &\qquad +O(2d-1).\end{aligned} \tag{2.5}$$

Proof. Returning to (2.1) for $(\hat{x},\hat{y})=\varphi_2(x,y)$, we get

$$\hat{x}_j=x_j-K_{y_j}(x,y)-\sum_{k=1}^{2}T_{y_jy_k}(x,y)T_{x_k}(x,y)+O(2d-2),$$
$$\hat{y}_j=y_j+K_{x_j}(x,y)+\sum_{k=1}^{2}T_{x_jy_k}(x,y)T_{x_k}(x,y)+O(2d-2).$$

Now $\hat{h}(\hat{x},\hat{y})$ is equal to

$$\begin{aligned}f(x,y) &= \hat{h}(x,y)+\sum_{j,k=1}^{2}\lambda_j\big[x_jT_{x_jy_k}(x,y)T_{x_k}(x,y)\\ &\qquad -y_jT_{y_jy_k}(x,y)T_{x_k}(x,y)\big]-\sum_{j=1}^{2}\lambda_jT_{x_j}(x,y)T_{y_j}(x,y)\\ &\qquad +\sum_{|\alpha|\geq 2}\alpha_j\hat{h}_{\alpha\alpha}x^{\alpha-\delta_j}y^{\alpha-\delta_j}\big(x_jS_{x_j}(x,y)-y_jS_{y_j}(x,y)\big)+O(2d-1).\end{aligned} \tag{2.6}$$

Here the term in the last summation is zero if $\alpha_j=0$. Note that, for each j,

$$x_jS_{x_j}(x,y)-y_jS_{y_j}(x,y)=\sum_{\alpha\beta}(\alpha_j-\beta_j)S_{\alpha\beta}x^{\alpha}y^{\beta}$$

does not contain terms of the form x^ay^a. Therefore,

$$\mathcal{N}\big\{\alpha_jx^{\alpha-\delta_j}y^{\alpha-\delta_j}\big(x_jS_{x_j}(x,y)-y_jS_{y_j}(x,y)\big)\big\}=0.$$

Applying the projection $\mathcal{N}$ to (2.6), we get (2.5). □

We now identify the small divisors that contribute to the divergence of a Birkhoff normal form. The way that the small divisors appear will be crucial in the proof of the theorem. We will carry out computations in two steps. The first step is the following.

LEMMA 2.2. *Keep nations and assumptions in Lemma 2.1. Let* $N+m=d$, $\alpha=(N,m-1)$, $a=(N,0)$ *and* $b=(0,m)$. *Assume that* $m\geq 1$. *Then*

$$\hat{h}_{\alpha\alpha}=f_{\alpha\alpha}+\frac{m^2(\lambda_1 N-\lambda_2)f_{ab}f_{ba}}{(\lambda\cdot(a-b))^2}+\frac{f_{ab}A_{N+m}(f,\mu)+f_{ba}B_{N+m}(f,\mu)}{\lambda\cdot(a-b)}+C_{N+m}(f,\mu), \tag{2.7}$$

where f_{ab} *are of the form* (2.4). *Also* $A_{N+m}(f,\mu)$, $B_{N+m}(f,\mu)$ *are linear combinations of* $\mu_{\alpha'\beta'}f_{\alpha'\beta'}$ *with* $|\alpha'|+|\beta'|=N+m$ *and* $(\alpha',\beta')\neq(\alpha,\beta),(\beta,\alpha)$. *And* $C_{N+m}(f,\mu)$ *is a linear combination of* $f_{\alpha'\beta'}f_{\alpha''\beta''}$ *with* $|\alpha'|+|\beta'|=|\alpha''|+|\beta''|=N+m$ *and* $(\alpha',\beta'),(\alpha'',\beta'')\neq(\alpha,\beta),(\beta,\alpha)$.

Proof. Write

$$T(x,y)=T_{ab}x_1^N y_2^m+T_{ba}x_2^m y_1^N+\sum_{(a',b')\neq(a,b),(b,a)}T_{a'b'}x^{a'}y^{b'}.$$

Then we obtain

$$\begin{aligned}\sum_{j,k}\lambda_j x_j T_{x_jy_k}T_{x_k}&=T_{ab}T_{ba}\big(\lambda_1 Nm^2(x_1y_1)^N(x_2y_2)^{m-1}\\&\qquad+\lambda_2 mN^2(x_1y_1)^{N-1}(x_2y_2)^m\big)+\cdots,\\ \sum_{j,k}\lambda_j y_j T_{y_jy_k}T_{x_k}&=0+\cdots,\\ \sum_j\lambda_j T_{x_j}T_{y_j}&=T_{ab}T_{ba}\big(\lambda_1 N^2(x_1y_1)^{N-1}(x_2y_2)^m\\&\qquad+\lambda_2 m^2(x_1y_1)^N(x_2y_2)^{m-1}\big)+\cdots.\end{aligned}$$

In the above and the next formula, the omitted terms have coefficients that are linear combinations with integer coefficients in $T_{ab}T_{a'b'}$, $T_{ba}T_{a''b''}$, and $T_{a'b'}T_{a''b''}$ with (a',b'), $(a'',b'')\neq(a,b)$, (b,a). Thus by (2.5)

$$\begin{aligned}\hat{h}(x,y)-f(x,y)=T_{ab}T_{ba}\big\{&(\lambda_1-\lambda_2 m)N^2(x_1y_1)^{N-1}(x_2y_2)^m\\&+(\lambda_2-\lambda_1 N)m^2(x_1y_1)^N(x_2y_2)^{m-1}\big\}+\cdots.\end{aligned}$$

By (2.3) where h is replaced by f, we have

$$T_{\alpha\beta}=\frac{f_{\alpha\beta}}{\lambda\cdot(\alpha-\beta)}.$$

Combining the last two identities gives us (2.7). □

In the next step, we want to use (2.4) to further express $\hat{h}_{\alpha\alpha}$ in terms of coefficients of h and the small divisors.

PROPOSITION 2.3. *Let* $h(x,y)=\lambda_1x_1y_1+\lambda_2x_2y_2+O(3)$ *be a real analytic function. Assume that* λ_1,λ_2 *are non-resonant. Let* φ *be any formal symplectic map so that* $\hat{h}(x,y)=h\circ\varphi^{-1}(x,y)=\lambda_1x_1y_1+\lambda_2x_2y_2+O(3)$ *is in the*

Birkhoff normal form. Then for $\alpha = (N, m-1)$, $a = (N, 0)$, $b = (0, m)$ *with* $m \geq 1$, *one has*

$$\hat{h}_{\alpha\alpha} = h_{\alpha\alpha} + \frac{m^2(\lambda_1 N - \lambda_2)(h_{ab} + Q_{ab}(h,\lambda))(h_{ba} + Q_{ba}(h,\lambda))}{(\lambda \cdot (a-b))^2} + \frac{h_{ab}A_{ab}(h,\lambda) + h_{ba}B_{ab}(h,\lambda) + C_{ab}(h,\lambda)}{\lambda \cdot (a-b)} + \hat{Q}_{ab}(h,\lambda). \tag{2.8}$$

Here Q_{ab} *is a polynomial in* $h_{\alpha'\beta'}, \frac{1}{\lambda\cdot(\alpha''-\beta'')}$ *with*

$$\alpha'' \neq \beta'', \qquad \max\{|\alpha'| + |\beta'|, |\alpha''| + |\beta''|\} < |a| + |b|;$$

$\hat{Q}_{ab}$ *is a polynomial in* $h_{\alpha'\beta'}, \frac{1}{\lambda\cdot(\alpha''-\beta'')}$ *with*

$$\alpha'' \neq \beta'', \qquad (\alpha'', \beta'') \neq (a,b), (b,a),$$
$$|\alpha'| + |\beta'| < 2|\alpha|, \qquad |\alpha''| + |\beta''| \leq |a| + |b|;$$

and A_{ab}, B_{ab}, C_{ab} *are polynomials in* $h_{\alpha'\beta'}$, $\frac{1}{\lambda\cdot(\alpha''-\beta'')}$ *with*

$$\alpha'' \neq \beta'', \qquad (\alpha'', \beta'') \neq (a,b), (b,a),$$
$$\max\{|\alpha'| + |\beta'|, |\alpha''| + |\beta''|\} \leq |a| + |b|.$$

Proof. We apply a symplectic map φ_1 of the form (2.1), in which

$$S(x, \hat{y}) = \sum_{\alpha\neq\beta, 3\leq|\alpha|+|\beta|<N+m} S_{\alpha\beta} x^\alpha \hat{y}^\beta,$$

so that $\tilde{h} = h \circ \varphi_1^{-1}$ satisfies $\tilde{h}_{\alpha\beta} = 0$ for all $\alpha \neq \beta$ and $|\alpha| + |\beta| < N + m$. We know that $\tilde{h}_{\alpha\beta} = h_{\alpha\beta} + D_{\alpha\beta}(h,\lambda)$, where $D_{\alpha\beta}(h,\lambda)$ depends on $h_{\alpha'\beta'}$ with $|\alpha'| + |\beta'| < |\alpha| + |\beta|$ and on $1/(\lambda\cdot(\alpha'' - \beta''))$ with $|\alpha''| + |\beta''| < |a| + |b|$, $\alpha'' \neq \beta''$. Apply a formal symplectic map φ_2 of the form (2.1) with

$$S(x, \hat{y}) = \sum_{\alpha\neq\beta, |\alpha|+|\beta|\geq N+m} S_{\alpha\beta} x^\alpha \hat{y}^\beta,$$

so that $\tilde{h} \circ \varphi_2^{-1}$ is in the Birkhoff normal form. By (2.4) where h is actually $\tilde{h}$ and by (2.7), we can write (with abuse of notation for $Q_{ab}(h,\lambda)$)

$$\tilde{h}_{\alpha\alpha} + C_{N+m}(\tilde{h},\lambda) = h_{\alpha\alpha} + \hat{Q}_{ab}(h,\lambda),$$
$$\tilde{h}_{ab} + Q_{ab}(\tilde{h},\lambda) = h_{ab} + Q_{ab}(h,\lambda),$$
$$\tilde{h}_{ab}A_{N+m}(\tilde{h},\lambda) + \tilde{h}_{ba}B_{N+m}(\tilde{h},\lambda) = h_{ab}A_{ab}(h,\lambda) + h_{ba}B_{ab}(h,\lambda) + C_{ab}(h,\lambda).$$

Here $C_{ab}(h,\lambda) = D_{ab}(\tilde{h},\lambda)A_{N+m}(\tilde{h},\lambda) + D_{ba}(\tilde{h},\lambda)B_{N+m}(\tilde{h},\lambda)$, $A_{ab}(h,\lambda) = A_{N+m}(\tilde{h},\lambda)$ and $B_{ab}(h,\lambda) = B_{N+m}(\tilde{h},\lambda)$.

We have obtained (2.8), via the above normalizing map $\varphi_2 \circ \varphi_1$. On the other hand the Birkhoff normal form $\hat{h}$, with the same quadratic form as h, is independent of the normalizing map. In other words, the right-hand side of

(2.8) is independent of φ. Since each $\hat{h}_{\alpha\alpha}$ is a polynomial with integer coefficients in $h_{\alpha'\beta'}$ and $\frac{1}{\lambda\cdot(\alpha''-\beta'')}$, we conclude that each term in (2.8) depends only on h and is a polynomial in the sought form. $\square$

We now restrict ourselves to $|h_{\alpha\beta}|\leq 2$ for all α,β. Then we have

$$(2.9)\qquad \big(|Q_{ab}|+|Q_{ba}|+|A_{ab}|+|B_{ab}|+|C_{ab}|+|\hat{Q}_{ab}|\big)(h,\lambda)\leq \delta_{ab}(\lambda)^{-\tau_{ab}},$$

where $\tau_{ab}>1$ is a constant independent of λ and

$$\delta_{ab}(\lambda)=\min\left\{\frac{1}{2},\big|\lambda\cdot(\alpha-\beta)\big| : \alpha\neq\beta, |\alpha|+|\beta|\leq|a|+|b|, (\alpha,\beta)\neq(a,b),(b,a)\right\}.$$

Put $\lambda_2=1$. Notice that for $a=(N,0), b=(0,m)$, one has $|a-b|=|a|+|b|$. Thus, we can choose an irrational $\lambda_1\in(0,1)$ so that

$$(2.10)\qquad \big|(a-b)\cdot\lambda\big|=|N\lambda_1-m|<\frac{\delta_{ab}(\lambda)^{\tau_{ab}}}{100(N+m)!},\quad a=(N,0),\ b=(0,m)$$

hold for a sequence $(N,m)=(N_j,m_j)$ with N_j,m_j being positive integers. We may assume that $N_{j+1}+m_{j+1}>2(N_j+m_j)$. Put $a_j=(N_j,0)$ and $b_j=(0,m_j)$. Note that the existence of λ_1 can be obtained easily by modifying Siegel's argument [12] for

$$\lambda_1=\sum_{k=1}^{\infty}2^{-L_k},$$

where L_k are suitable positive integers tending to ∞ rapidly to ensure λ_1 is irrational and satisfies (2.10).

We now complete the proof of the theorem.

We shall find h whose coefficients $h_{\alpha\beta}$ are real. We also require that $h_{\alpha\beta}=h_{\beta\alpha}$ to show the divergence of normal forms of another type of quadratic parts for h. Put $h_{\alpha\beta}=0$ for all α,β with $|\alpha|+|\beta|>2$ and $(\alpha,\beta)\neq(a_j,b_j),(b_j,a_j)$. Inductively, we shall choose $h_{a_jb_j}=h_{b_ja_j}=0,2$, or -2 as follows. Notice that if u_0,v_0 are real and $|u_0v_0|<1$, then either $(u_0+2)(v_0+2)\geq 2$ or $(u_0-2)(v_0-2)\geq 2$; otherwise, we would have both $u_0+v_0<-1/2$ and $u_0+v_0>1/2$, which is a contradiction. Therefore for two real numbers u_0,v_0, choosing (u,v) among $(0,0)$, $(2,2)$ and $(-2,-2)$ yields $|(u_0+u)(v_0+v)|\geq 1$. This shows that we can find $h_{a_jb_j}=h_{b_ja_j}=0$, 2 or -2, so that

$$(2.11)\qquad \big|\big(h_{ab}+Q_{ab}(h)\big)\big(h_{ba}+Q_{ba}(h)\big)\big|\geq 1,\quad a=a_j,\ b=b_j.$$

Here, we already used $N_{j+1}+m_{j+1}>2(N_j+m_j)$, which implies that if (2.11) holds for $a=a_j,b=b_j$ then it remains true no matter how a_{j+1},b_{j+1} are chosen. By (2.8) we have

$$\begin{aligned}|\hat{h}_{\alpha\alpha}|\geq \big|\lambda\cdot(a-b)\big|^{-2}\big\{&m^2\big|(\lambda_1N-\lambda_2)\big(h_{ab}+Q_{ab}(h)\big)\big(h_{ba}+Q_{ba}(h)\big)\big|\\&-\big|\lambda\cdot(a-b)\big|^2\big(|h_{\alpha\alpha}|+\big|\hat{Q}_{\alpha\alpha}(h)\big|\big)\\&-\big|\lambda\cdot(a-b)\big|\big(\big|h_{ab}A_{ab}(h)\big|+\big|h_{ba}A_{ba}(h)\big|+\big|C_{ab}(h)\big|\big)\big\}.\end{aligned}$$

Recall that $\lambda_1 \in (0,1)$, $\lambda_2 = 1$ and $|h_{\alpha\beta}| \leq 2$. When N is sufficiently large, we have $\lambda_1 N - \lambda_2 > 2$. Recall that $\delta_{ab}(\lambda) < 1 < \tau_{ab}$. Thus by (2.9)–(2.11) and for $(N,m) = (N_j, m_j)$ with j sufficiently large, we obtain

$$|\hat{h}_{\alpha\alpha}| \geq \big|\lambda \cdot (a-b)\big|^{-2} \Big\{ 2m^2 - \left(\frac{\delta_{ab}^{\tau_{ab}}(\lambda)}{100(N+m)!} \right)^2 \cdot 3\delta_{ab}^{-\tau_{ab}}(\lambda) - \frac{\delta_{ab}^{\tau_{ab}}(\lambda)}{100(N+m)!} \cdot 5\delta_{ab}^{-\tau_{ab}}(\lambda) \Big\} > \big|\lambda \cdot (a-b)\big|^{-2}.$$

Finally, we conclude that for j sufficiently large

$$|\hat{h}_{\alpha\alpha}| > \big|\lambda \cdot (a-b)\big|^{-2} > (N+m)!, \quad \alpha = (N_j, m_j - 1).$$

This shows the divergence of $\hat{h}$.

We reformulate our theorem to cover another case.

PROPOSITION 2.4. *There exist some non-resonant real numbers λ_1 and λ_2 with $\lambda_1\lambda_2 > 0$ and a real analytic function $h(x,y) = \lambda_1(x_1^2 + y_1^2) + \lambda_2(x_2^2 + y_2^2) + O(3)$ such that the Birkhoff normal form of h is divergent.*

Proof. Indeed, for the above analytic real function $h(x,y)$ on $\mathbf{R}^2 \times \mathbf{R}^2$, its complexification, denoted by $h(z,w)$, is holomorphic near $0 \in \mathbf{C}^2 \times \mathbf{C}^2$. Let φ be a formal symplectic map of $\mathbf{R}^4$, which is tangent to the identity, so that $h \circ \varphi^{-1}(x,y) = g(x_1y_1, x_2y_2)$ is in the normal form. Since φ preserves $\omega = dx_1 \wedge dy_1 + dx_2 \wedge dy_2$, its complexification, still denoted by φ, preserves $\omega^c = dz_1 \wedge dw_1 + dz_2 \wedge dw_2$.

Let $L(\xi,\eta) = (\xi + i\eta, \xi - i\eta)$. Notice that $L^*\omega^c = -2i(d\xi_1 \wedge d\eta_1 + d\xi_2 \wedge d\eta_2)$. Thus $\psi = L^{-1}\varphi L$ preserves $d\xi_1 \wedge d\eta_1 + d\xi_2 \wedge d\eta_2$. Also $\tilde{h} \circ \psi^{-1}(\xi,\eta) = g(\xi_1^2 + \eta_1^2, \xi_2^2 + \eta_2^2)$ for $\tilde{h} = h \circ L$. In other words, $\tilde{h} \circ \psi^{-1}$ is the (formal holomorphic) Birkhoff normal form with respect to the holomorphic symplectic 2-form $d\xi_1 \wedge d\eta_1 + d\xi_2 \wedge d\eta_2$. Notice that the quadratic form of $\tilde{h}$ is now $\lambda_1(\xi_1^2 + \eta_1^2) + \lambda_2(\xi_2^2 + \eta_2^2)$. Let e be the restriction of $\tilde{h}$ on $\mathbf{R}^2 \times \mathbf{R}^2 : \xi = \overline{\xi}, \eta = \overline{\eta}$. Since $h_{\alpha\beta} = \overline{h}_{\beta\alpha}$ by construction, then e is real-valued. Thus $e(\xi,\eta)$ is an analytic real function of the form $\lambda_1(\xi_1^2 + \eta_1^2) + \lambda_2(\xi_2^2 + \eta_2^2) + O(3)$, while $L^*\omega^c$, restricted to $\mathbf{R}^2 \times \mathbf{R}^2 : \xi = \overline{\xi}, \eta = \overline{\eta}$, is a constant multiple of the standard symplectic real 2-form. Therefore $\tilde{h} \circ \psi^{-1}$, restricted to $\xi = \overline{\xi}, \eta = \overline{\eta}$, is a real Birkhoff normal form of e; since h diverges, one readily sees the divergence of the restriction. □

Our theorem is valid for higher dimension $\mathbf{R}^{2n}$. Indeed such a real analytic function which has a divergent normal form can be achieved by adding suitable quadratic forms in the remaining variables in higher dimension. Furthermore, one can see, from the proof of the theorem, that the set of real analytic Hamiltonian functions with a divergent Birkhoff normal form is dense in a suitable topology.

We emphasize that our theorem does not deal with the case when the Hamiltonian functions on $\mathbf{R}^4$ have eigenvalues $\lambda_1, -\lambda_1, \lambda_2, \overline{\lambda}_2$ for which λ_2/λ_1 is not real. This is an interesting case since by a theorem of Moser [8] the real analytic Hamiltonian system can be solved real analytically. In fact, its Birkhoff normal form can be realized by a convergent symplectic transformation (see Bruno [1], pp. 228–229, and Giorgilli [3]).

References

[1] A. D. Brjuno, *Analytic form of differential equations I, II*, Tr. Mosk. Mat. Ob. **25** (1971), 119–262; ibid. **26** (1972), 199–239. MR 0377192

[2] L. H. Eliasson, *Normal forms for Hamiltonian systems with Poisson commuting integrals—elliptic case*, Comment. Math. Helv. **65** (1990), 4–35. MR 1036125

[3] A. Giorgilli, *Unstable equilibria of Hamiltonian systems*, Discrete Contin. Dynam. Systems **7** (2001), 855–871. MR 1849665

[4] X. Gong, *Conformal maps, monodromy transformations, and non-reversible Hamiltonian systems*, Math. Res. Lett. **7** (2000), 471–476. MR 1783625

[5] X. Gong, *Levi-flat invariant sets of holomorphic symplectic mappings*, Ann. Inst. Fourier (Grenoble) **51** (2001), 151–208. MR 1821073

[6] H. Ito, *Convergence of Birkhoff normal forms for integrable systems*, Comment. Math. Helv. **64** (1989), 412–461. MR 0998858

[7] J. Moser, *Nonexistence of integrals for canonical systems of differential equations*, Comm. Pure Appl. Math. **8** (1955), 409–436. MR 0079171

[8] J. Moser, *On the generalization of a theorem of A. Liapounoff*, Comm. Pure Appl. Math. **11** (1958), 257–271. MR 0096021

[9] R. Pérez-Marco, *Convergence or generic divergence of the Birkhoff normal form*, Ann. of Math. (2) **157** (2003), 557–574. MR 1973055

[10] H. Rüssmann, *Über das Verhalten analytischer Hamiltonscher Differentialgleichungen in der Nähe einer Gleichgewichtslösung*, Math. Ann. **154** (1964), 285–300. MR 0179409

[11] H. Rüssmann, *Über die Normalform analytischer Hamiltonscher Differentialgleichungen in der Nähe einer Gleichgewichtslösung*, Math. Ann. **169** (1967), 55–72. MR 0213679

[12] C. L. Siegel, *On integrals of canonical systems*, Ann. Math. (2) **42** (1941), 806–822. MR 0005818

[13] C. L. Siegel, *Über die Existenz einer Normalform analytischer Hamiltonscher Differntialgleichungen in der Nähe einer Gleichgewichtslösung*, Math. Ann. **128** (1954), 144–170. MR 0067298

[14] C. L. Siegel and J. K. Moser, *Lectures on celestial mechanics*, Translated from the German by C. I. Kalme, Reprint of the 1971 translation, Classics in Mathematics, Springer, Berlin, 1995. MR 1345153

[15] L. Stolovitch, *Singular complete integrability*, Inst. Hautes Études Sci. Publ. Math. **91** (2001), 133–210. MR 1828744

[16] J. Vey, *Sur certains systèmes dynamiques séperables*, Amer. J. Math. **100** (1978), 591–614. MR 0501141

Xianghong Gong, Department of Mathematics, University of Wisconsin, Madison, WI 53706, USA

E-mail address: gong@math.wisc.edu