ANNALES DE L'INSTITUT FOURIER

XIANGHONG GONG

Integrable analytic vector fields with a nilpotent linear part

Annales de l'institut Fourier, tome 45, nº 5 (1995), p. 1449-1470
<http://www.numdam.org/item?id=AIF_1995__45_5_1449_0>

Ann. Inst. Fourier, Grenoble
45, 5 (1995), 1449–1470

INTEGRABLE ANALYTIC VECTOR FIELDS WITH A NILPOTENT LINEAR PART

by Xianghong GONG

Introduction.

We are concerned with the normalization of an analytic vector field v given by

$$\frac{dx}{dt} = y + f(x,y), \qquad \frac{dy}{dt} = g(x,y), \tag{1.1}$$

where f, g are convergent power series starting with terms of order two. Since the matrix of the linear part in (1.1) is nilpotent, the Poincaré-Dulac normal form gives no simplification. In [7], F. Takens introduced a simplification for (1.1) by formal transformations as follows :

$$\frac{dx}{dt} = y + r(x), \qquad \frac{dy}{dt} = s(x). \tag{1.2}$$

The above system is subject to further classifications. Using representation theory of certain Lie algebras, A. Baider and J. C. Sanders [2] gave a complete classification of (1.2) under suitable non-degeneracy conditions. However, the normal form of Baider and Sanders excludes the case $s(x) \equiv 0$. We shall see that the vanishing of s corresponds to the case that v has a non-singular formal integral, i.e. a formal power series $H(x,y)$ such that $dH(0) \neq 0$ and $\langle v, \nabla H \rangle = 0$.

Partially supported by NSF grant DMS-9304580 at the Institute for Advanced Study.
Key words : Integrable vector field – Nilpotent linear part – Embeddability of mappings.
Math. classification : 58F36 – 32S65.

The bifurcation theory on vector fields in the form (1.1) was initially studied by F. Takens [7] and R. I. Bogdanov [3]. We also refer to [1] for the survey on vector fields with a nilpotent linear part. In this paper we shall deal with the convergence of the initial normalization given by Takens for integrable vector fields. We shall prove the following.

THEOREM 1.1. — *Let v be an analytic vector field given by (1.1). Assume that v has a non-singular formal integral. Then v can be transformed into (1.2) with $s \equiv 0$ through a convergent transformation.*

We cannot give a complete classification for (1.1) under the assumption of formal integrability. As a partial normalization, we have

THEOREM 1.2. — *Let v be an analytic vector field given as in Theorem 1.1. Assume that f, g in (1.1) are real power series. Then there is a real formal transformation which transforms (1.1) into*

$$\frac{dx}{dt} = y + r^*(x), \qquad \frac{dy}{dt} = 0, \tag{1.3}$$

where $r^* \equiv 0$, or

$$r^*(x) = \epsilon x^\sigma + \sum_{j>\sigma, \sigma \nmid j} r_j^* x^j,$$

and $\epsilon = 1$ for σ even, $\epsilon = \pm 1$ for σ odd. Furthermore, when σ is even, all r_j^* $(\sigma < j < 2\sigma)$ are invariants; when σ is odd, ϵ and the coefficients r_j^* $(\sigma < j < 2\sigma)$ are invariants, of which the first non-zero coefficient of even order is normalized to be positive.

We shall see that there are infinitely many invariants for (1.3) if $2 < \sigma < \infty$. We are unable to determine whether (1.3) can be realized by convergent transformations. Neither shall we deal with the convergence of the normal form of Baider-Sanders. In fact, the convergence proof for the vector fields considered here depends essentially on the assumption of the existence of non-singular integrals.

Theorem 1.1 demonstrates a significant difference between real analytic vector fields with a nilpotent linear part and real analytic mappings with a unipotent linear part, i.e. parabolic mappings. The parabolic mappings arise naturally from glancing hypersurfaces considered by R. B. Melrose [6], and also real analytic Lagrangian surfaces with a complex tangent studied by S. M. Webster [8]. In [5], it was shown that there exist real analytic transformations which are formally equivalent to the linear parabolic mapping $T(x, y) = (x + y, y)$, but they are not linearizable through any convergent transformation. On the other hand, Theorem 1.1 shows that a vector field (1.1) is linearizable by convergent transformations if and only if it is formally linearizable. To further understand the distinct nature of normalizing analytic vector fields with nilpotent linear parts and parabolic analytic mappings, we shall prove the following result.

COROLLARY 1.3. — *There exists a smoothly linearizable real analytic transformation $\varphi = T + O(2)$ which is not embeddable in any neighborhood of the origin as the time-1 mapping of any real analytic vector field (1.1).*

This paper is organized as follows. In section 2, we shall consider the formal theory of the integrable vector field (1.1). We shall also characterize the integrability in terms of the formal normal form as well as the singular points of the vector fields. The proof of Theorem 1.2 will be given in section 2. In section 3, we shall prove the convergence of solutions to the approximate equations arising from the normalizing the vector fields in Theorem 1.1. We complete the proof of Theorem 1 in section 4 through a KAM argument. In section 5, we shall discuss the embeddability of a parabolic mapping into the flow of a vector field (1.1). The proof of Corollary 1.3 will be presented in section 5.

We would like to thank the referee for bringing the article of D. Cerveau and R. Moussu [4] to our attention. In terms of the vector fields, the results in [4] contain a complete holomorphic (convergent) classification for all ideals of holomorphic vector fields in $\mathbb{C}^2$ which are formally equivalent to a given ideal of vector fields defined by $\omega = 0$, where the holomorphic 1-form ω is a certain perturbation of $ydy - x^n dx$.

2. Formal normalizations.

In this section, we shall first construct a formal transformation Φ which transforms (1.1) into (1.2). The convergence of Φ will be determined in section 4. We shall also discuss the formal integrability. Finally, we shall give a proof for Theorem 1.2.

Throughout the discussion of this paper, we shall decompose a power series $p(x, y)$ into the following form:

$$p(x, y) = p_0(x) + p_1(x, y), \qquad p_0(x) = p(x, 0).$$

Then the mapping $p \to p_1$ defines a projection $\Pi: \mathbb{C}[[x,y]] \to y\mathbb{C}[[x,y]]$, where $\mathbb{C}[[x,y]]$ is the ring of formal power series in x, y, and $y\mathbb{C}[[x,y]]$ is the ideal generated by y. For a non-zero power series $p(x,y)$, we denote by $\operatorname{ord} p$ the largest integer k such that $p_{\alpha,\beta} = 0$ for $\alpha + \beta < k$. We put $\operatorname{ord} p = \infty$ when $p \equiv 0$. We also write

$$p(x,y) = O(k),$$

if $\operatorname{ord} p \geq k$.

Let φ be a transformation defined by

$$x' = x + u(x,y), \qquad y' = y + v(x,y), \tag{2.1}$$

where u, v are power series starting with terms of order two. We say that φ is *normalized* if

$$u(0,y) \equiv 0 \equiv v(0,y). \tag{2.2}$$

One can see that φ is a normalized transformation if and only if φ preserves the y-axis, and its restriction to the y-axis is the identity mapping. Therefore, the normalized transformations form a group.

We shall seek a unique normalized transformation

$$\Phi: x' = x + U(x,y), \qquad y' = y + V(x,y),$$

which transforms (1.1) into (1.2). This leads to the following functional equations:

$$yU_x(x,y) - r(x) = V(x,y) - f(x,y) + E_1(x,y), \tag{2.3}$$

$$yV_x(x,y) - s(x) = -g(x,y) + E_2(x,y), \tag{2.4}$$

where

$$\begin{aligned} E_1(x,y) &= r(x+U) - r(x) - fU_x(x,y) - gU_y(x,y), \\ E_2(x,y) &= s(x+U) - s(x) - fV_x(x,y) - gV_y(x,y). \end{aligned} \tag{2.5}$$

We shall prove that under the normalizing condition (2.2), the equations (2.3) and (2.4) have a unique solution $\{U, V, r, s\}$. Let us denote by $E_{j;\alpha,\beta}$ the coefficient of the term $x^\alpha y^\beta$ of E_j. Then, it is easy to see that for given $\alpha + \beta = n$, $E_{j;\alpha,\beta}$ is a polynomial in $r_{\alpha'}, s_{\alpha'}$ $(\alpha' < n)$ and $f_{\alpha',\beta'}, g_{\alpha',\beta'}$ $(\alpha' + \beta' < n)$ with integer coefficients. Comparing the coefficient of $x^{\alpha-1}y^{\beta+1}$ on the both sides of (2.4), we get

$$V_{\alpha,\beta} = \frac{1}{\alpha}\left(-g_{\alpha-1,\beta+1} + E_{2;\alpha-1,\beta+1}\right), \quad 1 \leq \alpha \leq n. \tag{2.6}$$

By the normalizing condition (2.2), we also have

$$U_{0,n} = 0, \quad V_{0,n} = 0. \tag{2.7}$$

Next, we compare the coefficient of $x^{\alpha-1}y^{\beta+1}$ on the both sides of (2.3), and obtain

$$U_{\alpha,\beta} = \frac{1}{\alpha}\left(V_{\alpha-1,\beta+1} - f_{\alpha-1,\beta+1} + E_{1;\alpha-1,\beta+1}\right), \quad 1 \le \alpha \le n. \tag{2.8}$$

Finally, the coefficient of x^n on both sides of (2.3) and (2.4) gives us

$$r_n = f_{n,0} - E_{1;n,0} - V_{n,0}, \quad s_n = g_{n,0} - E_{2;n,0}. \tag{2.9}$$

Notice that $E_{j;\alpha,\beta} = 0$ for $\alpha + \beta = 2$. Thus, the coefficients of r, s, U, V of order 2 are uniquely determined by the coefficients of f, g of order 2 through formulas (2.6)-(2.9). By induction, one can show that the coefficients of r, s, U, V of order n are uniquely determined by the coefficients of f, g of order n. Therefore, there exists a unique solution $\{r, s, U, V\}$ to (2.3) and (2.4), of which U, V satisfy the condition (2.2). This proves that (1.1) can be transformed into (1.2) by a unique normalized formal transformation.

For the late use, we observe that if (2.3) and (2.4) are solvable for r, s, U, V with $s \equiv 0$, then g must satisfy the condition

$$\operatorname{ord} g_0 \ge \operatorname{ord} g_1. \tag{2.10}$$

To see this, we put $s \equiv 0$ in (2.4) and (2.5). From (2.5), we see that $\operatorname{ord} E_2 \ge \operatorname{ord} V + 1$. By (2.2), we have $\operatorname{ord} V = \operatorname{ord}\{yV_x(x,y)\}$. Applying Π to (2.4), we then get $\operatorname{ord} V \ge \operatorname{ord} g_1$. By comparing the orders on both sides of (2.4), one can see easily that (2.10) holds.

Next, we want to describe the integrability. We have

PROPOSITION 2.1. — *Let v be an analytic vector field defined by (1.1). Assume that (1.1) is transformed into (1.2) through a formal transformation φ. Then the following are equivalent:*

(a) $s(x) \equiv 0$.

(b) *v has a non-singular formal integral.*

(c) *The set of singular points of v is a curve through the origin.*

Remark 2.2. — There exist integrable vector fields with a nilpotent linear part, which have an isolated singular point at the origin. For instance, consider

$$\frac{dx}{dt} = y, \qquad \frac{dy}{dt} = x^2.$$

Then $H(x,y) = 2x^3 - 3y^2$ is an integral of the system. Hence, it is essential that the integral in Theorem 1.1 is non-singular.

From Theorem 1.1 and Proposition 2.1, we have the following.

COROLLARY 2.3. — *Let v be an analytic vector field given by (1.1). Then v is linearizable by convergent transformations if and only if it is formally linearizable.*

Proof of Proposition 2.1. — It is obvious that (a)$\Rightarrow$(b). To show that (b)$\Rightarrow$(a), we assume that $s(x)$ does not vanish identically and write

$$s(x) = s_\tau x^\tau + \dots, \qquad s_\tau \neq 0.$$

We want to prove that τ is the largest positive integer k such that there exists a formal power series H with

$$\text{(2.11)} \qquad \operatorname{ord}(\langle v, \nabla H\rangle) \geq k, \qquad dH(0) \neq 0.$$

First, the existence of such power series H does not depend on the choice of formal coordinates. Hence, we may assume that v is a formal vector field in the form (1.2). Moreover, if $H(x,y) = y$, then $\operatorname{ord}(\langle v, \nabla H\rangle) = \tau$.

Next, we want to show that there is no power series H such that (2.11) holds for some $k > \tau$. Assume that such a power series H exists, and put

$$H(x,y) = \sum_{j=0}^{\infty} H_j(y) x^j.$$

From (2.11), we get

$$yH_x(x,y) = -r(x)H_x(x,y) - s(x)H_y(x,y) + O(k).$$

Expanding both sides as power series in x and comparing the coefficient of x^{j-1}, we obtain

$$\text{(1.12)} \quad jyH_j(y) = -\sum_{l=2}^{j-1}(j-l)r_l H_{j-l}(y) - \sum_{l=\tau}^{j-1} s_l H'_{j-l}(y) + O(k-j+1)$$

for $1 \leq j \leq k$. In particular, we have

$$\text{(2.13)} \qquad H_j(y) = O(k-j)$$

for $j = 1, 2$. In fact, we want to show that (2.13) holds for $1 \leq j \leq \tau$. For the induction, we assume that (2.13) holds for $1 \leq j \leq \tau', \tau' < \tau$. Then for $2 \leq l \leq \tau'$, we get $H_{\tau'-l} = O(k - \tau' - 1 + l)$. It is clear that $k - \tau' - 1 + l \geq k - \tau' + 1$ for $l \geq 2$. Hence

$$H_{\tau'-l} = O(k - \tau' + 1).$$

Set $j = \tau' + 1$ in (2.12). Then the first summation on the right side of (2.12) can be replaced by $O(k - \tau' + 1)$. The second summation in (2.12) vanishes, since $\tau' < \tau$. Hence, (2.13) holds for $j = \tau'$. Therefore, we have verified by induction that (2.13) holds for $1 \leq j \leq \tau$. We now take $j = \tau + 1$ for (2.12), and get

$$(\tau + 1) y H_{\tau+1}(y) = -s_\tau H_0'(y) + O(k - \tau).$$

Since $dH(0) \neq 0$, we have $H_0'(0) \neq 0$. Hence, the above identity cannot hold if $k > \tau$. Therefore, the order τ of s is an invariant.

Now, the formal integrability implies that there is a formal power series H such that (2.11) holds for all integer k, which yields $s \equiv 0$.

(b)$\Leftrightarrow$(c). Notice that the origin is not an isolated singular point of v, if and only if $g(x, y)$ can be divided by $y + f(x, y)$. We have

$$\begin{pmatrix} y' + r(x') \\ s(x') \end{pmatrix} = D\varphi(x, y) \cdot \begin{pmatrix} y + f(x, y) \\ g(x, y) \end{pmatrix}, \tag{2.14}$$

where $D\varphi$ is the Jacobian matrix of the transformation $(x', y') = \varphi(x, y)$. We first assume that (b) holds. Then (2.14) implies that $y' + r(x')$ and $s(x')$ can be divided by $y + f(x, y)$ for $(x, y) = \varphi^{-1}(x', y')$. Since $y' + r(x')$ is irreducible, then $s(x')$ must be divided by $y' + r(x')$. Hence, we can write

$$s(x') = a(x', y')(y' + r(x')) \tag{2.15}$$

for some convergent power series $a(x', y')$. Assume for contradiction that $a \neq 0$. As power series in x', y', we compare orders on both sides of (2.15) and get

$$\operatorname{ord} s = \operatorname{ord} a + 1.$$

Next, we set $y = 0$ in (2.15). Then as power series in x alone, we obtain

$$\operatorname{ord} s \geq \operatorname{ord} a + 2.$$

Thus, the contradiction implies that $a \equiv 0$, i.e. $s \equiv 0$. Conversely, let us assume that (c) holds. Then from (2.14), we see that both $y + f(x, y)$ and $s(x)$ can be divided by $s \circ \varphi(x, y)$. Since $y + f(x, y)$ is irreducible, then $s(x)$ must be divided by $y + f(x, y)$. Hence, we get (b). This completes the proof of Proposition 2.1. $\square$

Next, we assume that $s(x) \equiv 0$. By a linear transformation

$$(x, y) \to (ax, ay),$$

one can achieve that

$$r_\sigma = \epsilon = \begin{cases} 1, & \text{if } \sigma \text{ is even,} \\ \pm 1, & \text{if } \sigma \text{ is odd.} \end{cases}$$

With the normalization for r_σ, we have $a = 1$ if σ is even, and $a = \pm 1$ if σ is odd.

From the proof of Proposition 2.1, we see that the vanishing order of $s(x)$ is an invariant of the system (1.2). One can also give a characterization for the vanishing order of $r(x)$ when $s \equiv 0$. Here, we need Theorem 1.1. First, Theorem 1.1 implies that v is actually integrable. The curve of the singular points of v is

$$S\colon y + f(x, y) = 0.$$

Let γ be the level curve of integral passing through the origin. Then σ is the order of contact of γ and S at the origin. When σ is even, S is located on one side of γ. If σ is odd and $\epsilon = -1$, then the orbits of the vector field are attracted to S. When σ is odd and $\epsilon = +1$, the orbits will leave S along level curves of integral.

Proof of Theorem 1.2. — We put $i + j\sigma$ to be the *weight* of $x^i y^j$. For a power series $p(x, y)$, we denote by $p_{n;j}$ the coefficient of $x^i y^j$ with weight n, and also by $\operatorname{wt} p$ the largest integer n such that all coefficients of p of weight less than n vanish.

Assume that φ is a transformation which transforms (1.2) into

$$\frac{d\xi}{dt} = \eta + \epsilon \xi^\sigma + \sum_{j=\sigma+1}^{\infty} r_j^* \xi^j, \qquad \frac{d\eta}{dt} = 0.$$

Then φ has the form

$$\begin{cases} \xi = ax + u(x, y), & \operatorname{wt} u \geq 2, \\ \eta = ay + v(y), \end{cases} \tag{2.16}$$

where $a = 1$ if σ is even, and $a = \pm 1$ if σ is odd. Notice that $a^{\sigma-1} = 1$. Then we have the following functional equation:

$$y u_x + \epsilon x^\sigma u_x - \epsilon \sigma x^{\sigma-1} u - r_0^*(ax) + a r_0(x) = v(y) + E(x, y), \tag{2.17}$$

where

$$r_0(x) = \sum_{j>\sigma} r_j x^j, \qquad r_0^*(x) = \sum_{j>\sigma} r_j^* x^j,$$

$$E(x, y) = r_0^*(ax + u) - r_0^*(ax) - r_0 u_x + \epsilon\left((ax + u)^\sigma - a x^\sigma - \sigma x^{\sigma-1} u\right).$$

We want to show that $\operatorname{wt} u \geq \sigma$. Let

$$u(x, y) = u_{k;0} x^k + \dots, \qquad 1 < k < \sigma.$$

Notice that the weight of terms in E is at least $k + \sigma$. On the other hand, the coefficient of $x^{k-1} y$ on the left side of (2.17) is $k u_{k;0}$. Hence, $u_{k;0} = 0$.

This shows that the weight of u is at least σ. By collecting terms in (2.17) with weight less than $2\sigma - 1$, we get

$$r_j^* a^j = a r_j, \qquad \sigma < j < 2\sigma - 1.$$

Next, by comparing the coefficients of terms of weight $2\sigma - 1$ in (2.17), we obtain

$$u_{\sigma;0} = \epsilon u_{\sigma;1}, \qquad r_{2\sigma-1}^* = r_{2\sigma-1}.$$

Therefore, the coefficients r_j $(\sigma + 1 \leq j \leq 2\sigma - 1)$ are invariants, if we restrict the first non-zero coefficient of even order to be positive when σ is odd.

We now assume that $a = 1$ in (2.16). To achieve that $r_{(m+1)\sigma}^* = 0$ for $m \geq 1$, we compare the coefficients of weight $(m+1)\sigma$ on both sides of (2.17), which gives

$$u_{n;0} = \frac{\epsilon}{\sigma - n}\left(r_{n+\sigma-1} - E_{n+\sigma-1;0}\right), \quad n = m\sigma + 1,$$
$$u_{n;j} = \frac{E_{n+\sigma-1;j} - (n - (j-1)\sigma)\, u_{n;j-1}}{\epsilon\,(n - (j+1)\sigma)}, \quad 1 \leq j \leq m,$$
$$v_{m+1} = u_{n;m} - E_{n+\sigma-1;m+1}.$$

Therefore, the coefficients of u of weight $m\sigma + 1$ and the coefficient of v of weight $(m+1)\sigma$ are uniquely determined by r_k $(k \leq (m+1)\sigma)$ and the coefficients of u with weight less than $m\sigma + 1$.

Next, consider the coefficients of u of weight

$$n = m\sigma + k, \qquad 0 \leq k < \sigma, \quad k \neq 1.$$

From (2.17), we get

$$k u_{n;m} = E_{n+\sigma-1;m+1}, \qquad k \geq 1,$$
$$u_{n;j-1} = \frac{E_{n+\sigma-1;j} - \epsilon\,(n - (j+1)\sigma)\, u_{n;j}}{n - (j-1)\sigma}, \quad 1 \leq j \leq m,$$
$$r_{n+\sigma-1}^* = r_{n+\sigma-1} + \epsilon\,(n - \sigma)\, u_{n;0} - E_{n+\sigma-1;0}.$$

Hence, $r_{(m+1)\sigma+k-1}^*$ and coefficients of u with weight $m\sigma + k$, except for $u_{m\sigma;m}$, are uniquely determined by r_j with $j \leq (m+1)\sigma + k - 1$ and the coefficients of u with weight less than $m\sigma + k$. Therefore, one can achieve that $r_{j\sigma}^* = 0$ for $j = 2, 3, \ldots, m$ through a formal transformation (2.16), of which the coefficients of u with weight up to $(m-1)\sigma + 1$ and coefficients of v with weight up to $m\sigma$ are uniquely determined by the coefficients $u_{\sigma;1}, u_{2\sigma;2}, \ldots, u_{(m-1)\sigma;m-1}$. □

Furthermore, by counting the number of coefficients, we see that the system (1.3) has infinitely many invariants when $\sigma \geq 3$.

3. Solutions to approximate equations.

The convergence of transformation Φ cannot be determined directly from the functional equations (2.3) and (2.4). In this section, we shall give some estimates of solutions to the approximate equations.

We shall consider the following approximate equations:

$$\begin{aligned}(3.1)\qquad -yu_x(x,y)+\widetilde{f}_0(x) &= f(x,y)+f_0(x)u_x(x,y)\\ &\quad - f_0'(x)u(x,y)-v(x,y),\\ (3.2)\qquad -yv_x(x,y)+\widetilde{g}_0(x) &= g(x,y)+f_0(x)v_x(x,y),\end{aligned}$$

in which $\tilde{f}_0$ and $\tilde{g}_0$ are added to adjust terms purely in x.

One can see that under the normalizing condition (2.2), power series $u, v, \tilde{f}_0, \tilde{g}_0$ are determined uniquely from (3.1) and (3.2). The proof can be given by an argument similar to the proof of the existence and the uniqueness of solutions r, s, U, V to (2.3) and (2.4). We left the details to the reader.

With the above solution $\{u, v\}$, we define a formal transformation φ by (2.1). Assume that φ transforms (1.1) into

$$(3.3)\qquad \frac{dx'}{dt}=y'+p(x',y'),\qquad \frac{dy'}{dt}=q(x',y').$$

Then we have the following identities:

$$(3.4)\quad p(x',y')=f(x,y)+(y+f(x,y))u_x(x,y)+g(x,y)u_y(x,y)-v(x,y),$$

$$(3.5)\quad q(x',y')=g(x,y)+(y+f(x,y))v_x(x,y)+g(x,y)v_y(x,y)$$

with $(x,y)=\varphi^{-1}(x',y')$. We denote

$$d_0=\min\{\operatorname{ord} f_1,\operatorname{ord} g\},\quad d_1=\min\{\operatorname{ord} p_1,\operatorname{ord} q\}.$$

We need the following.

LEMMA 3.1. — *Assume that f, g in (1.1) are holomorphic in Δ_r. Let u, v be solutions to (3.1) and (3.2), which satisfy the normalizing condition (2.2). Assume that φ defined by (2.1) transforms (1.1) into (3.3). If v defined by (1.1) has a non-singular formal integral, then*

$$(3.6)\qquad d_1\geq 2d_0-1.$$

Proof. — Applying Π to (3.2), we get

$$\operatorname{ord}\{yv_x\}\geq\min\{\operatorname{ord} g,\operatorname{ord}\{f_0v_x\}\}.$$

By (2.2), we also have $\operatorname{ord} v_x = \operatorname{ord} v - 1$. Notice that $\operatorname{ord} f \geq 2$. Hence

$$\operatorname{ord} v \geq d_0. \tag{3.7}$$

We now apply Π to (3.1) and get

$$\operatorname{ord} u \geq d_0. \tag{3.8}$$

Adding $yu_x(x,y) - f_0(x)$ to both sides of (3.1), we see that

$$\operatorname{ord}(\tilde{f}_0 - f_0) \geq d_0. \tag{3.9}$$

Eliminating v from (3.1) and (3.4), we get

$$p(x+u, y+v) = \tilde{f}_0(x) + f_0' u(x,y) + f_1 u_x(x,y) + g u_y(x,y).$$

From (3.7) and (3.8), we see that

$$p(x,y) = \tilde{f}_0(x) + O(d_0), \tag{3.10}$$

$$p(x+u, y+v) = \tilde{f}_0(x) + f_0'(x)u(x,y) + O(2d_0 - 1). \tag{3.11}$$

Combining (3.10) and (3.9), we obtain that $p(x,y) = f_0(x) + O(d_0)$, i.e.

$$p_1(x,y) = O(d_0), \qquad p_0(x) = f_0(x) + O(d_0).$$

Thus, we get

$$p(x+u, y+v) = p(x,y) + f_0'(x)u(x,y) + O(2d_0 - 1).$$

Now (3.11) yields

$$p(x,y) = \tilde{f}_0(x) + O(2d_0 - 1).$$

In particular, we have

$$\operatorname{ord} p_1 \geq 2d_0 - 1. \tag{3.12}$$

From (3.2) and (3.5), it follows that

$$q(x+u, y+v) = \tilde{g}_0(x) + O(2d_0 - 1).$$

This implies that $\operatorname{ord} q \geq d_0$, and

$$q(x+u, y+v) = q(x,y) + O(2d_0 - 1).$$

Therefore, we have $q(x,y) = \tilde{g}_0(x) + O(2d_0 - 1)$. In particular, we obtain that

$$\operatorname{ord} q_1 \geq 2d_0 - 1. \tag{3.13}$$

Notice that (3.3) also has a non-singular formal integral. From (2.10), it follows that

$$\operatorname{ord} q_0 \geq \operatorname{ord} q_1.$$

Hence, (3.12) and (3.13) give us (3.6). □

We now want to show the convergence of u and v. We need the following notations :

$$\begin{aligned} a_n &= |f_{n+1,0}|, & \mu_n &= \max_{\alpha+\beta=n,\alpha\geq 1}\{|u_{\alpha,\beta}|\}, \\ b_n &= \max_{\alpha+\beta=n,\beta\geq 1}\{|f_{\alpha,\beta}|,|g_{\alpha,\beta}|\}, & \nu_n &= \max_{\alpha+\beta=n,\alpha\geq 1}\{|v_{\alpha,\beta}|\}. \end{aligned}$$

Given two power series $p(x,y)$ and $q(x,y)$, we shall denote $p \prec q$, if $|p_{\alpha,\beta}| \leq q_{\alpha,\beta}$ for all $\alpha,\beta\geq 0$. We shall also denote

$$\hat{p}(x,y)=\sum |p_{\alpha,\beta}|x^\alpha y^\beta.$$

Comparing the coefficient of $x^\alpha y^\beta$ on both sides of (3.2), we have

$$-(\alpha+1)v_{\alpha+1,\beta-1}=g_{\alpha,\beta}+\sum_{\alpha'+\alpha''=\alpha}(\alpha'+1)v_{\alpha'+1,\beta}f_{\alpha'',0}$$

for $\beta\geq 1$ and $\alpha+\beta=n$. Let $\gamma=\alpha''-1$. Then $\alpha'+1+\beta=n-\gamma$. Hence

$$|v_{\alpha+1,\beta-1}|\leq b_n+\sum \nu_{n-\gamma}a_\gamma.$$

Therefore, we have

$$\nu(t)\prec a(t)\nu(t)+b(t),$$

which gives us

$$\nu(t)\prec \frac{b(t)}{1-a(t)}. \tag{3.14}$$

Solving (3.1) for u, one gets

$$\mu(t)\prec \nu(t)+2a(t)\mu(t)+b(t).$$

From (3.4), it follows that

$$\mu(t)\prec \frac{1}{1-2a(t)}\left(\frac{b(t)}{1-a(t)}+b(t)\right),$$

which yields

$$\mu(t)\prec \frac{2b(t)}{(1-2a(t))^2}. \tag{3.15}$$

Remark 3.2. — The results in this paper, except for Theorem 1.2 which needs obvious modifications, are valid for holomorphic vector fields. In fact, for the proof of Theorem 1.1, we shall introduce holomorphic coordinates.

From now on, we shall treat all variables as complex variables until we finish the proof of Theorem 1.1. Let us introduce

$$\|f\|_r = \max\{|f(x,y)|; (x,y) \in \Delta_r\}, \quad \Delta_r = \{(x,y) \in \mathbb{C}^2; |x| \le r, |y| \le r\}.$$

Assume that f and g are holomorphic on Δ_r. Denote

$$B_0 = \max\{\|f_1\|_r, \|g\|_r\}.$$

Assume also that

$$A_0 = \|\hat{f}_0\|_r \le \frac{r}{4}, \qquad 0 < r < 1, \quad 0 < \theta < 1/4. \tag{3.16}$$

The Cauchy inequalities give

$$|f_{\alpha,\beta}| \le \frac{\|f_1\|_r}{r^{\alpha+\beta}}, \qquad |g_{\alpha,\beta}| \le \frac{\|g_1\|_r}{r^{\alpha+\beta}}, \quad \beta \ge 1.$$

Hence, we get

$$\|b\|_{(1-\theta)r} \le \sum_{k=2}^{\infty} \frac{B_0}{r^k}((1-\theta)r)^k \le \frac{B_0}{\theta}. \tag{3.17}$$

We now have

$$\begin{aligned}\|\hat{u}\|_{(1-2\theta)r} &\le \sum_{k\ge 2}(k+1)\mu_k((1-2\theta)r)^k \le 2r\sum_{k\ge 2} k\mu_k((1-2\theta)r)^{k-1}\\ &\le 2r\|\mu'\|_{(1-2\theta)r} \le \frac{2\|\mu\|_{(1-\theta)r}}{\theta},\end{aligned}$$

in which the last inequality comes from the Cauchy formula. Now (3.15)–(3.17) yield

$$\|\hat{u}\|_{(1-2\theta)r} \le \frac{c_1 B_0}{\theta^2}, \tag{3.18}$$

where, and also in the rest of discussion, $c_j > 1$ stands for a constant. With a similar computation, one can also obtain the following estimate:

$$\|\hat{v}\|_{(1-2\theta)r} \le \frac{c_2 B_0}{\theta^2}. \tag{3.19}$$

We are ready to prove the following.

LEMMA 3.3. — *Let φ be as in Lemma 3.1. Suppose that, for $c_3 = \max\{c_1, c_2\}$,*

$$A_0 \le \frac{r}{4}, \qquad B_0 \le \frac{\theta^3 r}{4c_3}. \tag{3.20}$$

Then we have

$$\varphi\colon \Delta_{(1-2\theta)r} \to \Delta_{(1-\theta)r}, \qquad \varphi^{-1}\colon \Delta_{(1-4\theta)r} \to \Delta_{(1-3\theta)r}. \tag{3.21}$$

Proof. — From (3.18)–(3.20), it is easy to see that $\varphi: \Delta_{(1-2\theta)r} \to \Delta_{(1-\theta)r}$. To show the existence of the inverse mapping, we fix $(x', y') \in \Delta_{(1-4\theta)r}$ and consider the mapping

$$T(x,y) = (x' - u(x,y), y' - v(x,y)).$$

It is clear that T maps $\Delta_{(1-3\theta)r}$ into itself. From (3.18) and (3.20), we get

$$\|u_x\|_{(1-3\theta)r} \leq \frac{\|u\|_{(1-2\theta)r}}{\theta r} < \frac{1}{4}.$$

Similarly, we can verify that $\|u_y\|_{(1-3\theta)r}$, $\|v_x\|_{(1-3\theta)r}$ and $\|v_y\|_{(1-3\theta)r}$ are less than $1/4$. This implies that with the norm

$$\|(x,y)\| = \max\{|x|, |y|\},$$

T is a contraction mapping. By the fixed-point theorem, T has a unique fixed point (x,y) in $\Delta_{(1-3\theta)r}$, which is clearly $\varphi^{-1}(x', y')$. □

Let us keep the notations and assumptions given in Lemma 3.3. Fix $(x', y') \in \Delta_{(1-4\theta)r}$. Then $(x,y) = \varphi^{-1}(x', y') \in \Delta_{(1-4\theta)r}$. From (3.4), we have

$$\begin{aligned}|f(x', y') - f(x,y)| &= \left| \int_0^1 \frac{d}{dt} f(x' - tu(x,y), y' - tv(x,y))\, dt \right| \\ &\leq \|f_x\|_{(1-3\theta)r} \|u\|_{(1-3\theta)r} + \|f_y\|_{(1-3\theta)r} \|v\|_{(1-3\theta)r}.\end{aligned}$$

From (3.20), it follows that $\|f\|_r < r$. Now the Cauchy formula gives

$$\|f_x\|_{(1-3\theta)r} \leq \frac{1}{3\theta}.$$

A similar estimate also holds for f_y. From (3.18) and (3.19), we now get

$$|f(x', y') - f(x,y)| \leq \frac{(c_1 + c_2)B_0}{\theta^3}. \tag{3.22}$$

From (3.20), we have

$$|y + f(x,y)| \leq (1 - 3\theta)r + A_0 + B_0 < 3r.$$

Using (3.18) and the Cauchy formula, we get

$$\|u_x\|_{(1-3\theta)r} \leq \frac{c_1 B_0}{\theta^3 r}, \qquad \|u_y\|_{(1-3\theta)r} \leq \frac{c_1 B_0}{\theta^3 r}.$$

Hence

$$|(y + f(x,y)) u_x(x,y)| \leq \frac{3c_1 B_0}{\theta^3}. \tag{3.23}$$

We also have

$$|g(x,y) u_y(x,y)| \leq B_0 \frac{c_1 B_0}{\theta^3 r} \leq B_0, \tag{3.24}$$

in which the last inequality is obtained by using (3.20) to get rid of one of two B_0's.

Substituting (3.22), (3.23) and (3.24) into (3.4), we get

$$\|p - f\|_{(1-4\theta)r} \leq \frac{c_4 B_0}{\theta^3}.$$

In particular, we have the estimates

$$\|p_0 - f_0\|_{(1-4\theta)r} \leq \frac{c_4 B_0}{\theta^3}, \qquad \|p_1 - f_1\|_{(1-4\theta)r} \leq \frac{2c_4 B_0}{\theta^3}. \tag{3.25}$$

From (3.5), one also gets

$$\|q\|_{(1-4\theta)r} \leq \frac{c_5 B_0}{\theta^3}. \tag{3.26}$$

We are ready to prove the following.

PROPOSITION 3.4. — *Let f, g, p, q be as in Lemma 3.1. Then there exist two positive constants c_0 and ϵ_0 satisfying the following property. If*

$$A_0 = \|\hat{f}_0\|_r \leq r/4, \qquad B_0 = \max\{\|f_1\|_r, \|g\|_r\} \leq \epsilon_0 \theta^3 r, \tag{3.27}$$

then φ, defined by (2.1), transforms (1.1) into (3.3) such that

$$A_1 = \|\hat{p}_0\|_{(1-5\theta)r} \leq A_0 + \frac{c_0}{\theta^4} B_0, \tag{3.28}$$

$$B_1 = \max\{\|p_1\|_{(1-5\theta)r}, \|q\|_{(1-5\theta)r}\} \leq \frac{c_0}{\theta^3} B_0 (1-\theta)^{d_1}, \tag{3.29}$$

in which $d_1 = \max\{\operatorname{ord} p_1, \operatorname{ord} q\}$.

Proof. — We choose $\epsilon_0 = 1/(4c_3)$. Then (3.27) implies that (3.20) holds. From (3.25) and Cauchy inequalities, we get

$$|p_{k,0} - f_{k,0}| \leq \frac{c_4 B_0}{\theta^3} \cdot \frac{1}{((1-4\theta)r)^k}.$$

Hence

$$\|\widehat{p_0 - f_0}\|_{(1-5\theta)} \leq \frac{c_4 B_0}{\theta^3} \sum_{k=2}^{\infty} \left(\frac{1-5\theta}{1-4\theta}\right)^k \leq \frac{c_4 B_0}{\theta^4}.$$

This gives us (3.28), if we choose $c_0 \geq c_4$.

From (3.25), we have

$$\|p_1\|_{(1-4\theta)r} \leq \|f_1\|_{(1-4\theta)r} + \|p_1 - f_1\|_{(1-4\theta)r} \leq \frac{3c_4 B_0}{\theta^3}.$$

Now, the Schwarz lemma yields

$$\|p_1\|_{(1-5\theta)r} \leq \|p_1\|_{(1-4\theta)r} \left(\frac{1-5\theta}{1-4\theta}\right)^{d_1}.$$

Notice that $(1-5\theta)/(1-4\theta) < 1-\theta$. Therefore, we get

$$\|p_1\|_{(1-5\theta)r} \leq \frac{3c_4 B_0}{\theta^3}(1-\theta)^{d_1}.$$

From (3.26) and the Schwarz inequality, we also have

$$\|q\|_{(1-5\theta)r} \leq \frac{c_5 B_0}{\theta^3}(1-\theta)^{d_1}.$$

Let us put $c_0 = \max\{3c_4, c_5\}$. Then the last two inequalities yield (3.29). □

4. A KAM argument.

In this section, we shall first construct a sequence of formal transformations Φ_n such that (1.1) is transformed into (1.2) under the limit transformation of $\{\Phi_n\}$. We shall use a KAM argument to show the convergence of the sequence Φ_n.

We shall construct a sequence of systems

$$\frac{dx}{dt} = y + p_n(x,y), \qquad \frac{dy}{dt} = q_n(x,y),$$

where $p_0(x,y) = f(x,y)$ and $q_0(x,y) = g(x,y)$ give the initial system (1.1). Recursively, p_n, q_n are obtained through the transformation φ_n constructed through the approximate equations in section 3.

Let us decompose

$$p_n(x,y) = p_{n;0}(x) + p_{n;1}(x,y), \qquad p_{n;0}(x) = p_n(x,0).$$

Put $d_n = \min\{\operatorname{ord} p_{n;1}, \operatorname{ord} q_n\}$. We have $d_0 \geq 2$. From 3.1, we know that

$$d_n \geq 2^n + 1.$$

We now put

$$r_n = \frac{1}{2}\left(1 + \frac{1}{n+1}\right) r_0, \quad n = 0, 1, \ldots,$$

in which $r_0 < 1$ will be determined late. Rewrite

$$r_{n+1} = (1 - 5\theta_n) r_n, \qquad \theta_n = \frac{1}{5(n+2)^2}, \quad n = 0, 1, \ldots$$

We need to choose r_0 so small that the following norms are well-defined:

$$A_n = \|\hat{p}_{n;0}\|_{r_n}, \qquad B_n = \max\{\|p_{n;1}\|_{r_n}, \|q_n\|_{r_n}\}.$$

Let us first prove a numerical result.

LEMMA 4.1. — *Let r_n, θ_n, d_n be given as above, and let ϵ_0, c_0 be as in Lemma 3.3. Then there exists $\epsilon_1 < \epsilon_0$, which is independent of r_0, such that for two sequences of non-negative numbers $\{A_n^*\}_{n=0}^\infty$ and $\{B_n^*\}_{n=0}^\infty$, we have*

$$A_n^* \leq r_n/4, \qquad B_n^* \leq \frac{\epsilon_0 \theta_n^4 r_n}{c_0 2^{n+2}}, \quad n = 1, 2, \ldots, \tag{4.1}$$

provided that for all n

$$A_{n+1}^* \leq A_n^* + \frac{c_0 B_n^*}{\theta_n^4}, \qquad B_{n+1}^* \leq \frac{c_0 B_n^*}{\theta_n^3} (1-\theta_n)^{d_{n+1}}, \tag{4.2}$$

$$A_0^* \leq r_0/16, \qquad B_0^* \leq \epsilon_1 \theta_0^4 r_0. \tag{4.3}$$

Proof. — We put

$$\hat{B}_n = \frac{\epsilon_0 \theta_n^4 r_n}{c_0 2^{n+2}}.$$

Clearly, we see that $\hat{B}_{n+1}/\hat{B}_n \to 1$ as $n \to \infty$. On the other hand, $d_n > 2^n$ implies that for large n, one has $d_{n+1} > \theta_n^{-2}$. Hence, we have

$$(1-\theta_n)^{d_{n+1}} < (1-\theta_n)^{1/\theta_n^2} < (1/2)^{1/\theta_n^2},$$

if n is sufficiently large. Now, it is easy to see that

$$\frac{c_0}{\theta_n^3} (1-\theta_n)^{d_{n+1}} \to 0.$$

Hence, it follows from (4.2) that there exists n_0 independent of the choice of r_0 such that

$$B_{n+1}^* \leq B_n^* \frac{\hat{B}_{n+1}}{\hat{B}_n}, \qquad \text{for } n \geq n_0. \tag{4.4}$$

Choose ϵ_1 so small that if B_0^* satisfies the condition (4.3), then

$$B_n^* \leq \hat{B}_n, \qquad 0 \leq n \leq n_0.$$

Thus, (4.4) yields the estimate of B_n^* in (4.1).

As for the estimate of A_n^*, we have

$$A_n^* \leq A_0^* + \sum_{j=0}^{n} \frac{c_0 B_j^*}{\theta_j^4}.$$

Using the estimate of B_n^* obtained above, we get

$$A_n^* \le \frac{r_0}{16} + \sum_{j=0}^{n} \frac{\epsilon_0 r_n}{2^{n+2}}.$$

Notice that $r_0/2 < r_n < r_0$. Hence

$$\sum_{j=0}^{n} \frac{r_n}{2^n} < r_0 \sum_{j=0}^{\infty} \frac{1}{2^j} = 2r_0 < 4r_n.$$

Thus, we obtain

$$A_n^* \le r_n/8 + \epsilon_0 r_n.$$

We may assume that ϵ_0, chosen in the proof of Lemma 3.3, is less than $1/8$. Therefore, we obtain the desired estimate of A_n^*. □

Proof of Theorem 1.1. — In order to apply Lemma 4.1 to $\{A_n\}_{n=0}^{\infty}$ and $\{B_n\}_{n=0}^{\infty}$, we need to choose r_0. Since $\operatorname{ord} f_0 \ge 2$, we may choose r_0 so small that $\hat{f}_0$ converges for $|x| \le r_0$, and

$$\|\hat{f}_0\|_{r_0} \le r_0/16. \tag{4.5}$$

By choosing a smaller r_0 if it is necessary, we may also assume that f_1, g are holomorphic functions on Δ_{r_0}, and

$$\|f_1\|_{r_0} \le \epsilon_1 \theta_0^4 r_0, \qquad \|g\|_{r_0} \le \epsilon_1 \theta_0^4 r_0. \tag{4.6}$$

From (4.5) and (4.6), it follows that (3.27) is satisfied. Hence, Proposition 3.4 says that $\{A_0, A_1\}$ and $\{B_0, B_1\}$ satisfy (4.2). One also sees that two initial conditions in (4.3) follow from (4.5) and (4.6). Now, Lemma 4.1 implies that

$$A_1 \le r_1/4, \qquad B_1 \le \frac{\epsilon_0 \theta_1^4 r_1^2}{c_0 2^3}.$$

In particular, this gives us two initial conditions in (3.20) for the new system defined by p_1 and q_1. Hence, we may apply Proposition 3.4 again. By repeating this process, we can prove that

$$A_n \le r_n/4, \qquad B_n \le \frac{\epsilon_0 \theta_n^4 r_n}{c_0 2^{n+2}}, \quad n = 0, 1, \ldots \tag{4.7}$$

Therefore, Lemma 3.3 gives us

$$\varphi_n^{-1} \colon \Delta_{r_{n+1}} \to \Delta_{r_n}, \qquad n = 0, 1, \ldots$$

Notice that $r_n \ge r_0/2$. We have

$$\Phi_n^{-1} = \varphi_0^{-1} \circ \varphi_1^{-1} \circ \ldots \circ \varphi_n^{-1} \colon \Delta_{\frac{1}{2}r_0} \to \Delta_{r_0}.$$

Hence

$$\|\Phi_{n+1}^{-1} - \Phi_n^{-1}\|_{\frac{1}{2}r_0} \leq 2r_0,$$

where the norm on the left side is defined to be the maximum of norms of two components. On the other hand, we know from (3.9) that each component of $\Phi_{n+1}^{-1} - \Phi_n^{-1}$ vanishes with order at least $d_n > 2^n$. From Schwarz Lemma, it follows that

$$\|\Phi_{n+1}^{-1} - \Phi_n^{-1}\|_{\frac{1}{4}r_0} \leq r_0 \left(\frac{1}{2}\right)^{2^n}.$$

Therefore, the sequence Φ_n^{-1} converges to a transformation Φ_∞^{-1}.

It is clear that Φ_∞ transforms the system (1.1) into a system of the form (1.2). Since the normalized transformations form a group, we see that Φ_∞ is still a normalized transformation. In section 2, we have seen that Φ is the unique formal transformation which transforms (1.1) into (1.2). Therefore, we obtain that $\Phi = \Phi_\infty$, so Φ is a convergent transformation. □

5. Embeddability of parabolic mappings.

In this section, we shall investigate the relation between the embeddability of parabolic mappings as time-1 mappings and the convergence of normalization for parabolic transformations. We shall first show that a parabolic mapping is formally linearizable if and only if it is embeddable as a time-1 mapping of a formally linearizable vector field.

Let us put

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \qquad A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Rewrite (1.1) as

$$\frac{dx}{dt} = Ax + \sum_{|I|\geq 2} F_I x^I \equiv F(x),$$

where each F_I is a constant matrix of 2 by 1, and $x^I = x_1^\alpha x_2^\beta$ for $I = (\alpha, \beta)$.

We first assume that $F(x)$ is only given by formal power series. Let φ_t be a family of formal transformations generated by the formal vector field (5.1), i.e.

$$\frac{d}{dt}\varphi_t = F \circ \varphi_t, \qquad \varphi_0 = \mathrm{Id}. \tag{5.2}$$

Since the linear part of φ_t with respect to x_1 and x_2 is determined by the matrix A, then we have the expansion

$$\varphi_t(x) = e^{At}x + \sum_{|I|\geq 2} B_I(t)x^I, \qquad B_I(0) = 0,$$

where $B_I(t)$ is a matrix of 2 by 1 given by formal power series, and

$$e^{At} = \sum_{n=0}^{\infty} \frac{A^n}{n!}t^n = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

Now, (5.2) takes the form

$$B_I'(t) = AB_I(t) + b_I(t), \qquad B_I(0) = 0, \tag{5.3}$$

where $b_I(t)$ depends only on A and $B_J(t)$ with $|J| < |I|$. More precisely, we have

$$b_I(t) = \left\{ \sum_{|J|=|I|} F_J(e^{At}x) \right\}_I + \tilde{b}_I(t), \tag{5.4}$$

where $\{\cdot\}_I$ denotes the matrix of coefficients for x^I, and

$$\tilde{b}_I(t) \equiv 0, \qquad \text{if } F_J = 0, \text{ for all } |J| < |I|.$$

The solution B_I to (5.3) is given by

$$B_I(t) = e^{At}\int_0^t e^{-At}b_I(t)\,dt. \tag{5.5}$$

Hence, $B_I(t)$ are real analytic functions defined on the whole real line. In particular, we see that φ_t is a family of formal transformations defined for $-\infty < t < \infty$. One also notices that if $\tilde{v} = \Psi_* v$ for a formal transformation Ψ, then the 1-parameter family of formal transformations generated by $\tilde{v}$ are given by $\Psi \circ \varphi_t \circ \Psi^{-1}$.

We need the following lemma.

LEMMA 5.1. — *Let v be a formal vector field defined by (5.1), and φ_t the 1-parameter family of formal transformations generated by v. Assume that (5.1) is not a linear system. Then φ_t is not a linear transformation for all $t \neq 0$.*

Proof. — We assume that there is $I_0 = (\alpha_0, \beta_0)$ such that

$$F_{I_0} \neq 0, \qquad F_J = 0,$$

for $|J| < |I_0|$, or for $J = (\alpha, |I_0| - \alpha)$ with $\alpha > \alpha_0$. Then from (5.4), it follows that

$$b_{I_0}(t) = F_{I_0} \neq 0.$$

Next, we use the formula (5.5) and get

$$B_{I_0}(t) = \begin{pmatrix} t & t^2/2 \\ 0 & t \end{pmatrix} b_{I_0} \neq 0, \qquad \text{for } t \neq 0,$$

which implies that φ_t is not a linear transformation for each $t \neq 0$. □

We now consider a parabolic transformation

$$\varphi(x,y) = T(x,y) + O(2), \qquad T(x,y) = (x+y,y). \tag{5.6}$$

In [5], it was proved that there exist real analytic parabolic mappings which are not linearizable by any convergent transformation. In fact, the parabolic mappings are constructed through a pair of real analytic glancing hypersurfaces. On the other hand, Melrose [6] showed that a pair of smooth glancing hypersurfaces can always be put into a certain normal form by smooth transformations; and consequently, the parabolic mappings coming from a pair of smooth glancing hypersurfaces are always linearizable by smooth transformations. Therefore, we can state the following.

THEOREM 5.2. — *There exists a smoothly linearizable real analytic transformation φ of the form (5.6), which cannot be transformed into T by any convergent transformation.*

Now, we see that Corollary 1.3 follows from the following.

PROPOSITION 5.3. — *Let φ be a real analytic transformation of the form (5.6). Assume that φ is formally equivalent to T. Then φ is a time-1 mapping of a real analytic vector field of the form (5.1), if and only if φ can be transformed into T through a convergent transformation.*

Proof. — Obviously, φ is embeddable if it is linearizable through convergent transformations. We now assume that $\varphi = \varphi_1$ for a 1-parameter family of transformations φ_t generated by a real analytic v of the form (5.1). Let Φ be a formal transformation which linearizes φ. This implies that the time-1 mapping of the formal vector field $\Phi_* v$ is a linear transformation. From Lemma 5.1, it follows that $\Phi_* v$ is a linear vector field. Now, Theorem 1.1 implies that v is linearizable by a convergent transformation ϕ. Using Lemma 5.1 again, we know that φ is also linearizable by the same transformation ϕ. □

BIBLIOGRAPHY

[1] V.I. ARNOL'D and Yu. S. IL'YASHENKO, Ordinary differential equations, in "Dynamical Systems I, EMS" vol. 1, Springer-Verlag, Berlin, 1990.

[2] A. BAIDER and J. C. SANDERS, Further reduction of the Takens-Bogdanov normal form, J. Diff. Equations, 99 (1992), 205–244.

[3] R.I. BOGDANOV, Versal deformation of a singularity of a vector field on the plane in the case of zero eigenvalues, Seminar Petrovski (1976), and Selecta Math. Soviet, n° 4, 1 (1981), 389–421.

[4] D. CERVEAU and R. MOUSSU, Groupes d'automorphismes de $(\mathbb{C}, 0)$ et équations différentielles $y\,dy + \ldots = 0$, Bull. Soc. Math. France, 116 (1988), 459–488.

[5] X. GONG, Divergence for the normalization of real analytic glancing hypersurfaces, Commun. Partial Diff. Equations, 19, n° 3 & 4 (1994), 643–654.

[6] R.B. MELROSE, Equivalence of glancing hypersurfaces, Invent. Math., 37 (1976), 165–191.

[7] F. TAKENS, Singularities of vector fields, Publ. Math. I.H.E.S., 43 (1974), 47–100.

[8] S.M. WEBSTER, Holomorphic symplectic normalization of a real function, Ann. Scuola Norm. Sup. di Pisa, 19 (1992), 69–86.

Manuscrit reçu le 2 février 1995,
accepté le 22 mars 1995.

Xianghong GONG,
The Institute for Advanced Study
School of Mathematics
Princeton, NJ 08540 (USA).
Principal address:
Mathematical Sciences Research Institute
1000 Centennial Drive
Berkeley, CA 94720 (USA).
gong@msri.org.