JOURNAL OF DIFFERENTIAL EQUATIONS 2, 182-194 (1966)

Linear Normal Forms of Differential Equations*

RICHARD J. VENTI

Sandia Laboratory, Albuquerque, New Mexico

1. INTRODUCTION

Consider the system of real differential equations

$$\dot{x} = \Lambda x + X(x), \qquad \left(\dot{x} = \frac{dx}{dt}\right) \tag{1}$$

where x, $X(x)$ are n-vectors, Λ is a constant square matrix of order n, X is of class C^k on a neighborhood of $x = 0$ for some integer $k \geqslant 1$, and $X(x) = o(|x|)$ as $x \to 0$ ($|x|$ = Euclidean norm). This paper generalizes a result of Sternberg and a result of Hartman concerning the existence of a C^k change of variables

$$y = x - \varphi(x), \qquad \text{where} \qquad \varphi(x) = o(|x|) \qquad \text{as} \qquad x \to 0, \tag{2}$$

which is defined on a neighborhood of $x = 0$ and transforms (1) into the linear system

$$\dot{y} = \Lambda y. \tag{3}$$

If such a change of variables exists, (1) is said to be C^k equivalent to (3), and (3) is called the linear normal form of (1).

For $k \geqslant 2$ it can be shown ([*1*], Lemma 12.1, p. 258) that if the eigenvalues $\lambda_1, \cdots, \lambda_n$ of Λ satisfy

$$\lambda_i \neq \sum_{j=1}^{n} m_j \lambda_j, \qquad i = 1, \cdots, n, \tag{4}$$

for all sets of non-negative integers $m_1, \cdots, m_n$ such that

$$2 \leqslant \sum_{j=1}^{n} m_j \leqslant k,$$

* This research was supported by the Office of Naval Research and by the United States Atomic Energy Commission. The results here are an extension of the author's thesis [*9*], which was written under the direction of Professor S. P. Diliberto. Reproduction in whole or in part is permitted for any purpose of the U. S. Government.

then there exists an analytic change of variables (2) defined for small $|x|$ and transforming (1) into

$$\dot{y} = Ay + Y(y),$$

where $Y \in C^k$ on a neighborhood of $y = 0$ and $Y(y) = o(|y|^k)$ as $y \to 0$. Moreover, if (4) is violated then (1) need not be C^k equivalent to (3), i.e., there exists an $X = X(x)$ satisfying the conditions of paragraph one and such that (1) is not C^k equivalent to (3). In fact, an analytic X may be chosen (for an example, see [2], p. 812).

Sternberg [2] has proven the following theorem.

THEOREM A. *In (1), suppose that X is of class C^k on a neighborhood of $x = 0$ for some $k \geqslant 2$ and $X(x) = o(|x|^k)$ as $x \to 0$. If the eigenvalues of A satisfy*

$$\operatorname{Re} \lambda_i < 0, \qquad i = 1, \cdots, n \tag{5}$$

$$\max |\operatorname{Re} \lambda_i| / \min |\operatorname{Re} \lambda_i| < k, \tag{6}$$

then (1) is C^k equivalent to (3). (The case where the Re $\lambda_i > 0$ *may be reduced to the case where (5) holds by making the change of variables $t = -s$ in (1).)*

The following result, for the case $k = 1$, is due to Hartman [3].

THEOREM B. *In (1), suppose that X is of class C^1 in a neighborhood of $x = 0$ and $X(x) = o(|x|)$ as $x \to 0$. If the derivatives of X are uniformly Lipschitz continuous for small $|x|$ and the eigenvalues of A satisfy (5), then (1) is C^1 equivalent to (3).*

The main result of this paper, Theorem 3, generalizes both of the above results and shows that the seemingly wide difference between the hypotheses of Theorem A and the hypotheses of Theorem B disappears when the problem of C^k linearizations is considered in the framework of $C^k[x, \mu]$ systems.

Condition (5) will be essential to all arguments in this paper. We first prove (Lemma 1) that a partially linear system, satisfying certain conditions, has a smooth manifold of solutions which may be used (Theorem 1) to put the system into a particular nonlinear form. Then we show (Theorem 2) that this nonlinear form can be further linearized. Repeated application of Theorems 1 and 2 yield Theorem 3, which roughly states that (1) is C^k equivalent to (3) provided that the Re λ_i satisfy a certain spacing condition. To what extent the spacing condition, which depends on the smoothness of X, can be weakened remains undecided. In the last section, we construct some C^k systems (1) which are not C^k equivalent to (3); these examples show that condition (5) together with $X(x) = o(|x|^k)$ as $x \to 0$ is not enough to ensure the existence of a C^k linearizing map (2).

Sufficient conditions for the existence of smooth linearizing maps (2) in the absence of condition (5) have been given by Nagumo and Isé [*4*], Sternberg [*5*], Hartman [*3*], Chen [*6*], and Brjuno [*7*].

In the analytic case the problem of normal forms has been considered by Poincaré, Siegel and others. (For references, see [*1*], pp. 271-272.)

I would like to express my sincere gratitude to Professor S. P. Diliberto for suggesting this topic and for advising and encouraging me throughout the course of this work. I am also grateful to Dr. Allen Kelley for many helpful conversations.

2. Preliminaries

Let $F(x) = (f_{ij}(x))_{pxq}$ be a real C^1 matrix function on real x-space, where x is finite dimensional, and define F_i, $\mathbf{F}$, F_p, F^x by

$$F_i = i\text{th row of } F \qquad \mathbf{F} = (F_1 \cdots F_p)_{1xpq},$$

$$F_x = \text{the Jacobian matrix } \partial_x \mathbf{F}, \qquad F^x = (F^T)_x,$$

where F^T denotes the transpose of F. For $r = 1, 2, \cdots$, define $\alpha_r(F)$, $\beta_r(F)$ by

$$\alpha_r(F) = \begin{pmatrix} F & 0 & 0 & \cdots & 0 \\ 0 & F & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & 0 & F \end{pmatrix}_{rpxrq} \qquad \beta_r(F) = \begin{pmatrix} \alpha_r(F_1) \\ \vdots \\ \alpha_r(F_p) \end{pmatrix}_{rpxrq}.$$

Then if $F = DE$ where D, E are pxs, sxq C^1 matrix functions, respectively, it is easily verified that

$$F_x = \beta_q(D)\, E^x + \alpha_p(E^T)\, D_x$$
$$F^x = \alpha_q(D)\, E^x + \beta_p(E^T)\, D_x. \tag{7}$$

For a real constant rectangular matrix Γ define the norm $|\Gamma|$ to be the max $|\Gamma\xi|$ for $|\xi| = 1$, where ξ is a real vector of appropriate dimension. With this norm it follows that

$$|\alpha_r(\Gamma)| = |\beta_r(\Gamma)| = |\Gamma| \tag{8}$$

for every positive integer r.

Let f be a real continuous function defined on a set N of real (x, y)-space, where x, y are finite dimensional, and let μ be a real number satisfying $0 \leqslant \mu \leqslant 1$. We say that f is uniformly μ-Holder continuous on N with respect to x, if there exists a $\gamma > 0$ such that

$$|f(x, y) - f(\bar{x}, y)| \leqslant \gamma(|x - \bar{x}|^{\mu})$$

for every (x, y), $(\bar{x}, y) \in N$. If on some neighborhood of $(x, y) = 0$, $f \in C^k$, $f(x, y) = o(|x, y|^k)$ as $(x, y) \to 0$ and the kth order derivatives of f are uniformly ν-Holder continuous with respect to x and uniformly μ-Holder continuous with respect to y, then f is said to be of class $C^k[x, y; \nu, \mu]$ (if the variable y is absent, we simply write $f \in C^k[x; \nu]$). A vector function is of class $C^k[x, y; \nu, \mu]$ if each of its components is.

In the next section we shall need the following simple fact:

$$\text{If} \quad 0 < \alpha < 1, \quad p_n, \quad q_n \geqslant 0, \quad q_n \to 0 \quad \text{as} \quad n \to \infty \quad \text{and}$$
$$p_{n+1} \leqslant \alpha p_n + q_n \quad \text{for} \quad n = 1, 2, \cdots, \quad \text{then} \quad p_n \to 0 \quad \text{as} \quad n \to \infty. \tag{9}$$

3. A Nonlinear Normalization

Consider the real system of differential equations

$$\dot{x} = Ax$$
$$\dot{y} = By + Y(x, y), \tag{10}$$

where x, y, $Y(x, y)$ are vectors of dimension m, n, n, respectively, and A, B are constant square matrices of order m, n, respectively.

Lemma 1. *Suppose that* $Y \in C^k[x, y; \nu, \mu]$, *where* $k \geqslant 1$ *and* $0 < \nu \leqslant \mu \leqslant 1$, *and* A, B *satisfy*

$$\max \{\text{Re}\, \lambda : \lambda = \text{eigenvalues of } A \text{ or } B\} < 0, \tag{11}$$

$$\max_{\lambda = \text{e.v. of } B} \text{Re}\, \lambda < (k + \nu) \min_{\lambda = \text{e.v. of } A} \text{Re}\, \lambda. \tag{12}$$

Then there exists an n-vector function $\varphi = \varphi(x)$ *of class* $C^k[x; \nu]$ *such that for some* $\epsilon > 0$ *the manifold* $M = \{(x, \varphi(x)) : |x| \leqslant \epsilon\}$ *has the property that every solution of (10) which is on* M *at* $t = 0$ *remains on* M *for all* $t > 0$.

Proof. Put $C = e^{-A}$, $D = e^B$; then in view of (11), (12) we may assume, by making a linear change of variables, if necessary, that

$$|C| > 1, \qquad |C^T| > 1, \qquad |D| < 1, \qquad |D|\,|C|^{(1+\nu)}\,|C^T|^{(k-1)} < 1, \tag{13}$$

and

$$\frac{d}{dt}|e^{tA}x| < 0, \qquad (\text{for all } x \neq 0, -\infty < t < \infty). \tag{14}$$

Thus, by (14), for every $\epsilon > 0$ there exists a unique non-negative analytic function $t = t(x)$ defined by

$$|e^{-t(x)A}x| = \epsilon, \qquad (0 < |x| \leqslant \epsilon), \tag{15}$$

i.e., $t(x)$ is just the time it takes for the positive half-trajectory of $\dot{x} = -Ax$ starting at x to reach $\{x : |x| = \epsilon\}$.

By (11) we may choose an $\epsilon_0 > 0$ such that for $|x, y| \leqslant \epsilon_0$, $t \geqslant 0$ there exists a unique solution $(e^{tA}x, \eta(t, x, y))$ of (10) satisfying $\eta(0, x, y) \equiv y$ and $\eta(t, x, y) \to 0$ as $t \to \infty$. Let ϵ_0 be chosen so small that all the kth order derivatives of V defined by

$$V(x, y) = \int_0^1 e^{(1-\tau)B}\, Y(e^{\tau A}x, \eta(\tau, x, y))\, d^\tau \qquad (|x, y| \leqslant \epsilon_0) \tag{16}$$

exist and are uniformly ν-Holder continuous with respect to (x, y); clearly, $V(x, y) = o(|x, y|^k)$ as $(x, y) \to 0$.

Further, there exists an $\epsilon > 0$ such that the vector function $\varphi = \varphi(x)$ defined by

$$\varphi(x) = \eta(t(x), e^{-t(x)A}x, 0) \qquad (0 < |x| \leqslant \epsilon),$$

where $t(x)$ is given by (15), is of class C^k and satisfies the inequality

$$|x, \varphi(x)| \leqslant \epsilon_0 \qquad (0 < |x| \leqslant \epsilon).$$

Clearly, $\varphi(x) \to 0$ as $x \to 0$ and the manifold $\{(x, \varphi(x)) : 0 < |x| \leqslant \epsilon\}$ consists of all positive half-trajectories of (10) which start on $\{(x, 0) : |x| = \epsilon\}$.

We shall now show that the derivatives of φ through order k tend to zero as $x \to 0$. Let

$$w^0 = \{x : x = e^{tA}\bar{x}, 0 \leqslant t \leqslant 1, |\bar{x}| = \epsilon\},$$

$$w^r = \{x : x = e^{A}\bar{x}, \bar{x} \in w^{r-1}\}, \qquad \text{for} \qquad r \geqslant 1,$$

thus if $x \in w^r$ for some $r \geqslant 1$, it follows that

$$\begin{aligned} \varphi(x) &= \eta(1, e^{-A}x, \varphi(e^{-A}x)) \\ &= \eta(1, Cx, \varphi(Cx)) \\ &= D\varphi(x) + \int_0^1 e^{(1-\tau)B}\, Y(e^{\tau A}Cx, \eta(\tau, Cx, \varphi(Cx)))\, d\tau, \end{aligned}$$

and by (16),

$$\varphi(x) = D\varphi(Cx) + V(Cx, \varphi(Cx)), \tag{17}$$

where $Cx \in w^{r-1}$.

Taking the Jacobian of both sides of (17) we get

$$\varphi_x(x) = (D + V_y{}^0)\, \varphi_x{}^0 C + V_x{}^0 C, \tag{18}$$

where the superscript ^{0} on φ_x and V_x, V_y means that the arguments are Cx and $(Cx, \varphi(Cx))$ respectively; therefore

$$|\varphi_x(x)| \leqslant |D + V_y{}^0|\, |\varphi_x{}^0|\, |C| + |V_x{}^0 C|.$$

By (13) there exists a $\delta > 0$ such that

$$\alpha = \max_{|x| \leqslant \delta} | D + V_y(x, \varphi(x)) | \, | C |^{(1+\nu)} \, | C^T |^{(k-1)} < 1.$$

Let

$$p_r = \max_{x \in w^r} | \varphi_x(x) | , \qquad q_r = \max_{x \in w^r} | V_x(x, \varphi(x)) \, C | .$$

Since

$$p_{r+1} \leqslant \alpha p_r + q_r$$

for r sufficiently large and $q_r \to 0$ as $r \to \infty$, it follows from (9) that $p_r \to 0$ as $r \to \infty$ and thus $\varphi_x(x) \to 0$ as $x \to 0$.

If $k > 1$, $x \in w^r$, $r \geqslant 1$, then by (7), (18) we get

$$\begin{aligned} \varphi_{xx}(x) &= \beta_m(D + V_y{}^0) \, (\varphi_x{}^0 C)^x + \alpha_n((\varphi_x{}^0 C)^T) \, (V_y{}^0)_x + (V_x{}^0 C)_x \\ &= \beta_m(D + V_y{}^0) \, \beta_n(C^T) \, \varphi_{xx}^0 C + T_2(Cx, \varphi^0, \varphi_x{}^0), \end{aligned} \tag{19}$$

where

$$T_2(Cx, \varphi^0, \varphi_x{}^0) = \alpha_n((\varphi_x{}^0 C)^T) \, (V_y{}^0)_x + (V_x{}^0 C)_x \, .$$

Let $\partial^1 \varphi = \varphi_x$ and let $\partial^\ell \varphi = (\partial^{\ell-1}\varphi)_x$ for $\ell > 1$. Also, if P is a p-by-p matrix, define $\lambda(P)$ by

$$\lambda(P) = \beta_m(P) \, \beta_p(C^T).$$

Then (19) becomes

$$\partial^2 \varphi(x) = \lambda(D + V_y{}^0) \, \partial^2 \varphi^0 C + T_2(Cx, \varphi^0, \partial^1 \varphi^0).$$

In fact, for $\ell = 2, \cdots, k$, $x \in w^r$, $r \geqslant 1$, we have

$$\partial^\ell \varphi(x) = \lambda^{\ell-1}(D + V_y{}^0) \, \partial^\ell \varphi^0 C + T_\ell(Cx, \varphi^0, \partial^1 \varphi^0, \cdots, \partial^{\ell-1} \varphi^0),$$

where $T_\ell \to 0$ as its argument tends to zero.

Let

$$p_r{}^\ell = \max_{x \in w^r} | \partial^\ell \varphi(x) | ,$$

$$q_r{}^\ell = \max_{x \in w^r} | T_\ell(x, \varphi(x), \partial^1 \varphi(x), \cdots, \partial^{\ell-1} \varphi(x)) | .$$

Then for sufficiently large r

$$p_r^\ell \leqslant \alpha p_{r-1}^\ell + q_{r-1}^\ell \, .$$

Since $\varphi(x)$, $\partial^1 \varphi(x) \to 0$ as $x \to 0$, it follows that $q_r{}^2 \to 0$ as $r \to \infty$; therefore $p_r{}^2 \to 0$ as $r \to \infty$ and $\partial^2 \varphi(x) \to 0$ as $x \to 0$. Suppose $\partial^1 \varphi(x), \cdots, \partial^{\ell-1} \varphi(x) \to 0$ as $x \to 0$; then $q_r{}^\ell \to 0$ as $r \to \infty$, and thus $p_r{}^\ell \to 0$ as $r \to \infty$ and $\partial^\ell \varphi(x) \to 0$ as $x \to 0$. The fact that the kth order derivatives of φ are uniformly ν-Holder continuous with respect to x on a neighborhood of $x = 0$ follows by a similar argument.

THEOREM 1. *Under the assumptions of Lemma 1 there exists a change of variables*

$$x = x, \qquad u = y - \varphi(x) \tag{20}$$

defined in a neighborhood of $x = 0$ *and transforming (10) into*

$$\dot{x} = Ax$$
$$\dot{u} = Bu + U(x, u)$$

where

$$\varphi \in C^k[x, \nu], \qquad U \in C^k[x, y; \nu, \mu] \qquad \textit{and} \qquad U(x, 0) \equiv 0.$$

Proof. Let φ and M be as in Lemma 1, and consider the change of variables (20). Differentiating the equation $u = y - \varphi(x)$ with respect to t, we get

$$\begin{aligned} \dot{u} &= \dot{y} - \varphi_x(x)\,\dot{x} \\ &= By + Y(x, y) + B\varphi(x) - B\varphi(x) - \varphi_x(x)\,Ax \\ &= Bu + Y(x, u + \varphi(x)) + B\varphi(x) - \varphi_x(x)\,Ax \\ &= Bu + U(x, u). \end{aligned}$$

Since the change of variables (20) maps M into the plane $u = 0$, it follows that $U(x, 0) \equiv 0$ for small $|\,x\,|$ and

$$U(x, u) = Y(x, u + \varphi(x)) - Y(x, \varphi(x))$$

for small $|\,x, u\,|$.

4. A Partial Linearization and Main Result

Consider the real vector differential equations

$$\begin{aligned} \dot{x} &= Ax \\ \dot{y} &= By + Y(x, y, z) \\ \dot{z} &= Cz + Z(x, y, z), \end{aligned} \tag{21}$$

where x, y, z are finite dimensional, A, B, C are constant square matrices of appropriate order, $Y, Z \in C^k[x, (y, z); \nu, \mu]$, where $k \geqslant 1$ and $0 \leqslant \nu \leqslant \mu \leqslant 1$, and $Y(x, 0, 0)$, $Z(x, 0, 0)$ vanish identically. Suppose the eigenvalues of A, B,

C satisfy

$$\max_{\lambda = \text{e.v. of } C} \operatorname{Re} \lambda \leqslant \max_{\lambda = \text{e.v. of } B} \operatorname{Re} \lambda \leqslant \max_{\lambda = \text{e.v. of } A} \operatorname{Re} \lambda < 0,$$

and put

$$a = \left| \max_{\lambda = \text{e.v. of } A} \operatorname{Re} \lambda \right|, \qquad b = \left| \max_{\lambda = \text{e.v. of } B} \operatorname{Re} \lambda \right|.$$

For small $| x, y, z |$ and $t \geqslant 0$ there exists a unique solution

$$\zeta(t, x, y, z) = (e^{tA}, \eta(t, x, y, z))$$

of (21) such that $\zeta(0, x, y, z) = (x, y, z)$. By a straightforward extension of the arguments in ([*8*], p. 315), estimates of the derivatives of η with respect to the x, y, z coordinates can be obtained which yield the following result:

LEMMA 2. *Let the above assumptions hold; let*

$$F(t, x, y, z) = (Y(\zeta(t, x, y, z)), Z(\zeta(t, x, y, z)));$$

let $F^{(k)}$ be any kth *order derivative of F with respect to the x, y, z coordinates and suppose that*

$$L = \min \{(k - 1 + \nu)\, a, ka - (1 - \mu)\, b\} > 0.$$

Then for every $\epsilon > 0$, there exist positive numbers γ, δ, α such that

$$| F^{(k)}(t, x, y, z) | \leqslant \gamma e^{(-b-L+\epsilon)t}$$

and

$$| F^{(k)}(t, x, y, z) - F^{(k)}(t, \bar{x}, \bar{y}, \bar{z}) |$$
$$\leqslant \gamma(| x - \bar{x} |^{\alpha\nu} + | (y, z) - (\bar{y}, \bar{z}) |^{\mu})\, e^{(-b-L+\epsilon)t}$$

for $| x, y, z |$, $| \bar{x}, \bar{y}, \bar{z} | \leqslant \delta$ and $t \geqslant 0$.

The proof of Lemma 2 is essentially contained in ([*9*], Lemmas 1.1-1.5); there it is shown that the derivatives of $\eta(t, x, y, z)$ with respect to the x, y, z coordinates through order k are $0(e^{(-b+\epsilon)t})$ as $t \to \infty$ uniformly on some neighborhood of $(x, y, z) = 0$.

THEOREM 2. *Let the hypotheses of Lemma 1 hold and suppose that*

$$\max_{\lambda, \bar{\lambda} = \text{e.v. of } B} | \operatorname{Re} \lambda - \operatorname{Re} \bar{\lambda} | < L. \tag{22}$$

Then there exists a change of variables

$$x = x, \qquad u = y - \varphi(x, y, z), \qquad z = z$$

defined in a neighborhood of $(x, y, z) = 0$ *and transforming (21) into*

$$\begin{aligned}\dot{x} &= Ax\\ \dot{u} &= Bu\\ \dot{z} &= Cz + \tilde{Z}(x, u, z),\end{aligned}$$

where $\varphi \in C^k[x, (y, z); \alpha\nu, \mu]$ *and* $\tilde{Z} \in C^k[x, (u, z); \alpha\nu, \mu]$ *for some* $\alpha > 0$.

Proof. Put $w = (x, y, z)$ and rewrite (21) as $\dot{w} = W(w)$, where

$$W(w) = (Ax, By + Y(x, y, z), Cz + Z(x, y, z)).$$

Differentiating both sides of $u = y - \varphi(w)$ with respect to t we get

$$\begin{aligned}\dot{u} &= \dot{y} - \varphi_w(w)\,\dot{w}\\ &= By + Y(w) - \varphi_w(w)\,W(w)\\ &= By - B\varphi(w) + B\varphi(w) + Y(w) - \varphi_w(w)\,W(w)\\ &= Bu + B\varphi(w) + Y(w) - \varphi_w(w)\,W(w).\end{aligned}$$

Thus a necessary and sufficient condition that $\dot{u} = Bu$ is that

$$\varphi_w(w)\,W(w) = B\varphi(w) + Y(w). \tag{23}$$

By Lemma 2 and (22) there exists an open neighborhood N of $w = 0$ such that $\zeta(t, w) = \zeta(t, x, y, z)$ is defined for all $w \in N$, $t \geqslant 0$ and the vector function φ defined by

$$\varphi(w) = -\int_0^\infty e^{-\tau B}\, Y(\zeta(\tau, w))\, d\tau, \qquad (w \in N), \tag{24}$$

is continuously differentiable and of class $C^k[x, (y, z); \alpha\nu, \mu]$ for some $\alpha > 0$. We shall now show that φ satisfies (23).

Let $Q = Q(t, w)$ be defined by

$$Q(t, w) = e^{tB}\left[\varphi(w) + \int_0^t e^{-\tau B}\, Y(\zeta(\tau, w))\, d\tau\right], \qquad (w \in N, t \geqslant 0).$$

Note that Q satisfies

$$\dot{Q}(t, w) = BQ(t, w) + Y(\zeta(t, w)), \qquad Q(0, w) = \varphi(w)$$

and

$$Q(t, w) = -\int_0^\infty e^{-\tau B}\, Y(\zeta(t + \tau, w))\, d\tau; \tag{25}$$

hence for $w \in N$, $t \geqslant 0$, $\zeta(t, w) \in N$, it follows from (24) and (25) that $Q(t, w) = \varphi(\zeta(t, w))$.

Thus for $w \in N$, $t \geqslant 0$, $\zeta(t, w) \in N$,

$$\begin{aligned}\frac{d}{dt}\varphi(\zeta(t, w)) &= \dot{Q}(t, w)\\ &= BQ(t, w) + Y(\zeta(t, w))\\ &= B\varphi(\zeta(t, w)) + Y(\zeta(t, w))\end{aligned}$$

and

$$\begin{aligned}\frac{d}{dt}\varphi(\zeta(t, w)) &= \varphi_w(\zeta(t, w))\,\dot{\zeta}(t, w)\\ &= \varphi_w(\zeta(t, w))\,W(\zeta(t, w));\end{aligned}$$

therefore setting $t = 0$ we obtain (23).

Let

$$x = x, \qquad y = u + \psi(x, u, z), \qquad z = z$$

be the inverse change of variables; then

$$\tilde{Z}(x, u, z) = Z(x, u + \psi(x, u, z), z).$$

Remark 1. If Y, Z are independent of x, then Theorem 2 holds when L is defined by $L = (k - 1 + \mu)\, b$.

Let $a_1, \cdots, a_n$ be real numbers satisfying

$$a_n \leqslant a_{n-1} \leqslant \cdots \leqslant a_1 < 0,$$

and let $S = \{a_1, \cdots, a_n\}$. We shall say that S is (k, μ)-*spaced* if S is the disjoint union of subsets $S_1, \cdots, S_\ell$ with the following properties:

(P_1) if $a \in S_j$, $\bar{a} \in S_{j+1}$, $j \neq \ell$, then $\bar{a} < ka$;

(P_2) if $a, \bar{a} \in S_j$, $j \neq 1$, then $|a - \bar{a}| \leqslant (k - 1)\,|a_1|$;

(P_3) if $a, \bar{a} \in S_j$, $j = 1, \cdots, \ell$, then

$$|a - \bar{a}| < k\,|a_1| - (1 - \mu)\,|\max S_j|.$$

Remark 2. By (P_3), a necessary condition for S to be (k, μ)-spaced is that

$$ka_1 < (1 - \mu)\, a_n.$$

This condition is sufficient for $k = 1$; to see this, take the S_j to be the distinct singletons $\{a_i\}$.

THEOREM 3. *In (1), let $X \in C^k[x, \mu]$ for some k, μ satisfying $k \geqslant 1$, $0 \leqslant \mu \leqslant 1$. If Λ satisfies (5) and the set $S = \{\operatorname{Re} \lambda : \lambda = \text{eigenvalue of } \Lambda\}$ is (k, μ)-spaced, then (1) is C^k equivalent to (3).*

Proof. Let S be the disjoint union of $S_1, \cdots, S_\ell$ satisfying the three properties above; then, by making a linear change of variables, if necessary, (1) may be written as

$$\frac{dx^1}{dt} = A^1 x^1 + X^1(x^1, \cdots, x^\ell)$$
$$\vdots$$
$$\frac{dx^\ell}{dt} = A^\ell x^\ell + X^\ell(x^1, \cdots, x^\ell)$$

where the real parts of the eigenvalues of A^j belong to S_j. By Remark 1, the x^1-equations may be linearized. If $\ell \geqslant 2$, then by Remark 2 and (P_1) it follows that $\mu > 0$; hence we can make a change of variables of the type indicated in Theorem 1 and then linearize the x^2-equations by again applying Theorem 2. Continuing in this way, first applying Theorem 1 and then Theorem 2, we can linearize all the x^j-equations.

Remark 3. If $k \geqslant 2$ and the eigenvalues of A satisfy (5), then the condition that the set $\{\mathrm{Re}\,\lambda : \lambda = \text{e.v. of } A\}$ be $(k, 0)$-spaced is equivalent to condition (6); thus, for $k \geqslant 2$ and $\mu = 0$, Theorem 3 reduces to Theorem A. By Remark 2, it follows that Theorem B is also a special case of Theorem 1. Since every set $\{a_1, a_2\}$, where $a_1, a_2 < 0$, is $(k, 1)$-spaced for all k, it follows that every two dimensional system (1), where (5) holds and $X \in C^k[x, 1]$, is C^k equivalent to (3).

5. C^k Systems Which Are Not C^k Equivalent to Their Linear Parts

Let $\epsilon = \epsilon(t)$ be a real valued continuous function on $(-\infty, \infty)$ satisfying

$$\epsilon(t) \to 0 \qquad \text{as} \qquad t \to 0,$$
$$\int_1^x \frac{\epsilon(\tau)}{\tau}\, d\tau \to \infty \qquad \text{as} \qquad x \to 0^+, \tag{26}$$

and define $Y = Y(x)$ by

$$Y(x) = \int_0^x \epsilon(\tau)\, d\tau.$$

Then $Y \in C^1$ on $(-\infty, \infty)$ and $Y(x) = o(|x|)$ as $x \to 0$.

Consider the two dimensional system

$$\begin{aligned} \dot{x} &= -x \\ \dot{y} &= -y - Y(x). \end{aligned} \tag{27}$$

We shall show that there is no C^1 change of variables

$$\begin{aligned} u &= x - \varphi_1(x, y) \qquad \varphi_1(x, y), \varphi_2(x, y) = o(|x, y|) \quad \text{as} \quad (x, y) \to 0 \\ v &= y - \varphi_2(x, y) \end{aligned} \tag{28}$$

which is defined for small $| x, y |$ and transforms (27) into

$$\begin{aligned} \dot{u} &= -u \\ \dot{v} &= -v. \end{aligned} \tag{29}$$

Suppose such a change of variables exists. Let $T : (x, y) \to (u, v)$ denote the transformation (28) and let T^{-1} denote the inverse transformation. Choose $\delta > 0$ so that T^{-1} is of class C^1 on $B = \{(u, v) : | u, v | \leqslant \delta\}$ and let M be the intersection of the line $v = 0$ with B. Since M consists of positive half-trajectories of (29), it follows that $T^{-1}(M)$ consists of positive half-trajectories of (27). Since T^{-1} is of the form

$$\begin{aligned} x &= u + \psi_1(u, v) \qquad \psi_1(u, v), \psi_2(u, v) = o(| u, v |) \quad \text{as} \quad (u, v) \to 0 \\ y &= v + \psi_2(t, v) \end{aligned}$$

it follows that

$$T^{-1}(M) = \{(u + \psi_1(u, 0), \psi_2(u, 0)) : | u | \leqslant \delta\}$$

and hence there exists a $\gamma > 0$ and a real valued function $Q = Q(x)$ such that $Q \in C^1$ for $| x | \leqslant \gamma$ and $\{(x, Q(x)) : | x | \leqslant \gamma\} \subset T^{-1}(M)$. By (27), if $0 < | x | \leqslant \gamma$, then

$$\frac{dQ(x)}{dx} = \frac{Q(x)}{x} + \frac{Y(x)}{x}.$$

Thus if $0 < x \leqslant \gamma$, then

$$Q(x) = \frac{Q(\gamma)}{\gamma} x + x \int_\gamma^x \frac{Y(\tau)}{\tau^2} d\tau$$

and

$$\frac{dQ(x)}{dx} = \frac{Q(\gamma)}{\gamma} + \int_\gamma^x \frac{Y(\tau)}{\tau^2} d\tau + \frac{Y(x)}{x}.$$

But

$$\int_\gamma^x \frac{Y(\tau)}{\tau^2} d\tau = -\frac{1}{\tau} Y(\tau) \Big|_\gamma^x + \int_\gamma^x \frac{\epsilon(\tau)}{\tau} d\tau \to \infty \qquad x \to 0^+$$

by (26), and therefore $Q \notin C^1$ in any neighborhood of $x = 0$.

Let $Y^1(x) = Y(x)$ and let

$$Y^k(x) = \int_0^x Y^{k-1}(\tau) \, d\tau \qquad \text{for} \qquad k \geqslant 2.$$

Then by a similar argument the C^k system

$$\dot{x} = -x$$
$$\dot{y} = -ky + Y^k(x)$$

is not C^k equivalent to its linear part. Note that $Y^k \in C^k[x; \mu]$ only for $\mu = 0$, but the numbers -1, $-k$ are (k, μ)-spaced only for $\mu > 0$; therefore Theorem 3 does not apply.

References

1. Hartman, P. "Ordinary Differential Equations." Wiley, New York (1964).
2. Sternberg, S. Local contractions and a theorem of Poincaré. *Amer. J. Math.* **79** (1957), 809-824.
3. Hartman, P. On local homeomorphisms of Euclidean spaces. *Bol. Soc. Mat. Mexicana* **5** (1960), 220-241.
4. Nagumo, M., and Isé, K. On the normal forms of differential equations in the neighborhood of an equilibrium point. *Osaka Math. J.* **9** (1957), 221-234.
5. Sternberg, S. On the structure of local homeomorphisms of Euclidean *n*-space. II, *Amer. J. Math.* **80** (1958), 623-631.
6. Chen, K. T. Equivalence and decomposition of vector fields about an elementary critical point. *Amer. J. Math.* *85* (1963), 693-722.
7. Brjuno, A. D. Normal form of differential equations. *Dokl. Akad. Nauk SSSR* **157** (1964), 1276-1279.
8. Coddington, E. A., and Levinson, N. "Theory of Ordinary Differential Equations." McGraw-Hill, New York (1955).
9. Venti, R. On the problem of normal forms for a system of ordinary differential equations at a singular point. O.N.R. Report prepared under Contract Nonr 222(88), University of California, Berkeley, August 1964.