MATHEMATICAL CONTROL THEORY
BANACH CENTER PUBLICATIONS, VOLUME 14
PWN – POLISH SCIENTIFIC PUBLISHERS
WARSAW 1985

GEOMETRIC METHODS IN LINEARIZATION OF CONTROL SYSTEMS

WITOLD RESPONDEK

Institute of Mathematics, Technical University of Warsaw, Warsaw, Poland

1. Introduction

Consider a nonlinear control system of the form

$$\dot{x} = f(x) + \sum_{i=1}^{k} u_i g_i(x), \tag{1.1}$$

where $x \in M$ (an n-dimensional manifold), f and $g_1, \ldots, g_k$ are smooth or analytic vector fields on M and u_i's are controls which belong to a certain class of admissible controls. We will study the problem of local linearization of system (1.1) around a given point $x_0 \in M$, using different groups of transformations including changes of coordinates in the state and input spaces and feedback. We prove some linearization results in the case of changing the coordinates in the state space only. We also formulate and compare recent results on linearization using other classes of transformations [2], [4], [5], [8], [9].

We define three groups H, F, and G of transformations of system (1.1).

The group H, which was studied by Krener [6], is given by

(i) changes of coordinates in the state space M:

$$f, g_i \mapsto \varphi_* f, \varphi_* g_i,$$

where φ: $M \to M$ is a diffeomorphism and $\varphi_* f, \varphi_* g_i$ are vectors fields f and g_i in new coordinates given by φ, respectively.

The group F was defined and studied by Brockett [2] and is generated by

(i) changes of coordinates in the state space M (as in the case of H), (ii)′ linear changes of coordinates in the input space $\boldsymbol{R}^k$: $g_i \mapsto \sum_{j=1}^{k} d_{ji} g_j$ where $D = (d_{ji})$ is a nonsingular matrix,

(iii) feedbacks of the form:

$$f \mapsto f + \sum_{j=1}^{k} a_j g_j,$$

where a_j are smooth functions on M.

The largest group G was defined by Jakubczyk and Respondek [5]. It is generated by:

(i) changes of coordinates in the state space M (like in the cases of H and F),

(ii)″ linear changes of coordinates in the input space $\boldsymbol{R}^k$, nonlinearly depending on x:

$$g_i \mapsto \sum_{j=1}^{k} h_{ji} g_j,$$

where $H(x) = h_{ji}(x)$ is a $k \times k$-matrix of smooth functions, nonsingular at $x_0 \in M$,

(iii) feedbacks of the form:

$$f \mapsto f + \sum_{j=1}^{k} a_j g_j$$

(like in the case of F).

Consider a linear control system of the form

$$\dot{x} = Ax + \sum_{i=1}^{k} u_i b_i, \tag{1.2}$$

where $x \in \boldsymbol{R}^n$ and matrix A and vector fields b_i, $i = 1, \ldots, k$, are constant. In this paper we give necessary and sufficient conditions for local linearization of system (1.1), using transformations from H, F or G, to a controllable linear form (1.2). We will make a natural assumption that $f(x_0) = 0$, and so we assume also that φ in (i) and a_j's in (iii) satisfy $\varphi(x_0) = 0 \in \boldsymbol{R}^n$ and $a_j(x_0) = 0$, respectively.

Krener gave in [6] necessary and sufficient conditions for two analytic systems to be locally H-equivalent. A similar problem, but in a different framework, was also studied by Nagano [7]. Krener studied also the linearization problem but his condition is not correct. We formulate and prove necessary and sufficient conditions for H-linearization in both smooth and analytic cases.

Brockett [2] gave a necessary and sufficient condition for F-linearization in the scalar case, i.e., for $k = 1$. We formulate an analogous result for any natural k.

G-equivalence was introduced by Jakubczyk and Respondek [6], where necessary and sufficient conditions for G-linearization were given. We compare the linearization theorems in all three cases and give examples.

Let us mention that in the linear case the problem of classification by using feedback was solved by Brunovsky [3], who found the canonical forms. If we consider a linear system, then the groups F and G do not differ and they are equal to the group studied by Brunovsky.

Consider the set $V(M)$ of all C^∞-vector fields on M as a Lie algebra with the Lie product $[a, b]$, $a, b \in V(M)$. We will also use the notation $\mathrm{ad}_a b = [a, b]$ and inductively $\mathrm{ad}_a^i b = [a, \mathrm{ad}_a^{i-1} b]$. For any $A \subset V(M)$ we will write $A(x) = \{h(x) \mid h \in A\}$. Denote $L^j = \{\mathrm{ad}_f^q g_i \mid 1 \leqslant i \leqslant k, 0 \leqslant q \leqslant j\}$.

2. H-linearization

In this section we give necessary and sufficient conditions for H-linearization in the C^∞-case. Then we analyse the analytic case, and so we assume the diffeomorphism φ in (i) to be analytic. We formulate the Krener–Nagano condition for the H-equivalence and study the problem of linearization.

We will denote by $\mathscr{L}$ the Lie algebra generated by $L = \{f, g_1, \ldots, g_k\}$, i.e., the smallest Lie algebra which contains L.

DEFINITION 2.1. We say that system (1.1) *satisfies the Lie rank condition* if $\dim \mathscr{L}(x) = n$. This condition is sometimes called *accessibility* or *weak controllability*.

It is obvious that if we want system (1.1) to be locally H-equivalent to a controllable linear system (1.2), then the Lie rank condition and $f(x_0) = 0$ have to be satisfied.

THEOREM 2.1. *Let system* (1.1) *satisfy the Lie rank condition and let* $f(x_0) = 0$. *System* (1.1) *is locally H-equivalent to a controllable linear system* (1.2) *at* $x_0 \in M$ *and* $0 \in \mathbf{R}^n$, *respectively, if and only if it satisfies the conditions*

$$[g_s, \mathrm{ad}_f^j g_t] = 0 \tag{2.1}$$

in a neighbourhood of x_0 *for any* $1 \leqslant s, t \leqslant k$ *and any* $0 \leqslant j \leqslant 2n-1$.

Remark. From the proof (see Section 5) it follows that under (2.1) the Lie rank condition is equivalent to the weaker one:

$$\dim \operatorname{span}\{\mathrm{ad}_f^q g_i(x) \mid 0 \leqslant q \leqslant n-1, 1 \leqslant i \leqslant k\} = n.$$

Thus the above theorem can be expressed in the following form, which is similar to the theorems in Section 3: a system for which $f(x_0) = 0$ is, locally around x_0, H-equivalent to a controllable linear system (1.2) if and only if it satisfies:

(H1) $[\mathrm{ad}_f^p g_s, \mathrm{ad}_f^q g_t] = 0$ for any $1 \leqslant s$, $t \leqslant k$ and any $0 \leqslant p \leqslant n-1$, $0 \leqslant q \leqslant n$,

(H2) $\dim \mathrm{span}\{\mathrm{ad}_f^q g_i(x) \mid 0 \leqslant q \leqslant n-1, 1 \leqslant i \leqslant k\} = n$,
in a neighbourhood of x_0.

Now we study the analytic case, i.e., we assume M to be an n-dimensional analytic manifold and $f, g_1, \ldots, g_k$ to be analytic vector fields on M. In this case, studying the values of certain Lie brackets at the point x_0 only, we will be able to answer the question if (1.1) and (1.2) are equivalent.

At first we formulate the Krener–Nagano result on H-equivalence in the analytic case. Assume we are given another analytic control system

$$\dot{y} = \hat{f}(y) + \sum_{i=1}^{k} u_i \hat{g}_i(y), \qquad y \in N, \tag{2.2}$$

where N is an analytic manifold. Denote $f = g_0$ and $\hat{f} = \hat{g}_0$.

THEOREM 2.2 (Krener [6], Nagano [7]). *Systems* (1.1) *and* (2.2) *are locally H-equivalent at $x_0 \in M$ and $y_0 \in N$, respectively, if and only if there exists a linear isomorphism L: $T_{x_0}M \to T_{y_0}N$ such that*

$$Lg_i(x_0) = \hat{g}_i(y_0) \quad \textit{for any } i = 0, \ldots, k$$

and

$$L\big[g_{i_1}, [\ldots [g_{i_{p-1}}, g_{i_p}] \ldots]\big](x_0) = \big[\hat{g}_{i_1}, [\ldots [\hat{g}_{i_{p-1}}, \hat{g}_{i_p}] \ldots]\big](y_0), \tag{2.3}$$

$$\textit{for any } p \geqslant 2 \textit{ and } 0 \leqslant i_j \leqslant k.$$

For the linear system (1.2) we have $\hat{f} = Ay$ and $\hat{g}_i = b_i$, $i = 1, \ldots, k$, and so $\mathrm{ad}_{\hat{f}}^q \hat{g}_i = A^q b_i$ and the Lie bracket of two constant vector fields is equal to zero. Thus the H-invariance of Lie brackets and the above theorem give

COROLLARY. *System* (1.1), *for which* $f(x_0) = 0$, *is, locally around* $x_0 \in M$, *H-equivalent to a controllable linear system* (1.2) *if and only if it satisfies*

$$\dim \mathrm{span}\{\mathrm{ad}_f^q g_s(x_0) \mid q \geqslant 0, 1 \leqslant s \leqslant k\} = n$$

and

$$\big[\mathrm{ad}_f^{q_1} g_{i_1}, [\ldots [\mathrm{ad}_f^{q_{p-1}} g_{i_{p-1}}, \mathrm{ad}_f^{q_p} g_{i_p}] \ldots]\big](x_0) = 0 \quad \textit{for any } p \geqslant 2,$$

$$q_j \geqslant 0, 1 \leqslant i_j \leqslant k.$$

In other words, the second of the above conditions means that all iterative brackets

$$\left[g_{i_1}, \left[\ldots [g_{i_{p-1}}, g_{i_p}] \ldots\right]\right], \qquad i_j \geqslant 0,$$

which include at least two g_i's, $1 \leqslant i \leqslant k$, have to vanish at x_0 (see (2.3)).

The above condition can be replaced by a weaker one in the following way:

THEOREM 2.3. *System* (1.1), *for which* $f(x_0) = 0$, *is, locally around* $x_0 \in M$, *H-equivalent to a controllable linear system* (1.2) *if and only if it satisfies*

$$\dim\operatorname{span}\{\operatorname{ad}_f^q g_s(x_0) \mid q \geqslant 0, 1 \leqslant s \leqslant k\} = n$$

and

$$\left[\operatorname{ad}_f^{q_1} g_{i_1}, \left[\ldots [\operatorname{ad}_f^{q_{p-1}} g_{i_{p-1}}, \operatorname{ad}_f^{q_p} g_{i_p}] \ldots\right]\right](x_0) = 0 \tag{2.4}$$

for any $p \geqslant 2, 1 \leqslant i_j \leqslant k, 0 \leqslant q_j \leqslant n-1$ *when* $1 \leqslant j \leqslant p-1$ *and* $0 \leqslant q_p \leqslant n$.

3. F- and G-linearizations

In this section we will study the problem of the linearization of system (1.1) by means of the transformations in the groups F and G. We consider the smooth case, i.e., the state space for system (1.1) is an n-dimensional C^∞-manifold, f and g_i, $i = 1, \ldots, k$ are C^∞-vector fields on M.

Consider the H-linearization condition $[g_s, \operatorname{ad}_f^j g_t] = 0$. Assume for a moment the scalar input case and denote $g_1 = g$. For $j = 1$ we have $[g, \operatorname{ad}_f g] = 0$. Use feedback and replace f by $\tilde{f} = f + ag$. For any vector fields X, Y and smooth functions a, β we have

$$[aX, \beta Y] = a\beta[X, Y] + a\big((\nabla\beta)X\big)Y - \beta\big((\nabla a)Y\big)X;$$

thus

$$\operatorname{ad}_{\tilde{f}} g = [f + ag, g] = \operatorname{ad}_f g - \big((\nabla a)g\big)g$$

and

$$[g, \operatorname{ad}_{\tilde{f}} g] = \big[g, \operatorname{ad}_f g - \big((\nabla a)g\big)g\big] = [g, \operatorname{ad}_f g] + \beta g$$

for a suitable smooth function β. Therefore $[g, \operatorname{ad}_f g] = 0$ implies that $[g, \operatorname{ad}_{\tilde{f}} g] = \beta g$, and we see that even in the scalar control case (H1) is neither F- nor G-invariant. We have to find more general conditions which are invariants for F or G. This was done in [2], [5] and [8]. The problem of G-linearization was solved in [5] in the following way:

THEOREM 3.1. *System* (1.1) *is, locally around* $x_0 \in M$, *G-linearizable to a controllable linear system* (1.2) *if and only if it satisfies in a neighbourhood of* x_0 *the following conditions*:

(G1) *For any* $0 \leqslant p \leqslant j \leqslant n-1$ *and* $1 \leqslant s, t \leqslant k$ *there exist functions* $a_{iq} \in C^\infty(M)$ *such that*

$$[\mathrm{ad}_f^p g_s, \mathrm{ad}_f^j g_t] = \sum_{\substack{0 \leqslant q \leqslant j \\ 1 \leqslant i \leqslant k}} a_{iq} \mathrm{ad}_f^q g_i,$$

i.e., L^j *are involutive for* $j = 0, \ldots, n-1$.

(G2) $\dim\operatorname{span}\{\mathrm{ad}_f^q g_i(x) \mid 0 \leqslant q \leqslant j, 1 \leqslant i \leqslant k\} = r_j(x) = \text{const}$.

(G3) $\dim\operatorname{span}\{\mathrm{ad}_f^q g_i(x) \mid 0 \leqslant q \leqslant n-1, 1 \leqslant i \leqslant k\} = r_{n-1}(x) = n$.

We refer the reader to [5] for the proof of the above theorem. In very recent papers Hunt, Su and Meyer [4], [9] have given weaker conditions for G-linearization in the following way. As we have noticed, any controllable linear system (1.2) is G-equivalent to its Brunovsky canonical form based on its Kronecker indices (see [3], [10]). This implies that we can as well ask the question when (1.1) is G-equivalent to (1.2) given in the Brunovsky canonical form. Suppose we are given the integers (the Kronecker indices) $k_1, k_2, \ldots, k_m$ satisfying $\sum_{i=1}^{m} k_i = n$ and $k_1 \geqslant k_2 \geqslant \ldots \geqslant k_m > 0$. Denote $\tilde{L} = L_1^{k_1-1} \cup L_2^{k_2-1} \cup \ldots \cup L_m^{k_m-1}$, where $L_i^j = \{\mathrm{ad}_f^q g_i \mid 0 \leqslant q \leqslant j\}$. Hunt, Su and Meyer [4] proved (for notational convenience we write $k = m$)

THEOREM 3.2. *System* (1.1) *is, locally around* $x_0 \in M$, *G-linearizable to a controllable linear system* (1.2) *if and only if it satisfies in a neighbourhood of* x_0 *the following conditions*:

(T1) *For any* $j = k_m-2, k_{m-1}-2, \ldots, k_2-2, k_1-2$, *any* $0 \leqslant p \leqslant j$ *and any* $1 \leqslant s, t \leqslant k$ *there exist functions* $a_{iq} \in C^\infty(M)$ *such that*

$$[\mathrm{ad}_f^p g_s, \mathrm{ad}_f^j g_t] = \sum_{\substack{0 \leqslant q \leqslant j \\ 1 \leqslant i \leqslant m}} a_{iq} \mathrm{ad}_f^q g_i,$$

i.e., L^j *are involutive for* $j = k_m-2, k_{m-1}-2, \ldots, k_2-2, k_1-2$.

(T2) $\operatorname{span} L^{k_r-2}(x) = \operatorname{span}(L^{k_r-2} \cap \tilde{L})(x)$ *for any* $r = 1, \ldots, m$.

(T3) $\dim\operatorname{span}\tilde{L}(x) = n$.

The reader is referred to [4] for the proof of the above theorem and some comments.

The problem of F-linearization was studied by Brockett in [2]. He gave necessary and sufficient conditions for F-linearization in the scalar control case, i.e., for $k = 1$. In the multi-input case the following theorem holds. Define $\mathrm{ad}_f^{-1} g = 0$.

THEOREM 3.3. *System* (1.1) *is, locally around* $x_0 \in M$, *F-linearizable to a controllable linear system* (1.2) *if and only if it satisfies in a neighbourhood of* x_0 *the following conditions*:

(F1) *For any* $0 \leqslant p \leqslant j \leqslant n-1$ *and* $1 \leqslant s, t \leqslant k$ *there exist functions* $a_{iq} \in C^\infty(M)$ *such that*

$$[\mathrm{ad}_f^p g_s, \mathrm{ad}_f^j g_t] = \sum_{\substack{-1 \leqslant q \leqslant j-1 \\ 1 \leqslant i \leqslant k}} a_{iq} \mathrm{ad}_f^q g_i.$$

(F2) $\dim \operatorname{span}\{\mathrm{ad}_f^q g_i(x) \mid 0 \leqslant q \leqslant j, 1 \leqslant i \leqslant k\} = r_j(x) = \text{const}.$

(F3) $\dim \operatorname{span}\{\mathrm{ad}_f^q g_i(x) \mid 0 \leqslant q \leqslant n-1, 1 \leqslant i \leqslant k\} = r_{n-1}(x) = n.$

(F4) *For any* $1 \leqslant m \leqslant k$ *and* $0 \leqslant j \leqslant n-1$ *such that*

$$\mathrm{ad}_f^j g_m(x_0) = \sum_{\substack{i \neq m \\ 1 \leqslant i \leqslant k}} c_i \mathrm{ad}_f^j g_i(x_0) + \sum_{\substack{0 \leqslant q \leqslant j-1 \\ 1 \leqslant i \leqslant k}} c_{iq} \mathrm{ad}_f^q g_i(x_0),$$

where $c_i, c_{iq} \in \boldsymbol{R}$ *we have*

$$\mathrm{ad}_f^j g_m = \sum_{\substack{i \neq m \\ 1 \leqslant i \leqslant k}} c_i \mathrm{ad}_f^j g_i + \sum_{\substack{0 \leqslant q \leqslant j-1 \\ 1 \leqslant i \leqslant k}} a_{iq} \mathrm{ad}_f^q g_i$$

for some smooth functions a_{iq}.

Remark 1. Conditions (F2) and (F3) in the above theorem and conditions (G2) and (G3) in Theorem 3.1 are the same. If we compare conditions (G1) and (F1), we see that in (F1) the sum on the right-hand side does not contain any term corresponding to j. This is connected with the fact that the functions d_{ji} in (ii)′ are constant while the functions h_{ji} in (ii)″ may depend on the state (see the definitions of F and G in Section 1). Condition (F1) implies for $j = 0$ that $[g_s, g_t] = 0$, $1 \leqslant s, t \leqslant k$ while for $j \geqslant 1$ the presence of the terms $\mathrm{ad}_f^{-1} g_i$ on the right-hand side of the sum in (F1) does not change anything.

Remark 2. Condition (F4) has a meaning only in the multi-input case. It is connected with the fact that functions d_{ji} in (ii)′ are constant.

The proof of the above theorem is given in [8].

4. Examples

In this section we give some examples of H-, F- and G-linearizable systems.

EXAMPLE 4.1. Consider the system on $\boldsymbol{R}^3$

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} 0 \\ x \\ d(y, z) \end{bmatrix} + u \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = f + ug. \tag{4.1}$$

We want to find the conditions on the function $d(y, z)$ under which the above system is locally, around $0 \in \boldsymbol{R}^3$, H-linearizable. Compute

$$\mathrm{ad}_f g = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}, \quad \mathrm{ad}_f^2 g = \begin{bmatrix} 0 \\ 0 \\ \dfrac{\partial d}{\partial y} \end{bmatrix}, \quad [g, \mathrm{ad}_f g] = 0, \quad [\mathrm{ad}_f g, \mathrm{ad}_f^2 g] = \begin{bmatrix} 0 \\ 0 \\ -\dfrac{\partial^2 d}{\partial y^2} \end{bmatrix}.$$

Condition (H1) in the Remark after Theorem 2.1 implies that $\frac{\partial^2 d}{\partial y^2} = 0$; so $d(y, z) = a(z)y + b(z)$, where $a(z)$, $b(z)$ are certain smooth functions. Condition (H2) gives $a(0) \neq 0$. We have

$$\mathrm{ad}_f^3 g = \begin{bmatrix} 0 \\ 0 \\ c(z) \end{bmatrix}, \quad \text{where} \quad c = ab' - ba'$$

and for the functions a, b, c we denote by a', b', c', respectively, the derivatives with respect to z. Now we compute

$$[g, \mathrm{ad}_f^3 g] = 0, \quad [\mathrm{ad}_f g, \mathrm{ad}_f^3 g] = 0, \quad [\mathrm{ad}_f^2 g, \mathrm{ad}_f^3 g]$$

$$= \left[\begin{bmatrix} 0 \\ 0 \\ a \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ c \end{bmatrix} \right] = \begin{bmatrix} 0 \\ 0 \\ ac' - ca' \end{bmatrix}.$$

Condition (H1) gives $ac' - ca' = 0$. Therefore the necessary and sufficient linearization condition for (4.1) is that $d(y, z)$ satisfies $d(y, z) = a(z)y + b(z)$, where $a(0) \neq 0$ and

$$a(ab' - ba')' - a'(ab' - ba') = 0. \tag{4.2}$$

Now we find the coordinate system in which (4.1) takes a linear form. It follows from the proof of Theorem 2.1 that we have to introduce a new coordinate system $(\bar{x}, \bar{y}, \bar{z})$, in which

$$g = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathrm{ad}_f\, g = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathrm{ad}_f^2 g = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Therefore put

$$\bar{x} = x, \quad \bar{y} = y, \quad \bar{z} = \int_0^z \frac{1}{a(t)}\, dt.$$

In the new coordinates we have

$$f = \begin{bmatrix} 0 \\ \bar{x} \\ \bar{y} + \frac{b}{a}\big(z(\bar{z})\big) \end{bmatrix}, \quad g = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

We show that $\frac{b}{a}$ is a linear function of $\bar{z}$, i.e., $\frac{d^2}{d\bar{z}^2}\left(\frac{b}{a}\right) = 0$.

Compute

$$\frac{d}{d\bar{z}}\left(\frac{b}{a}\right) = \frac{d}{dz}\left(\frac{b}{a}\right)\cdot\frac{dz}{d\bar{z}} = \frac{ab'-ba'}{a^2}\cdot a = \frac{c}{a} \quad \text{and} \quad \frac{d^2}{d\bar{z}^2}\left(\frac{b}{a}\right)$$

$$= \frac{d}{dz}\left(\frac{c}{a}\right)\cdot\frac{dz}{d\bar{z}} = \frac{ac'-ca'}{a}.$$

The last term is equal to zero because of (4.2), thus system (4.1) takes a linear form in coordinates $(\bar{x}, \bar{y}, \bar{z})$.

EXAMPLE 4.2. We want to find the conditions under which system (4.1) from Example 4.1 is F-linearizable. From condition (F1) we have

$$[\mathrm{ad}_f g, \mathrm{ad}_f^2 g] = \begin{bmatrix} 0 \\ 0 \\ -\dfrac{\partial^2 d}{\partial y^2} \end{bmatrix} = a\begin{bmatrix}1\\0\\0\end{bmatrix} + \beta\begin{bmatrix}0\\1\\0\end{bmatrix}$$

for certain smooth functions a, β. This gives $\dfrac{\partial^2 d}{\partial y^2} = 0$; so $d(y,z) = a(z)y + b(z)$. Condition (F3) implies that $a(0) \neq 0$; conditions (F2) and (F4) are in the scalar input case satisfied automatically.

We find a change of coordinates and a feedback to bring (4.1) into a linear form. Put

$$\bar{x} = x + (ay+b)\frac{d}{dz}\left(\frac{b}{a}\right),$$

$$\bar{y} = y + \frac{b}{a},$$

$$\bar{z} = \int_0^z \frac{1}{a(t)}\,dt.$$

We have

$$f = \begin{bmatrix} m(\bar{x},\bar{y},\bar{z}) \\ \bar{x} \\ \bar{y} \end{bmatrix}, \quad g = \begin{bmatrix}1\\0\\0\end{bmatrix},$$

where m is a certain smooth function. Using feedback we can eliminate m by putting $a_1 = -m$ in (iii) (see the definition of F), i.e., we take new $\hat{f} = f - mg = \begin{bmatrix}0\\\bar{x}\\\bar{y}\end{bmatrix}$.

EXAMPLE 4.3. Consider again system (4.1) and study the problem of G-linearization. Compute $g = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathrm{ad}_f g = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$ and $\mathrm{ad}_f^2 g = \begin{bmatrix} 0 \\ 0 \\ \frac{\partial d}{\partial y} \end{bmatrix}$. Condition (G3) gives $\frac{\partial d}{\partial y}(0) \neq 0$, and if we assume that, then (G1) is satisfied.

We find transformations in G to bring (4.1) to a linear form under the above condition. Introduce the new coordinates

$$\begin{aligned} \bar{x} &= x\frac{\partial d}{\partial y} + d\frac{\partial d}{\partial z}, \\ \bar{y} &= d(y, z), \\ \bar{z} &= z, \end{aligned}$$

in which we have

$$f = \begin{bmatrix} m(\bar{x}, \bar{y}, \bar{z}) \\ \bar{y} \\ \bar{x} \end{bmatrix}, \quad g = \begin{bmatrix} \frac{\partial d}{\partial y} \\ 0 \\ 0 \end{bmatrix},$$

where m is certain smooth function. Using (ii)″ (see the definition of G) we can replace g by $\hat{g} = \frac{1}{\frac{\partial d}{\partial y}} \cdot g = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. Finally, using feedback (iii), we replace f by $\hat{f} = f - m\hat{g}$ to obtain $f = \begin{bmatrix} 0 \\ \bar{x} \\ \bar{y} \end{bmatrix}$.

5. Proofs

In this section we give proofs of Theorems 2.1 and 2.3.

Proof of Theorem 2.1. From the Jacobi identity $[f, [g_1, g_2]] = [g_1, \mathrm{ad}_f g_2] + [\mathrm{ad}_f g_1, g_2]$ we have by an induction argument

$$\mathrm{ad}_f^j [g_1, g_2] = \sum_{p_1+p_2=j} \frac{j!}{(p_1)!(p_2)!} [\mathrm{ad}_f^{p_1} g_1, \mathrm{ad}_f^{p_2} g_2] \quad \text{for any } f, g_1, g_2 \in V(M).$$

This implies that condition (2.1) is equivalent to the following one:

$$[\mathrm{ad}_f^p g_s, \mathrm{ad}_f^j g_t] = 0 \quad \text{for any } 1 \leqslant s, t \leqslant k \text{ and any } 0 \leqslant p \leqslant n-1,\ 0 \leqslant j \leqslant n. \tag{5.1}$$

Write

$$L_0 = \{\mathrm{ad}_f^q g_i \mid q \geqslant 0, 1 \leqslant i \leqslant k\},$$

and denote by $\mathscr{L}_0$ the Lie algebra generated by L_0. Note that $\mathscr{L} = \operatorname{span}\{\mathscr{L}_0, f\}$ and, since $f(x_0) = 0$, the Lie rank condition means that $\dim \mathscr{L}_0(x_0) = n$. We prove that the condition $\dim \mathscr{L}_0(x_0) = n$ and (2.1) imply that

$$\dim\operatorname{span}\{\mathrm{ad}_f^q g_i(x_0) \mid 0 \leqslant q \leqslant n-1, 1 \leqslant i \leqslant k\} = n.$$

Take a neighbourhood U_{x_0} of x_0 in which $\dim \mathscr{L}_0(x) = n$. For any $1 \leqslant i \leqslant k$ let $p_i(x)$ denote the largest non-negative number such that $g_i(x), \mathrm{ad}_f g_i(x), \ldots, \mathrm{ad}_f^p g_i(x)$ are linearly independent. Denote $p(x) = \max\limits_{1 \leqslant i \leqslant k} p_i(x)$. It is obvious that $0 \leqslant p(x) \leqslant n-1$. Since all the functions $p_i(x)$, $i = 1, \ldots, k$ are lower semicontinuous, $p(x)$ is lower semicontinuous and therefore $p(x)$ is locally constant on an open dense subset $\mathcal{O} \subset U_{x_0}$. Take a connected component $\mathcal{O}_a \subset \mathcal{O}$ on which $p(x)$ is constant and equal to p. Define the family of vector fields

$$L^p = \{\mathrm{ad}_f^q g_i \mid 1 \leqslant i \leqslant k, 0 \leqslant q \leqslant p\}.$$

We show that $\dim\operatorname{span} L^p(x) = n$ on a certain subset of $\mathcal{O}_a$.

Denote $N(x) = \dim\operatorname{span} L^p(x)$. Since the function $N(x)$ is lower semicontinuous, it is locally constant on a certain open dense subset $\pi \subset \mathcal{O}_a$. Take a connected component $\pi_b \subset \pi$ on which $N(x)$ is constant and equal to N. Thus on π_b the distribution D spanned by vector fields from L^p is N-dimensional and involutive (see (5.1)).

The definition of $p(x)$ implies that on π_b we have

$$\mathrm{ad}^{p+1} g_i = \sum_{0 \leqslant q \leqslant p} a_{iq} \mathrm{ad}_f^q g_i$$

for suitable smooth functions a_{iq}. For any vector fields X, Y and smooth functions α, β we have

$$[\alpha X, \beta Y] = \alpha\beta[X, Y] + \alpha\big((\nabla\beta)X\big)Y - \beta\big((\nabla\alpha)Y\big)X.$$

This and the above expression for $\mathrm{ad}_f^{p+1} g_i$ imply that for any $1 \leqslant i \leqslant k$ and $0 \leqslant j \leqslant p$ the Lie bracket $[\mathrm{ad}_f^{p+1} g_i, \mathrm{ad}_f^j g_s]$ belongs to D. Using an induction argument, we find that $\mathrm{ad}_f^q g_i$ belongs to D for any $q \geqslant 0$ and any $1 \leqslant i \leqslant k$. Thus the involutivity of D gives $h \in D$ for any $h \in \mathscr{L}_0$. Therefore from the Lie rank condition we have $\dim\operatorname{span} L^p(x) = n$ for $x \in \pi_b$.

Take a point $x \in \pi_b$. Choose from the family L^p vector fields $h_1, h_2, \ldots$ $\ldots, h_n$ in such a way that $h_1(x), \ldots, h_n(x)$ span an n-dimensional space.

From (5.1) we have $[h_i, h_j] = 0$ for any $1 \leqslant i, j \leqslant n$. Thus there exists (see Bishop–Crittenden [1]) a coordinate system $(x_1, \ldots, x_n)$ around $x \in \pi_b$ in which h_i takes the form $h_i = \frac{\partial}{\partial x_i}$, $i = 1, \ldots, n$. From (5.1) we have $[h_i, g_s] = \left[\frac{\partial}{\partial x_i}, g_s\right] = 0$ for any $1 \leqslant i \leqslant n$ and $1 \leqslant s \leqslant k$ and this means that in the coordinate system $(x_1, \ldots, x_n)$ the vector fields g_s, $s = 1, \ldots, k$, are constant. Since $h_i = \mathrm{ad}_f^q g_s$, $h_j = \mathrm{ad}_f^r g_t$ for certain $0 \leqslant q, r \leqslant p$ and $1 \leqslant s, t \leqslant k$, in the coordinate system $(x_1, \ldots, x_n)$ we have

$$-[\mathrm{ad}_f^q g_s, \mathrm{ad}_f^{r+1} g_t] = [\mathrm{ad}_f^q g_s, [\mathrm{ad}_f^r g_t, f]] = [h_i, [h_j, f]] = \left[\frac{\partial}{\partial x_i}, \left[\frac{\partial}{\partial x_j}, f\right]\right] = 0$$

because of (5.1). This means that in the above coordinate system the vector field f takes the affine form $f = Ax + C$, where the matrix A and the vector field C are constant.

This implies that on π_b we have $[\mathrm{ad}_f^q g_s, \mathrm{ad}_f^j g_t] = 0$ for any $0 \leqslant q, j$ and $1 \leqslant s, t \leqslant k$ as the Lie bracket of two constant vector fields. Since $\pi = \bigcup \pi_b$ is dense and open in $\mathcal{O}_a$ and $\mathcal{O} = \bigcup \mathcal{O}_a$ is dense and open in U_{x_0}, we have $[\mathrm{ad}_f^q g_s, \mathrm{ad}_f^j g_t] = 0$ on U_{x_0}. Therefore the Lie rank condition implies that $\dim \operatorname{span} L_0(x_0) = n$ and from the family L_0 we can choose n vector fields $h_1, \ldots, h_n$ linearly independent at x_0 and commuting. Take a coordinate system $(x_1, \ldots, x_n)$ around x_0 in which h_i are of the form $h_i = \frac{\partial}{\partial x_i}$, $i = 1, \ldots, n$. Repeat the above arguments to show that in this coordinate system the vector fields g_i, $i = 1, \ldots, k$, are constant and, since $f(x_0) = 0$, it follows that $f = Ax$ for a certain constant matrix A. ∎

To prove Theorem 2.3 we need some notation. For any diffeomorphism $\varphi\colon M \to M$ let φ_* denote its tangent map, i.e., for any vector field $h \in V(M)$ we have

$$\varphi_* h(x) = D\varphi|_{\varphi^{-1}(x)} h(\varphi^{-1}(x)),$$

where $D\varphi|_{\varphi^{-1}(x)}$ means the differential of φ taken at $\varphi^{-1}(x)$. For any $h \in V(M)$ we will write $\operatorname{Exp}(th)x$ for a (local) flow of h, i.e.,

$$\frac{d}{dt}\operatorname{Exp}(th)x = h(\operatorname{Exp}(th)x) \quad \text{and} \quad \operatorname{Exp}(0h)x = x.$$

For a fixed t, the map $x \mapsto \operatorname{Exp}(-th)x$ is a diffeomorphism from a neighbourhood of $\operatorname{Exp}(th)x$ on a neighbourhood of x and has the tangent map $\operatorname{Exp}(th)_*$. The derivative of the curve $t \mapsto \operatorname{Exp}(-th)_* g(x)$ at $t = 0$ is equal to $[h, g](x)$ (see Bishop–Crittenden [1]). Therefore in the analytic case we obtain

$$\operatorname{Exp}(-th)_* g(x) = \sum_{j=0}^{\infty} \frac{t^j}{j!} \mathrm{ad}_h^j g(x).$$

Proof of Theorem 2.3. We now show that under the condition $f(x_0) = 0$ the assumption $\dim\operatorname{span} L_0(x_0) = n$ implies $\dim\operatorname{span} L^{n-1}(x_0) = n$ (see the notation in the proof of Theorem 2.1). Notice that if $\mathrm{ad}_f^j g_m(x_0)$ is linearly dependent on the other vectors $\mathrm{ad}_f^q g_i(x_0)$, $q \leqslant j$, $1 \leqslant i \leqslant k$, then for any $p > j$ the vector $\mathrm{ad}_f^p g_m(x_0)$ is linearly dependent on the other vectors of the form $\mathrm{ad}_f^q g_i(x_0)$, $q \leqslant p$, $1 \leqslant i \leqslant k$. Indeed, let

$$\mathrm{ad}_f^j g_m(x_0) = \sum_{i\neq m} c_i \mathrm{ad}_f^j g_i(x_0) + \sum_{\substack{0\leqslant q\leqslant j-1\\ 1\leqslant i\leqslant k}} c_{iq}\mathrm{ad}_f^q g_i(x_0) \quad \text{for certain } c_i, c_{iq} \in \mathbf{R}.$$

We will write a for $\mathrm{ad}_f^j g_m$ and b_α, $\alpha \in A$ for the vector fields $c_i \mathrm{ad}_f^j g_i$ and $c_{iq}\mathrm{ad}_f^q g_i$; thus $a(x_0) = \sum_{\alpha\in A} b_\alpha(x_0)$. Obviously we have

$$\operatorname{span}\{a(x_0), b_\alpha(x_0) \mid \alpha \in A\} = \operatorname{span}\Big\{\Big(a - \sum_{\alpha\in A} b_\alpha\Big)(x_0), b_\alpha(x_0) \mid \alpha \in A\Big\},$$

but $\mathrm{ad}_f(a - \sum_{\alpha\in A} b_\alpha)(x_0) = 0$ as the Lie bracket of two vector fields vanishing at x_0. This implies that

$$\mathrm{ad}_f^{j+1} g_m(x_0) \in \operatorname{span}\{\mathrm{ad}_f b_\alpha(x_0) \mid \alpha \in A\}.$$

An induction argument gives the desired property for any $p > 1$. Since $\dim\operatorname{span} L_0(x_0) = n$, it follows that $\dim\operatorname{span} L^{n-1}(x_0) = n$ and from the family L^{n-1} we can choose n vector fields $h_1, \ldots, h_n$ linearly independent at x_0. We show that $h_1, \ldots, h_n$ commute in a certain neighbourhood of x_0.

Let $\mathcal{O}$ be a certain neighbourhood of $0 \in \mathbf{R}^n$ and define a map $a\colon \mathcal{O} \to M$ by the formula

$$\mathcal{O} \ni (s_1, \ldots, s_n) = s \mapsto a(s) = \operatorname{Exp}(s_1 h_1) \ldots \operatorname{Exp}(s_n h_n) x_0.$$

Since $\dfrac{\partial a}{\partial s_i}(0) = h_i(x_0)$, a maps $\mathcal{O}$ onto a neighbourhood U_{x_0} of x_0. Notice that for a fixed $s = (s_1, \ldots, s_n)$ a is a diffeomorphism and its differential will be denoted by Da. Taking if needed, a smaller $\mathcal{O}$, we may assume a to be a diffeomorphism on $\mathcal{O}$. We show that $[h_i, h_j] = 0$ on U_{x_0}. We have

$$\begin{aligned}[h_i, h_j](x) = [h_i, h_j]\big(a(s)\big) &= Da|_{x_0} Da^{-1}|_{a(s)}[h_i, h_j]\big(a(s)\big)\\ &= Da|_{x_0} a_*^{-1}[h_i, h_j](x_0).\end{aligned}$$

Therefore to conclude that h_i and h_j commute on U_{x_0} we have to prove that $a_*^{-1}[h_i, h_j] = 0$. In fact, we have

$$\begin{aligned}a_*^{-1}[h_i, h_j](x_0) &= \big(\operatorname{Exp}(-s_1h_1)\operatorname{Exp}(-s_2h_2)\ldots\operatorname{Exp}(-s_nh_n)\big)_*[h_i, h_j](x_0)\\ &= \operatorname{Exp}(-s_1h_1)_*\operatorname{Exp}(-s_2h_2)_*\ldots\operatorname{Exp}(-s_nh_n)_*[h_i, h_j](x_0)\end{aligned}$$

$$= \mathrm{Exp}(-s_1 h_1)_* \ldots \mathrm{Exp}(-s_{n-1} h_{n-1})_* \sum \frac{(s_n)^{r_n}}{(r_n)!} \cdot$$
$$\cdot \mathrm{ad}_{h_n}^{r_n} [h_i, h_j](x_0)$$
$$= \sum \frac{(s_1)^{r_1}}{(r_1)!} \mathrm{ad}_{h_1}^{r_1} \Big(\ldots \Big(\sum \frac{(s_n)^{r_n}}{(r_n)!} \mathrm{ad}_{h_n}^{r_n} [h_i, h_j] \Big) \ldots \Big)(x_0).$$

In view of analyticity, the above term can be expressed as

$$\sum \frac{(s_1)^{r_1}}{(r_1)!} \Big(\ldots \Big(\sum \frac{(s_n)^{r_n}}{(r_n)!} \mathrm{ad}_{h_1}^{r_1} \ldots \mathrm{ad}_{h_n}^{r_n} [h_i, h_j] \Big) \ldots \Big)(x_0)$$

and is equal to zero because of (2.4).

Therefore $h_1, \ldots, h_n$ are vector fields commuting and linearly independent at x_0 and we can choose a coordinate system $(x_1, \ldots, x_n)$ around x_0 in which $h_i = \dfrac{\partial}{\partial x_i}$, $i = 1, \ldots, n$. We show that in this coordinate system the vector fields f and g_i are of the form (1.2). In fact, for any $1 \leqslant j = j_1 + j_2 + \ldots + j_n$ we have

$$\frac{\partial^j g_i}{\partial x_n^{j_n} \ldots \partial x_1^{j_1}}(x_0) = \mathrm{ad}_{h_n}^{j_n} \ldots \mathrm{ad}_{h_1}^{j_1} g_i(x_0) = 0$$

because of (2.4) for any $1 \leqslant i \leqslant k$. This means that all derivatives of the coordinates of vector fields g_i, $i = 1, \ldots, k$, are equal to zero at x_0. Analyticity shows that they are constant.

The vector field f takes in this coordinate system a linear form $f = Ax$. In fact, for any $1 \leqslant i \leqslant n$ and $1 \leqslant j = j_1 + j_2 + \ldots + j_n$ we have

$$\frac{\partial^{j+1} f}{\partial x_n^{j_n} \ldots \partial x_1^{j_1} \partial x_i}(x_0) = \mathrm{ad}_{h_n}^{j_n} \ldots \mathrm{ad}_{h_1}^{j_1} \mathrm{ad}_f h_i(x_0) = 0$$

because of (2.4). Thus, by analyticity, the matrix $\dfrac{\partial f}{\partial x_i}$ is constant and, since $f(x_0) = 0$, we have $f = Ax$.

References

[1] R. L. Bishop and R. J. Crittenden, *Geometry of manifolds*, Academic Press, New York 1964.
[2] R. W. Brockett, *Feedback invariants for nonlinear systems*, in: Proc. IFAC Congress, Helsinki 1978.
[3] P. Brunovsky, *A classification of linear controllable systems*, Kybernetika (1970), 173–180.
[4] L. R. Hunt, R. Su and G. Meyer *Multi-input nonlinear systems*, in: *Differential Geometric Control Theory* (ed. R. W. Brockett, R. S. Millman and H. J. Sussmann) Birkhäuser, Boston 1983.

[5] B. Jakubczyk and W. Respondek, *On linearization of control systems*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. 28, 9–10 (1980), 517–522.

[6] A. J. Krener, *On the equivalence of control systems and linearization of nonlinear systems*, SIAM J. Control Optimization 11 (1973), 670–676.

[7] T. Nagano, *Linear differential systems with singularities and an application to transitive Lie algebras*, J. Math. Soc. Japan 18 (1966), 398–404.

[8] W. Respondek, *Geometric methods in linearization and decomposition of nonlinear control systems*, Ph. D. Thesis, Inst. of Math. Pol. Acad. Sci., Warsaw 1981 [in Polish].

[9] R. Su, *On the linear equivalents of nonlinear systems*, Systems and Control Letters 2 (1982), 48–52.

[10] W. M. Wonham and A. S. Morse, *Feedback invariants of linear multivariable systems*, Automatica–J. IFAC 8 (1972), 93–100.