Systems & Control Letters 8 (1986) 69–74
North-Holland

Conditioned invariant and locally conditioned invariant distributions

Arthur J. KRENER *

Department of Mathematics, University of California, Davis, CA 95616, USA

Received 23 January 1986
Revised 10 April 1986

Abstract: We define the concepts of conditioned invariant and locally conditioned invariant distributions for nonlinear system and show how they can be used to solve the problem of tracking an output signal of a nonlinear system in spite of unknown disturbances.

Keywords: Nonlinear systems, Invariant distributions, (h, f) invariance, Conditioned invariance, Nonlinear tracking.

1. Introduction

The mathematically dual concepts of (A, B) (or controlled) invariance and (C, A) (or conditioned) invariance play an important role in the geometric theory of linear systems. These concepts were introduced by Basile and Marro [1] and Wonham and Morse [12] and were used by them and others [2,9,11,13] to study the problems of disturbance rejection and tracking for linear systems in state space form.

The concept of (A, B) invariance was generalized to nonlinear systems by Isidori, Krener, Gori-Giorgi and Monaco [6] and Hirschorn [5] and used by them and others [3,4,8,10] to solve problems of disturbance rejection and noninteracting control for nonlinear systems. This concept, called (f, g) invariance in [6], seemed at first to be a straightforward generalization of the linear one, utilizing nonlinear state feedback in the obvious way. But later it becomes apparent that there are at least two distinct generalizations of (A, B) invariance called (f, g) invariance and local (f, g) invariance [7]. Both are needed for a complete treatment of the problem of nonlinear disturbance rejection. If one considers singular distributions then even more distinctions must be made [8]. For the sake of brevity we shall not consider singular distributions in this paper.

Progress on the nonlinear generalizations of (C, A) invariance has been slower because there is no obvious way of extending the concept of output injection to nonlinear systems. In [6], (h, f) invariance was defined and used to solve problems of disturbance rejection using static and dynamic output feedback. That work followed the linear paradigm found in [11]. This paper, which builds on [6], proposes two generalizations of (C, A) invariance which we call (h, f) invariance and local (h, f) invariance. The latter is the (h, f) invariance of [6]. We show that both concepts are needed for a complete treatment of the problem of nonlinear tracking.

Van der Schaft [14] has discussed the problem of constructing observers and partial observers for nonlinear systems. He showed that the existence of a partial observer without stability requirements is locally equivalent to the existence of a conditioned invariant distribution. The partial observer problem can be viewed as the tracking problem when the external noise is absent.

The rest of the paper is organized as follows. Section 2 reviews the linear geometric theory of tracking. This is generalized to nonlinear systems in Section 3. The conclusion follows in Section 4.

2. Linear paradigm

To motivate the definition of (h, f) (or conditioned) invariance for nonlinear systems we review the linear theory. A simple (perhaps too simple) problem which motivates this concept is the tracking of a signal in the presence of unknown noise. Consider the linear autonomous system

$$\dot{x} = Ax + Bu + Ew, \tag{2.1a}$$

$$y = Cx, \tag{2.1b}$$

* Research supported in part by NSF under MCS8300884 and NASA under NAG 2-268.

$$z = Dx, \tag{2.1c}$$

$$x(0) = x^0, \tag{2.1d}$$

where the state variable or memory of the process is $x(t)$, the control is $u(t)$, the noise is $w(t)$, the observation is $y(t)$ and the signal is $z(t)$. The problem is to track the signal $z(t)$ exactly from knowledge of the system, i.e. the matrices A, B, C, D, E, the control $u(t)$, the observation $y(t)$ and the initial condition x^0 in spite of the effects of the unknown noise $w(t)$.

The type of solution that we are looking for is another linear system

$$\dot{\hat{x}} = \hat{A}\hat{x} + \hat{B}u + \hat{K}y, \tag{2.2a}$$

$$\hat{z} = \hat{D}\hat{x} + \hat{L}y, \tag{2.2b}$$

$$\hat{x}(0) = \hat{x}^0 = Px^0, \tag{2.2c}$$

and a linear mapping

$$P : x \to \hat{x}. \tag{2.2d}$$

Hopefully if (2.2) is chosen properly then $z(t) = \hat{z}(t)$ for all $t > 0$ and all x^0, $u(t)$ and $w(t)$. Notice that if a tracking system exists we need only know x^0 modulo the kernel of P, or in other words we need only know $\hat{x}^0 = Px^0$.

The following result is a standard application of the geometric method of Basile and Marro [2].

Theorem 2.1. *There exists a system of the form* (2.2) *which tracks the signal* $z(t)$ *of* (2.1) *iff there exists subspace of* $\mathscr{V} \subseteq \mathbb{R}^n$ *such that*

$$\mathscr{V} \text{ is } (C, A) \text{ invariant}, \tag{2.3a}$$

$$\mathscr{R}(E) \subseteq \mathscr{V}, \tag{2.3b}$$

$$\mathscr{V} \cap \mathscr{N}(C) \subseteq \mathscr{N}(D), \tag{2.3c}$$

where

$$\mathscr{R}(E) = \text{Range } E : \mathbb{R}^l \to \mathbb{R}^n$$

and

$$\mathscr{N}(D) = \text{Kernel } D : \mathbb{R}^n \to \mathbb{R}^q.$$

Recall that a subspace $\mathscr{V}$ is (C, A) invariant (also called conditioned invariant) if one of the following equivalent conditions holds, there exists an $n \times p$ matrix K such that

$$(A + KC)\mathscr{V} \subseteq \mathscr{V} \tag{2.4a}$$

or

$$A(\mathscr{V} \cap \mathscr{N}(C)) \subseteq \mathscr{V}. \tag{2.4b}$$

We refer to (2.4a) and (2.4b) as the local and global characterizations of (C, A) invariance for reasons that will become apparent when we generalize to nonlinear systems.

The proof of Theorem 2.1 is a straightforward exercise in geometric linear systems theory. One shows that the global characterization of (C, A) invariance (2.4a) and conditions (2.3b) and (2.3c) imply the existence of a tracking system (2.2). From $\mathscr{V}$ one can explicitly compute the tracking system (2.2) by choosing the appropriate state coordinates as we describe in the next section (3.5). The next step is to show that if a tracking system exists then there exists a subspace $\mathscr{V}$ which satisfies the local characterization of (C, A) invariance (2.4b) and conditions (2.3b) and (2.3c). The final step is to show the equivalence of the two characterizations (2.4a,b) of (C, A) invariance. It is clear that the global characterization implies the local, the reverse implication is only a little harder to show.

As the proof of the theorem suggests, the global characterization of (C, A) invariance is more useful in constructing a tracking system. On the other hand the local characterization is more useful in finding a subspace $\mathscr{V}$ satisfying (2.3b,c). It is readily apparent that the collection of all subspaces of the state space which satisfy the local characterization of (C, A) invariance forms a semilattice under inclusion and intersection. Since (2.3b) is a lower bound on $\mathscr{V}$, it is convenient to construct the minimal (C, A) invariant subspace $\mathscr{V}_*$ containing $\mathscr{R}(E)$. This can be found by standard techniques [2,13]. We define an expanding sequence of subspace

$$\mathscr{V}^0 = \mathscr{R}(E), \tag{2.5a}$$

$$\mathscr{V}^{r+1} = \mathscr{V}^r + A(\mathscr{V}^r \cap \mathscr{N}(C)), \tag{2.5b}$$

and finally

$$\mathscr{V}_* = \bigcup_{r \geq 0} \mathscr{V}^r. \tag{2.5c}$$

Because the state space is finite dimensional the algorithm terminates in a finite number of steps, i.e. $\mathscr{V}_* = \mathscr{V}^r$ for some r. This construction allows us to express Theorem 2.1 as follows.

Theorem 2.2. *There exists a system of the form* (2.2) *which tracks the signal* $z(t)$ *of* (2.1) *iff the minimal* (C, A) *invariant subspace* $\mathscr{V}_*$ *containing* $\mathscr{R}(E)$ *satisfies* (2.3c).

3. Nonlinear conditioned invariance

Consider the following nonlinear generalization of (2.1):

$$\dot{x} = f(x, u, w) = g^0(x) + g(x)u + q(x)w, \tag{3.1a}$$

$$y = h(x), \tag{3.1b}$$

$$z = k(x), \tag{3.1c}$$

$$x(0) = x^0. \tag{3.1d}$$

The state space is a smooth (C^∞) manifold M and the state is described in local coordinates $x(t)$ on M. For convenience we assume that the control $u(t)$ and the noise $w(t)$ enter the dynamics linearly and that $u(t) \in \mathbb{R}^m$ and $w(t) \in \mathbb{R}^l$. The observable take values in a smooth p-dimensional manifold N and is given in local coordinates by $y(t)$. The signal $z(t) \in \mathbb{R}^q$. One can easily extend what follows to the more complicated situations where all of the above spaces are nonlinear and where the control and noise enter the dynamics in a nonlinear fashion.

The vector fields described in local coordinates by $g^0, g^1, \ldots, g^m, q^1, \ldots, q^l$ are assumed to be smooth as are the outputs maps h and k.

As in the linear case, we wish to track the signal $z(t)$ from knowledge of the system (3.1), the control $u(t)$, the observable $y(t)$ and the initial condition x^0, but without knowledge of the noise $w(t)$.

More specifically we are interested in finding a nonlinear system which tracks (3.1) of the form

$$\dot{\hat{x}} = \hat{f}(\hat{x}, y, u) = \hat{g}^0(\hat{x}, y) + \hat{g}(\hat{x}, y)u, \tag{3.2a}$$

$$\hat{z} = \hat{k}(\hat{x}, y), \tag{3.2b}$$

$$\hat{x}(0) = \hat{x}^0 = \Pi(x^0), \tag{3.2c}$$

$$\Pi : M \to \hat{M}, \tag{3.2d}$$

where $\hat{x}$ denotes local coordinates on a smooth manifold $\hat{M}$ of dimension d and Π is a smooth mapping. Notice that if the tracking system exists, we need only know $\hat{x}^0 = \Pi(x^0)$.

Notice that the vector fields $\hat{g}^0, \hat{g}^1, \ldots, \hat{g}^m$ depend both on the state $\hat{x}$ of (3.2) and on the observable y of (3.1). In other words they are smooth sections of the vector bundle $(T\hat{M}) \times N$ with base space $\hat{M} \times N$ and fibers the various tangent spaces of $\hat{M}$. Hopefully if (3.2) is properly chosen then $z(t) = \hat{z}(t)$ for all $t > 0$ and all x^0, $u(t)$ and $w(t)$.

The nonlinear generalization of Theorem 2.1 is not as straightforward. First, we shall give rather strong sufficient conditions for the existence of a tracking system (3.2). These conditions include a generalization of the global characterization of (C, A) invariance which we call (h, f) invariance (or conditioned invariance). Next we shall give weaker necessary conditions for the existence of a tracking system. These involve the concept of local (h, f) invariance (or local conditioned invariance). Then we show that the necessary and sufficient conditions are equivalent locally around a generic point of the state space. Finally as in the linear case we present an algorithm which facilitates checking the necessary conditions.

A distributions $\mathscr{D}$ on the state space M of (3.1) is (h, f) invariant if these exists a smooth submersion

$$\Pi : M \to \hat{M}, \quad \Pi : x \mapsto \hat{x} \tag{3.3a}$$

($\hat{M}$ is another smooth manifold with local coordinates $\hat{x}$) and vector fields $\hat{g}^0(\hat{x}, y), \ldots, \hat{g}^m(\hat{x}, y)$ parameterized by $y \in N$ such that

$$\Pi_* g^j(x) = \hat{g}^j(\Pi(x), h(x)), \quad j = 0, \ldots, m, \tag{3.3b}$$

for all $x \in M$ and

$$\mathscr{D} = \mathscr{N}(\Pi_*). \tag{3.3c}$$

See [8] for the definition of a distribution and other background material. This definition makes the next theorem almost a tautology.

Theorem 3.1. *There exists a system* (3.2) *which tracks the signal* $z(t)$ *of* (3.1) *if there exists a distribution* $\mathscr{D}$ *such that*

$$\mathscr{D} \text{ is } (h, f) \text{ invariant}, \tag{3.4c}$$

$$\mathscr{R}(q) \subseteq \mathscr{D}, \tag{3.4b}$$

and there exists a function $\hat{k} : \hat{M} \times N \to \mathbb{R}^q$ *such that for all* $x \in M$,

$$k(x) = \hat{k}(\Pi(x), h(x)). \tag{3.4c}$$

$\mathscr{R}(q)$ is the distribution spanned over smooth functions by the vector fields of q.

Proof. We define the system (3.2) using (3.3) and (3.4). Let $e(t)=z(t)-\hat{z}(t)$, then from (3.4c) we see that $e(t)$ will be zero if the solution $\hat{x}(t)$ of (3.2a,c) equals $\Pi(x(t))$ where $x(t)$ is the solution of (3.1a,d). But this is assured by (3.3b) and (3.4b). □

We would like to emphasize that (h, f) invariance (3.3) is a generalization of the global characterization (2.4a) of (C, A) invariance. Given K and $\mathscr{V}$ satisfying (2.4a) we can choose coordinates $x=(x_1, x_2)$ so that $\mathscr{V}=\{x_1=0\}$. Let $\tilde{A}=A+KC$; then in these coordinates (2.1a) becomes (with $w=0$)

$$\begin{pmatrix}\dot{x}_1\\ \dot{x}_2\end{pmatrix}=\begin{pmatrix}\tilde{A}_{11} & 0\\ \tilde{A}_{21} & \tilde{A}_{22}\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}+\begin{pmatrix}K_1\\ K_2\end{pmatrix}\begin{pmatrix}C_1 & C_2\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}+\begin{pmatrix}B_1\\ B_2\end{pmatrix}u. \quad (3.5a)$$

Let $\hat{M}=\mathbb{R}^n/\mathscr{V}$ and $\Pi: x\mapsto\hat{x}=x_1$. Then Π carries (3.5a) onto

$$\dot{\hat{x}}=\tilde{A}_{11}\hat{x}+K_1y+B_1u \quad (3.5b)$$

as in (3.3b). Moreover condition (3.4c) is a global and nonlinear generalization of (2.3c).

A distribution $\mathscr{D}$ on the state space M of (3.1) is locally (h, f) invariant if

$$\operatorname{ad}(g^j)(\mathscr{D}\cap\mathscr{N}(\mathrm{d}h))\subseteq\mathscr{D} \quad (3.6)$$

for $j=0,\ldots,m$. ($\mathscr{N}(\mathrm{d}h)$ is the distribution of all vector fields in the kernel of $\mathrm{d}h=h_*$.) In other words if $X\in\mathscr{D}\cap\mathscr{N}(\mathrm{d}h)$,

$$\operatorname{ad}(g^j)X=[g^j, X]\in\mathscr{D}.$$

This is a generalization of the local characterization of (C, A) invariance (2.4b). In [6] this property (3.6) was called (h, f) invariance, we have added the adverb 'locally' to distinguish it from (3.3).

Theorem 3.2. *If there exists a system* (3.2) *which tracks the signal $z(t)$ of* (3.1) *then these exists an distribution $\mathscr{D}$ on M such that*

$\mathscr{D}$ is involutive and locally (h, f) invariant, (3.7a)

$$\mathscr{R}(q)\subseteq\mathscr{D}, \quad (3.7b)$$

$$\mathscr{D}\cap\mathscr{N}(\mathrm{d}h)\subseteq\mathscr{N}(\mathrm{d}k). \quad (3.7c)$$

Proof. Suppose (3.2) exists, and consider the combined system on $\tilde{M}=M\times\hat{M}$ with local coordinates $\tilde{x}=(x, \hat{x})$.

$$\dot{\tilde{x}}=\tilde{g}^0(\tilde{x})+\tilde{g}(\tilde{x})u+\tilde{q}(\tilde{x})w, \quad (3.8a)$$

$$\begin{pmatrix}\dot{x}\\ \dot{\hat{x}}\end{pmatrix}=\begin{pmatrix}g^0(x)\\ \hat{g}^0(\hat{x}, h(x))\end{pmatrix}+\begin{pmatrix}g(x)\\ \hat{g}(\hat{x}, h(x))\end{pmatrix}u+\begin{pmatrix}q(x)\\ 0\end{pmatrix}w, \quad (3.8b)$$

$$e=\tilde{k}(x, \hat{x})=k(x)-\hat{k}(\hat{x}, h(x)), \quad (3.8c)$$

$$\tilde{x}(0)=\begin{pmatrix}x^0\\ \hat{x}^0\end{pmatrix}=\begin{pmatrix}x^0\\ \Pi(x^0)\end{pmatrix}. \quad (3.8d)$$

We define an expanding sequence of codistributions on $\tilde{M}$ by

$$\tilde{\mathscr{E}}^0=\mathscr{R}(\mathrm{d}\tilde{k}), \quad (3.9a)$$

$$\tilde{\mathscr{E}}^{r+1}=\tilde{\mathscr{E}}^r+\sum_{j=0}^{m}L_{\tilde{g}^j}(\tilde{\mathscr{E}}^r), \quad (3.9b)$$

$$\tilde{\mathscr{E}}=\bigcup_{r\geq 0}\tilde{\mathscr{E}}^r, \quad (3.9c)$$

and a distribution $\tilde{\mathscr{D}}$ by

$$\tilde{\mathscr{D}}=\mathscr{N}(\tilde{\mathscr{E}}). \quad (3.10)$$

$\mathscr{R}(\mathrm{d}\tilde{k})$ denotes all C^∞ combinations of the one forms $\mathrm{d}\tilde{k}_1,\ldots,\mathrm{d}\tilde{k}_q$. $\mathscr{N}(\tilde{\mathscr{E}})$ is the distribution of vector fields on $\tilde{M}$ which are annihilated by all the one forms of $\tilde{\mathscr{E}}$. It is not hard to see that at least locally $\tilde{\mathscr{E}}$ is spanned by exact one forms hence $\tilde{\mathscr{D}}$ is involutive. By definition $\tilde{\mathscr{D}}$ is invariant under $\tilde{g}^j$, $j=0,\ldots,m$,

$$\operatorname{ad}\tilde{g}^j(\tilde{\mathscr{D}})\subseteq\tilde{\mathscr{D}}, \quad (3.11a)$$

and $\tilde{\mathscr{D}}$ is the maximal such distribution satisfying

$$\tilde{\mathscr{D}}\subseteq\mathscr{N}(d\tilde{k}). \quad (3.11b)$$

Since tracking is achieved, w is decoupled from e. Hence

$$\mathscr{R}(\tilde{q})\subseteq\tilde{\mathscr{D}}. \quad (3.11c)$$

Next we define a distribution $\mathscr{D}$ on M as

$$\mathscr{D}=\left\{X(x):\begin{pmatrix}X(x)\\0\end{pmatrix}\in\tilde{\mathscr{D}}\right\}.$$

Since $\tilde{\mathscr{D}}$ is involutive so is $\mathscr{D}$. Conditions (3.11a) implies (3.7a), (3.11b) implies (3.7b) and (3.11c) implies (3.7c). □

The next two lemmas relates the necessary conditions of the last theorem with the sufficient conditions of Theorem 3.1.

Lemma 3.3. *Suppose $\mathscr{D}$ is an (h, f) invariant distribution on M; then $\mathscr{D}$ is involutive and locally (h, f) invariant. Suppose $\mathscr{D}$ is involutive and locally (h, f) invariant; then for an open and dense subset of $x^0 \in M$ there exists an open neighborhood $\mathscr{U}$ of x^0 such that restricted to $\mathscr{U}$, $\mathscr{D}$ is (h, f) invariant.*

Proof. Suppose $\mathscr{D}$ is (h, f) invariant (3.3); then, as the kernel of a map (3.3c) $\mathscr{D}$ is involutive.

We choose local coordinates $x=(x_1, x_2)$ around any x^0 so that $x_1=\hat{x}\circ\Pi$. In these coordinates,

$$\dot{x}_1=\hat{g}_1^0(x_1, h(x))+\hat{g}_1(x_1, h(x))u,$$

$$\dot{x}_2=g_2^0(x)+g_2(x)u.$$

If $X(x)\in\mathscr{D}\cap\mathscr{N}(\mathrm{d}h)$ then

$$X(x)=\begin{pmatrix}0\\X_2(x)\end{pmatrix}$$

and

$$[g^j(x), X(x)]=\left[\begin{pmatrix}\hat{g}_1^j(x_1, h(x))\\g_2^j(x)\end{pmatrix},\begin{pmatrix}0\\X_2(x)\end{pmatrix}\right]$$

$$=\begin{pmatrix}\dfrac{\partial\hat{g}_1^j}{\partial y}\dfrac{\partial h}{\partial x}X\\ *\end{pmatrix}=\begin{pmatrix}0\\ *\end{pmatrix}$$

as desired.

On the other hand suppose $\mathscr{D}$ is an involutive and locally (h, f) invariant distribution. Let x^0 be a point where $\mathscr{D}$, $\mathscr{N}(\mathrm{d}h)$ and $\mathscr{D}\cap\mathscr{N}(\mathrm{d}h)$ are of constant dimension in a neighborhood of x^0. Such points are open and dense in M. We can choose local coordinates $x=(x_1, x_2)$ around x^0 where $\mathscr{D}=\mathscr{N}(\mathrm{d}x_1)$. Since $\mathscr{D}$, $\mathscr{N}(\mathrm{d}h)$ and $\mathscr{V}\cap\mathscr{N}(\mathrm{d}h)$ are nonsingular around x^0 it is easy to verify that the local (h, f) invariant condition (3.6) is equivalent to

$$L_{g^j}(\mathscr{D}^\perp)\subseteq\mathscr{D}^\perp+\mathscr{R}(\mathrm{d}h). \tag{3.12a}$$

$\mathscr{D}^\perp$ is the codistribution of one forms which annihilates the vector fields of $\mathscr{D}$. In the above local coordinates $\mathscr{D}^\perp=\mathscr{R}(\mathrm{d}x_1)$. From (3.12a) we have

$$\mathrm{d}g_i^j(x)=L_{g^j}(\mathrm{d}x_1)\in\mathscr{R}(\mathrm{d}x_1)+\mathscr{R}(\mathrm{d}h) \tag{3.12b}$$

where

$$\dot{x}_1=g_1^0(x)+\hat{g}_1(x)u.$$

The implicit function and (3.12b) ensure that locally there exists $g^j(x_1, y)$ such that

$$g_1^j(x)=\hat{g}_1^j(x_1, y).\qquad\square$$

We omit the proof of the next result since it is similar to the above.

Lemma 3.4. *Suppose $\mathscr{D}$ is (h, f) invariant and there exists a function $\hat{k}$ satisfying* (3.4c); *then* (3.7c) *is satisfied. Suppose $\mathscr{D}$ is an involutive locally (h, f) invariant distribution and* (3.7c) *is satisfied; then for an open dense subset of $x^0=M$ there exists an open neighborhood $\mathscr{U}$ of x^0 such that* (3.4c) *is satisfied on $\mathscr{U}$.*

Notice that it is considerably more difficult to find the $\hat{f}$, $\hat{k}$, etc. of a nonlinear tracking system than it is to find the $\hat{A}$, $\hat{D}$, etc., of a linear tracking system. This is because the latter problem involves inverses and generalized inverses of linear mappings while the former involves inverses and generalized inverses of nonlinear mappings.

As in the linear case, the collection of involutive and locally (h, f) invariant distributions forms a semilattice under inclusion and intersection. Therefore there exists a minimal such distribution containing $\mathscr{R}(q)$ which can be found by a straightforward generalization of (2.5). We define an expanding family of involutive distribution (where $\overline{\mathscr{D}}$ denotes the involutive closure of $\mathscr{D}$) by

$$\mathscr{D}^0=\overline{\mathscr{R}(q)}, \tag{3.13a}$$

$$\mathscr{D}^{r+1}=\overline{\mathscr{D}^r+\sum_{j=0}^{m}\mathrm{ad}(g^j)(\mathscr{D}^r\cap\mathscr{N}(\mathrm{d}h))}, \tag{3.13b}$$

and finally

$$\mathscr{D}_* = \bigcup_{r \geq 0} \mathscr{D}^r. \tag{3.13c}$$

Of course this algorithm need not terminate in a finite number of steps at all $x \in M$ but it will do so at a generic x. We can express Theorem 3.2 as follows.

Theorem 3.5. *If there exists a system* (3.2) *which tracks the signal* $z(t)$ *of* (3.1) *then the minimal involutive locally* (h, f) *invariant distribution* $\mathscr{D}_*$ *containing* $\mathscr{R}(q)$ *satisfies* (3.7c).

4. Conclusion

We have presented a nonlinear generalization of a linear tracking problem. The solution is very similar to the linear theory but differs from it in two essential ways. The first is that the nonlinear necessary and sufficient conditions are not equivalent as are the linear conditions. While this is important it is not a crucial defect of the nonlinear theory because locally around a generic state the conditions are equivalent.

The second difference is more substantial, namely that we have not treated the stability of the error of tracking system. If there is any noise in the observation, or inaccuracies in the model and its initial condition the tracking system will not exactly track the signal. Of course, stability is always one of the most difficult aspects of nonlinear systems theory. In linear theory, one has available convenient pole placement techniques and perhaps this is the principle advantage of linearity. It is clear that more work must be done on nonlinear stability in the geometric context of invariant foliations.

In closing we note that some of our earlier work with Isidori, Gori-Giorgi and Monaco [6] can be reformulated along the lines of Section 3. For example consider the problem of decoupling the noise $w(t)$ from the signal $z(t)$ in (3.1) by static state feedback. Sufficient conditions involve the concept of (f, g) invariance, necessary conditions involve the concept of local (f, g) invariance and locally around a generic point these conditions are equivalent. See also [8].

References

[1] G. Basile and G. Marro, Controlled and conditioned invariant subspaces in linear systems theory, *J. Optim. Theory Appl.* **3** (1969) 306–315.

[2] G. Basile and G. Marro, On the observability of linear, time-invariant systems with unknown inputs, *J. Optim. Theory Appl.* **3** (1969) 410–415.

[3] D. Claude, Decouplage des systemes non lineaires, series, generatrices non commutatives at algebres de Lie, *SIAM J. Control Optim.* **24** (1986) 562–578.

[4] J.W. Grizzle, Controlled invariance for discrete-time nonlinear systems with an application to the disturbance decoupling problem, *IEEE Trans. Automat. Control* **30** (1985) 868–874.

[5] R. Hirschorn, (A, B) invariant distributions and the disturbance decoupling of nonlinear systems, *SIAM J. Control Optim* **19** (1981) 1–19.

[6] A. Isidori, A.J. Krener, C. Gori-Giorgi and S. Monaco, Nonlinear decoupling via feedback, a differential geometric approach, *IEEE Trans. Automat. Control* **26** (1981) 331–345.

[7] A. Isidori, A.J. Krener, C. Gori-Giorgi and S. Monaco, locally (f, g) invariant distributions, *Systems Control Lett.* **1** (1981) 12–15.

[8] A.J. Krener, (ad f, g), (ad f, g) and locally (ad f, g) invariant and controllability distributions, *SIAM J. Control Optim.* **23** (1985) 523–549.

[9] A.S. Morse, and W.M. Wonham, Status of noninteracting control, *IEEE Trans. Automat. Control* **15** (1971) 568–581.

[10] H. Nijmeijer, Controlled invariance for affine control systems, *Internat. J. Control* **34** (1981) 825–833.

[11] J.M. Schumacher, Compensator synthesis using (C, A, B) pairs, *IEEE Trans. Automat. Control* **25** (1980) 1133–1137.

[12] W.M. Wonham and A.S. Morse, Decoupling and pole assignment in linear multivariable systems; a geometric approach, *SIAM J. Control* **8** (1970) 1–18.

[13] W.M. Wonham, *Linear Multivariable Control* (Springer, New York, 1974).

[14] A.J. van der Schaft, On nonlinear observers, *IEEE Trans. Automat. Control* **30** (1985) 1254–1256.