Nonlinear Decoupling via Feedback: A Differential Geometric Approach

ALBERTO ISIDORI, MEMBER, IEEE, ARTHUR J. KRENER, MEMBER, IEEE, CLAUDIO GORI-GIORGI, AND SALVATORE MONACO

***Abstract*—The paper deals with the nonlinear decoupling and noninteracting control problems. A complete solution to those problems is made possible via a suitable nonlinear generalization of several powerful geometric concepts already introduced in studying linear multivariable control systems. The paper also includes algorithms concerned with the actual construction of the appropriate control laws.**

Manuscript received June 25, 1979; revised October 19, 1979 and August 25, 1980. Paper recommended by M. Vidyasagar, Past Chairman of the Stability, Nonlinear, and Distributed Systems Committee. This work was supported by the University of Rome under Grant 7.8*P/1978 and by the National Research Council of Italy.

A. Isidori, C. Gori-Giorgi, and S. Monaco are with the Istituto di Automatica, University of Rome, Rome, Italy.

A. J. Krener is with the Department of Mathematics, University of California, Davis, CA 95616.

I. Introduction

CONSIDER a nonlinear system of the form

$$\dot{x}=f(x)+g(x)u \tag{1.1a}$$

$$x(0)=x^0 \tag{1.1b}$$

$$y=h(x) \tag{1.1c}$$

where the input u, the output y, and the state x are l, m, and n dimensional, f and h are vector-valued differentiable functions, and g is a matrix-valued differentiable function, all of the appropriate dimensions. We shall be more precise later on. The input–output behavior of such a system can

0018-9286/81/0400-0331$00.75

be altered by feedback to obtain some desired goal. Feedback means letting all or part of the input depend on the state or output in either a static or dynamic way.

For example, in the *disturbance decoupling problem* the input splits into two parts:

$$\dot{x}=f(x)+g(x)u+p(x)w$$

where u is the part that can be controlled and w is an uncontrollable noise or disturbance. We may seek to modify the dynamics via static state feedback

$$u=\alpha(x)$$

to remove the effect of the disturbance w on the output y.

More generally, the output may also split into two parts

$$y=h(x)$$
$$z=k(x)$$

and we may seek to remove the effect of the disturbance w on the output z using static feedback depending only on the output y:

$$u=\alpha(y).$$

In some cases, in order to achieve this goal, it may be desirable to allow dynamic feedback, i.e., to let

$$u=\alpha(\xi, y)$$
$$\dot{\xi}=\varphi(\xi, y)$$
$$\xi(0)=\xi^0.$$

A similar problem is that of *noninteracting control*. Here we have a system with the same number of input and output channels, $u_1,\cdots,u_m$ and $y_1,\cdots,y_m$, and we wish to use static (or dynamic) feedback in order to obtain a system in which each input u_i affects only the corresponding output y_i in a nontrivial fashion.

We give necessary and sufficient conditions for the existence of solutions to the nonlinear decoupling, noninteracting control, and similar problems. The linear autonomous versions of these problems have been treated extensively in the literature [4], [15], and our work is a generalization of those. In particular, our results include linear but nonautonomous systems which to our knowledge have not been completely discussed before. Related work on nonlinear systems can be found in [9]. We work in the geometric spirit of Wonham and Morse [15]. The key tools are the nonlinear generalizations of the notions of (A, B) and (C, A) invariant subspaces, introduced by Basile and Marro [1] and Wonham and Morse [15]. These generalizations are extremely powerful and can be used to solve numerous variations of the above-mentioned problems which are of practical interest. They are also related to the study of controllability, observability, minimality, decomposability, and invertibility for nonlinear systems which will be discussed later.

The rest of the paper is organized as follows. Section II sets notation and introduces the key concepts. In Section III we deal with disturbance decoupling. Section IV is concerned with construction of maximal and minimal invariant distributions and feedback control laws. Finally, Section V discusses noninteracting control.

II. Invariant Distributions

Let us be more precise about the system (1.1) which we are discussing. We assume the state x evolves on a C^∞ (or analytic) manifold M of dimension n, i.e., a Hausdorff topological space with a countable cover $(\mathrm{U}^\alpha, x_\alpha)$ of coordinate charts such that:

1) U^α is an open subset of M

2) $x_\alpha=\mathrm{col}(x_{\alpha 1},\cdots,x_{\alpha n})$: $\mathrm{U}^\alpha\rightarrow\mathbb{R}^n$ is a homeomorphism onto its range

3) If $(\mathrm{U}^\alpha, x_\alpha)$ and $(\mathrm{U}^\beta, x_\beta)$ are two such coordinate charts, then the change of coordinates $x_\beta\circ x_\alpha^{-1}$: $x_\alpha(\mathrm{U}^\alpha\cap \mathrm{U}^\beta)\rightarrow x_\beta(\mathrm{U}^\alpha\cap\mathrm{U}^\beta)$ is C^∞ (analytic).

Whenever possible we work in local coordinates, identifying points of M with their coordinates relative to some fixed coordinate system. It is in this spirit that (1.1) is to be interpreted. More precisely, $f(x)$ and each column of $g(x)$ is the local coordinate representation of a C^∞ (analytic) vector field on M, i.e., a mapping which assigns to each point $x\in M$ a tangent vector in T_xM, the tangent space to M at x. The output h is a C^∞ (analytic) mapping from M to $\mathbb{R}^m$.

If M, f, g, and h are C^∞ (analytic), the system is said to be smooth (analytic). All systems treated in this paper are implicitly assumed to be smooth; occasionally, as is needed, we make an explicit assumption of analyticity. In most cases the assumption of infinite differentiability is not essential, but only invoked to avoid having to count the degree of differentiability needed in a particular argument. Moreover, most of our results can easily be generalized to systems nonlinear in the control

$$\dot{x}=f(x,u).$$

In the linear systems theory, linear subspaces of the state space which have particular properties are studied. Because one distinguishes between points and tangent vectors in a nonlinear setting, the concept of subspace generalizes to two different ones. A k-dimensional *regularly imbedded submanifold* M' of M is a subset of M such that around each point of M' there exists a coordinate neighborhood U such that $M'\cap\mathrm{U}$ is given by $\{x_j=c_j: j=1,\cdots,n-k\}$ where $c_1,\cdots,c_{n-k}$ are constants. On $M'\cap\mathrm{U}$, one has local coordinates given by $(x_{n-k+1},\cdots,x_n)$. A *distribution* Δ on M is a mapping which assigns to each $x\in M$ a subspace $\Delta(x)$ of T_xM in a smooth (analytic) fashion. If each of these subspaces is of dimension k, Δ is said to be of rank k. The connection between the two concepts is the following: M' is an *integral submanifold* of Δ if for every $x\in M'$, $\Delta(x)=T_xM'\subseteq T_xM$. In other words, $\Delta(x)$ is the tangent space to M' at x.

It is useful to identify tangent vectors with directional differentations, i.e., if $\tau(x)\in T_xM$ and is given in local coordinates by $\tau(x)=\mathrm{col}(\tau_1(x),\cdots,\tau_n(x))$, then $\tau(x)$ can be identified with the directional differentiation

$$\sum_{i=1}^{n}\tau_i(x)\frac{\partial}{\partial x_i}.$$

Under this identification, $\partial/\partial x_i$ is the unit vector in the x_i direction. Hence, if M' is an integral manifold of Δ, then at each $x\in M'$

$$\Delta(x)=\mathrm{span}\left\{\frac{\partial}{\partial x_i}: i=n-k+1,\cdots,n\right\}.$$

When, for clarity, it is important to emphasize differentiation with respect to a vector field τ, we use the Lie differentiation symbol

$$L_\tau=\sum_{i=1}^{n}\tau_i(x)\frac{\partial}{\partial x_i}.$$

Frequently we confuse a distribution Δ with the space of vector fields which are pointwise in Δ. In other words, a vector field τ belongs to Δ if $\tau(x)\in\Delta(x)$ for every x.

A distribution Δ is involutive if it is closed under the Lie bracket, i.e., if $\tau,\sigma\in\Delta$ then $[\tau,\sigma]\in\Delta$ where

$$[\tau,\sigma](x)=\frac{\partial\sigma}{\partial x}(x)\tau(x)-\frac{\partial\tau}{\partial x}(x)\sigma(x).$$

The *involutive closure* $\bar{\Delta}$ of a distribution Δ is the minimal involutive distribution containing Δ.

Theorem (Frobenius): Let Δ be an involutive distribution of rank k on M. There then exists a partition of M into maximal integral submanifolds of Δ, each of dimension k.

Theorem (Hermann–Nagano): Let Δ be an analytic involutive distribution on an analytic manifold M. There then exists a partition of M into maximal integral submanifolds of Δ of varying dimension.

The submanifolds of these theorems may not be regularly imbedded, for example, the winding line on a torus. *A regularly imbedded submanifold* of dimension k is one which is given locally as the intersection of level sets of $n-k$ independent functions. Even if they are regularly imbedded the quotient may still not be a C^∞ manifold. For that reason we make the following definition. A distribution Δ is said to be *regular* if it is involutive, of constant rank, and partitions M into regularly imbedded submanifolds where the quotient admits a C^∞ structure such that the canonical projection is a submersion.

In the linear case, a subspace of $\mathbb{R}^n$ induces a distribution on $\mathbb{R}^n$ by assigning to each $x\in\mathbb{R}^n$ precisely that subspace of tangent directions. Any such distribution is involutive and of constant rank. The maximal integral submanifolds that the distribution generates are the translates of this subspace.

A distribution Δ is *invariant* under the dynamics (1.1) if

$$[f,\Delta]\subseteq\Delta \tag{2.1}$$

$$[g,\Delta]\subseteq\Delta. \tag{2.2}$$

By this we mean the bracket of f and any column of g with any vector field of Δ is again a vector field of Δ. We constantly abuse notation in this fashion, using the symbol for a collection of objects to stand for a generic object in the collection.

Lemma 2.1: Let $\Delta^1\subseteq\Delta^2$ be distributions satisfying

$$[f,\Delta^1]\subseteq\Delta^2$$

$$[g,\Delta^1]\subseteq\Delta^2;$$

then

$$[f,\bar{\Delta}^1]\subseteq\Delta^3$$

$$[g,\bar{\Delta}^1]\subseteq\Delta^3$$

where Δ^3 is the ideal generated by Δ^1 in Δ^2.

Proof: Let σ and τ be vector fields of Δ^1; then

$$[f,\sigma]\in\Delta^2$$

$$[f,\tau]\in\Delta^2;$$

thus,

$$[\tau,[f,\sigma]]\in[\Delta^1,\Delta^2]$$

$$[\sigma,[f,\tau]]\in[\Delta^1,\Delta^2].$$

By subtracting and applying the Jacobi identity, we obtain

$$[f,[\tau,\sigma]]=[\tau,[f,\sigma]]-[\sigma,[f,\tau]]\in[\Delta^1,\Delta^2].$$

A similar argument holds for g. By induction the lemma follows. ■

If we apply this lemma with $\Delta^1=\Delta^2=\Delta$ and $\Delta^3=\bar{\Delta}$, we see that the involutive closure of an invariant distribution is also invariant.

The importance of invariant distributions is that the system (1.1) projects onto a lower dimensional system. In local coordinates it is easy to see. Let Δ be an invariant regular distribution of rank k. Choose local coordinates $x=(x_1,x_2)$ where x_1 and x_2 are $n-k$ and k-dimensional, respectively, and Δ is spanned by

$$\frac{\partial}{\partial x_2}=\left(\frac{\partial}{\partial x_{2,1}},\cdots,\frac{\partial}{\partial x_{2,k}}\right).$$

In these coordinates, (1.1) splits

$$\dot{x}_1=f_1(x_1,x_2)+g_1(x_1,x_2)u \tag{2.3}$$

$$\dot{x}_2=f_2(x_1,x_2)+g_2(x_1,x_2)u; \tag{2.4}$$

but since

$$\left[f,\frac{\partial}{\partial x_2}\right]=-\left(\frac{\partial f_1}{\partial x_2}\frac{\partial}{\partial x_1}+\frac{\partial f_2}{\partial x_2}\frac{\partial}{\partial x_2}\right)\subseteq\mathrm{span}\left\{\frac{\partial}{\partial x_2}\right\}$$

we see that

$$\frac{\partial f_1}{\partial x_2}=0,$$

i.e., f_1 does not depend on x_2. Similarly, neither does g_1, thus (2.3) becomes

$$\dot{x}_1=f_1(x_1)+g_1(x_1)u. \tag{2.5}$$

The global version of (2.5) requires regularity.

Lemma 2.2: Let Δ be an invariant regular distribution. Then Δ induces a regular equivalence relation on M such that the dynamics passes to the quotient denoted by $M'=M/\Delta$.

M' is not Hausdorff in general, but it is if we assume some sort of controllability of the system (see [6] for details). Similar cascade decompositions have been extensively treated by Krener [14].

Consider the linear system

$$\dot{x}=Ax+Bu \tag{2.6}$$

$$y=Cx. \tag{2.7}$$

The concept corresponding to an invariant distribution is that of an *invariant subspace*, i.e., a subspace V satisfying

$$A\mathrm{V}\subseteq \mathrm{V}. \tag{2.8}$$

This is a specialization of (2.1); the bracket of the linear vector field Ax with a constant vector field of V. The second condition (2.2) is trivially satisfied, as the bracket of any of the constant vector fields making up the columns of B with a constant vector field of V is 0 and V is spanned by constant vector fields. Lemma 2.1 is trivial in this case for any subspace describes an involutive distribution. Lemma 2.2 corresponds to the existence of a linear change of coordinates in which the dynamics (2.6) becomes

$$\dot{x}_1=A_{11}x_1+B_1u \tag{2.9}$$

$$\dot{x}_2=A_{21}x_1+A_{22}x_2+B_2u. \tag{2.10}$$

Basile and Marro [1] and independently Wonham and Morse [15] have introduced the concept of (A, B) invariant subspace, that is, a subspace V which is invariant after suitable modification of the dynamics (2.6) by linear feedback $u=Fx+v$:

$$\dot{x}=(A+BF)x+Bv. \tag{2.11}$$

More precisely, a subspace V is (A, B) *invariant* if there exists a $l\times n$ matrix F such that

$$(A+BF)\mathrm{V}\subseteq \mathrm{V}. \tag{2.12a}$$

An equivalent formulation of this concept is that V is (A, B) invariant if

$$A\mathrm{V}\subseteq \mathrm{V}+\mathrm{B} \tag{2.12b}$$

where B is the subspace spanned by the columns of B.

The dual concept was introduced by Basile and Marro [1]: a subspace V is (C, A) *invariant* if there exists an $n\times m$ matrix F such that

$$(A+FC)\mathrm{V}\subseteq \mathrm{V} \tag{2.13a}$$

or equivalently

$$A(\mathrm{V}\cap \ker C)\subseteq \mathrm{V}. \tag{2.13b}$$

In (A, B) invariance, one is limited in changing the dynamics via feedback through B but with complete information about the state. In the dual concept, (C, A) invariance, one is not limited in the way the feedback enters the equation for $\dot{x}$ but that it uses only the information available from the output. The new dynamics is

$$\dot{x}=(A+FC)x+Bu. \tag{2.14}$$

We would like to point out the obvious generalization of both. A subspace V is (C, A, B) *invariant* if there exists a $l\times m$ matrix F such that

$$(A+BFC)\mathrm{V}\subseteq \mathrm{V}. \tag{2.15}$$

This corresponds to output feedback, $u=Fy+v$,

$$\dot{x}=(A+BFC)x+Bv. \tag{2.16}$$

It is obvious that a (C, A, B) invariant subspace is both (C, A) and (A, B) invariant. With a little work one can prove the converse, namely, if V is both (C, A) and (A, B) invariant, then for suitable F it is (C, A, B) invariant. Hence, (2.15) is equivalent to (2.12b) and (2.13b).

Notice that, in the above concepts, we were only concerned with modifying the drift term via feedback; we did not consider linear change of coordinates in the input space nor the deletion of inputs. This is because it does not affect the invariance of V. As we noted before, only the linear version of (2.1) is important; the linear version of (2.2) is trivially satisfied. In the nonlinear generalization of these concepts one must also study nonlinear change of coordinates in the input space and the possibility of modifying the dynamics by the deletion of inputs.

A distribution Δ is (f, g) *invariant* if there exist α and β which are $l\times 1$ and $l\times r$ valued functions of x such that

$$[\tilde{f}, \Delta]\subseteq \Delta \tag{2.17a}$$

$$[\tilde{g}, \Delta]\subseteq \Delta \tag{2.17b}$$

where

$$\tilde{f}(x)=f(x)+g(x)\alpha(x) \tag{2.18a}$$

$$\tilde{g}(x)=g(x)\beta(x). \tag{2.18b}$$

This corresponds to Δ being invariant under the new dynamics

$$\dot{x}=\tilde{f}(x)+\tilde{g}(x)v.$$

For smooth (analytic) systems we require α and β to be smooth (analytic).

Of course this is a local description, in global terms we require the new dynamics be given by (2.18) where the vector fields $\tilde{f}(x)-f(x)$ and the columns of $\tilde{g}(x)$ are pointwise in the span of $g(x)$.

The new control v is of dimension $r \leq l$, if $l=r$ and if $\beta(x)$ is nonsingular or equivalently if $\tilde{g}(x)$ spans the same subspace of the tangent space as $g(x)$ for every x, then we say Δ is (f,g) *invariant with full control*, otherwise Δ is (f,g) *invariant with partial control.* By Lemma 2.1, if Δ is (f,g) invariant, then the involutive closure $\bar{\Delta}$ is (f,g) invariant also.

There is a local version of (f,g) invariance analogous to (2.12b). Δ is *locally* (f,g) *invariant* on an open subset $U \subseteq M$ if for every $x \in U$

$$[f,\Delta](x) \subseteq \Delta(x) + \text{span}\, g(x) \tag{2.19a}$$

$$[g,\Delta](x) \subseteq \Delta(x) + \text{span}\, g(x). \tag{2.19b}$$

Using Lemma 2.1, one can easily show that if Δ is locally (f,g) invariant, then the involutive closure $\bar{\Delta}$ is (f,g) invariant also.

The following lemma is a generalization of Hirschorn [16].

Lemma 2.3: Suppose Δ is locally (f,g) invariant on a simply connected open subset $U \subseteq M$, and on U the dimension of $\Delta(x) \cap \text{span}\, g(x)$ is constant. Then there exists $\alpha(x)$, $\beta(x)$ such that (2.17) is satisfied on U.

Since it is closer to (f,g) invariance, we define an (h,f,g) invariance next. Consider Δ^0, the distribution of all vector fields on M which annihilate the output $y=h(x)$ under directional differentiation

$$\Delta^0(x) = \ker dh(x) = \bigcap_{i=1}^{m} \ker dh_i(x)$$

where dh_i is the differential of the ith component h_i of h, i.e., $dh_i = \text{row}\,(\partial h_i/\partial x_1, \cdots, \partial h_i/\partial x_n)$.

A distribution Δ is (h,f,g) *invariant* if there exist functions α, β such that (2.17) holds, and in addition

$$\Delta^0(x) \subseteq \ker d\alpha(x) \tag{2.20a}$$

$$\Delta^0(x) \subseteq \ker d\beta(x). \tag{2.20b}$$

Intuitively this means that locally α and β only depend on y. Again, one makes a distinction between full and partial control depending on the invertibility of β. Unfortunately, there seems to be no simple local formulation of (h,f,g) invariance analogous to (2.19) not explicitly involving α and β. Once again from Lemma 2.1, if Δ is (h,f,g) invariant, its involutive closure $\bar{\Delta}$ is (h,f,g) invariant.

Finally, we define (h,f) invariance, the nonlinear generalization of (C,A) invariance. A distribution Δ is (h,f) *invariant* if

$$[f, \Delta \cap \Delta^0] \subseteq \Delta \tag{2.21a}$$

$$[g, \Delta \cap \Delta^0] \subseteq \Delta. \tag{2.21b}$$

Notice we implicitly assume full control in (2.21b). It is straightforward to verify that a distribution which is (h,f,g) invariant with full control is both (f,g) and (h,f) invariant with full control, but the converse does not seem to be true. The following example shows that if Δ is (h,f) invariant, the involutive closure of Δ need not be.

Example: Let $\Delta(x)$ be spanned by two vector fields $\sigma(x)=\text{col}(1,0,0,0)$, $\tau(x)=\text{col}(0,1,x_1,0)$, and $y=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Then $\Delta \cap \Delta^0 = 0$; hence Δ is trivially (h,f) invariant. However, $[\sigma,\tau]=\text{col}(0,0,1,0) \in \Delta^0$ and if we define $f(x)=\text{col}(0,0,0,x_3)$, then $[f,[\sigma,\tau]]=\text{col}(0,0,0,-1) \notin \bar{\Delta}$ so that $\bar{\Delta}$ is not (h,f) invariant.

To see how an (h,f) invariant Δ is useful, suppose we choose local coordinates $x=(x_1,x_2,x_3)$ such that

$$\Delta \cap \Delta^0 = \text{span}\left\{\frac{\partial}{\partial x_3}\right\}$$

$$\Delta = \text{span}\left\{\frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}\right\}.$$

In these coordinates the system becomes

$$\dot{x}_1 = f_1(x_1,x_2) + g_1(x_1,x_2)u \tag{2.22a}$$

$$\dot{x}_2 = f_2(x_1,x_2,x_3) + g_2(x_1,x_2,x_3)u \tag{2.22b}$$

$$\dot{x}_3 = f_3(x_1,x_2,x_3) + g_3(x_1,x_2,x_3)u \tag{2.22c}$$

$$y = h(x_1,x_2) \tag{2.22d}$$

since (2.21) implies that $\partial f_1/\partial x_3 = 0$ and $\partial g_1/\partial x_3 = 0$. Locally, the map $(x_1,x_2) \to (x_1,y)$ is one-to-one and hence has a left inverse: therefore, $x_2 = x_2(x_1,y)$. Hence, if $x_1(0)$, $u(t)$, and $y(t)$ are known, then $x_1(t)$ can be computed via the differential equation

$$\dot{x}_1 = f_1(x_1, x_2(x_1,y)) + g_1(x_1, x_2(x_1,y))u$$

regardless of any noise affecting the differential equations for x_2 and x_3. As will be demonstrated in the next section, this property is essential for decoupling via dynamic output feedback.

Notice that in defining (f,g) invariance, one uses the generalization of definition (2.12a) which is more restrictive than the generalization of (2.12b). In defining (h,f) invariance, one generalizes (2.13b) which is less restrictive than the generalization of (2.13a).

III. Disturbance Decoupling

Suppose we have the following system where both input and output have been split into two channels:

$$\dot{x} = f(x) + g(x)u + p(x)w \tag{3.1a}$$

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

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

The input w represents a noise or disturbance over which we have no control and the output z represents the quantities which we wish to insulate from the effects of the noise. We intend to do this via feedback: in the least general situation, we allow static feedback depending on the state.

That is, we seek smooth (analytic) feedback functions $\alpha(x)$ and $\beta(x)$ which make the output z independent of noise w. In contrast to the linear case, this does not mean that the output z is zero or even constant, but rather that it is the same for all possible noise inputs. If this is possible, we say that the *static, state feedback, noise decoupling problem* is solvable. If $\beta(x)$ is invertible, then the modified system has lost none of its controllability and we say the problem is solvable *with full control*. We assume that the controls $u(t)$ and $v(t)$ are piecewise smooth (piecewise analytic) but that the noise $w(t)$ need only be bounded and measurable.

Suppose there exists $\alpha(x)$ and $\beta(x)$ such that in suitable local coordinates (3.1) becomes

$$\dot{x}_1=\tilde{f}_1(x_1)+\tilde{g}_1(x_1)v \quad (3.2a)$$

$$\dot{x}_2=\tilde{f}_2(x_1,x_2)+\tilde{g}_2(x_1,x_2)v+p_2(x_1,x_2)w \quad (3.2b)$$

$$z=k(x_1) \quad (3.2c)$$

where

$$\tilde{f}(x)=f(x)+g(x)\alpha(x) \quad (3.3a)$$

$$\tilde{g}(x)=g(x)\beta(x) \quad (3.3b)$$

then clearly the static, state feedback noise decoupling problem is solvable at least locally. The global version of this is that there exists a projection π: $M\to M'$, a manifold of lower dimension, given in the local coordinates of (3.2) by

$$\pi(x_1,x_2)=x_1. \quad (3.4)$$

The feedback modified system projects to a system on M' given by (3.2a) and (3.2c) with the same input (v,w)/output z behavior as (3.2). In the projected system, the vector fields multiplied by the noise w are identically zero; hence the system is unaffected by the noise. If there exist feedback functions α, β and such a projection π, then we say that the *static, state feedback, noise decoupling problem is solvable in a regular fashion*.

Theorem 3.1: The static, state feedback, noise decoupling problem is solvable in a regular fashion (with full control) iff there exists a regular distribution Δ such that

1) Δ is (f,g) invariant (with full control)
2) $p\subseteq\Delta\subseteq\ker dk$.

Theorem 3.2: For analytic systems the static, state feedback, noise decoupling problem is solvable (with full control) iff there exists a distribution Δ satisfying 1) and 2) of Theorem 3.1.

We defer the proof of these results to consider the mathematically more general situation of requiring the feedback α and β only to depend on the other output y [see (3.1b)]. To be more precise, α and β are functions of x as before but their directional derivatives are zero whenever the directional derivative of h is zero:

$$\ker dh\subseteq\ker d\alpha\cap\ker d\beta. \quad (3.5)$$

This is equivalent to (2.20).

We refer to this case as the *static, output feedback, noise decoupling problem*. Although we have restricted the available feedback functions, this latter problem is more general because if we take the output to be the state $y=x$, then we have the static, state feedback, noise decoupling problem. Hence, Theorems 3.1 and 3.2 are corollaries of the next results.

Theorem 3.3: The static, output feedback, noise decoupling problem is solvable in a regular fashion (with full control) if and only if there exists a regular distribution Δ such that

1) Δ is (h,f,g) invariant (with full control)
2) $p\subseteq\Delta\subseteq\ker dk$.

Proof: Suppose a regular distribution Δ exists satisfying 1), then there exist α,β satisfying (3.5). If we define $\tilde{f}$ and $\tilde{g}$ by (3.3), then Δ is invariant. Invoking Lemma 2.2, we obtain the local decomposition

$$\dot{x}_1=\tilde{f}_1(x_1)+\tilde{g}_1(x_1)v+p_1(x_1,x_2)w \quad (3.6a)$$

$$\dot{x}_2=\tilde{f}_2(x_1,x_2)+\tilde{g}_2(x_1,x_2)v+p_2(x_1,x_2)w \quad (3.6b)$$

$$z=k(x_1,x_2). \quad (3.6c)$$

But from 2) we see that $p_1=0$ and $\partial k/\partial x_2=0$ so that (3.6) reduces to (3.2) and we have solved the static, output feedback, noise decoupling problem in a regular fashion.

On the other hand, if the problem is solvable in a regular fashion then there exists feedback functions α,β satisfying (3.5) and a projection π given locally by (3.4) such that the modified dynamics is given locally by (3.2). Clearly, $\ker d\pi$ is a regular distribution on M satisfying 1) and 2). ■

Theorem 3.4: For analytic systems the static, output feedback, noise decoupling problem is solvable (with full control) iff there exists a distribution Δ satisfying 1) and 2) of Theorem 3.3.

Proof: Suppose Δ exists satisfying 1) and 2), then there exist α,β satisfying (3.5) such that Δ is left invariant [see (2.17)] by the modified dynamics (2.18). Since Δ is analytic, by the Hermann–Nagano theorem it partitions M into maximal integral submanifolds and since $\Delta\subseteq\ker dk$, the value of k is constant on each of these submanifolds.

Let $v(t)$ be a piecewise analytic control and $w(t)$ be bounded and measurable noise defined for $t\in[0,\tau]$. For any $0\leq t,s\leq\tau$, define

$$w(t;s)=\begin{cases}w(t) & t\leq s\\ 0 & t>s\end{cases}$$

and let $x(t;s)$ be the family of solutions of the modified dynamics with inputs $v(t)$ and $w(t;s)$

$$\frac{\partial x}{\partial t}(t;s)=\tilde{f}(x(t;s))+\tilde{g}(x(t;s))v(t)+p(x(t;s))w(t;s)$$

$$x(0;s)=x^0.$$

Let M' be the maximal integral submanifold of Δ containing $x(\tau;0)$. Using the Campbell–Baker–Hausdorff formula [17, p. 47], one can expand $\partial x(\tau;s)/\partial s$ in a series whose coefficients are constructed from Lie brackets of $\tilde{f}$

and $\tilde{g}$ with p evaluated at $x(\tau; s)$. From this we see that

$$\frac{\partial x}{\partial s}(\tau; s) \in \Delta(x(\tau; s))$$

and hence the curve $s \mapsto x(\tau; s)$ lies in M'. This implies that $k(x(\tau; 0)) = k(x(\tau; \tau))$. In other words, the output $z(\tau)$ with noise $w(t)$ is exactly the same as the output $z(\tau)$ with zero noise and hence the problem is solved.

On the other hand, suppose this problem is solvable, i.e., the output z is independent of the noise w when feedback functions α, β satisfying (2.20) are implemented. Define $\tilde{f}$ and $\tilde{g}$ by (2.18) and let Δ be smallest distribution which is invariant under $\tilde{f}$ and $\tilde{g}$ and which contains p. Clearly Δ is (h, f, g) invariant; thus we need only show that Δ is contained in the kernel of dk. Consider the dynamics of the modified system

$$\dot{x} = \tilde{f}(x) + \tilde{g}(x)v + p(x)w \tag{3.7a}$$

$$z = k(x). \tag{3.7b}$$

Fix an initial state, $x(0) = x^0$, a time $t > 0$ and an input $v(\cdot)$. Let $A(x^0, t; v(\cdot))$ be the set of points accessible at time t from x^0 using input $v(\cdot)$ and any bounded measurable noise input $w(\cdot)$. We have the following lemma, a generalization of Sussmann–Jurdjevic [14].

Lemma 3.5: $A(x^0, t; v(\cdot))$ is contained in an integral submanifold of Δ. For any $x \in M$ there exist x^0, t, and piecewise analytic $v(\cdot)$ such that x is contained in the interior of $A(x^0, t; v(\cdot))$ in the topology of that submanifold.

We defer the proof of the lemma for the moment. Suppose that at some x, $\Delta(x) \not\subseteq \ker dk(x)$. Then k varies on the integral submanifold of Δ through x. By proper choice of x^0, t, and $v(\cdot)$, we can have $A(x^0, t; v(\cdot))$ living on that integral submanifold with nonempty interior containing x. As we vary $w(\cdot)$, the output z changes, contradicting our decoupling assumption. Hence, $\Delta(x) \subseteq \ker dk(x)$ for all $x \in M$. ■

Proof of Lemma 3.5: From Sussmann and Jurdjevic we know that $A(x^0, t; v(\cdot))$ is contained in an integral submanifold of Δ. We would like to show that for any x^0 there exists a $t > 0$ and a piecewise analytic control $v(\cdot)$ such that $A(x^0, t; v(\cdot))$ has a nonempty interior in that submanifold. Choose any constant v^1 and $t_1 > 0$. Let $x^1 = x(t_1)$ be the endpoint of the trajectory of (3.7a) starting at x^0 using $v(t) = v^1$, $w(t) = 0$. At x^1, choose w^1 such that

$$p(x^1)w^1 \neq 0 \tag{3.8}$$

and consider the family of trajectories indexed by s_1 generated by the family of controls

$$\begin{aligned} v(t) &= v^1 \qquad 0 \leq t \leq t_1 \\ w(t) &= 0 \qquad 0 \leq t < t_1 - s_1 \\ w(t) &= w^1 \qquad t_1 - s_1 \leq t \leq t_1. \end{aligned}$$

Let $x(t_1; s_1)$ denote the locus of endpoints of these trajectories. From (3.8), as s_1 varies over small positive values, $x(t_1; s_1)$ describes a one-dimensional submanifold M^1 contained in $A(x^0, t_1; v(\cdot))$. Choose for some value of $s_1 > 0$, such that the point $x^2 = x(t_1; s_1) \in M^1$ and suppose there exists w^2 such that $p(x^2)w^2$ is not tangent to M^1. Define a family of controls indexed by $0 < s_2 < s_1 < t_1$:

$$\begin{aligned} v(t) &= v^1 \qquad 0 \leq t \leq t_1 \\ w(t) &= 0 \qquad 0 \leq t < t_1 - s_1 \\ w(t) &= w^1 \qquad t_1 - s_1 \leq t < t_1 - s_2 \\ w(t) &= w^2 \qquad t_1 - s_2 \leq t \leq t_1 \end{aligned}$$

and let $x(t_1; s_1, s_2)$ be the locus of endpoints of the corresponding trajectories. As we vary s_1 and s_2, we sweep out a two-dimensional submanifold M^2 of $A(x^0, t_1; v(\cdot))$. We continue in this fashion for as long as we can find a new w^j such that $p(x^j)w^j$ is not tangent to M^{j-1}. In this way we generate a k-dimensional submanifold M^k of $A(x^0, t_1; v(\cdot))$ such that at every $x \in M^k$ the columns of $p(x)$ are tangent to M^k.

Next choose any analytic input $v^2(t)$ defined on $[t_1, t_2]$ for some $t_2 > t_1$. Let $x(t_2; s_1, s_2, \cdots, s_k)$ be the endpoint of the trajectory generated by (3.7a) with $v(t)$ and $w(t)$ as before on $[0, t_1)$ and $v(t) = v^2(t)$, $w(t) = 0$ on $[t_1, t_2]$. Let $M^k_{t_2}$ be the k-dimensional manifold swept out by $x(t_2; s_1, \cdots, s_k)$ as we vary $s_1, \cdots, s_k$. Either $p(x)$ is tangent to $M^k_{t_2}$ for every $v^2(\cdot), t_2$, and at every $x \in M^k_{t_2}$ or it is not.

Suppose the former. From the Campbell–Baker–Hausdorff formula, it follows that for any column $p_j(x)$ and column $\tilde{g}_{i_r}(x)$, $i_r \in \{0, \cdots, l\}$ [where for convenience of notation $\tilde{g}_0(x) = \tilde{f}(x)$] we have that $[g_{i_1}[g_{i_2} \cdots [g_{i_p}, p_j] \cdots](x)$ is tangent to M^k at every $x \in M^k$. Moreover, the bracket of vector fields tangent to M^k is also tangent to M^k. Using the Jacobi identity it is easy to see that Δ is generated by vector fields of the above form, and thus $\Delta(x)$ is precisely the tangent space of M^k at every $x \in M^k$. Since M^k is contained in an integral manifold of Δ it must have nonempty interior in the topology of that manifold and the lemma is proved.

If there exists a $v(\cdot)$ and t_2 such that for some $x \in M^k_{t_2}$, some column of $p(x)$ is not tangent to $M^k_{t_2}$, we repeat at t_2 the construction done at t_1, obtaining a larger dimensional manifold. The proof continues as above until the former holds.

We have shown for any x^0 there exists a t and $v(\cdot)$ such that $A(x^0, t; v(\cdot))$ has nonempty interior in the integral mainfold of Δ. We wish to show that given any x^1 there exists x^0, t, and $v(\cdot)$ such that x^1 is an element of the interior of $A(x^0, t; v(\cdot))$. The latter follows from the former by taking x^1 as the initial condition of the time reversed system and choosing x^0 in the interior of the set of accessible points of that system at time t. ■

Remark: For linear systems (2.6) and (2.7), one restricts attention to linear feedback functions $u = Fx + v$ and distributions which are the translates of linear subspaces. Since such systems are analytic and such distributions are always regular, the distinction between Theorems 3.3 and 3.4 [(or 3.1 and 3.2)] vanishes. To our knowledge the linear version of Theorem 3.3 (or 3.4) has never been explicitly stated, i.e., the state, output feedback, noise decoupling problem is solvable for a linear system by linear feedback iff there exists a (C, A, B) invariant subspace satisfying the linear analog of 2). We leave the proof of this result to the interested reader.

Finally we consider the dynamic, *output feedback, noise decoupling problem*. The goal is the same as before, to insulate the output z from the input w, but now we allow the feedback functions α and β to depend dynamically on the output. In other words, we are allowed to add to our system (3.1) the additional system whose dynamics is

$$\dot{\xi}=\varphi(\xi,\mu) \tag{3.9}$$

and whose output is equal to the state and then decouple by feedback from the outputs (y,ξ) to the inputs (u,μ). In essence one uses this additional dynamics to track some of the state variables of the original system which are not available through y using other state variables which are available through y. For this reason we must assume that the initial state of the system is known

$$x(0)=x^0. \tag{3.10}$$

We say that the *dynamic, output feedback, noise decoupling problem is solvable in a regular fashion* if there exists a system (3.9), feedback functions $\alpha(y,\xi)$ and $\beta(y,\xi)$, and a projection π: $M\to M'$ such that in local coordinates (3.2) and (3.4) hold. The next two results are generalizations of theorems of Laschi and Marro [18], Basile and Marro [19], and Schumacker [12] for linear systems.

Theorem 3.6: Consider the system (3.1) and (3.10) and suppose the following:

1) there exists Δ^1, a regular (f,g) invariant distribution (with full control)
2) there exists Δ^2, a regular (h,f) invariant distribution
3) the distribution $\Delta^3=\Delta^2\cap\Delta^0$ is regular
4) $p\subseteq\Delta^2\subseteq\Delta^2\subseteq\ker dk$. (3.11)

Then the dynamic, output feedback, noise decoupling problem is solvable in a regular fashion (with full control).

The following theorem is almost the converse of the last result.

Theorem 3.7: Suppose the dynamic, output feedback, noise decoupling problem is solvable in a regular fashion with full control; there then exists a regular distribution Δ such that

1) Δ is locally (f,g) invariant on M
2) Δ is (h,f) invariant
3) $p\subseteq\Delta\subseteq\ker dk$.

Proof of Theorem 3.6: Given any $x^1\in M$, choose a coordinate cube (U^1,x) centered at x^1 such that $x=(x_0,x_1,x_2,x_3)$ and

$$\Delta^3=\operatorname{span}\left\{\frac{\partial}{\partial x_3}\right\} \tag{3.12a}$$

$$\Delta^2=\operatorname{span}\left\{\frac{\partial}{\partial x_2},\frac{\partial}{\partial x_3}\right\} \tag{3.12b}$$

$$\Delta^1=\operatorname{span}\left\{\frac{\partial}{\partial x_1},\frac{\partial}{\partial x_2},\frac{\partial}{\partial x_3}\right\}. \tag{3.12c}$$

Let $\bar{\mathsf{U}}^{-1}$ be the slice of U^1 given by $x_3=x_3^1$.

Define an equivalence relation D on M, x^iDx^j if there exists a piecewise smooth curve $\gamma(t)$ joining x^i to x^j such that $\dot{\gamma}(t)\in\Delta^3(\gamma(t))$. For $\mathsf{V}\subseteq M$, let $D(\mathsf{V})$ be the set of points D-related to a point of V. Then $D(\bar{\mathsf{U}}^1)=D(\bar{\mathsf{U}}^1)$ and since the distribution Δ^3 is regular, each $x\in D(\mathsf{U}^1)$ is D-related to a unique point of $\bar{\mathsf{U}}^1$. Call this mapping π^1:

$$\pi^1: D(\mathsf{U}^1)\to\bar{\mathsf{U}}^1$$

it is a smooth projection. In this way x_0, x_1, and x_2 extend to functions on all of $D(\mathsf{U}^1)$, $x_i=x_i\circ\pi^1$.

Let $\tilde{\alpha}$ and $\tilde{\beta}$ be the feedback functions making $\Delta^1(f,g)$ invariant, define feedback α^1,β^1 on $D(\mathsf{U}^1)$ by

$$\alpha^1=\tilde{\alpha}\circ\pi^1$$

$$\beta^1=\tilde{\beta}\circ\pi^1$$

trivially

$$\Delta^3(x)\subseteq\ker d\alpha^1(x)\cap\ker d\beta^1(x) \tag{3.13}$$

on $D(\mathsf{U}^1)$.

We would like to show α^1 and β^1 leave Δ^1 invariant on $D(\mathsf{U}^1)$, i.e., for $x\in D(\mathsf{U}^1)$

$$[f+g\alpha^1,\Delta^1](x)\subseteq\Delta^1(x) \tag{3.14a}$$

$$[g\beta^1,\Delta^1](x)\subseteq\Delta^1(x). \tag{3.14b}$$

Since Δ^2 is (h,f) invariant

$$[f,\Delta^3]\subseteq\Delta^2 \tag{3.15a}$$

$$[g,\Delta^3]\subseteq\Delta^2. \tag{3.15b}$$

From (3.13) and (3.15) we have

$$[f+g\alpha^1,\Delta^3](x)\subseteq\Delta^2(x)\subseteq\Delta^1(x) \tag{3.16a}$$

$$[g\beta^1,\Delta^3](x)\subseteq\Delta^2(x)\subseteq\Delta^1(x) \tag{3.16b}$$

for all $x\in D(\mathsf{U}^1)$. Moreover, since $\alpha^1=\tilde{\alpha}$ on $\bar{\mathsf{U}}^1$, for $x\in\bar{\mathsf{U}}^1$ and $i=1$ or 2

$$\left[f+g\alpha^1,\frac{\partial}{\partial x_1}\right](x)\subseteq\Delta^1(x). \tag{3.17}$$

For arbitrary x in $D(\mathsf{U}^1)$, consider a piecewise smooth curve $\gamma(t)$ tangent to Δ^3 joining x to $\pi^1(x)\in\bar{\mathsf{U}}^1$; by (3.16) and the Jacobi identity

$$\begin{aligned}\frac{d}{dt}\left[f+g\alpha^1,\frac{\partial}{\partial x_1}\right](\gamma(t))&\subseteq\left[\frac{\partial}{\partial x_3},\left[f+g\alpha^1,\frac{\partial}{\partial x_1}\right]\right](\gamma(t))\\&=\left[\frac{\partial}{\partial x_i}\left[f+g\alpha^1,\frac{\partial}{\partial x_3}\right]\right](\gamma(t))\\&\subseteq\left[\frac{\partial}{\partial x_1},\Delta^1\right](\gamma(t))\subseteq\Delta^1(\gamma(t));\end{aligned} \tag{3.18}$$

therefore, (3.14a) follows from (3.17) and (3.18). In a similar way (3.14b) is shown.

From definition (3.12) of x_2 and x_3, we see that the Jacobian $\partial y/\partial x_2$ must be of full rank equal to the dimension of x_2. Therefore, the map $(x_0, x_1, x_2) \mapsto (x_0, x_1, y)$ locally is injective and has a left inverse

$$x_2 = x_2(x_0, x_1, y).$$

Suppose this map is well defined on U^1; since $\partial y/\partial x_3 = 0$, it is well defined on $D(\mathrm{U}^1)$. This allows us to eliminate the dependence of α^1 and β^1 on x_2:

$$\alpha^1(x, y) = \alpha^1(x_0, x_1, y) = \alpha^1(x_0, x_1, x_2(x_0, x_1, y)),$$

$$\beta^1(x, y) = \beta^1(x_0, x_1, y) = \beta^1(x_0, x_1, x_2(x_0, x_1, y)).$$

In the above fashion we choose neighborhoods U^i such that $D(\mathrm{U}^i)$ cover M. On each $D(\mathrm{U}^i)$ we construct local feedback functions α^i and β^i which depend on y, leave Δ^1 invariant, and are annihilated by Δ^2.

Let $\pi: M \to M/D$ be the canonical projection and choose a partition of unity $\{\gamma_i\}$ on the M/D subordinate to the neighborhood system $\{\pi(\mathrm{U}^i)\}$. Define global feedback functions $\alpha(x, y)$, $\beta(x, y)$:

$$\alpha(x, y) = \sum_i \gamma_i(x_0, x_1, x_2(x_0, x, y))\alpha^i(x, y)$$

$$\beta(x, y) = \sum_i \gamma_i(x_0, x_1, x_2(x_0, x, y))\beta^i(x, y).$$

Notice that α^i does not depend on x_3 or x_2 (but rather y); hence α and β are annihilated by Δ^2. To see that α and β leave Δ^1 invariant, we compute. Since $\Sigma_i(\gamma_i \circ \pi) = 1$

$$[f + g\alpha, \Delta^i] = [f, \Delta^1] + \sum_i (\gamma_i \circ \pi)[g\alpha^i, \Delta^i] - \sum_i \Delta^1(\gamma_i \circ \pi) g\alpha^i$$

$$= \sum_i (\gamma_i \circ \pi)[f + g\alpha^i, \Delta^1] - \sum_i \Delta^1(\gamma_i \circ \pi) g\alpha^i.$$

From (3.14) the first summation is in Δ^1; as for the second, for simplicity of notation assume there are only two nonzero terms in the partition of unity γ_1 and γ_2 corresponding to neighborhoods U^1 and U^2 of M. Let $\mathrm{U} = \mathrm{U}^1 \cap \mathrm{U}^2$; since $\gamma_1 + \gamma_2 = 1$, for each vector field $\tau \in \Delta^1$, $\tau(\gamma_1 \circ \pi) = -\tau(\gamma_2 \circ \pi)$. Thus, we need to show that on $D(\mathrm{U})$

$$g(x)(\alpha^1 - \alpha^2)(x) \in \Delta^1(x) \tag{3.19}$$

where $\alpha^i(x, y) = \tilde{\alpha}(x_0, x_1, x_2, x_3^i)$ and $x_2 = x_2(x_0, x_1, y)$.

By assumption, Δ^1 is invariant under feedback $\tilde{\alpha}$ so that

$$\left[f + g\tilde{\alpha}, \frac{\partial}{\partial x_3}\right] = \left[f, \frac{\partial}{\partial x_3}\right] + \left[g, \frac{\partial}{\partial x_3}\right]\tilde{\alpha} - g\frac{\partial \tilde{\alpha}}{\partial x_3} \subseteq \Delta^1$$

but from (3.15) this implies

$$g\frac{\partial \tilde{\alpha}}{\partial x_3} \subseteq \Delta^1. \tag{3.20}$$

For $x = (x_0, x_1, x_2, x_3)$ and $x_2 = x_2(x_0, x_1, y)$, consider, as a function of ξ_3, the quantity

$$g(x_0, x_1, x_2, x_3)(\tilde{\alpha}(x_0, x_1, x_2, x_3) - \tilde{\alpha}(x_0, x_1, x_2, \xi_3)).$$

At $\xi_3 = x_3$, this is zero, hence trivially in Δ^1, and hence (3.20) implies, for all $x \in D(\mathrm{U})$ and $\xi_3 = x_3^i$,

$$g(x)(\tilde{\alpha}(x) - \alpha^i(x)) \in \Delta^1(x).$$

The difference of these expressions for $i = 1,2$ yields (3.19). In a similar fashion one can show that

$$[g\beta, \Delta^1] \subseteq \Delta^1.$$

Summing up, we have constructed feedback functions α and β which are annihilated by Δ^2 and which leave Δ^1 invariant. Let

$$\hat{f} = f + g\alpha$$

$$\hat{g} = g\beta.$$

Then

$$[\hat{f}, \Delta^1] \subseteq \Delta^1 \tag{3.21a}$$

$$[\hat{g}, \Delta^1] \subseteq \Delta^1 \tag{3.21b}$$

and in particular, from (3.15),

$$[\hat{f}, \Delta^3] \subseteq \Delta^2 \tag{3.22a}$$

$$[\hat{g}, \Delta^3] \subseteq \Delta^2. \tag{3.22b}$$

From (3.11), (3.21), and (3.22), the system in local coordinates (3.12) becomes

$$\dot{x}_0 = \hat{f}_0(x_0) + \hat{g}_0(x_0)v \tag{3.23a}$$

$$\dot{x}_1 = \hat{f}_1(x_0, x_1, x_2) + \hat{g}_1(x_0, x_1, x_2)v \tag{3.23b}$$

$$\dot{x}_2 = \hat{f}_2(x_0, x_1, x_2, x_3) + \hat{g}_2(x_0, x_1, x_2, x_3)v + p_2(x_0, x_1, x_2, x_3)w \tag{3.23c}$$

$$\dot{x}_3 = \hat{f}_3(x_0, x_1, x_2, x_3) + \hat{g}_3(x_0, x_1, x_2, x_3)v + p_3(x_0, x_1, x_2, x_3)w \tag{3.23d}$$

$$y = h(x_0, x_1, x_2) \tag{3.23e}$$

$$z = k(x_0). \tag{3.23f}$$

Clearly the new feedback functions α and β isolate the noise w from the output z as did the original $\tilde{\alpha}$ and $\tilde{\beta}$. The advantage of the new ones is that they do not depend on x_2 or x_3:

$$\alpha(x, y) = \alpha(x_0, x_1, y) \tag{3.24a}$$

$$\beta(x, y) = \beta(x_0, x_1, y). \tag{3.24b}$$

In (3.23a)–(3.23d), the dependence on y has been eliminated via (3.23e).

The next step is to construct a copy of the dynamics (3.23) on a copy of M. Let $\theta: M \to M^1$ be a diffeomorphism. On M^1, define local coordinates $\xi = (\xi_0, \xi_1, \xi_2, \xi_3)$ corresponding to (3.12) by

$$\xi_i = x_i \circ \theta^{-1}.$$

Let φ and ψ be the vector fields on M^1 corresponding to $\hat{f}$ and $\hat{g}$

$$\varphi=\theta_*(\hat{f}\circ\theta^{-1})$$
$$\psi=\theta_*(\hat{g}\circ\theta^{-1}).$$

On M^1 in local coordinates we have the dynamics

$$\dot{\xi}_0=\varphi_0(\xi_0)+\psi_0(\xi_0)v \tag{3.25a}$$
$$\dot{\xi}_1=\varphi_1(\xi_0,\xi_1,\xi_2)+\psi_1(\xi_0,\xi_1,\xi_2)v \tag{3.25b}$$
$$\dot{\xi}_2=\varphi_2(\xi_0,\xi_1,\xi_2,\xi_3)+\psi_2(\xi_0,\xi_1,\xi_2,\xi_3)v \tag{3.25c}$$
$$\dot{\xi}_3=\varphi_3(\xi_0,\xi_1,\xi_2,\xi_3)+\psi_3(\xi_0,\xi_1,\xi_2,\xi_3)v. \tag{3.25d}$$

Notice we have not included the noise term of (3.23) because w is unknown to us. If (3.23) is initialized at x^0, $x(0)=x^0$, and (3.25) at $\xi^0=\theta(x^0)$, then $\xi_0(t)=x_0(t)$ for all t because $\xi_0^0=x_0^0$. However, the other coordinates will not be equal because of the noise term. We can modify (3.25) by replacing ξ_2 by x_2, given as a function of x_0, x_1 and y:

$$\dot{\xi}_0=\varphi_0(\xi_0)+\psi_0(\xi_0)v \tag{3.26a}$$
$$\dot{\xi}_1=\varphi_1(\xi_0,\xi_1,x_2(x_0,x_1,y))+\psi_1(\xi_0,\xi_1,x_2(x_0,x_1,y))v \tag{3.26b}$$
$$\dot{\xi}_2=\varphi_2(\xi_0,\xi_1,x_2(x_0,x_1,y),\xi_3)+\psi_2(\xi_0,\xi_1,x_2(x_0,x_1,y),\xi_3)v \tag{3.26c}$$
$$\dot{\xi}_3=\varphi_3(\xi_0,\xi_1,x_2(x_0,x_1,y),\xi_3)+\psi_3(\xi_0,\xi_1,x_2(x_0,x_1,y),\xi_3)v. \tag{3.26d}$$

Notice that at any x and ξ satisfying $x_0=\xi_0$ and $x_1=\xi_1$, now $\dot{x}_1=\dot{\xi}_1$, so that $x_1(t)=\xi_1(t)$ for all t.

This allows us to eliminate the feedback dependence on x_0 and x_1:

$$\alpha(\xi,y)=\alpha(\xi_0,\xi_1,y) \tag{3.27a}$$
$$\beta(\xi,y)=\beta(\xi_0,\xi_1,y) \tag{3.27b}$$

and the dependence of ξ_0, ξ_1 on x_0 and x_1:

$$\dot{\xi}_0=\varphi_0(\xi_0)+\psi_0(\xi_0)v \tag{3.28a}$$
$$\begin{aligned}\dot{\xi}_1&=\varphi_1(\xi_0,\xi_1,y)+\psi_1(\xi_0,\xi_1,y)v\\&=\varphi_1(\xi_0,\xi_1,x_2(\xi_0,\xi_1,y))+\psi_1(\xi_0,\xi_1,x_2(\xi_0,\xi_1,y))v\end{aligned} \tag{3.28b}$$

and we have achieved our goal of noise decoupling via dynamic output feedback. The projection π is given in the local coordinates (3.12) by

$$\pi(x_0,x_1,x_2,x_3)=(x_0,x_1).$$

Notice that dynamic compensator (3.28) evolves on a copy of the quotient manifold M/Δ^2. It is not necessarily linear in y. In other words, in order to solve the disturbance decoupling problem for a system (3.1) which is linear in the control, it may be necessary to resort to a dynamic compensator (3.9) where the control enters nonlinearly. ■

Proof of Theorem 3.7: By assumption there exists an auxiliary system (3.9), feedback functions $\alpha(\xi,y)$, $\beta(\xi,y)$, and a projection π given locally by (3.4) such that the feedback modified dynamics are given locally by (3.2). Define a distribution $\Delta=\ker d\pi$, in the local coordinates of (3.2)

$$\Delta=\operatorname{span}\left\{\frac{\partial}{\partial x_2}\right\}.$$

From (3.2) we see immediately that

$$[\tilde{f},\Delta]\subseteq\Delta \tag{3.29a}$$
$$[\tilde{g},\Delta]\subseteq\Delta \tag{3.29b}$$

or

$$[f,\Delta]+[g,\Delta]\alpha+g\Delta(\alpha)\subseteq\Delta \tag{3.30a}$$
$$[g,\Delta]\beta+g\Delta(\beta)\subseteq\Delta. \tag{3.30b}$$

Since β is assumed to be invertible this implies that (2.19) holds.

Since $\Delta^0(\alpha)=0$, $\Delta^0(\beta)=0$, from (3.30) we have

$$[f,\Delta\cap\Delta^0]+[g,\Delta\cap\Delta^0]\alpha\subseteq\Delta \tag{3.31a}$$
$$[g,\Delta\cap\Delta^0]\beta\subseteq\Delta \tag{3.31b}$$

and the invertibility of β implies 2). The last conclusion 3) follows immediately from (3.2). ■

IV. Construction of Invariant Distributions

In this section we discuss the problem of finding (f,g) and (h,f) invariant distributions and also the relationship between invariant distributions and nonlinear controllability and observability. Before we do this we must develop some more mathematical background.

A *codistribution* θ (or Pfaffian system) on M is a mapping which assigns to each $x\in M$ a subspace $\theta(x)$ of the cotangent space T_x^*M in a smooth (analytic) fashion. Recall T_x^*M is the space of linear functionals on T_xM. If one considers a tangent vector $\tau=\operatorname{col}(\tau_1,\cdots,\tau_n)$ as a column vector, then a contangent vector $\omega=(\omega_1,\cdots,\omega_n)$ is a row vector and the pairing $\langle\omega,\tau\rangle$ is given by

$$\omega(\tau)=\langle\omega,\tau\rangle=\omega\tau=\sum\omega_i\tau_i.$$

Tangent vectors are viewed as directional differentiation

$$\tau=\sum\tau_i\frac{\partial}{\partial x_i}$$

cotangent vectors are viewed as gradients (evaluated at a point)

$$\omega=\sum\omega_i\,dx_i.$$

Corresponding to a vector field $\tau(x)$ is a *one form* $\omega(x)$. If f and τ are vector fields on M, the *Lie derivative* $L_f(\tau)$ of τ by f is another vector field given by the Lie bracket

$$L_f(\tau)(x)=[f,\tau](x)=\frac{\partial\tau}{\partial x}(x)f(x)-\frac{\partial f}{\partial x}(x)\tau(x). \tag{4.1}$$

In a similar fashion we define the Lie derivative $L_f\omega$ of a one form ω by f as the one form given by

$$L_f\omega(x)=\left(\frac{\partial\omega^t}{\partial x}(x)f(x)\right)^t+\omega(x)\frac{\partial f}{\partial x}(x) \tag{4.2}$$

where superscript t denotes transpose. We have already discussed the Lie derivative of functions, if φ is a function on M then

$$L_f(\varphi)=\frac{\partial\varphi}{\partial x}(x)f(x). \tag{4.3}$$

The pairing $\langle\omega,\tau\rangle(x)$ of a one form $\omega(x)$ and vector field $\tau(x)$ is a function and it is straightforward to verify that

$$L_f\langle\omega,\tau\rangle(x)=\langle\omega,L_f\tau\rangle(x)+\langle L_f\omega,\tau\rangle(x). \tag{4.4}$$

Finally we note that if φ is a function then the gradient $d\varphi$ is a one form and

$$L_f(d\varphi)=d(L_f\varphi). \tag{4.5}$$

Given any distribution Δ on M one can define a codistribution $\Delta^\perp$ as the annihilator of Δ:

$$\Delta^\perp(x)=\{\omega\in T_x^*M:\ \langle\omega,\Delta(x)\rangle=0\}. \tag{4.6}$$

In a similar fashion we define the annihilator $\theta^\perp$ of a codistribution θ:

$$\theta^\perp(x)=\{\tau\in T_xM:\ \langle\theta(x),\tau\rangle=0\}. \tag{4.7}$$

Of course $\Delta^{\perp\perp}=\Delta$, $\theta^{\perp\perp}=\theta$, and $\Delta^1\subseteq\Delta^2$ iff $\Delta^{1\perp}\supseteq\Delta^{2\perp}$.

As before we confuse a codistribution θ with the space of all one forms which are pointwise in it. A codistribution θ is *regular* if $\theta^\perp$ is regular. Alternately, θ is regular iff for every x there exists a neighborhood U and functions $\varphi_1,\cdots,\varphi_k$ whose gradients $d\varphi_1,\cdots,d\varphi_k$ are linearly independent on U and span θ. The integral manifolds of $\theta^\perp$ are the intersection of the level sets of $\varphi_1,\cdots,\varphi_k$.

A codistribution θ is *invariant* under the dynamics (1.1) if

$$L_f(\theta)\subseteq\theta \tag{4.8a}$$

$$L_g(\theta)\subseteq\theta. \tag{4.8b}$$

Using (4.4) it is easy to verify that θ is *invariant* under the dynamics (1.1) iff $\Delta=\theta^\perp$ is invariant. The following lemma is straightforward and gives an alternate characterization of local (f,g) and (h,f) invariance.

Lemma 4.1: $\Delta^1\subseteq\Delta^2$ and $[f,\Delta^1]\subseteq\Delta^2$, $[g,\Delta^1]\subseteq\Delta^2$ iff $\Delta^{2\perp}\subseteq\Delta^{1\perp}$, and $L_f(\Delta^{2\perp})\subseteq\Delta^{1\perp}$, $L_g(\Delta^{2\perp})\subseteq\Delta^{1\perp}$.

Corollary 4.2: Δ is locally (f,g) invariant on U iff on U

$$L_f\left(\Delta^\perp\cap(\text{span}\,g(x))^\perp\right)\subseteq\Delta^\perp$$

$$L_g\left(\Delta^\perp\cap(\text{span}\,g(x))^\perp\right)\subseteq\Delta^\perp.$$

Corollary 4.3: Δ is (h,f) invariant iff

$$L_f(\Delta^\perp)\subseteq\Delta^\perp+\text{span}\,dh$$

$$L_g(\Delta^\perp)\subseteq\Delta^\perp+\text{span}\,dh.$$

Let us take a look at the connection between invariant distributions/codistributions and nonlinear controllability and observability. Recall the following definitions from [6]. A state x^0 is indistinguishable from x^1 (denoted x^0Ix^1) if for any admissible input the outputs starting at the two initial conditions x^0 and x^1 are identical. A state x^0 is strongly indistinguishable from x^1 (denoted x^0SIx^1) if there exists a piecewise smooth curve $x(t)$ joining x^0 to x^1 such that $x^0Ix(t)$. A state x^1 is accessible from x^0 (denoted by x^1Ax^0) if there exists a trajectory of the system going from x^0 to x^1 in the positive time direction. A state x^1 is weakly accessible from x^0 (denoted by x^1WAx^0) if x^1 can be reached from x^0 by following a union of trajectories of the system traversed in either time direction. We then have the following theorems [6].

Theorem 4.4: Let θ^0 be the codistribution spanned by dh and let θ^1 be the minimal codistribution invariant under the dynamics (1.1) which contains θ^0. Assume θ^1 is of constant rank and let $\Delta=\theta^{1\perp}$. Δ is then a regular distribution on M and if M' is the maximal integral manifold of Δ through x^0 then

$$M'=\{x:\ xSIx^0\}.$$

Theorem 4.5 (Frobenius and Chow): Let $\Delta^0=\text{span}\{f,g\}$ and let Δ be the minimal distribution invariant under the dynamics (1.1) which contains Δ^0. Assume Δ is of constant rank (or the system is analytic). Let M' be the maximal integral submanifold of Δ through x^0; then

$$M'=\{x:\ xWAx^0\}.$$

Next we turn to the question of maximal (f,g) and minimal (h,f) invariant distributions. From the definition (2.21) it is clear that the intersection of (h,f) invariant distributions is again (h,f) invariant and therefore there always exists a minimal (h,f) invariant distribution containing a given distribution.

Motivated by Theorem 3.6, we consider the problem of finding the minimal (h,f) invariant involutive distribution containing the distribution $\Delta^1(x)=\text{span}\,p(x)$. We define a nondecreasing sequence of distributions

$$\Delta^{j+1}(x)=\bar\Delta^j(x)+[f,\Delta^0\cap\bar\Delta^j](x)+[g,\Delta^0\cap\bar\Delta^j](x) \tag{4.9}$$

whose involutive limit is

$$\Delta^*(x)=\bigcup_{j\geq1}\bar\Delta^j(x). \tag{4.10}$$

Suppose Δ is an involutive distribution containing Δ^1 which is (h,f) invariant, then Δ contains $\bar\Delta^1$. From the definition (2.21), Δ contains the right side of (4.9) for $j=1$; hence Δ contains Δ^2, and so on. Hence, $\Delta^*\subseteq\Delta$ and is the desired distribution.

By construction, Δ^* is involutive but it need not be regular or even of constant rank.

If the system is analytic, Δ^* is of constant rank on an open dense subset U of M and is given by

$$\Delta^*(x)=\bar{\Delta}^n(x).$$

To see this, let r^j be the maximal dimension of $\bar{\Delta}^j(x)$ as x varies over M. Since $\bar{\Delta}^j \subseteq \bar{\Delta}^{j+1} \subseteq TM$, the r^j's are monotone nondecreasing and bounded above by n = dimension M. For analytic systems there exists an open dense subset U^j of M on which the dimension of $\bar{\Delta}^j(x)=r^j$. If $r^j=r^{j+1}$, $\bar{\Delta}^j=\bar{\Delta}^{j+1}$ on an open dense subset $U=U^j$. But then $\bar{\Delta}^j=\bar{\Delta}^{j+i}$ on U and $r^j=r^{j+i}$ for all $i\geq 0$. Therefore, r^j must reach its maximal value for some $j\leq n$.

From the definition (2.17) of (f,g) invariant distributions, it does not follow that the sum of two (f,g) invariant distributions is again (f,g) invariant, and thus it is not clear whether maximal elements exist. However, the sum of two locally (f,g) invariant distributions is again locally (f,g) invariant, thus we can look for maximal elements there. Taking our cue from Theorem 3.6, we seek a maximal locally (f,g) invariant distribution contained in the distribution $\Delta^1=\ker dk$. It is more convenient to work with the dual characterization of local (f,g) invariance as provided by Corollary 4.2. We define the codistribution $\theta^1=\Delta^{1\perp}=\operatorname{span} dk$ and inductively

$$\theta^{j+1}(x)=\theta^j(x)+L_f\big(\theta^j\cap(\operatorname{span} g)^\perp\big)(x) + L_g\big(\theta^j\cap(\operatorname{span} g)^\perp\big)(x). \tag{4.11}$$

This yields a nondecreasing sequence of codistribution, the limit of which

$$\theta^*(x)=\bigcup_{j\geq 1}\theta^j(x) \tag{4.12}$$

is easily seen to be the minimal codistribution containing span dk which satisfies

$$L_f\big(\theta\cap(\operatorname{span} g)^\perp\big)\subseteq\theta$$
$$L_g\big(\theta\cap(\operatorname{span} g)^\perp\big)\subseteq\theta.$$

Hence, by Corollary 4.2, $\Delta^*=(\theta^*)^\perp$ is the maximal locally (f,g) invariant distribution contained in $\ker dk$. By construction, Δ^* must be involutive or else its involutive closure would be a larger locally (f,g) invariant distribution contained in $\ker dk$.

For analytic systems, an argument similar to the above shows that on an open dense subset of M

$$\theta^*(x)=\theta^n(x).$$

Sometimes the maximal (f,g) invariant distribution contained in $\ker dk$ and the appropriate feedback functions α and β can be explicitly computed. For each component k_i of the output k, let ρ_i denote the largest integer such that for all $r<\rho_i$ and all $x\in M$

$$L_gL_f^rk_i=0.$$

If no such integer exists, then $\rho_i=\infty$. Assuming $\rho_i<\infty$, let $A(x)$ denote the matrix whose i,j entry, $a_{ij}(x)$ is

$$a_{ij}(x)=L_{g_j}L_f^{\rho_i}k_i(x)$$

and $b(x)$, the column whose ith entry $b_i(x)$ is

$$b_i(x)=L_f^{\rho_i+1}k_i(x).$$

Theorem 4.6: Suppose $\rho_i<\infty$ for every i and the rank of $A(x)$ is equal to the dimension of $z=k(x)$ for every $x\in M$, the maximal locally (f,g) invariant distribution contained in $\ker dk$ is

$$\Delta^*(x)=\bigcap_i\bigcap_{r\leq\rho_i}\ker dL_f^rk_i(x) \tag{4.13}$$

and the feedback functions $\alpha(x)$ and $\beta(x)$ which leave this distribution invariant are any solutions of

$$A(x)\alpha(x)=-b(x) \tag{4.14a}$$
$$A(x)\beta(x)=C\ \text{(constant matrix)}. \tag{4.14b}$$

Remark: One seeks an invertible solution to (4.14b) so that one has full control. If the dimension of Δ^* is constant, then Δ^* is regular since its integral manifolds are the intersection of the level sets of the functions $L_f^rk_i(x)$ for $r\leq\rho_i$. If the rank condition holds on an open subset U of M, then the conclusion holds on U.

Proof: If the rank condition holds, then $A(x)$ has no rows which are all zeros, or in other words, for every x and i there exists j such that $L_{g_j}L_f^{\rho_i}k_i(x)\neq 0$. This allows us to explicitly compute the sequence of codistributions θ^j defined by (4.11)

$$\theta^1(x)=\operatorname{span}\{dk_i(x)\}.$$

Now, using (4.5),

$$L_f\big(\theta^1\cap(\operatorname{span} g)^\perp\big)(x)=\operatorname*{span}_{\rho_i\geq 1}\{dL_fk_i(x)\}$$
$$L_g\big(\theta^1\cap(\operatorname{span} g)^\perp\big)(x)=\operatorname*{span}_{\rho_i\geq 1}\{dL_gk_i(x)\}=0$$

so

$$\theta^2(x)=\theta^1(x)+\operatorname*{span}_{\rho_i\geq 1}\{dL_fk_i(x)\}.$$

In general,

$$\theta^{j+1}(x)=\theta^j(x)+\operatorname*{span}_{\rho_i\geq j}\{dL_f^jk_i(x)\}.$$

Therefore, $\Delta^*(x)$ [see (4.13)] is the annihilator of $\theta^*(x)$ [see (4.12)] and hence is the maximal locally (f,g) invariant distribution contained in dk.

Let α and β be as above [see (4.14)] and define $\tilde{f}$ and $\tilde{g}$ by (2.18). For any r and k_i,

$$L_{\tilde f}\big(dL_f^rk_i\big)=d\big(L_f^{r+1}k_i+\big(L_gL_f^rk_i\big)\alpha\big) \tag{4.15a}$$
$$L_{\tilde g}\big(dL_f^rk_i\big)=d\big(\big(L_gL_f^rk_i\big)\beta\big). \tag{4.15b}$$

If $r<\rho_i$, (4.15b) is zero and (4.15a) reduces to

$$dL_f^{r+1}k_i$$

which is in θ^{r+2}. If $r=\rho_i$, (4.14) implies (4.15) is zero; thus we have

$$L_{\tilde f}(\theta^j)\subseteq\theta^{j+1} \tag{4.16a}$$

$$L_{\tilde g}(\theta^j)\subseteq\theta^{j+1}. \tag{4.16b}$$

This implies that θ^* as defined by (4.12) is invariant (4.8) under $\tilde f$ and $\tilde g$, hence $\Delta^*=(\theta^*)^\perp$ is (f,g) invariant. Moreover, Δ^* must be the maximal (f,g) invariant distribution contained in $\ker dk$. ■

The above discussion indicates that maximal (locally) (f,g) invariant and minimal (h,f) invariant distributions are not terribly difficult to find, whereas the problem of finding arbitrary (f,g) or (h,f) invariant distributions might be quite difficult. This is also true in the case of linear systems as has been emphasized by Morse and Wonham [10]. Finding maximal (A,B) or minimal (C,A) invariant subspaces is essentially a linear problem, i.e., it can be solved using linear algebra, whereas finding arbitrary (A,B) or (C,A) invariant subspaces involves solving nonlinear equations.

The above remark emphasizes the importance of the theorems of Basile, Laschi, Marro, and Schumacker. Instead of having to overcome the nonlinear problem of finding a (C,A,B) invariant subspace in order to solve the linear, static, output feedback, noise decoupling problem, one can solve two linear problems; finding maximal (A,B) and minimal (C,A) invariant subspaces and apply Schumacker's result to decouple by linear dynamic, output feedback. Similar considerations hold in the nonlinear case, i.e., Theorem 3.6 is much easier to verify than Theorem 3.3.

V. Noninteracting Control

We now consider the problem of using feedback to transform a system of the form (1.1) into a new system in which each input controls a single output without influencing the others. The solution of this problem has been widely investigated for linear systems, e.g., [10], and some interesting results have also been obtained for nonlinear systems [11], [5], [13]. We consider only the *static, state feedback noninteracting control problem*. We seek a control law of the form

$$u=\alpha(x)+\beta(x)v$$

such that in suitable local coordinates, (1.1) becomes

$$\begin{aligned}\dot x_1&=\tilde f_1(x_1)+\tilde g_1(x_1)v_1\\ &=\vdots\\ \dot x_m&=\tilde f_m(x_m)+\tilde g_m(x_m)v_m\\ \dot x_{m+1}&=\tilde f_{m+1}(x)+\tilde g_{m+1}(x)v\end{aligned} \tag{5.1a}$$

$$\begin{aligned}y_1&=h_1(x_1)\\ &=\vdots\\ y_m&=h_m(x_m)\end{aligned} \tag{5.1b}$$

where $x=(x_1,\cdots,x_{m+1})$, $v=(v_1,\cdots,v_m)$, and $y=(y_1,\cdots,y_m)$, with each x_i, v_i, and y_i being possibly a vector.

In global terms we seek projections $\pi_1,\cdots,\pi_m$; $\pi_i\colon M\to M^i$ given in local coordinates by $\pi_i(x)=x_i$ such that π_i carries (5.1) onto a system on M^i with the same input v_i/output y_i behavior. The system on M^i is given in local coordinates corresponding to (5.1) by

$$\dot x_i=\tilde f_i(x_i)+\tilde g_i(x_i)v_i \tag{5.2a}$$

$$y_i=h_i(x_i). \tag{5.2b}$$

If such α, β, and $\pi_1,\cdots,\pi_m$ exist, we say that the *static, state feedback noninteracting control problem is solvable in a regular fashion*. Of course it is desirable to have as much control over (5.1) as over the original system. The problem is solvable *with full control* if $\beta(x)$ is invertible.

A family of (f,g) invariant distributions $\Delta_1,\cdots,\Delta_m$ are *compatible* if there exists feedback functions α and β which leave each one invariant

$$[f+g\alpha,\Delta_i]\subseteq\Delta_i \tag{5.3a}$$

$$[g\beta,\Delta_i]\subseteq\Delta_i. \tag{5.3b}$$

Theorem 5.1: The static, state feedback, noninteracting control problem is solvable in a regular fashion (with full control) iff there exists a family $\Delta_1,\cdots,\Delta_m$ of compatible (f,g) invariant distributions (with full control) such that

1) each Δ_i is regular
2) $\tilde g_j\subseteq\Delta_i\subseteq\ker dh_i$, $\forall i\neq j$
3) if I and J are any disjoint nonempty subsets of $\{1,\cdots,m\}$, then

$$\Big(\bigcap_{i\in I}\Delta_i\Big)+\Big(\bigcap_{j\in J}\Delta_j\Big)=TM.$$

Proof: Suppose the static, state feedback, noninteracting control problem is solvable in a regular fashion (with full control), then there exists α, β, and $\pi_1,\cdots,\pi_m$ such that in suitable local coordinates, (5.1) is valid. Let $\Delta_i=\ker d\pi_i$. It is straightforward to verify that $\Delta_1,\cdots,\Delta_m$ are compatible (f,g) invariant distributions (with full control) satisfying 1)–3).

As for the converse, suppose $\Delta_1,\cdots,\Delta_m$ are compatible (f,g) invariant distributions (with full control) satisfying 1)–3), then α and β are given and we define π_i as the canonical projection $\pi_i\colon M\to M/\Delta_i$. Define codistributions $\theta_i=\Delta_i^\perp$; since Δ_i is regular, there exists a vector of functions $\xi_i(x)$ such that locally the vector of gradients $d\xi_i(x)$ is a basis for $\theta_i(x)$:

$$\theta_i(x)=\operatorname{span} d\xi_i(x)$$

$$\Delta_i(x)=\ker d\xi_i(x).$$

By 3), if I and J are disjoint nonempty subsets of $\{1,\cdots,m\}$,

$$\left(\sum_{i\in I}\theta_i\right)\cap\left(\sum_{j\in J}\theta_j\right)=0.$$

This implies that gradients

$$d\xi_1(x),\cdots,d\xi_m(x)$$

are linearly independent and hence we can choose a vector of functions $\xi_{m+1}(x)$ such that the gradients

$$d\xi_1(x),\cdots,d\xi_m(x),d\xi_{m+1}(x)$$

are locally a linearly independent set of dimension $n=$ dimension M. Therefore, the map $x\rightarrow\xi(x)$ defines a local change of coordinates. Using 2) and the fact that each θ_i is invariant [see (4.8)] under $f,\tilde{g}$ dynamics, it is easy to see that in the ξ-coordinates the system is of the form (5.1). ■

Next let us consider the case when the dimension of u and y is exactly m so that each of the noninteracting components (5.2) of (5.1) is a scalar input/scalar output system. In this case, we have an alternate formulation. Let ρ_i be defined as before; the largest integer such that for all $r<\rho_i$ and all $x\in M$

$$L_gL_f^rh_i(x)=0.$$

Theorem 5.2: Let the dimensions of u and y be the same. The static, state feedback, noninteracting scalar input/scalar output control problem is solvable in a regular fashion if

a) the $m\times m$ matrix $A(x)$ defined by

$$a_{ij}(x)=L_{g_j}L_f^{\rho_i}h_i(x)$$

is nonsingular for every x

b) for each i the dimension of the codistribution $\theta_i(x)$ defined by

$$\theta_i(x)=\text{span}\left\{dh_i(x),dL_fh_i(x),\cdots,dL_f^{\rho_i}h_i(x)\right\}$$

is constant for all x

c) for any disjoint nonempty subsets I and J of $\{1,\cdots,m\}$

$$\left(\sum_{i\in I}\theta_i\right)\cap\left(\sum_{j\in J}\theta_j\right)=\phi.$$

Proof: We wish to show that a)–c) imply the existence of compatible (f,g) invariant distribution $\Delta_1,\cdots,\Delta_m$ with full control satisfying 1)–3). Define

$$\Delta_i=\theta_i^{\perp}.$$

By applying Theorem 4.6 with $k=h_i$, we have that Δ_i is the maximal (f,g) invariant distribution contained in ker dh_i and is invariant under any feedback functions $\alpha_i(x)$, $\beta_i(x)$ satisfying

$$L_gL_f^{\rho_i}h_i(x)\alpha_i(x)=L_f^{\rho_i+1}h_i(x)$$

$$L_gL_f^{\rho_i}h_i(x)\beta_i(x)=\text{constant matrix.}$$

Therefore, since $A(x)$ is invertible, there exists an m vector $\alpha(x)$ and an invertible $m\times m$ matrix $\beta(x)$ such that

$$A(x)\alpha(x)=\text{col}\left(L_f^{\rho_1+1}h_1(x),\cdots,L_f^{\rho_m+1}h_m(x)\right) \tag{5.4a}$$

$$A(x)\beta(x)=I(m\times m\text{ identity matrix}). \tag{5.4b}$$

Using these feedback functions, $\Delta_1,\cdots,\Delta_m$ are compatible (f,g) invariant distributions with full control.

Since θ_i is spanned by the gradients of functions and is of constant dimension, its annihilator Δ_i is regular, thus 1) is satisfied.

As for 2), we have already noted that $\Delta_i\subseteq\ker dh_i$; to see that $\tilde{g}_j\in\Delta_i$ for $i\neq j$ we show that $\tilde{g}_j$ annihilates θ_i. From the definition of ρ_i, for every $r<\rho_i$

$$L_gL_f^rh_i=0$$

thus,

$$\langle dL_f^rh_i,\tilde{g}_j\rangle=0.$$

For ρ_i,

$$\langle dL_f^{\rho_i}h_i,\tilde{g}_j\rangle(x)=\sum_r\left(L_{g_r}L_f^{\rho_i}h_i(x)\right)\beta_{rj}(x)$$

which is zero for $i\neq j$ by (5.4b).

Finally 3) follows immediately from c). ■

A few remarks about the scalar input/scalar output noninteracting control problem are in order. Notice that Theorems 5.1 and 5.2 are not equivalent for this problem because Theorem 5.2 deals with the maximal (f,g) invariant distributions contained in ker dh_i, whereas theorem 5.1 deals with any (f,g) invariant distributions.

It is also interesting to note that if Theorem 5.2 is satisfied, then for each input v_i/output y_i channel the condition

$$L_{\tilde{g}_i}L_f^{\rho_i}h_i(x)=1\neq0$$

is satisfied and hence each channel is strongly invertible in the sense of Hirschorn [7].

Suppose for each i and some x

$$dh_i(x),\cdots,dL_f^{\rho_i}h_i(x)$$

form a basis for $\theta_i(x)$, then in a neighborhood this is true. Following the construction of Theorem 5.1, we define

$$\xi_i(x)=\left(h_i(x),\cdots,L_f^{\rho_i}h_i(x)\right).$$

In the local coordinates $\xi(x)=(\xi_1(x),\cdots,\xi_m(x),\xi_{m+1}(x))$, the system becomes almost linear, i.e., each input/output channel (5.2) is of the form

$$\dot{\xi}_i=A_i\xi_i+B_iv_i$$

$$y_i=C_i\xi_i$$

where A_i is $(\rho_i+1)\times(\rho_i+1)$, B_i is $(\rho_i+1)\times1$, and C_i is

$1 \times (\rho_i + 1)$:

$$A_i = \begin{pmatrix} 0 & 1 & & 0 \\ & & \ddots & \\ & & & 1 \\ 0 & & & 0 \end{pmatrix}, \qquad b_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

$$c_i = (1 \quad 0 \quad \cdots \quad 0).$$

The rest of the system, which is unobservable, is nonlinear:

$$\dot{\xi}_{m+1} = \tilde{f}_{m+1}(\xi) + \tilde{g}_{m+1}(\xi)v.$$

This is related to work of Brockett [3], on when a nonlinear system may be modified by feedback so as to be locally diffeomorphic to a linear system.

Finally we note that the algebraic condition a) generalizes the well-known Falb and Wolovich [4] condition for the decoupling of linear systems and has already been used by Freund [5] and Sinha [13] in nonlinear decoupling problem.

ACKNOWLEDGMENT

The authors would like to thank M. Hautus for his helpful discussions during the development to this work.

REFERENCES

[1] G. Basile and G. Marro, "Controlled and conditioned invariant subspaces in linear system theory," *J. Optimiz. Theory Appl.*, vol. 3, no. 5, pp. 306–315, 1969.
[2] W. M. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry.* New York: Academic, 1975.
[3] R. W. Brockett, "Feedback invariants for nonlinear systems," *Preprints of 7th IFAC World Congress*, Helsinki, June 1978.
[4] P. L. Falb and W. A. Wolovich, "Decoupling in the design and synthesis of multivariable control systems," *IEEE Trans. Automat. Contr.*, vol. AC-12, pp. 651–659, Dec. 1967.
[5] E. Freund, "The structure of decoupled non-linear systems," *Int. J. Contr.*, vol. 21, no. 3, pp. 443–450, 1975.
[6] R. Hermann and A. J. Krener, "Nonlinear controllability and observability," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 728–740, Oct. 1977.
[7] R. M. Hirschorn, "Invertibility of nonlinear control systems," *SIAM J. Contr. Optim.*, vol. 17, no. 2, pp. 289–297, 1979.
[8] A. J. Krener, "A decomposition theory for differentiable systems," *SIAM J. Contr. Optim.*, vol. 15, no. 5, pp. 813–829, Aug. 1977.
[9] S. H. Mikhail and W. H. Wonham, "Local decomposability and the disturbance decoupling problem in nonlinear autonomous systems," in *Proc. 16th Allerton Conf. Commun. Contr. Comput.*, Oct. 1978, pp. 664–669.
[10] A. S. Morse and W. H. Wonham, "Status of noninteracting control," *IEEE Trans. Automat. Contr.*, vol. AC-16, no. 6, pp. 568–581, Dec. 1971.
[11] W. A. Porter, "Decoupling of and inverses for time-varying linear systems," *IEEE Trans. Automat. Contr.*, vol. AC-14, pp. 378–380, Aug. 1969.
[12] J. M. Schumacker, "(C,A) invariant subspaces by some fact and uses," Vrije Universiteit Amsterdam, Rep. 110.
[13] P. K. Sinha, "State feedback decoupling of nonlinear systems," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 487–489, June 1977.
[14] H. J. Sussmann and V. J. Jurdjevic, "Controllability of nonlinear systems," *J. Differ. Equat.*, vol. 12, pp. 95–116, 1972.
[15] W. M. Wonham and A. S. Morse, "Decoupling and pole assignment in linear multivariable systems: A geometric approach," *SIAM J. Contr.* vol. 8, pp. 1–18, Feb. 1970.
[16] R. M. Hirschorn, "(A, B)-invariant distributions and the disturbance decoupling of nonlinear systems," to be published.
[17] A. J. Krener, "A generalization of Chow's theorem and the bang-bang theorem to nonlinear control problems," *SIAM J. Contr.*, vol. 12, 1974.
[18] R. Laschi and G. Marro, "Alcune considerazioni sull'osservabilità dei sistemi dinamici con ingressi inaccessibili," in *Proc. 70th Ann. Meeting Elec. Eng. Assoc. Italy*, 1969, Paper 1.1.06.
[19] G. Basile and G. Marro, "L'invarianza rispetto ai disturbi studiata nello spazio degli stati," in *Proc. 70th Ann. Meeting Elec. Eng. Assoc. Italy*, 1969, Paper 1.4.01.

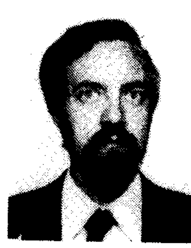
Alberto Isidori (M'80) was born in Rapallo, Genoa, on January 24, 1942. He received the doctor degree in electrical engineering from the University of Rome, Rome, Italy, in 1965.

Since 1970 he has been at the Istituto di Automatica, University of Rome, where since 1975 he has been a Professor of Control Theory. He has held visiting positions at the University of Florida, Gainesville, and Washington University, St. Louis, MO. His current research interests include nonlinear systems and control theory and realization theory.

Dr. Isidori is an Associate Editor of the journals *Systems and Control Letters* and *Ricerche di Automatica*.

Arthur J. Krener (M'77) was born in Brooklyn, NY, on October 8, 1942. He received the B.S. degree from Holy Cross College, Worcester, MA, in 1964, and the M.A. and Ph.D. degrees from the University of California, Berkeley, in 1967 and 1971, all in mathematics.

Since 1971 he has been at the University of California, Davis, where he is currently Professor of Mathematics. He has held visiting positions at Harvard University, Cambridge, MA, the University of Rome, Rome, Italy, and Imperial College, London, England. His current interests are nonlinear systems theory, stochastic processes, and acausal linear systems.



Claudio Gori-Giorgi received the doctorate degree in nuclear engineering in 1969 and the postgraduate Diploma in automatic control in 1971, both from the University of Rome, Rome, Italy.

In 1970 he joined the Istituto di Automatica of the University of Rome. Since 1972 he has been a Research Consultant of the National Research Council of Italy. In 1973 he was appointed a part-time Professor of Automatic Control Theory at the University of Firenze, Italy. His major research interest lies in mathematical system theory and its applications.

Salvatore Monaco was born in Udine, Italy, on January 17, 1951. He received the B.S. degree in electrical engineering from the University of Rome, Rome, Italy, in 1974.

Since 1977 he has been an Assistant Professor of System Theory at the University of Rome. Since 1978 he has been a part-time Professor of Control Theory at the University of L'Aquila. His present research interests are in the areas of nonlinear systems, modeling, and infinite dimensional systems.