BILINEAR AND NONLINEAR REALIZATIONS OF INPUT-OUTPUT MAPS*

ARTHUR J. KRENER†

Abstract. Given a nonlinear realization of an input-output map, sufficient conditions are given for the existence of an equivalent bilinear realization for small t. It is also shown that every nonlinear realization can be approximated by a bilinear realization, with an error that grows like an arbitrary power of t.

1. Introduction. In recent years there has been considerable interest in bilinear control systems. This interest can be attributed to the fact that this class of systems is general enough to model many physical and biological processes and at the same time, it is specific enough to support a rich mathematical structure [1], [2], [3], [4]. We would like to propose another reason for considering such systems, namely, that in a sense to be made precise later, every nonlinear system with control entering linearly is locally almost bilinear.

Given an input-output map, $u(t) \mapsto w(t)$, a bilinear realization of this is

(1.1)
$$\dot{x}(t) = \left(A_0 + \sum_{i=1}^h u_i(t)A_i\right)x(t),$$

$$w(t) = Cx(t),$$

$$x(0) = x^0, \quad u(t) \in \Omega,$$

where $x = (x_1, \dots, x_m)$, $w = (w_1, \dots, w_l)$. A_0, \dots, A_h are $m \times m$ matrices, C is an $l \times m$ matrix and $u(t) = (u_1(t), \dots, u_h(t))$ is a measurable control with values in $\Omega = \{u: |u_i| \le 1, i = 1, \dots, h\}$. The differential equation and map $x \mapsto w$ are called the *dynamics* and the *output map* of the realization and in this case are bilinear and linear respectively.

A nonlinear realization of an input-output map, $u(t) \mapsto z(t)$, is

$$\dot{y}(t) = b_0(y) + \sum_{i=1}^h u_i(t)b_i(y),$$

$$z(t) = f(y(t)),$$

$$y(0) = y^0, \qquad u(t) \in \Omega,$$

where $y=(y_1,\cdots,y_n),\ z=(z_1,\cdots,z_l),\ b_0(y),\cdots,b_h(y)$ are C^∞ *n*-dimensional vector fields, f is a C^∞ \mathbb{R}^l -valued function and $u(t)=(u_1(t),\cdots,u_n(t))$ a measurable control with values in $\Omega=\{u:|u_i|\leq 1,\ i=1,\cdots,h\}$. Here both the *dynamics* and output map, z=f(y), are *nonlinear*.

In Theorem 1, a necessary and sufficient condition is given for there to exist a change of state which bilinearizes the dynamics of (1.2) for small t. As a corollary using a technique of Brockett, we obtain sufficient conditions for the existence of a

^{*} Received by the editors March 7, 1974, and in revised form June 17, 1974.

[†] Department of Mathematics, University of California-Davis, Davis, California 95616.

bilinear realization, (1.1), such that every input, u(t), gives the same output, w(t) = z(t), for small t.

Theorem 2 shows that for any integer, $\mu \ge 0$, there exists a system with bilinear dynamics which approximates the dynamics of (1.2) with error $O(t^{\mu+1})$. As a corollary there exists for every nonlinear realization, (1.2), a bilinear realization, (1.1), such that every input gives approximately the same output, $w(t) = z(t) + O(t^{\mu+1})$, for small t.

2. Preliminaries. Instead of considering (1.1), it is useful to consider the matrix bilinear system

(2.1)
$$\dot{X}(t) = \left(A_0 + \sum_{i=1}^h u_i A_i\right) X(t),$$

$$W(t) = CX(t),$$

$$X(0) = I, \qquad u(t) \in \Omega,$$

where X(t) takes values in the group, $Gl(m, \mathbb{R})$, of all invertible $m \times m$ matrices.

Each column of the matrix equation (2.1) is a system of the form (1.1). Therefore instead of considering the problems of replacing or approximating (1.2) by (1.1), we study the equivalent problem of replacing or approximating (1.2) by (2.1).

The advantage of considering (2.1) over (1.1) is that $Gl(m, \mathbb{R})$ is a Lie group and each A_j defines a right invariant vector field, A_jX , on this group, hence a member of the associated Lie algebra, $gl(m, \mathbb{R})$, of all $m \times m$ real matrices. This algebra is finite-dimensional over the field, \mathbb{R} , and the multiplication is defined by the Lie bracket

$$[A_i, A_j] = A_j A_i - A_i A_j.$$

This is a noncommutative and nonassociative operation which instead satisfies the skew-symmetry and Jacobi relations,

$$[A_i, A_j] = -[A_j, A_i],$$

and

$$[A_i[A_j, A_k]] = [[A_i, A_j]A_k] + [A_j[A_i, A_k]].$$

For further discussion of Lie groups and algebras we refer the reader to [5], [6], [7].

There is a unique subalgebra, g, of $gl(m, \mathbb{R})$ generated by $\{A_0, \dots, A_h\}$ under bracketing and corresponding to this is a closed Lie subgroup, G, of $Gl(m, \mathbb{R})$. This subgroup is the set of all products of the form

$$\exp\left(t_{i_1}A_{i_1}\right)\cdots\exp\left(t_{i_k}A_{i_k}\right)$$

for all $k \ge 0$ and $t_{i_j} \in \mathbb{R}$, [8]. Another characterization of G is that it is the set of all accessible matrices of

$$\dot{X}(t) = \left(\sum_{i=0}^{h} u_i(t)A_i\right)X(t),$$

$$X(0) = I, \qquad |u_i| \le 1, \qquad i = 0, \dots, h.$$

This follows from the theorem of Chow [9].

The dimension of G as a submanifold of $Gl(m, \mathbb{R})$ is precisely the dimension of the Lie subalgebra, g. Furthermore, it has been shown [10], [11] that the set of accessible matrices of (2.1) is a subset of G with nonempty interior in the relative topology of G, hence G is the smallest subgroup of $Gl(m, \mathbb{R})$ containing all accessible matrices of (2.1). For this reason G is said to carry (2.1).

The corresponding situation for (1.2) is more complicated because of the nonlinearity. We restrict our discussion of this system to some neighborhood, \mathcal{V} , of y^0 in \mathbb{R}^n . If $b_i(y)$, $b_j(y)$ are C^{∞} -vector fields defined on \mathcal{V} , then the Lie bracket, $[b_i, b_j](y)$, is another C^{∞} -vector field defined on \mathcal{V} by

$$[b_i, b_j](y) = \frac{\partial b_j}{\partial y}(y)b_i(y) - \frac{\partial b_i}{\partial y}(y)b_j(y).$$

Once again the skew symmetry and Jacobi relations hold.

The set, V(Y), of all C^{∞} -vector fields on Y becomes a Lie algebra over \mathbb{R} with this definition, however it, in general, is infinite-dimensional. Let W(Y) denote the smallest subalgebra of V(Y) containing $\{b_0, \dots, b_h\}$. In many cases, but not in general, there is a submanifold \mathcal{N} of \mathcal{V} corresponding to W(Y), and containing y^0 . To be more precise, let W(y) be the linear subspace of \mathbb{R}^n formed by evaluating the vector fields of W(Y) at y. A submanifold \mathcal{N} of \mathcal{V} is an integral manifold of W(Y) if for every $y \in \mathcal{N}$, W(y) is precisely the tangent space to \mathcal{N} at y. We define the rank of W(Y) at y to be the dimension of W(y). Then there exists an integral manifold \mathcal{N} of W(Y) containing y^0 if the rank of W(Y) is constant (Frobenius) [12] or if $b_0(y), \dots, b_h(y)$ are analytic [13]. Other sufficient conditions are found in [12] and [14].

Henceforth we shall assume that \mathcal{N} exists, the dimension of \mathcal{N} is the same as the rank of $W(\mathcal{V})$ at y^0 and by Chow's theorem, is the set of all points in \mathcal{N} accessible from y^0 under the system

$$\dot{y}(t) = \sum_{i=0}^{h} u_i(t)b_i(y),$$

$$y(0) = y^0, \qquad |u_i| \le 1, \qquad i = 0, \dots, h.$$

The set of all points in \mathcal{N} accessible from y^0 by (1.2) is again a subset of \mathcal{N} with nonempty relative interior [10], so \mathcal{N} is said to *locally carry* (1.2).

3. Bilinearization. The problem of replacing a nonlinear realization by a bilinear one can be broken into two parts. The first is: when does there exist a change of state which linearizes the vector fields $b_0(y), \dots, b_h(y)$, resulting in a system with bilinear dynamics and nonlinear output map? The second is: given a realization of this hybrid type, when can it be converted into a bilinear realization?

As for the first question, Guillemin and Sternberg [15] have shown that a family of vector fields, $b_0(y), \dots, b_h(y)$, can be converted to linear vector fields, $A_0x \dots, A_hx$, by a change of coordinates, x = x(y), in some neighborhood of y^0 if the vector fields are analytic, all vanish at y^0 and generate a finite-dimensional semisimple Lie algebra. Hermann [16] gave a formal power series construction of the change of coordinates. However, these results are not directly applicable to our questions, since if all the vector fields vanish at y^0 , then the system, (1.2), is trivial.

Asking for a change of coordinates to linearize the vector fields in some neighborhood of y^0 is actually too restrictive for our purposes. Assuming (1.2) is carried locally by \mathcal{N} , what we would like is a system (2.1) carried by G, a neighborhood, \mathcal{M} , of I in G and a differentiable map $\lambda : \mathcal{M} \to \mathcal{N}$ which preserves solutions, that is, $\lambda(X(t)) = y(t)$ for each u(t). The map need not be a local diffeomorphism from \mathcal{M} onto \mathcal{N} , for the dimension of \mathcal{M} could be greater than that of \mathcal{N} ; however, it should be onto since \mathcal{N} carries (1.2). Hartman dealt with a similar question in studying the structural stability of a single vector field about a critical point [21].

If such a λ exists, then its differential, λ_* , is a Lie algebra homomorphism from the Lie algebra, g, generated by A_0, \dots, A_h onto the Lie algebra, $W(\mathcal{N})$, generated by b_0, \dots, b_h restricted to \mathcal{N} . Therefore a necessary condition for λ to exist is that $W(\mathcal{N})$ be a finite-dimensional Lie algebra. This also turns out to be sufficient and we have the following theorem.

Theorem 1. Suppose that $b_0(y), \dots, b_h(y)$ of (1.2) are analytic and the system is carried locally by \mathcal{N} . There exists a system (2.1) carried locally by \mathcal{M} in $Gl(m, \mathbb{R})$ and an analytic map $\lambda: \mathcal{M} \to \mathcal{N}$ preserving solutions if and only if the Lie algebra generated by $b_0(y), \dots, b_h(y)$ is finite-dimensional when restricted to \mathcal{N} .

Proof. Assume $W(\mathcal{N})$ is finite-dimensional. Then by Ado's theorem [17] there exists a Lie subalgebra, g, of $gl(m, \mathbb{R})$ for some m and a Lie algebra isomorphism $\varphi:W(\mathcal{N})\to g$. Define a system with matrix bilinear dynamics, (2.1), by letting $A_i=\varphi(h_i)$. Let e be the evaluation map, $e:W(\mathcal{N})\to W(y^0)$, defined by $e(c)=c(y^0)$ for $c\in W(\mathcal{N})$. Then the map $l=e\circ\varphi^{-1}$ satisfies the following

$$l([A_{i_1}\cdots[A_{i_{v-1}},A_{i_v}]\cdots])=[b_{i_1}\cdots[b_{i_{v-1}},b_{i_v}]\cdots](y^0)$$

for any v and $0 \le i_1, \dots, i_v \le h$.

It follows from a theorem of the author [18] (generalized by Sussmann [19]) that there exist a neighborhood \mathcal{M} of I and a map $\lambda: \mathcal{M} \to \mathcal{N}$ preserving solutions. O.E.D.

Remark 1. In general, the map λ is locally a projection from \mathcal{M} onto \mathcal{N} . However, if the evaluation map, $e: W(\mathcal{N}) \to W(y^0)$, is a vector space isomorphism, then so is l and the abovementioned theorem implies λ is a local diffeomorphism.

Remark 2. A Lie algebra homomorphism $\varphi:W(\mathcal{N})\to gl(m,\mathbb{R})$ is called a representation of $W(\mathcal{N})$ and is said to be faithful if φ is 1-1, and hence an isomorphism onto its range. If $W(\mathcal{N})$ is of dimension m, then the adjoint representation, $\operatorname{ad}:W(\mathcal{N})\to gl(m,\mathbb{R})$, can always be constructed as follows. Choose a basis d_1,\dots,d_m for $W(\mathcal{N})$, and for each $c\in W(\mathcal{N})$ let $\operatorname{ad}(c)$ be the matrix $B=[B_{ij}]$ defined by

$$[c, d_j] = \sum_{i=1}^m B_{ij} d_i.$$

The Jacobi relation implies this is a Lie algebra homomorphism.

The kernel of ad is the *center* of $W(\mathcal{N})$, i.e., the set of all c such that [c, d] = 0 for all $d \in W(\mathcal{N})$. If the center is empty, then ad is faithful and this representation can be used in Theorem 1. If $W(\mathcal{N})$ is semisimple, then the center is empty.

Remark 3. If the center is not empty but is contained in the kernel of the evaluation map, e, then the adjoint representation can still be used. In this case, l is constructed by the standard homomorphism theorem as illustrated in Fig. 1.

As for the second step in bilinearization, we have a theorem of Brockett [4] which states that every realization with bilinear dynamics and polynomial output map is equivalent to a realization with bilinear dynamics and linear output map. This results in the following.

COROLLARY 1. Given: any nonlinear realization (1.2) of the input-output, $u(t) \mapsto z(t)$, satisfying the hypothesis of Theorem 1. If the map $f \circ \lambda: X \to z$ is a polynomial, then there exists a bilinear realization (1.2) of $u(t) \mapsto w(t)$ and a constant T > 0 such that for any input, u(t), the corresponding outputs satisfy w(t) = z(t) for $t \in [0, T]$. (Polynomial here means each component of z is a polynomial in the components of X.)

4. Approximation of nonlinear systems by bilinear systems. If the Lie algebra, $W(\mathcal{X})$, is not finite-dimensional, then Theorem 1 does not hold; however, we can ask whether (1.2) can be approximated by systems of type (2.1). To be more precise, given (2.1) carried locally by \mathcal{M} and (1.2) carried locally by \mathcal{N} , a C^{∞} -map $\lambda: \mathcal{M} \to \mathcal{N}$ preserves solutions to order μ if there exists a T>0 and $K\geq 0$ such that for any solution, X(t) and y(t), of (2.1) and (1.2) using the same control

$$|\lambda(X(t)) - y(t)| \leq Kt^{\mu+1}$$

for $t \in [0, T]$.

THEOREM 2. Suppose that $b_0(y), \dots, b_n(y)$ of (1.2) are C^{∞} and the system is carried locally by \mathcal{N} . Then for any $\mu \geq 0$ there exists a system (2.1) carried locally by \mathcal{M} in $Gl(m, \mathbb{R})$ and a C^{∞} -map $\lambda: \mathcal{M} \to \mathcal{N}$ preserving solutions to order μ .

Proof. An abstract Lie algebra, g, is a vector space over \mathbb{R} with a multiplication which satisfies the skew symmetry and Jacobi relations. Suppose a_0, \dots, a_h are elements of g; then we call $[a_{i_1} \cdots [a_{i_{v-1}}, a_{i_v}] \cdots]$ a bracket of order v of a_0, \dots, a_h . One way to construct an abstract Lie algebra, g, is to consider a_0, \dots a_h to be elements of the algebra and linearly independent over \mathbb{R} . Then treat all the brackets of these up to and including order v as new elements of g which are linearly independent except for those relations implied by the skew symmetry and Jacobi relations. All brackets of order greater than v are taken to be 0. The result is a finite-dimensional Lie algebra which we shall call the canonical algebra of order v with v 1 generators.

By Ado's theorem, this algebra is isomorphic to a subalgebra of $gl(m, \mathbb{R})$ which we also denote by g. Under this identification, each a_i becomes a $m \times m$ matrix, A_i , and these are used to construct (2.1). We call the resulting system the canonical system of order μ with h controls.

Next we define a linear map $l:g \to \mathbb{R}^n$ by setting

$$l([A_{i_1}\cdots[A_{i_{v-1}},A_{i_v}]\cdots])=[b_{i_1}\cdots[b_{i_{v-1}},b_{i_v}]\cdots](y^0).$$

It then follows from a theorem of the author [20] that there exists a neighborhood, \mathcal{M} , of I in the subgroup, G, of $Gl(m, \mathbb{R})$ carrying (2.1), a neighborhood, \mathcal{N} , of y^0 in

the submanifold carrying (1.2) and a C^{∞} -map $\lambda: \mathcal{M} \to \mathcal{N}$ which preserves solutions to order μ . Q.E.D.

Remark 1. Once again λ is locally a projection; however, it need not be onto A unless the brackets of b_1, \dots, b_h up to order μ span the tangent space to A at y^0 . Of course if I is 1-1, then so is λ .

Remark 2. The adjoint representation of the canonical algebra of order μ is not a faithful representation because the center consists of all brackets of a_0, \dots, a_h of order μ . However, for precisely this reason, the adjoint representation of the canonical algebra of order $\mu + 1$ is isomorphic to the canonical algebra of order μ and can be used to construct (2.1).

Remark 3. Since the canonical algebra is nilpotent, the algebra generated by A_0, \dots, A_h will be nilpotent, and hence any matrix exponential solution of (2.1) for piecewise constant u is a finite series with no terms of order greater than μ .

Remark 4. The dimension of (2.1) is the dimension of the Lie algebra generated by A_0, \dots, A_h and not m or m^2 . The system (2.1) as constructed in the theorem may not be of minimal dimension among bilinear systems which preserve solutions of (1.2) to order μ . A smaller system can be constructed as follows.

Since l is only a linear map, the kernel of l need not be an ideal of g. However, it is a subalgebra of g because l preserves brackets to order μ and all higher order brackets are 0 in g. Let h denote the largest ideal of g contained in the kernel of l. Then the nilpotent Lie algebra g/h can be used in place of g in the construction of the theorem. There may also be Lie algebras of smaller dimension than g/h which need not be nilpotent that can be used to construct (1.1), for example, if (1.2) generates a finite-dimensional Lie algebra.

COROLLARY 2. Given any nonlinear realization, (1.2), of the input-output map, $u(t) \mapsto z(t)$, and any integer $\mu \ge 0$, there exists a bilinear realization (1.1) of $u(t) \mapsto w(t)$ and constants M and T > 0 such that for any input, u(t), the corresponding outputs satisfy

$$|w(t) - z(t)| \le Mt^{\mu+1}$$
 for $t \in [0, T]$.

Proof. Using Theorem 2, we construct a system with the matrix bilinear dynamics and a map $\lambda: \mathcal{M} \to \mathcal{N}$ which preserves solutions to order μ . We define a polynomial output map, ψ , for this system by letting ψ be the power series expansion around I of $f \circ \lambda$ up to and including terms of order μ . Using Brockett's technique [4], an equivalent system with bilinear dynamics and linear output map can always be constructed, so all we need show is that our system with bilinear dynamics and polynomial output map approximates (1.2) as required.

By passing to smaller neighborhoods if necessary, we can assume \mathcal{M} and are compact; then there exist constants K_1 and K_2 such that

$$|f \circ \lambda(X) - \psi(X)| \le K_1 |X - I|^{\mu + 1}$$
 for any $X \in \mathcal{M}$

$$|f(y^1) - f(y^2)| \le K_2|y^1 - y^2|$$
 for any $y^1, y^2 \in \mathcal{N}$.

Let X(t) and y(t) be the solutions of our matrix bilinear system and (1.2) for the same control |u(t)|. Then since λ preserves solution to order μ , there exists a

 K_3 and T > 0 such that for $t \in [0, T]$,

$$|\lambda(X(t)) - y(t)| \le K_3 t^{\mu+1}.$$

By a standard argument there exists a constant, K_4 , such that for $t \in [0, T]$,

$$|X(t) - I| \leq K_{\Delta}t$$
.

Putting it all together, we have

$$\begin{aligned} |\psi(X(t)) - f(y(t))| \\ &\leq |\psi(X(t)) - f \circ \lambda(X(t))| + |f \circ \lambda(X(t)) - f(y(t))| \\ &\leq K_1 |X(t) - I|^{\mu+1} + K_2 |\lambda(X(t)) - y(t)| \\ &\leq (K_1 K_4^{\mu+1} + K_2 K_3) t^{\mu+1}. \end{aligned}$$
 Q.E.D.

Acknowledgment. I would like to thank Roger Brockett for suggesting this problem to me.

REFERENCES

- [1] R. R. MOHLER, Bilinear Control Processes with Applications to Engineering, Ecology and Medicine Academic Press, New York, 1973.
- [2] C. Bruni, G. Dipillo and G. Koch, Bilinear systems: An appealing class of "nearly linear" systems in theory and applications, IEEE Trans. Automatic Control, AC 19 (1974), to appear.
- [3] P. D'ALESSANDRO, A. ISIDORI AND A. RUBERTI, Realization and structure theory of bilinear dynamical systems, this Journal, 12 (1974), pp. 517-535.
- [4] R. W. BROCKETT, On the algebraic structure of bilinear systems, Theory and Applications of Variable Structure Systems, R. R. Mohler and A. Ruberti, eds., Academic Press, New York, 1972, pp. 153-168.
- [5] J. BELINFANTE AND B. KOLMAN, A Survey of Lie Groups and Lie Algebras with Applications and Computational Methods, Society for Industrial and Applied Mathematics, Philadelphia, 1972
- [6] H. SAMELSON, Notes on Lie Algebras, Van Nostrand, New York, 1969.
- [7] J. E. Humphreys, Introduction to Lie Algebras and Representation Theory, Springer-Verlag, New York, 1972.
- [8] R. W. BROCKETT, Systems theory on group manifolds and coset spaces, this Journal, 10 (1972), pp. 265-284.
- [9] W. L. CHOW, Über Systeme von linearen partiellen Differentialgleichungen erster Ordnung, Math. Ann., 117 (1939), pp. 98–105.
- [10] A. J. Krener, A generalization of Chow's theorem and the bang-bang theorem to nonlinear control problems, this Journal, 12 (1974), pp. 43-52.
- [11] H. J. SUSSMANN AND V. J. JURDJEVIC, Controllability of non-linear systems, J. Differential Equations, 12 (1972), pp. 95–116.
- [12] R. HERMANN, On the accessibility problem in control theory, International Symposium, Non-linear Differential Equations and Nonlinear Mechanics, Academic Press, New York, 1963, pp. 325-332.
- [13] T. Nagano, Linear differential systems with singularities and an application to transitive Lie algebras, J. Math. Soc. Japan, 18 (1966), pp. 398–404.
- [14] C. LOBRY, Contrôlabilitié des systèmes non linéaires, this Journal, 8 (1970), pp. 573-605.
- [15] V. W. GUILLEMIN AND S. STERNBERG, Remarks on a paper of Hermann, Trans. Amer. Math. Soc., 130 (1968), pp. 110-116.
- [16] R. HERMANN, The formal linearization of a semisimple Lie algebra of vector fields about a singular point, Ibid., 130 (1968), pp. 105–109.
- [17] I. D. Ado, *The representation of Lie algebras by matrices*, Uspehi Mat. Nauk (22), 3 (1947), no. 6, pp. 159–173; English transl., Amer. Math. Soc. Transl. no. 2, 1949.

- [18] A. J. Krener, On the equivalence of control systems and the linearization of nonlinear systems, this Journal, 11 (1973), pp. 670–676.
- [19] H. J. Sussmann, An extension of a theorem of Nagano on transitive Lie algebras, Proc. Amer. Math. Soc., to appear.
- [20] A. J. Krener, Local approximation of control systems, J. Differential Equations, to appear.
- [21] P. HARTMAN, A lemma in the structural stability of differential equations, Proc. Amer. Math. Soc., 11 (1960), pp. 610–620.