Reprinted from PROCEEDINGS OF THE 22ND IEEE CONFERENCE ON DECISION & CONTROL, December 1983

PARTIAL AND ROBUST LINEARIZATION BY FEEDBACK

Arthur J. Krener
Department of Mathematics
University of California
Davis, California 95616 USA

Alberto Isidori
Instituto di Automatica
Universita di Roma
Via Eudossiana 18
00184 Roma Italy

Witold Respondek
Institute of Mathematics
Technical University of Warsaw
Pl. Jednosci Robotniczej 1
00-661 WARSAW POLAND

1. Introduction

While no system is truly linear, in many circumstances linear models do suffice for controller design. If the nonlinearities are mild and the controller is sufficiently stabilizing then the inaccuracies of the linear model can be safely ignored. We elaborate on this point in section 2. Recently, motivated by systems where the nonlinearities are severe, interest has focused on the linearization of systems by change of state coordinates and nonlinear feedback. We discuss this procedure and its robustness in section 3. In the last section we focus on the question of partial linearization via the same transformations. We find that there always are maximally linearizing transformations which are not necessarily unique. Motivated by robustness considerations we study the spectrum of all such transformed systems. For simplicity we consider only single input systems.

2. Infinitesmal linear approximation

Consider the C^∞ nonlinear system

$$\dot{\xi} = f(\xi,\nu) \qquad \xi \in \mathbb{R}^n \quad \nu \in \mathbb{R} \tag{2.1}$$

around some nominal operating point (ξ^o,ν^o) which is a critical point of the controlled differential equation, i.e.

$$f(\xi^o,\nu^o) = 0. \tag{2.2}$$

A standard approach to such systems is to replace (2.1) by its infinitesmal approximation

$$\dot{x} = Ax + Bu \tag{2.3a}$$

where

$$A = \frac{\partial f}{\partial \xi}(\xi^o,\nu^o) \qquad B = \frac{\partial f}{\partial u}(\xi^o,\nu^o) \tag{2.3b}$$

$$x \simeq \xi - \xi^o \qquad u = \nu - \nu^o \tag{2.3c}$$

around $(x^o,u^o) = (0,0)$. It is well-known that as long as $(\xi(t),\nu(t))$ stays close to (ξ^o,ν^o) then the approximation is fairly accurate, but this can only happen if the input $\nu(t)$ remains close to ν^o and (ξ^o,ν^o) is a stable critical point, i.e., the spectrum of A lies strictly in the left half plane.

If this is not true but the infinitesmal approximation (2.3) is controllable (or stabilizable) then one can use state feedback

$$u(x,v) = Fx + Gv \tag{2.4}$$

to transform (2.3) to

$$\dot{x} = (A+BF)x + BGv \tag{2.5}$$

The controllability (or stabilizability) of (A,B) allow us to choose F so that spectrum of $(A+BF)$ is strictly in the left half plane.

On the other hand we could modify the original nonlinear system by nonlinear state feedback

$$\nu = \nu(\xi,\mu) \tag{2.6}$$

to obtain

$$\dot{\xi} = f(\xi,\nu(\xi,\mu)) \tag{2.7}$$

If $\nu(\xi^o,\nu^o) = \nu^o$ and

$$\frac{\partial \nu}{\partial \xi}(\xi^o,\nu^o) = F \qquad \frac{\partial \nu}{\partial \mu}(\xi^o,\nu^o) = G$$

then the infinitesmal approximation of (2.7) at $(\xi,\mu) = (\xi^o,\nu^o)$ is just (2.5). In other words we can think (2.5) as either the linear feedback (2.4) modification of the infinitesmal approximation (2.3) of the original system (2.1) as the infinitesmal approximation of the nonlinear feedback modification (2.6) of the original system (2.1).

This is extremely fortunate for it means that a feedback chosen to stabilize the infinitesmal approximation (2.3) also stabilizes the original nonlinear system and greatly improves the accuracy of the approximation of (2.5) to (2.7). This is the fundamental mathematical fact that allows one to use linear models obtained by infinitesmal approximation (after being stabilized by feedback) to model nonlinear processes in the small. Moreover even if (2.3) is not exactly the infinitesmal approximation of (2.1) because (2.3b) does not exactly hold we can expect the eigenvalues to remain strictly stable and hence the approximation still will be accurate.

3. Linearization by Feedback

In this section we describe a different technique for linearizing a system developed independently by Hunt-Su [1] and Jakubczyk-Respondek [2] following earlier work of Brockett [3]. It is usually described as linearization by state feedback but a more accurate description would be linearization by change of state and input coordinates. This change of coordinates has a triangular structure, the new state coordinates depend only on the old state coordinates and the new input coordinates depend on the old state and input coordinates. As was mentioned before we only consider systems with a one dimensional input, m=1, the general case is described in the above references.

The first step is to transform (2.2) into a system which is linear in the control

Research supported in part by NSF (MCS-8063263) NASA (N CA2-OOY-18 0-251) and DOE (DE-A G01-80RA 50421)

0191-2216/83/0000-0126 $1.00 © 1983 IEEE

$$\dot{\xi} = g^o(\xi) + g(\xi)\mu. \qquad (3.1)$$

This amounts to a state dependent change of input coordinates $\mu = \mu(\xi,\nu)$ given by the relation

$$g^o(\xi) + g(\xi)\mu(\xi,\nu) = f(\xi,\nu) \qquad (3.2)$$

A necessary condition for such a change of input coordinates to be valid in that the range of the map $\nu \rightarrow f(\xi,\nu)$ lie on a line. This is not sufficient for this map may not be 1-1 or onto the line. The system may be subject to saturation, etc.

But these problems are always present when trying to approximate a real plant by a mathematical model. If one attempts to incorporate all the peculiarities of the plant into the model from the beginning the result typically is a mathematically untractable model so one starts by analyzing a simplified model for which it is possible to do a rational controller design, and then one tests this design on the real plant. If the simplifications of the model do not overly degrade the real performance then one has a satisfactory model. If the degradation is too severe, then a more complicated and hopefully more accurate model is employed. This process is iterated until the design is successful or the model is untractable.

Proceeding from (3.1) we seek a change of state coordinates $x=x(\xi)$ and change of input coordinates (or state feedback) $u = \alpha(\xi) + \beta(\xi)\mu$ which transform (3.1) to

$$\dot{x} = Ax + Bu \qquad (3.3)$$

where (A,B) is a controllable pair.

The domain of these transformations is some open set encompassing the nominal operating point (ξ^o,μ^o) which should transform to $(x^o,u^o) = (0,0)$. One could consider transformations taking (ξ^o,μ^o) to other (x^o,u^o) but we shall not do so. Since the right side of (3.3) is 0 at $(x^o,u^o) = 0$, an immediate necessary condition is that right side of (3.1) be zero at (ξ^o,μ^o)

$$0 = g^o(\xi^o) + g(\xi^o)\mu^o \ . \qquad (3.4)$$

Since the class of transformations which we are considering includes linear change of state coordinates and linear state feedback and (A,B) is assumed to be controllable, we can without loss of generality assume (3.3) to be in Brunovsky canonical form, i.e.,

$$A = \begin{pmatrix} 0 & 1 & & 0 \\ & & \ddots & \\ & & & 1 \\ 0 & & & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \ . \qquad (3.5)$$

Let's define

$$\mathcal{D}^k = C\{g,\ldots,ad^{k-1}_{g^o} g\} \qquad (3.6a)$$

$$\mathcal{D}^k(\xi) = \{X(\xi) : X \in \mathcal{D}^k\} \qquad (3.6b)$$

where $C\{\cdot\}$ denotes all C^∞ linear combinations of the enclosed vector fields. The controllability of (A,B) leads to the second necessary condition, for all ξ in the domain of linearization

$$\dim \mathcal{D}^n(\xi) = n \qquad (3.7)$$

This follows from the fact that if the transformations exist then $\mathcal{D}^k(\xi)$ must be transformed onto

$$R\{B,AB,\ldots,A^{k-1}B\}$$

where R denotes the span over the reals.

The third necessary condition, which together with the previous two form a set of sufficient conditions, is obtained as follows. Consider a pseudo-output

$$y = Cx \qquad C = (1\ 0\ \ldots\ 0) \qquad (3.8)$$

for the linear system (3.3). Since $C = dCx$ annihilates the vector fields $\{B,\ldots,A^{n-2}B\}$ which transforms to $\mathcal{D}^{n-2}$, there must be some function $h(\xi) = y$ such that dh annihilates $\mathcal{D}^{n-2}$. The existence of such a function where $dh \neq 0$ is equivalent to the involutiveness $\mathcal{D}^{k-2}$. In other words if $\bar{\mathcal{D}}^{k-2}$ denotes the involutive closure of $\mathcal{D}^{k-2}$; then

$$\bar{\mathcal{D}}^{k-2} = \mathcal{D}^{k-2} \ . \qquad (3.9)$$

(The involutive closure of $\mathcal{D}^{k-2}$ is the space of all vector fields obtained by repeated bracketing of vector fields of $\mathcal{D}^{k-2}$.) This condition is just the integrability condition for the solvability of the underdetermined system of partial differential equations.

$$L_{ad^j_{g^o} g}(h) = 0 \quad j=0,\ldots,n-2 \qquad (3.10)$$

where the Lie differentiation is defined by

$$L_{ad^j_{g^o} g}(h) = \langle dh, ad^j_{g^o} g\rangle = \frac{\partial h}{\partial g}\, ad^j_{g^o} g \ .$$

Summarizing we have sketched half of the proof of the following:

<u>Theorem</u> Hunt-Su [1]/Jakubcyzk-Respondek [2]. There exist change of coordinates $x=x(\xi)$, $u = \alpha(\xi)+\beta(\xi)\mu$ carrying (ξ^o,μ^o) to $(x^o,u^o) = (0,0)$ and transforming the nonlinear system (3.1) into the linear system (3.3) and (3.5) iff (3.4), (3.7) and (3.9) hold.

The other half is proved as follows. From (3.9) and (3.7) we deduce the solvability of the partial differential equation (3.10) using Frobenius Theorem. Let $h(\xi)$ be a nontrivial solution, $dh \neq 0$, and define

$$y = h(\xi), \text{ and } x_j = L_f^{j-1} h(\xi),\ j=1,\ldots,n.$$

The nontriviality of $h(\xi)$, (3.10) and the controllability condition (3.7) ensure that $(x_1,\ldots,x_n)$ are n independent coordinates and

$$\dot{x}_j = \begin{cases} x_j + 1 \\ L^n_{g^o}(h)(\xi) + L_g L^{n-1}_{g^o}(h)(\xi)\mu \end{cases} \qquad (3.11)$$

where $L_g L^{n-1}_{g^o}(h)(\xi) \neq 0$. Therefore the desired feedback is $\alpha(\xi) = L^n_{g^o}(h)(\xi)$ and $\beta(\xi) = L_g L^n_{g^o}(h)(\xi)$. It follows from (3.4) that (ξ^o,μ^o) transforms to $(x^o,u^o) = (0,0)$ for

$$x_j^o = L_{g^o}^{j-1}(h)(\xi^o) = L_{g^o(\xi^o)+g(\xi^o)\mu^o} L_{g^o}^{j-2}(h) = 0$$

$$u^o = L_{g^o(\xi^o)+g(\xi^o)\mu^o} L_{g^o}^{n-1}(h) = 0 \quad .$$

A systems interpretation of the above is as follows. Notice that the linear system (3.3) (3.5) and (3.8) has no finite zeroes, or n infinite zeroes. In other words one must time differentiate the output $y(t)$ n times before the effect of control $u(t)$ is directly seen. We define the nonlinear system (3.1) with output $y=h(\xi)$ to have no finite zeroes if one must differentiate $y(t)$ n times before directly seeing the effect of the control $\mu(t)$. (For a more precise definition of the zeroes of a nonlinear system see [4]). The above theorem can be loosely paraphrased as follows. A system (3.1) is linearizable by state feedback and change of coordinates iff it admits an output for which there are no finite zeroes. Of course every linear system admits such an output, a fact which follows immediately from controllable companion form.

If the underdetermined PDE (3.10) is solvable on an open set $\mathcal{U}$ around ξ^o then linearization is achievable on $\mathcal{U}$, but how robust is it. If we perturb a nonlinear system (3.1) which satisfies the theorem we will not usually obtain a system which satisfies the theorem. Condition (3.7) is robust but (3.4) and (3.9) are not. It is reasonable to expect that (3.4) will be satisfied at a perturbed operating point, but (3.9) will almost always fail to be true. Suppose as part of design procedure we linearize by the above technique but our actual plant is perturbation of the model (3.1). If we stopped at this point we could not expect the linearized model (3.3) to accurately reflect the behaviour of the plant under feedback. However we do not stop at this point but we implement a second feedback, linear in the x coordinates, which stabilizes the linear model. For the reason discussed in the last section the stabilizing feedback applied both to the model and the plant improves the accuracy of the model. Therefore we see that in the total context of a stable feedback design the linearization technique described above is robust.

4. Partial Linearization

The infinitesmal approximation of a linear system is the same linear system therefore the approximation is exact. But the infinitesmal approximation is coordinate dependent. If we transform the linear system (3.3) into (3.1) by change of state and input coordinates, the infinitesmal approximation to (3.1) will not be exact. The H-S/J-R theorem describes those nonlinear systems for which there exist state and input coordinates in which the infinitesmal approximation is exact. In this section we shall consider a generalization of this, given a nonlinear system find coordinates in which the infinitesmal linearization is as accurate as possible. As stated this is a somewhat vague goal, as one way of making this precise is to consider the existence of coordinate transformations $\xi = \xi(x,\xi)$ and $u = \alpha(\xi) + \beta(\xi)\mu$ which transforms (3.1) into the partially linear system

$$\dot{x} = Ax + Bu \qquad (4.1a)$$

$$\dot{\zeta} = \varphi(x,\zeta) + \psi(x,\zeta)u \qquad (4.1b)$$

where (A,B) is a controllable pair. Again the transformations are local and (ξ^o,μ^o) goes to $(x^o,\zeta^o,u^o) = (0,\zeta^o,0)$. The dimensions of x and ζ are r and $n-r$ respectively. Clearly the infinitesmal linearization of (4.1) is exact relative to the x-coordinates hence one would like to make r as large as possible.

<u>Definition</u> Let ξ^o be a critical point of (3.1), $g^o(\xi^o) = 0$. The index of linearity ρ of the nonlinear system (3.1) at ξ^o is ρ if (i) the dimension $\mathcal{D}^\rho(\xi^o)$ is n. (ii) for every ξ in some neighborhood of ξ^o the dimension of $\bar{\mathcal{D}}^{\rho-1}(\xi)$ is n-d where $d > 0$.

<u>Theorem 1</u>. Let ρ be the index of linearity of (3.1) around ξ^o. If (3.1) can be transformed to (4.1) then $r \leq \rho$. Moreover there exist transformations carrying (3.1) into (4.1) where $r=\rho$.

<u>Proof</u>. The proof is very similar to our proof of the H-S/J-R theorem given in the last section. Suppose (3.1) can be transformed to (4.1) then we can assume (A,B) is in Brunovsky form (3.5) and define $h(\xi) = x_1$. A straightforward calculation shows that

$$L_g L_{g^o}^j(h) = 0 \qquad j=0,\ldots,r-2 \qquad (4.2)$$

and this is equivalent to

$$L_{ad_{g^o}^j g}(h) = 0 \qquad j=0,\ldots,r-2 \qquad (4.3)$$

hence $dh \perp \mathcal{D}^{r-1}$. This implies that $dh \perp \bar{\mathcal{D}}^{r-1}$ and so $r \leq \rho$.

On the other hand since $\bar{\mathcal{D}}^{\rho-1}$ is of positive codimension d, there exists a function h such that $dh \neq 0$ and $dh \perp \bar{\mathcal{D}}^{\rho-1}$. If we define $x_j = L_{g^o}^{j-1}(h)$ $j=1,\ldots,\rho$, $\alpha = L_{g^o}^\rho(h)$, $\beta = L_g L_{g^o}^{\rho-1}(h)$ and ξ as any complimentary coordinates to $(x_1,\ldots,x_\rho)$ then we obtain (4.1). Q.E.D.

This result was suggested by that of Isidori and Krener [5], where the output h is known a-priori and the index of linearity is to be determined. Notice that an effect of the feedback is to make the nonlinear part (4.1b) unobservable from the pseudo-output, $h(\xi)$.

The transform of (3.1) with maximal linear part is not necessarily unique for we can choose any h such that $dh \perp \bar{\mathcal{D}}^{\rho-1}$. We did not make any assumption of controllability for (3.1) (except the implicit one that ρ is finite) the linear part of transformed system (4.1a) is controllable. This is not a contradiction for the transformation $u \to u = \alpha(\xi) + \beta(\xi)\mu$ may not be invertible because $\beta(\xi) = L_g L_{g^o}^{\rho-1}(h)(\xi) = 0$. Of course there is some freedom in our choice of h, there are d independent functions annihilating $\bar{\mathcal{D}}^{\rho-1}$, denote them by $h_1,\ldots,h_d$. While it is possible that $L_g L_{g^o}^{\rho-1}(h_i)(\xi^o) = 0$ for $i=1,\ldots,d$, it is not possible that $L_g L_{g^o}^{\rho-1}(h_i)(\xi) = 0$ for $i=1,\ldots,d$ and for all ξ in some neighborhood of ξ^o. For if this where the case then each dh_i annihilates $\mathcal{D}^\rho$ hence $\bar{\mathcal{D}}^\rho$ which is a contradiction. Therefore at almost every ξ^o we can choose h such that $\beta(\xi^o) = L_g L_{g^o}^{\rho-1}(h) \neq 0$.

Suppose we chosen such an h and transformed (3.1) to (4.1). The obvious next step is to replace the nonlinear part by its infinitesmal linear approximation. For convenience let us change notation and denote (4.1a) by,

$$\dot{x}_1 = A_{11}x_1 + B_1 u \quad . \tag{4.4a}$$

The infinitesmal approximation to (4.1b) is

$$\dot{x}_2 = A_{21}x_1 + A_{22}x_2 + B_2 u \tag{4.4b}$$

where in the notation of (4.1)

$$A_{21} = \frac{\partial\varphi}{\partial x}(0,\xi^o) \quad A_{22} = \frac{\partial\varphi}{\partial\xi}(0,\zeta^o) \quad B_2 = \psi(0,\zeta^o).$$

A natural question is whether one can stabilize (4.4) by linear state feedback which preserves the block triangular structure of the system, $u = Fx+u$ where

$$F = (F_1 \quad 0) \tag{4.5}$$

$$\tilde{A} = A+BF = \begin{pmatrix} A_{11} + B_1F_1 & 0 \\ A_{21} + B_2F_1 & A_{22} \end{pmatrix} \tag{4.6}$$

The reason for this is that these feedbacks leave the x_1 coordinates exact.

Since (A_{11}, B_1) is a controllable pair, clearly one can set the spectrum of $\tilde{A}_{11}$ arbitrarily up to complex conjugation. On the other hand nothing can be done by feedback about the spectrum of $\tilde{A}_{22}$.

For those familiar with the geometric theory of linear systems the above is not surprising. The first coordinate of x_1 is the pseudo-output $h(\xi)$. Relative to this the maximal (A,B) invariant subspace in the null space of C is $\mathcal{V}^* = \{x_1=0\}$ and the maximal (A,B) controllability subspace is $\mathcal{R}^* = \{0\}$. The spectrum of $\tilde{A}$ on $\mathcal{V}^*/\mathcal{R}^*$ is the transmission zeroes and they are invariant under feedback.

Of course we can always choose a different pseudo-output $h(\xi)$ which annihilates $\bar{\mathcal{D}}^{\rho-1}$ and ask how this might change the spectrum of the infinitesmal approximation to the nonlinear part. This spectrum is precisely the transmission zeroes of the linear infinitesmal approximation to (3.1) with output $y = Cx$ where $C = dh(\xi^o)$.

Assume that the linear infinitesmal approximation is controllable ((3.7) is satisfied) and given by (A,B) which we take in controllable companion form

$$A = \begin{pmatrix} 0 & 1 & & 0 \\ & & & \\ & & & 1 \\ * & & & * \end{pmatrix} \qquad B = \begin{pmatrix} 0 \\ \\ 0 \\ 1 \end{pmatrix} \tag{4.7}$$

Let $h_1,\ldots,h_d$ be independent functions annihilating $\mathcal{D}^{\rho-1}$ and assume the generic situation that for at least one h_i, $L_g L_{g^o}^{\rho-1}(h_i)(\xi^o) \neq 0$. Then without loss of generality we can assume that

$$L_g L_{g^o}(h_i)(\xi^o) = \begin{cases} 0 & i=1,\ldots,d-1 \\ 1 & i=d \end{cases} \tag{4.8}$$

$C_i = dh_i(\xi^o)$, then

$$C = \begin{pmatrix} C_1 \\ \\ C_d \end{pmatrix} = \left(\begin{array}{c|c} \begin{matrix} & 0 \\ * & \\ & 1 \end{matrix} & 0 \end{array}\right) \tag{4.9}$$

where the first block has $n-\rho$ columns and the second $\rho-1$ columns.

If we choose $y = \sum_{i=1}^{d} \lambda_i C_{ix}$ as output where $\lambda_i \in \mathbb{R}$, $\lambda d=1$, then the transmission zeroes are the $n-\rho$ roots of the polynomial

$$p(s) = \sum_{i=1}^{d} \sum_{j=1}^{n-\rho+1} \lambda_i c_{ij} s^{j-1} \quad .$$

Since the rows of C are linearly independent we have the following results.

<u>Theorem 4.1.</u> Suppose the index of linearity of (3.1) at ξ^o is ρ, the controllability assumption (3.7) is satisfied and $d-1 = n-\rho$, then the spectrum of the infinitesmal approximation to (4.16) can be set arbitrarily by proper choice of output.

<u>Theorem 4.2.</u> If the polynomials $p_i(s) = \sum_{j=1}^{n-\rho+1} c_{ij} s^{j-1}$ have a common zero s^o then s^o is always in the spectrum of infinitesmal approximation to (4.1b).

By suitable row operations C can be transformed to

$$C = \left(\begin{array}{ccccccccccc|c} * & * & 1 & 0 & . & . & . & & & 0 & \\ * & . & . & . & * & 1 & 0 & . & . & . & 0 & 0 \\ * & . & . & . & . & * & . & . & . & & 1 & \end{array}\right) \tag{4.10}$$

Let $K(C)$ be the convex cone in $\mathbb{R}^{n-\rho+1}$ spanned by the first part of the rows of C.

<u>Theorem 4.3.</u> If $K(C)$ does not intersect the nonnegative orthant in $\mathbb{R}^{n-\rho+1}$ then some part of the spectrum of the infinitesmal approximation to (4.1b) must be unstable.

5. Conclusions

We have discussed how a system can be partially and totally linearized by state feedback and coordinate change. Since the robustness of this linearization is dependent on the ability to stabilize the infinitesmal approximation of the transformed system we have studied this question in some depth. Clearly situations can arise where one must sacrifice linearizability to obtain stabilizability.

References

[1] L.R. Hunt and R. Su, Linear equivalents of nonlinear time varying systems, Int. Symp. Math. Theory Networks and Systems, Santa Monica (1981) pp. 119-123.

[2] B. Jakubczyk and W. Respondek, On the linearization of control systems, Bull. Acad. Polon. Sci. Ser. Math. Astronom. Phys. 28(1980), pp. 517-522.

[3] R.W. Brockett, Feedback invariants for nonlinear systems, International Federation for Auto. Cont. Helsinki, (1978).

[4] A.J. Krener and A. Isidori, Nonlinear zero distributions, 19th IEEE CDC, Albuquerque, 1980, pp. 665-668.

[5] A. Isidori and A.J. Krener, On the feedback equivalence of nonlinear systems, Systems and Control Letters, 2(1982), pp. 118-121.