**Proceedings of the 28th Conference on Decision and Control
Tampa, Florida • December 1989**

A NOTE ON THE LINEARIZATION AROUND AN EQUILIBRIUM

Jan C. Willems
Mathematics Institute
University of Groningen
9700 AV Groningen
The Netherlands

Françoise Lamnabhi-Lagarrigue
Laboratoire des Signaux et Systèmes
Ecole Supérieure d'Electricité
91190 Gif-sur-Yvette
France

ABSTRACT

In this note conditions are derived under which the linearization of a system described by the set of behavioral equations $f(w,w^{(1)},\dots,w^{(L)}) = 0$ around the equilibrium point $\omega^* = 0$ is indeed $\frac{\partial f}{\partial\omega_0}(0)\Delta + \frac{\partial f}{\partial\omega_1}(0)\Delta^{(1)} + \dots + \frac{\partial f}{\partial\omega_L}(0)\Delta^{(L)} = 0$.

1. INTRODUCTION

Linearization, the **raison d'être** of linear systems, is firmly established for systems described by state equations

$$\dot{x} = F(x,u) \ ; \quad y = H(x,u)$$

If $0 = F(0,0)$ and $0 = H(0,0)$, then the linearized system becomes

$$\dot{\Delta}_x = \frac{\partial F}{\partial x}(0,0)\Delta_x + \frac{\partial F}{\partial u}(0,0)\Delta_u$$

$$\Delta_y = \frac{\partial H}{\partial x}(0,0)\Delta_x + \frac{\partial H}{\partial u}(0,0)\Delta_u$$

Under suitable smoothness conditions it is readily established that this linear system defines indeed the linear part of the map $(x(0),u) \to y$ defined by the nonlinear differential equation.

However, more often than not models derived from first principles will come to us in the form of high order differential equations, for example as behavioral equations of the type

$$g(y,y^{(1)},\dots,y^{(L)},u,u^{(1)},\dots,u^{(L)}) = 0$$

The problem of writing state equations for such systems is usually studied using differential algebraic techniques. In this paper we will study the problem of approximating this system. If $g(0,0,\dots,0,0,0,\dots,0) = 0$ it is natural to look for a linear approximation.

In order to linearize around this equilibrium, we can do one of two things. Either first find a state representation and linearize. However, this is much easier said than done, a smooth state representation may not even exist ; if it does, it may be impossible to find it [3,4], worse yet, we may not need it. A second approach is to linearize with the audacity of the classical applied mathematician : simply expand the right hand side in a Taylor series and keep the linear term. This is, of course, much easier to do, but **is it justified** ? This is the problem addressed in this note. The results, by the way, are entirely as expected.

2. TECHNIQUE

We will study the linearization of a dynamical system described by a set of (high order) differential or difference equations. We will describe the setting in the case of differential equations. Let $f : (\mathbb{R}^q)^{L+1} \to \mathbb{R}^g$, assume that it is C^∞, and consider the dynamical system described by the behavioral differential equations

$$f(w,w^{(1)},\dots,w^{(L)}) = 0 \qquad \text{(NL)}$$

Suppose that $\omega^* \in \mathbb{R}^q$ (we will take $\omega^* = 0$) is an **equilibrium**, that is $f(0,0,\dots,0) = 0$. Linearizing (NL) around this operating point yields the **linearized** differential equations

$$f_0'\Delta + f_1'\Delta^{(1)} + \dots + f_L'\Delta^{(L)} = 0 \qquad \text{(L)}$$

where $\Delta : \mathbb{R} \to \mathbb{R}^q$, $f_0' = \frac{\partial f}{\partial\omega_0}(0,0,\dots,0)$, $f_1' = \frac{\partial f}{\partial\omega_1}(0,0,\dots,0),\dots,$ $f_L' = \frac{\partial f}{\partial\omega_L}(0,0,\dots,0)$ are the real $(g \times q)$ matrices obtained by taking the first partials of the map $(\omega_0,\omega_1,\dots,\omega_L) \to f(\omega_0,\omega_1,\dots,\omega_L)$.

The question arises when and, if so, in what sense (L) describes approximately (NL) in the neighborhood of ω^*. We will see that a sufficient condition for (L) to be a linearization of (NL) is the following. Let $\omega = (\omega^1,\omega^2,\dots,\omega^g,\omega^{g+1},\dots,\omega^q)$, $f = (f^1,f^2,\dots,f^g)$ denote the components of ω and f. Assume that there exist nonnegative integers $L_1,L_2,\dots,L_g$ such that locally around 0, the nonlinear equations (NL) may be solved as

$$\begin{bmatrix} w^{1(L_1)} \\ w^{2(L_2)} \\ \vdots \\ w^{g(L_g)} \end{bmatrix} = r(w,w^{(1)},\dots,w^{(L)})$$

with on the right hand side no derivative of w^i of order larger than (L_i-1), for $i = 1,2,\dots,g$. This condition, while representative, is only necessary, and our results are a bit stronger. Note that the order of the derivatives appearing in the variables $\omega^{g+1},\dots,\omega^q$ are arbitrary. In other words, the variables $\omega^1,\omega^2,\dots,\omega^g$ may or may not play the role of outputs in the conventional meaning of the term.

Examples :

1. $y^{(2)}+u^{(1)}y^{(1)}+\sin y = u$ is linearized by $\Delta^{(2)}+\Delta_y = \Delta_u$
2. $(y^{(2)})^2+y^{(1)}+\sin y = u^{(1)}$ is linearized by $\Delta_y^{(1)}+\Delta_y = \Delta_u^{(1)}$
3. It is not clear from our results what the linearization of $(y^{(2)})^2 + u^{(1)}y^{(1)} + \sin y = u$ is.

3. FORMALISM

Ca va sans dire that we will consider the framework for discussing dynamical systems introduced in [1]. **A dynamical system** is a triple (T,W,B) with $T \subset \mathbb{R}$ the **time-axis**, W the **signal space**, and $B \subset W^T$ the **behavior**. We will take $T = \mathbb{R}$ and $W = \mathbb{R}^q$. The dynamical system $(\mathbb{R},\mathbb{R}^q,B)$ is said to be **time-invariant** if $\sigma^t B = B$ for all $t \in \mathbb{R}$, where σ^t denotes the **t-shift**. It is said to be **linear** if B is a linear subspace of $(\mathbb{R}^q)^{\mathbb{R}}$.

The equations (NL) provide an example of a dynamical system described by **behavioral differential equations**. Define its behavior by

$$B_{NL} = \{w \in C^\infty(\mathbb{R};\mathbb{R}^q) \mid \text{(NL) is satisfied}\}$$

Clearly $\Sigma_{NL} = (\mathbb{R}, \mathbb{R}^q, B_{NL})$ is time-invariant. Defining B_L analogously from the equations (L) yields the dynamical system $\Sigma_L = (\mathbb{R}, \mathbb{R}^q, B_L)$, which is clearly linear and time-invariant.

Our notion of linearization depends on the notion of tangency. Let $\Sigma_i = (\mathbb{R}, \mathbb{R}^q, B_i)$, $i = 1,2$, be two time-invariant dynamical systems with $B_i \subseteq C^\infty(\mathbb{R}, \mathbb{R}^q)$. We will say that Σ_1 and Σ_2 are **tangent** at 0 if (i) $0 \in B_i$, $i = 1,2$, if (ii) for all sequences $w_k \in B_1$, $k \in \mathbb{N}$, such that $w_k \xrightarrow[k\to\infty]{} 0$ (in the C^∞-topology), there exists a sequence $\tilde{w}_k \in B_2$ such that for some $N \in \mathbb{Z}_+$ and all $T \geq 0$ there holds

$$\frac{\|w_k - \tilde{w}_k\|_{C^N([-T,T];\mathbb{R}^q)}}{\|w_k\|_{C^N([-T,T];\mathbb{R}^q)}} \xrightarrow[k\to\infty]{} 0$$

(where $\|f\|_{C^N([-T,T];\mathbb{R}^q)} = \max_{0\leq i\leq N} \sup_{t\in[-T,T]} \|f^{(i)}(t)\|$) and finally if (iii) the analogue of (ii) holds with the roles of B_1 and B_2 reversed.

If Σ_1 and Σ_2 are tangent at 0 and if Σ_2 is linear, then we will call Σ_2 the **linearization** of Σ_1. This makes precise what it means that (L) defines the linearization of (NL). It is easy to see that the linearization, if it exists, is unique.

Note that the linearized equations (L) deduced from (NL) depend very much on the behavioral equation used to represent B_{NL}. The representation (NL) is far from unique, however (in order to see this, replace $f = (f^1, f^2, \ldots, f^g)$ by $((f^1)^2, (f^2)^2, \ldots, (f^g)^2)$. This does not change the behavior, but the linearized equations will become trivial !). This lack of uniqueness can be alleviated to some extent by considering f as defining a variety in $(\mathbb{R}^q)^{L+1}$.

4. RESULT

Consider (L). Introduce the polynomial matrix $R(s) = f_0' + f_1's + \ldots + f_L's^L \in \mathbb{R}^{g\times q}[s]$. Let g' be the rank of R(s). Then there will exist a unimodular polynomial matrix $U(s) = U_0 + U_1 s + \ldots + U_{L'}s^{L'}$ such that UR is of the $\begin{bmatrix}\tilde{R}\\ \vdots \\ 0\end{bmatrix}$ with $\tilde{R}(s) \in \mathbb{R}^{g'\times q}[s]$ row proper, that is

$$\tilde{R}(s) = \begin{bmatrix} r_{1,L_1}s^{L_1} + r_{1,L_1-1}s^{L_1-1} + \ldots + r_{1,0} \\ r_{2,L_2}s^{L_2} + r_{2,L_2-1}s^{L_2-1} + \ldots + r_{2,0} \\ \vdots \\ r_{g',L_{g'}}s^{L_{g'}} + r_{g',L_{g'}-1}s^{L_{g'}-1} + \ldots + r_{L_{g'},0} \end{bmatrix}$$

is such that $\tilde{R}_{lead} \begin{bmatrix} r_{1,L_1} \\ r_{2,L_2} \\ \vdots \\ r_{g',L_{g'}} \end{bmatrix}$ is of full row rank.

Now assume that the distinct integers $1 \leq q_1, q_2, \ldots, q_{g'} \leq q$ are such that the corresponding columns of $\tilde{R}_{lead}$ are independent.

Now return to the differential equation (NL) : $f(w, w^{(1)}, \ldots, w^{(L)}) = 0$. For simplicity assume that f is defined as a map of $(\mathbb{R}^q)^{Z+}$ to $\mathbb{R}^g$ but that it depends only on a finite number of arguments. We will now define for each polynomial matrix $F(s) \in \mathbb{R}^{g\times g}[s]$ a new such map $F \square f$ as follows. For $F(s) = F_0$, define $(F\square f)(\omega_0, \omega_1, \ldots) = F_0 f(\omega_0, \omega_1, \ldots)$, for $F(s) = s$, define $(F.f)(\omega_0, \omega_1, \ldots) = \frac{\partial f}{\partial \omega_0}(\omega_0, \omega_1, \ldots)\omega_1 + \frac{\partial f}{\partial \omega_1}(\omega_0, \omega_1, \ldots)\omega_2 + \ldots + \frac{\partial f}{\partial \omega_k}(\omega_0, \omega_1, \ldots)\omega_{k+1} + \ldots$, for $F(s) = s^k$, define $(s^k \square f) = s\square(s^{k-1}\square f)$, for $F(s) = F_0 + F_1 s + \ldots + F_n s^n$, define

$$(F\square f) = F_0 \square f + F_1 \square (s\square f) + \ldots + F_n \square (s^n \square f).$$

Consider now the differential equation

$$\tilde{f}(w, w^{(1)}, \ldots, w^{(L')}) = 0$$

with $\tilde{f} = U\square f$. (It is easy to see that the linearization of $(\tilde{NL})$ around $\omega^* = 0$ is - if it exists - described by

$$\tilde{R}\left(\frac{d}{dt}\right)w = 0$$

with $UR = \begin{bmatrix}\tilde{R}\\ \vdots \\ 0\end{bmatrix}$). Note that $\tilde{f} : (\mathbb{R}^q)^{L'} \to \mathbb{R}^g$ and that L' need not be larger than L plus the degree of U(s).

Now assume that the map $\tilde{f} = (\tilde{f}^1, \tilde{f}^2, \ldots, \tilde{f}^g)$ satisfies the following conditions involving the integers $g', L_1, L_2, \ldots, L_{g'}$, and $q_1, q_2, \ldots, q_{g'}$ defined earlier on :

(i) either $g' = g$ or $\tilde{f}^{g'+1} = \ldots = \tilde{f}^g = 0$;

(ii) the $(g'\times g')$ matrix with (i,j)-th element $\frac{\partial \tilde{f}^i}{\partial \omega_{L_j}^{q_j}}(0, 0, \ldots, 0)$ is nonsingular ;

(iii) $\tilde{f}^i(\omega_0, \omega_1, \ldots)$ does not depend on $\omega_k^{q_j}$ for $i, j = 1, 2, \ldots, g'$ and $k > L_j$.

Equivalently for (ii) and (iii), assume that the equation $(\tilde{NL})$ can, locally around 0, be solved as

$$\begin{bmatrix} w^{q_1 (L_1)} \\ w^{q_2 (L_2)} \\ \vdots \\ w^{q_{g'} (L_{g'})} \end{bmatrix} = \tilde{r}(w, w^{(1)}, \ldots, w^{(L)})$$

with on the right hand side only derivatives in w^{q_j} up to order $L_j - 1$ for $1 \leq j \leq g'$.

Theorem. Let $R(s) = f_0' + f_1's + \ldots + f_L's^L$, and U(s) be unimodular and such that $U(s)R(s) = \begin{bmatrix}\tilde{R}(s)\\ \cdots \\ 0\end{bmatrix}$ with $\tilde{R}$ row proper. Assume that there exist distinct integers $1 \leq q_1, q_2, \ldots, q_{g'} \leq q$ satisfying the relevant conditions for $\tilde{R}$ and that $\tilde{f} = U\square f$ satisfies condition (i), (ii), and (iii) above. Then the system with behavioral equations (L) defines a linearization of the system with behavioral equations (NL).

5. PROOF

We will only give a global outline. The proof starts from the following observation : $w \in C^\infty$ satisfies $f(w, w^{(1)}, \ldots, w^{(L)}) = 0$ if and only if it satisfies $U(\frac{d}{dt})f(w, w^{(1)}, \ldots) = 0$. This latter equation is precisely $\tilde{f}(w, w^{(1)}, \ldots) = 0$ with $\tilde{f} = U.f$. Similarly $R(\frac{d}{dt})w = 0$ if and only if $\tilde{R}(\frac{d}{dt})w = 0$. This shows that it suffices to prove the claim made in the theorem for the system discussed in Section 2. The linearization in this case can be proven by associating with each $w \in B_L$ a $\tilde{w} \in \tilde{B}_{NL}$ and vice versa. Let $w \in B_L$. Take $\tilde{w}^{g+1} = w^{g+1}, \ldots, \tilde{w}^q = w^q$. Consider then the nonlinear differential equation $(\tilde{NL})$ with initial conditions $\tilde{w}^{i(j)}(0) = w^{i(j)}(0)$ for for $0 \leq j \leq L_i - 1$ and $1 \leq i \leq g$. In order to associate a

$\tilde{w} \in B_L$ with a $w \in B_{NL}$, *interchange the* roles of B_L and B_{NL} in this construction. Now prove, using standard facts from the theory of differential equations, that this association satisfies the required linearization condition. Consider for instance Example 1, with $w = (u,y)$, g=1, q=2 and $\frac{\partial f}{\partial w_2^2}(0,0) = 1$. Therefore $L_1=2$ and the linearization is then

$$\Delta_y^{(2)} + \Delta_y = \Delta_u.$$

Now consider Example 2, with $w = (u,y)$, $\frac{\partial f}{\partial w_1^2}(0,0) = 1$ so that $L_1=1$. The linearization is then $\Delta_y^{(1)} + \Delta_y = \Delta_u^{(1)}$ Note than here $\frac{\partial f}{\partial w_1^1}(0,0) = -1$.

6. LATENT VARIABLES

Often mathematical models obtained from first principles will contain other variables than those which we are interested in modelling. In the case of systems described by behavioral equations which are differential equations this will lead to a model of the type

$$f(w,w^{(1)},\ldots,w^{(L)},a,a^{(1)},\ldots,a^{(L)}) = 0 \qquad \text{(NL')}$$

Formalizing this yields a dynamical system with **latent variables** $\Sigma^\ell = (\mathbb{R},\mathbb{R}^q,\mathbb{R}^\ell,B_f)$ relating the **signal trajectory** $w : \mathbb{R} \to \mathbb{R}^q$ to the **latent variable trajectory** $a : \mathbb{R} \to \mathbb{R}^\ell$ via the ***full behavior*** $B_f \subseteq (\mathbb{R}^q \times \mathbb{R}^\ell)^{\mathbb{R}}$. If B_f is described by a behavioral differential equation as above then B_f consists of all solutions of this differential equation. Let us again consider the C^∞ case. The latent variable system Σ^ℓ induces the system $\Sigma = (\mathbb{R},\mathbb{R}^q,B)$ with **intrinsic behavior** $B = \{w : \mathbb{R} \to \mathbb{R}^q \mid \exists\, a : \mathbb{R} \to \mathbb{R}^\ell$ such that $(w,a) \in B_f\}$. If in the above differential equation f is linear then it can be shown that B will also be described by a linear differential equation. However, *for nonlinear systems this need not* be the case.

The question arises if we can obtain the linearization of B from the linearization of B_f. In particular, assume that $(\omega^*,a^*) \in \mathbb{R}^q \times \mathbb{R}^\ell$ defines an equilibrium of (NL') (assume $\omega^* = 0$ and $a^* = 0$). Linearizing (NL') around this equilibrium yields the linear differential equation

$$f'_{\omega_0}\Delta_\omega + f'_{N_1}\Delta_\omega^{(1)} + \ldots + f'_{\omega_L}\Delta_\omega^{(L)} + f'_{a_0}\Delta_a + f'_{a_1}\Delta_a^{(1)} + \ldots + f'_{a_L}\Delta_a^{(L)} = 0 \qquad \text{(L')}$$

(the notation is—hopefully—self-explaining). By introducing the polynomial matrices $R(s) = f'_{\omega_0}+f'_{\omega_1}s+\ldots+f'_{\omega_L}s^L$ and $M(s) = f'_{a_0}+f'_{a_1}s+\ldots+f'_{a_L}s^L$, we can write this as

$$R(\tfrac{d}{dt})w + M(\tfrac{d}{dt})a = 0$$

This equation describes, under conditions of the type spelled out in Section 4, the full behavior of the linearization of the latent variable system (NL'). It is well-known [2] that the intrinsic behavior of (L') can be expressed by a differential equation of the type

$$R'(\tfrac{d}{dt})w = 0 \qquad \text{(L")}$$

for some suitably chosen polynomial matrix R'(s). The question arises if this differential equation describes the linearization of the intrinsic behavior B associated with the nonlinear system (NL').

It turns out hat this is the case if the system (L') is **observable.** Observability is defined abstractly in [2]. For the system (L') it means that there must exist a polynomial matrix N(s) such that if (w,a) satisfies (L') there will hold $a = N(\frac{d}{dt})w$. Actually, as proven in [2], (L') will be observable **if and only if** rand $M(\lambda) = \ell$ for all $\lambda \in \mathbb{C}$.

There holds that **if (i) (L') is a linearization of (NL') (cfr. Section 4) and if (ii) (L') is observable, then (L") will define a linearization of the intrinsic behavior induced by (NL').**

Example : The linear state space system of the introduction is a linearization of the *nonlinear one*, if we consider in both cases the full behavior (with x as latent variable). The intrinsic behavior of the linear one will be a linearization of the intrinsic behavior of the nonlinear one if in addition the pair of matrices $(\frac{\partial F}{\partial x}(0,0), \frac{\partial H}{\partial x}(0,0))$ is observable. That observability is not superfluous in this may be seen from the example $\dot{x} = u$; $y = x^3$. The linearization, $\dot{\Delta}_x = \Delta_u$; $\Delta_y = 0$ *is not observable and the sequence of constant* trajectories $u = 0$, $y = \varepsilon_k$ in the intrinsic behavior of the nonlinear system can be made to converge to zero, but has no approximant in the linearized version!

7. DISCRETE TIME

For discrete-time systems, identical results are obtained in case the time axis in $\mathbb{Z}_+$ and the linearization condition is changed to

$$\frac{\sup\limits_{0 \le t \le T} \|w_k(t) - \tilde{w}_k(t)\|}{\sup\limits_{0 \le t \le T} \|w_k(t)\|} \xrightarrow[k \to \infty]{} 0$$

For discrete-time systems with time axis $\mathbb{Z}$, the polynomial matrix U(s) should be chosen such that $\tilde{R}(s)$ is a bilaterally row proper.

Examples :

4. $y_{t+2} = y_{t+1}u_t$, is linearized by $\Delta_{y_{t+2}} = 0$
5. $y_{t+1} = y_{t+2}u_t$, is linearized by $\Delta_{y_{t+1}} = 0$
6. What, the linearization of $y_{t+2}^2 = y_{t+1}u_t$ is not clear from our results.

8. REFERENCES

[1] J.C. Willems, "From Time Series to Linear System, Part I : Finite Dimensional Linear Time Invariant Systems", Automatica, Vol. 22, N° 5, pp. 561-580, 1986.

[2] J.C. Willems, "Models for Dynamics", Dynamics Reported, Vol. 2, 1989 (to appear).

[3] A.J. *Van der Schaft*, "On Realization of Nonlinear Systems described by Higher-Order Differential Equations", Math. Systems Theory, Vol. 19, pp. 239-275, 1987.

[4] P.E. Crouch and F. Lamnabhi-Lagarrigue, "State Space Realizations of Nonlinear Systems defined by Input-Output Differential Equations", In "Analysis and Optimization of Systems", Ed. A. Bensoussan and J.L. Lions, Lect. Notes in Control and Information Sciences, Vol. 111 , pp. 138-149, 1988.