A DIAGONALIZATION TECHNIQUE FOR THE COMPUTATION OF SENSITIVITY FUNCTIONS OF LINEAR TIME-INVARIANT SYSTEMS

by

T. L. Wuu, R. G. Becker, and E. Polak

Memorandum No. UCB/ERL M86/13

14 February 1986

COVER PAGE

A DIAGONALIZATION TECHNIQUE FOR THE COMPUTATION OF SENSITIVITY FUNCTIONS OF LINEAR TIME-INVARIANT SYSTEMS

by

T. L. Wuu, R. G. Becker, and E. Polak

Memorandum No. UCB/ERL M86/13

14 February 1986

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

TITLE PAGE

A DIAGONALIZATION TECHNIQUE FOR THE COMPUTATION OF SENSITIVITY FUNCTIONS OF LINEAR TIME-INVARIANT SYSTEMS

by

T. L. Wuu, R. G. Becker, and E. Polak

Memorandum No. UCB/ERL M86/13

14 February 1986

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

A Diagonalization Technique for the Computation of Sensitivity Functions of Linear Time-Invariant Systems

T.L. Wuu†, *R.G. Becker*‡ *and E. Polak*†

†Department of Electrical Engineering
and Computer Sciences
University of California
Berkeley, Ca. 94720
USA

‡Department of Electrical Engineering
Imperial College
London, SW7-2BT
ENGLAND

ABSTRACT

This note presents a method for computing the sensitivity functions of parametrized linear time-invariant systems, for the case where the the system matrix A is diagonalizable. The method is based on a formula, derived in the note, for the sensitivity of an exponential of a diagonalizable matrix.

Research sponsored in part by the National Science Foundation Grant ECS-8121149, Air Force Office of Scientific Research Grant 83-0361, Office of Naval Research Contract N00014-83-K-0602, The Semiconductor Research Corp. Contract SRC-82-11-008. The State of California MICRO Program and the General Electric Co.

1. Introduction

When designing a linear control system by semi-infinite optimization techniques, one is required to compute time and frequency domain responses as well as their sensitivities to the designable compensator parameters, see, e.g., [Pol.1]. While there is a considerable literature on the solution of state equations and the evaluation of frequency responses, see, e.g., [Lau.1, Lau.2], there are hardly any results available dealing with the efficient computation of response sensitivities.

This paper generalizes the unpublished results in [Bec.1] to obtain an efficient method for computing time domain response sensitivities for an important class of problems. Our method results from the following observations.

Consider a parametrized linear time-invariant system whose dynamics are given by:

$$\dot{x}(t,p) = A(p)\,x(t,p) + B(p)\,u(t), \tag{1.1a}$$

$$y(t,p) = C(p)\,x(t,p) + D(p)\,u(t), \tag{1.1b}$$

where $x(t,p) \in \mathbb{R}^n$ is the state, $p \in \mathbb{R}^N$ is the design parameter vector, $A : \mathbb{R}^N \to \mathbb{R}^{n\times n}$, $B : \mathbb{R}^N \to \mathbb{R}^{n\times k}$, $C : \mathbb{R}^N \to \mathbb{R}^{m\times n}$ and $D : \mathbb{R}^N \to \mathbb{R}^{m\times k}$ are continuously differentiable matrices, and the input $u : \mathbb{R} \to \mathbb{R}^k$ is of the form $u(t)=\sum_{i=1}^{m} t^{\alpha_i} e^{\lambda_i t} c_i$, where the $c_i \in \mathbb{R}^k$, the α_i are nonnegative integers and the $\lambda_i \in \mathbb{C}$.

Since the solution of (1.1a) is given by

$$x(t,p) \quad = e^{A(p)t}x(0,p) + \int_0^t e^{A(p)(t-\tau)}B(p)u(\tau)d\tau, \tag{1.2a}$$

its partial derivatives are given by

$$\frac{\partial x(t,p)}{\partial p^i} = \frac{\partial e^{A(p)t}}{\partial p^i}x(0,p) + e^{A(p)t}\frac{\partial x(0,p)}{\partial p^i} + \tag{1.2b}$$

$$\int_0^t \frac{\partial}{\partial p^i}(e^{A(p)(t-\tau)})B(p)u(\tau)d\tau + \int_0^t e^{A(p)(t-\tau)}\frac{\partial B(p)}{\partial p^i}u(\tau)d\tau ,$$

for all $i \in \underline{N} \triangleq \{1,2, \cdots ,N\}$. When the matrix $A(p)$ is diagonalizable and the input $u(t)$ is componentwise polynomial, (1.2a) represents a viable method for computing the state response. In this case, the efficient computation of partial derivatives $\frac{\partial x(t,p)}{\partial p^i}$ requires an efficient method for computing $\frac{\partial e^{A(p)t}}{\partial p^i}$. In [Bec.1] we find such a technique, based on Lie bracket decompositions, for the case where $A(p)$ has distinct eigenvalues. In this paper, the results in [Bec.1] are extended to include the case where the matrix $A(p)$ has repeated eigenvalues, but is diagonalizable. We shall comment in the conclusion on possible ways of dealing with the nondiagonalizable case.

2. A Formula for the Matrix $\frac{\partial e^{A(p)t}}{\partial p^i}$

If we are to use (1.2b) to compute $\frac{\partial x(t,p)}{\partial p^i}$, we must evaluate $\frac{\partial e^{A(p)t}}{\partial p^i}$.

Proposition 1 : Let $A(p)$ be defined as in equation (1.1a), then

$$\frac{\partial e^{A(p)t}}{\partial p^i} = e^{A(p)t} \int_0^t e^{-A(p)\tau}\frac{\partial A(p)}{\partial p^i}e^{A(p)\tau}d\tau . \tag{2.1}$$

Proof : By assumption, $A(\cdot)$ is continuously differentiable. Hence $e^{A(p)t}$ is continuously differentiable in (p,t), and it therefore follows that (see [Mar.1])

$$\frac{d}{dt}\left[\frac{\partial e^{A(p)t}}{\partial p^i}\right] = \frac{\partial}{\partial p^i}\left[\frac{de^{A(p)t}}{dt}\right] = \frac{\partial}{\partial p^i}\left[A(p)\cdot e^{A(p)t}\right]$$

$$= A(p)\cdot\frac{\partial e^{A(p)t}}{\partial p^i} + \frac{\partial A(p)}{\partial p^i}e^{A(p)t} \tag{2.2}$$

Integrating the linear differential equation (2.2) from 0 to t, we obtain that

$$\frac{\partial e^{A(p)t}}{\partial p^i} = e^{A(p)t}\left[\frac{\partial e^{A(p)t}}{\partial p^i}\right]_{t=0} + e^{A(p)t}\int_0^t e^{-A(p)\tau}\frac{\partial A(p)}{\partial p^i}e^{A(p)\tau}d\tau \tag{2.3}$$

Since $\left[\frac{\partial e^{A(p)t}}{\partial p^i}\right]_{t=0} = 0$, (2.1) follows directly.

∎

The evaluation of $e^{A(p)t}$ in (2.1) is relatively easy and can be carried out by some of the methods described in [Mol.1], [Par.1], [Lau.1]. In general, the evaluation of the second term, $\int_0^t e^{-A(p)\tau} \frac{\partial A(p)}{\partial p^i} e^{A(p)\tau} d\tau$, in (2.1) is more problematic. However, the following two observations lead to the conclusion that there may be cases where this term can be computed without resorting to numerical integration.

(a) Let $V \in \mathbb{R}^{n \times n}$ be such that $A(p)V - V A(p) = 0$. Then $e^{A(p)t} V = V e^{A(p)t}$ and

$$\int_0^t e^{-A(p)\tau} V e^{A(p)\tau} d\tau = tV. \tag{2.4}$$

(b) For any $U \in \mathbb{R}^{n \times n}$,

$$e^{-A(p)t}\{A(p)U - UA(p)\}e^{A(p)t} = -\frac{d}{dt}(e^{-A(p)t} U e^{A(p)t}). \tag{2.5}$$

Proposition 2 : Let $A(p)$ be defined as in equation (1.1a). For any $p \in \mathbb{R}^N$, if there exist $V \in \mathbb{R}^{n \times n}$ and $U \in \mathbb{R}^{n \times n}$, such that

$$\frac{\partial A(p)}{\partial p^i} = V + \{A(p)U - U A(p)\}, \tag{2.6a}$$

$$A(p)V - V A(p) = 0, \tag{2.6b}$$

then

$$\int_0^t e^{-A(p)\tau} \frac{\partial A(p)}{\partial p^i} e^{A(p)\tau} d\tau = tV + e^{-A(p)t}\{e^{A(p)t} U - U e^{A(p)t}\}. \tag{2.7}$$

Proof : Since

$$\int_0^t e^{-A(p)\tau} \frac{\partial A(p)}{\partial p^i} e^{A(p)\tau} d\tau = \int_0^t e^{-A(p)\tau}\{V + [A(p)U - U A(p)]\}e^{A(p)\tau} d\tau$$

$$= tV + e^{-A(p)t}\{e^{A(p)t}\; U - U\, e^{A(p)t}\}. \tag{2.8}$$

■

In establishing the existence and uniqueness of solutions (U, V) for Equation (2.6), we shall make use of the following result which can be found in [Tay.1].

Proposition 3 : Let $\mathcal{A}$ be a linear operator mapping a finite dimensional linear space $\mathcal{V}$ into itself and let $R(\mathcal{A})$ and $N(\mathcal{A})$ denote the range space and the null space of $\mathcal{A}$. Given any scalar λ, if q is the smallest nonnegative integer such that $N[(\mathcal{A}-\lambda)^q] = N[(\mathcal{A}-\lambda)^{q+1}]$, then

$$\mathcal{V} = N[(\mathcal{A}-\lambda I)^q] \oplus R[(\mathcal{A}-\lambda I)^q]. \tag{2.9}$$

■

Corollary 4 : If $\mathcal{A}$ is diagonalizable, then $\mathcal{V} = R(\mathcal{A}) \oplus N(\mathcal{A})$.

Proof : If $\mathcal{A}$ is diagonalizable, then $N[(\mathcal{A}-\lambda I)] = N[(\mathcal{A}-\lambda I)^2]$ for any scalar λ.

■

Proposition 5 : Let $A \in \mathbb{R}^{n \times n}$ and let the Lie bracket type operator $\mathcal{A}$: $\mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$ be defined by

$$\mathcal{A}(X) = [A, X] \triangleq AX - XA. \tag{2.10}$$

Then $\mathbb{R}^{n \times n} = R(\mathcal{A}) \oplus N(\mathcal{A})$, if and only if A is diagonalizable.

Proof : "<==": Suppose that A is diagonalizable. Let $\lambda_1, \lambda_2, \cdots, \lambda_n$ be the eigenvalues of A, let $u_1, u_2, \cdots, u_n$ be a set of linearly independent corresponding right eigenvectors of A and let $v_1^T, v_2^T, \cdots, v_n^T$ be a set linearly independent of corresponding left eigenvectors of A. Then (see [Gan.1]) $(\lambda_i - \lambda_j)$ for all $i,j \in \underline{n}$ is an eigenvalue of $\mathcal{A}$, and $u_i v_j^T$ is an eigenvector of $\mathcal{A}$ corresponding to $(\lambda_i - \lambda_j)$ for $i,j \in \underline{n}$. Since the set $\{\, u_i v_j^T \mid i,j \in \underline{n} \,\}$ contains n^2 linearly independent eigenvectors, $\mathcal{A}$ is diagonalizable. By Corollary 4, $\mathbb{R}^{n \times n} = R(\mathcal{A}) \oplus N(\mathcal{A})$.

"==>": We give a proof by contraposition. Assume that A is not diagonalizable. Then, by the Jordan form theorem [Gan.1], there exists a nonsingular $n \times n$ matrix U such that

$$U^{-1}AU = \begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & J_p \end{bmatrix} = J \tag{2.11}$$

where the J_i are $n_i \times n_i$ (note that $n_i \neq m_i$) Jordan blocks associated with eigenvalue λ_i, and $n = \sum_{i=1}^{p} n_i$. If the columns of U are denoted by u_i, so that $U = [u_1, u_2, \cdots, u_n]$, and the rows of U^{-1} are denoted by v_i^T, so that $V = [\, v_1, v_2, \cdots, v_n \,]^T \triangleq U^{-1}$, then $u_1, u_2, \cdots, u_n$ are generalized right eigenvectors of A and $v_1^T, v_2^T, \cdots, v_n^T$ are generalized left eigenvectors of A, see [Gan.1]. Also the dyads in $\{u_i v_j^T \mid i, j \in \underline{n}\}$ form a basis for $\mathbb{R}^{n \times n}$. Because J is block diagonal, the equations $AU = UJ$ and $VA = JV$ decompose into equations of the form

$$A\, u_i = \lambda_i u_i + \nu_i u_{i-1} \, ; \qquad v_i^T A = \lambda_i v_i^T + \mu_i v_{i+1}^T \, , \tag{2.12}$$

where ν_i, μ_i can have only the values 0 or 1. Consider the n_1 equations in (2.12) corresponding to the first Jordan block J_1:

$$\begin{aligned} A\, u_1 &= \lambda_1 u_1 \, ; & v_1^T A &= \lambda_1 v_1^T + v_2^T \\ A\, u_2 &= \lambda_1 u_2 + u_1 \, ; & v_2^T A &= \lambda_1 v_2^T \\ &\cdots & &\cdots \\ A\, u_{n_1-1} &= \lambda_1 u_{n_1-1} + u_{n_1-2} \, ; & v_{n_1-1}^T A &= \lambda_1 v_{n_1-1}^T + v_{n_1}^T \\ A\, u_{n_1} &= \lambda_1 u_{n_1} + u_{n_1-1} \, ; & v_{n_1}^T A &= \lambda_1 v_{n_1}^T \end{aligned} \tag{2.13}$$

Hence $\mathcal{A}(u_1 v_{n_1}^T) = 0$ and $\mathcal{A}(u_1 v_{n_1-1}^T) = -u_1 v_{n_1}^T$. This implies that $u_1 v_{n_1}^T \in N(\mathcal{A}) \cap R(\mathcal{A})$ and therefore that $\mathbb{R}^{n \times n} \neq N(\mathcal{A}) \oplus R(\mathcal{A})$.

■

Corollary 6 : Let $A \in \mathbb{R}^{n \times n}$ be diagonalizable, then for any $M \in \mathbb{R}^{n \times n}$, there exist $V,\ U \in \mathbb{R}^{n \times n}$ such that

$$M = V + [A, U]; \qquad [A, V] = 0. \tag{2.14}$$

Furthermore, the solution V of $[A, V] = 0$ in (2.14) is unique.

Proof : By Proposition 5, for any $M \in \mathbb{R}^{n \times n}$ there exists $U \in R^{n \times n}$ and $V \in N(\mathcal{A})$ such that $M = V + \mathcal{A}(U)$. By the definition of direct sum, V and $\mathcal{A}(U)$ are unique.

■

Corollary 7 : Let $A(p)$ be defined as in equation (1.1). If $A(p)$ is diagonalizable for a given $p \in \mathbb{R}^N$, then

$$\frac{\partial e^{A(p)t}}{\partial p^i} = tVe^{A(p)t} + [e^{A(p)t}, U] \tag{2.15}$$

for $i \in \underline{N}$, where $V,\ U$ satisfy following equations:

$$\frac{\partial A(p)}{\partial p^i} = V + [A(p), U]; \qquad [A(p), V] = 0. \tag{2.16}$$

Proof : The proof follows directly from the results of Proposition 2 and Corollary 6.

■

3. A Procedure for Computing the Matrix $\frac{\partial e^{A(p)t}}{\partial p^i}$ via Diagonalization.

We shall now state our procedure, based on (2.15) and Corollaries 6 and 7, for computing $\frac{\partial e^{A(p)t}}{\partial p^i}$ when $A(p)$ is an $n \times n$, continuously differentiable, diagonalizable matrix. To simplify the notation we denote $\frac{\partial A(p)}{\partial p^i}$ by $M(p)$.

Procedure :

Data: A continuously differentiable, diagonalizable $n \times n$ matrix $A(p)$.

Step 1: Diagonalize $A(p)$ by computing a matrix of linearly independent eigenvectors $T(p) \in \mathbb{C}^{n \times n}$ and setting $T^{-1}(p)A(p)T(p) = \Lambda(p) = diag\{\lambda_1(p), \lambda_2(p), \cdots, \lambda_n(p)\}$.

Step 2: Compute $\widetilde{M}(p) = T^{-1}(p)M(p)T(p)$, and construct the matrices $\widetilde{U}(p), \widetilde{V}(p) \in \mathbb{C}^{n \times n}$ as follows:

$$\widetilde{V}_{ij}(p) = \begin{cases} \widetilde{M}_{ij}(p) & \text{if } \lambda_i(p) = \lambda_j(p) \\ 0 & \text{otherwise} \end{cases} \tag{3.1}$$

$$\widetilde{U}_{ij}(p) = \begin{cases} \dfrac{\widetilde{M}_{ij}(p)}{\lambda_i(p) - \lambda_j(p)} & \text{if } \lambda_i(p) \neq \lambda_j(p) \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

for all $i, j \in \underline{n}$.

Comment: Note that $[\Lambda, \widetilde{V}] = 0$, and

$$\{\widetilde{V} + [\Lambda, \widetilde{U}]\}_{ij} = \widetilde{V}_{ij} + \lambda_i \widetilde{U}_{ij} - \widetilde{U}_{ij}\lambda_j = \begin{cases} \widetilde{V}_{ij} & \text{if } \lambda_i = \lambda_j \\ \widetilde{M}_{ij} & \text{if } \lambda_i \neq \lambda_j \end{cases} \tag{3.3}$$

Hence $\widetilde{M} = \widetilde{V} + [\Lambda, \widetilde{U}]$.

Step 3: Compute $V(p) = T(p)\widetilde{V}(p)T(p)^{-1}$ and $U(p) = T(p)\widetilde{U}(p)T(p)^{-1}$, which solve equations (2.16), and set

$$\frac{\partial e^{A(p)t}}{\partial p^i} = tVe^{A(p)t} + [e^{A(p)t}, U]. \tag{3.4}$$

If the sensitivity of the diagonalized system is required, set

$$T^{-1}\frac{\partial e^{A(p)t}}{\partial p^i}T = t\widetilde{V}e^{\Lambda(p)t} + [e^{\Lambda(p)t}, \widetilde{U}]. \tag{3.5}$$

4. Conclusion

We have presented a procedure for the computation of the matrix $\frac{\partial e^{A(p)t}}{\partial p^i}$ which appears in (1.2b) for the case where $A(p)$ is diagonalizable. Since this procedure involved the diagonalization of the matrix $A(p)$, it should be clear that the completion of the computation of $\frac{\partial x(t,p)}{\partial p^i}$ is quite straightforward and does not require numerical integration for inputs whose components are sums of terms of the form $t^{\alpha}e^{\lambda t}$, with α a nonnegative integer. Our procedure is particularly efficient when the matrix $A(p)$ in (1.1a) is large and/or when evaluations of $e^{A(p)t}$ and $\frac{\partial e^{A(p)t}}{\partial p^i}$ must be carried out for many different values of t.

Since matrix diagonalization is unstable when $A(p)$ is defective or near-defective, we follow [Mol.1] and use the condition number, $cond(T) = \|T\|\,\|T^{-1}\|$, of the matrix of eigenvectors as a testing function for the defectiveness of $A(p)$. When $A(p)$ is nearly (exactly) defective, the $cond(T)$ is large (infinite). Hence any errors in $A(p)$, and roundoff errors in the eigenvalue computation, may be magnified in final result of the decomposition $A(p) = T(p)\Lambda(p)T^{-1}(p)$ by $cond(T)$. Consequently, when $cond(T)$ is large, the computed $e^{A(p)t}$ and $\frac{\partial e^{A(p)t}}{\partial p^i}$ will most likely be inaccurate.

When $cond(T)$ is large, we propose to abandon the formulae (1.1a,b) and solve instead the following equation,

$$\frac{d}{dt}\begin{bmatrix} x(t,p) \\ \frac{\partial x(t,p)}{\partial p^i} \end{bmatrix} = \begin{bmatrix} A(p) & 0 \\ \frac{\partial A(p)}{\partial p^i} & A(p) \end{bmatrix} \begin{bmatrix} x(t,p) \\ \frac{\partial x(t,p)}{\partial p^i} \end{bmatrix} + \begin{bmatrix} B(p) \\ \frac{\partial B(p)}{\partial p^i} \end{bmatrix} u(t). \tag{4.1}$$

Solution methods based on the *Schur* transformation [Par.1] or the *Padé* approximation (see [Mol.1, Lau.1]) appear to be particularly appropriate for this case.

Acknowledgment:

The authors would like to thank prof. C.A. Desoer for helpful discussions.

Reference:

[Bec.1] R. G. Becker, "Linear System functionals via diagonalization", Imperial College, CCD report No: 79/10, Oct. 1979.

[Gan.1] F. R. Gantmacher, *The Theory of Matrices*, Vol I, II, Chelsea Publ. Co., New York, 1959.

[Lau.1] A. J. Laub, "Numerical Linear Algebra Aspects of Design Computations", *IEEE* AC-30, p97-108, Feb., 1985.

[Mar.1] J.E. Marsden, *Elementary Classical Analysis*, Freeman, 1974.

[Mol.1] C. Moler, C. Van Loan, "Nineteen Dubious Ways to Compute the Exponential of a Matrix", *SIAM Review*, p801-836, Oct. 1978.

[Par.1] B. N. Parlett, K. C. Ng, "Development of An Accurate Algorithm for EXP(BT)", *Center for Pure and Applied Mathematics*, PAM-294, University of California, Berkeley.

[Pol.1] E. Polak, D. Q. Mayne and D. M. Stimler, "Control System Design via Semi-Infinite Optimization", *Proceedings of the IEEE*, pp 1777-1795, December 1984.

[Tay.1] A.E. Taylor, D.C. Lay, *Introduction to Functional Analysis*, John Wiley & Sons, 1980.