An Input Normal Form Homotopy for the L^2 Optimal Model Order Reduction Problem

Yuzhen Ge, Emmanuel G. Collins,
Layne T. Watson, and L. D. Davis

TR 93-16

Department of Computer Science
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

May 1, 1993

An Input Normal Form Homotopy for the L^2 Optimal Model Order Reduction Problem

Y. GE, E. G. COLLINS, JR., MEMBER, IEEE,
L. T. WATSON, FELLOW, IEEE, AND L. D. DAVIS

***Abstract*—In control system analysis and design, finding a reduced order model, optimal in the L^2 sense, to a given system model is a fundamental problem. The problem is very difficult without the global convergence of homotopy methods, and a homotopy based approach has been proposed. The issues are the number of degrees of freedom, the well posedness of the finite dimensional optimization problem, and the numerical robustness of the resulting homotopy algorithm. A homotopy algorithm based on the input normal form characterization of the reduced order model is developed here and is compared with the homotopy algorithms based on Hyland and Bernstein's optimal projection equations. The main conclusions are that the input normal form algorithm can be very efficient, but can also be very ill conditioned or even fail.**

***Index Terms*—homotopy method, input normal form, optimal projection equations, parameter optimization, reduced order model problem.**

I. INTRODUCTION.

The L^2 optimal model reduction problem, i.e., the problem of approximating a higher order dynamical system by a lower order one so that a quadratic model reduction criterion is minimized, is of significant importance and is under intense study. Several earlier attempts to apply homotopy methods to the L^2 optimal model order reduction problem were not entirely satisfactory. Richter and Collins [9]–[11] devised a homotopy approach which only estimated certain crucial partial derivatives and employed relatively crude curve tracking techniques. Žigić, Bernstein, Collins, Richter, and Watson [14]–[16] formulated the problem so that numerical linear algebra techniques could be used to explicitly calculate partial derivatives, and employed sophisticated homotopy curve tracking algorithms, but the number of variables made large problems intractable. We propose here a method to reduce the dimension of the homotopy map so that large problems are computationally feasible. Alternative numerical algorithms can be found in [2].

The problem can be formulated as: given the asymptotically stable, controllable, observable, time invariant, continuous time system

$$\begin{aligned} \dot{x}(t) &= A\,x(t) + B\,u(t), \\ y(t) &= C\,x(t), \end{aligned} \tag{1}$$

Manuscript received , 1992; revised , 1993. The work of Ge and Watson was supported in part by Department of Energy grant DE-FG05-88ER25068, and Air Force Office of Scientific Research grant F49620-92-J-0236. The work of Collins and Davis was sponsored in part by NASA contract NAS8-38575 and Sandia National Laboratories contract 54-7609.

Y. Ge and L. T. Watson are with the Departments of Computer Science and Mathematics, Virginia Polytechnic Institute & State University, Blacksburg, VA 24061-0106.

E. G. Collins, Jr. and L. D. Davis are with Harris Corporation, Government Aerospace Systems Division, Melbourne, FL 32902.

where $A \in \mathbf{R}^{n\times n}$, $B \in \mathbf{R}^{n\times m}$, $C \in \mathbf{R}^{l\times n}$, the goal is to find a reduced order model

$$\begin{aligned} \dot{x}_m(t) &= A_m\, x_m(t) + B_m\, u(t), \\ y_m(t) &= C_m\, x_m(t), \end{aligned} \tag{2}$$

where $A_m \in \mathbf{R}^{n_m\times n_m}$, $B_m \in \mathbf{R}^{n_m\times m}$, $C_m \in \mathbf{R}^{l\times n_m}$, $n_m < n$ which minimizes the cost function

$$J(A_m, B_m, C_m) \equiv \lim_{t\to\infty} E\left[(y-y_m)^T R(y-y_m)\right], \tag{3}$$

where the input $u(t)$ is white noise with symmetric and positive definite intensity V and R is a symmetric and positive definite weighting matrix.

The optimal projection equations of Hyland and Bernstein [5], [6], described in [5], are basis independent and correspond to the maximum number of degrees of freedom one could plausibly use. Richter and Collins [11] use this maximum number, and Žigić [14] reduced it somewhat. At the other extreme, the minimum number of degrees of freedom corresponds to the input normal form described in Section II, and developed into a probability-one homotopy algorithm in Sections III and IV. Comparisons between the input normal form and the optimal projection equations approach are given in Section V.

II. Input normal form formulation.

The following theorem is needed to present the homotopy method for the input normal form.

Theorem 1 [7]. *Suppose $\bar{A}_m$ is asymptotically stable. Then for every minimal $(\bar{A}_m, \bar{B}_m, \bar{C}_m)$, i.e., $(\bar{A}_m, \bar{B}_m)$ is controllable and $(\bar{A}_m, \bar{C}_m)$ is observable, there exist a similarity transformation U and a positive definite matrix $\Omega = \mathrm{diag}(\omega_1, \cdots, \omega_{n_m})$ such that $A_m = U^{-1}\bar{A}_m U$, $B_m = U^{-1}\bar{B}_m$, and $C_m = \bar{C}_m U$ satisfy*

$$\begin{aligned} 0 &= A_m + A_m^T + B_m V B_m^T, \\ 0 &= A_m^T \Omega + \Omega A_m + C_m^T R C_m. \end{aligned} \tag{4}$$

In addition,

$$\begin{aligned} (A_m)_{ii} &= -\frac{1}{2}(B_m V B_m^T)_{ii}, \\ \omega_i &= \frac{(C_m^T R C_m)_{ii}}{(B_m V B_m^T)_{ii}}, \\ (A_m)_{ij} &= \frac{(C_m^T R C_m)_{ij} - \omega_j (B_m V B_m^T)_{ij}}{\omega_j - \omega_i}, \qquad \text{if } \omega_i \neq \omega_j. \end{aligned} \tag{5}$$

Definition 1. The triple (A_m, B_m, C_m) satisfying (4) or (5) is said to be in *input normal form.*

Note that generically $\omega_i \neq \omega_j$ for $i \neq j$, and this is assumed henceforth. Under the assumption that a solution (A_m, B_m, C_m) in input normal form is sought, the only independent variables are B_m and C_m, and in this case the domain is

$$\left\{(A_m, B_m, C_m):\ A_m \text{ is stable, } (A_m, B_m, C_m) \text{ is minimal and in input normal form}\right\}.$$

Assuming (A_m, B_m, C_m) is in input normal form, the cost function (3) can be written as

$$J(A_m, B_m, C_m) = \text{tr } (\tilde{Q}\tilde{R}) \tag{6}$$

where $\tilde{Q}$ is a symmetric and positive definite matrix satisfying

$$\tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^T + \tilde{V} = 0, \tag{7}$$

and

$$\tilde{A} = \begin{pmatrix} A & 0 \\ 0 & A_m \end{pmatrix}, \qquad \tilde{R} = \begin{pmatrix} C^T RC & -C^T RC_m \\ -C_m^T RC & C_m^T RC_m \end{pmatrix}, \qquad \tilde{V} = \begin{pmatrix} BVB^T & BVB_m^T \\ B_m VB^T & B_m VB_m^T \end{pmatrix}. \tag{8}$$

$\tilde{Q}$ can be written as

$$\tilde{Q} = \begin{pmatrix} \tilde{Q}_1 & \tilde{Q}_{12} \\ \tilde{Q}_{12}^T & \tilde{Q}_2 \end{pmatrix}, \tag{9}$$

where $\tilde{Q}_1 \in \mathbf{R}^{n\times n}$, $\tilde{Q}_{12} \in \mathbf{R}^{n\times n_m}$, and $\tilde{Q}_2 \in \mathbf{R}^{n_m\times n_m}$.

The goal of minimizing (6) under the constraints (4) and (7) leads to the Lagrangian

$$\begin{aligned} L(A_m, B_m, C_m, \Omega, \tilde{Q}) = \text{tr}\big[&\tilde{Q}\tilde{R} + (A_m + A_m^T + B_m V B_m^T) M_c \\ &+ (A_m^T \Omega + \Omega A_m + C_m^T RC_m) M_o + (\tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^T + \tilde{V})\tilde{P}\big], \end{aligned}$$

where the symmetric matrices M_o, M_c, and $\tilde{P}$ are Lagrange multipliers.

Setting $\partial L/\partial \tilde{Q} = 0$ gives

$$\tilde{A}^T\tilde{P} + \tilde{P}\tilde{A} + \tilde{R} = 0, \tag{10}$$

where $\tilde{P}$ is symmetric positive definite and can be partitioned as

$$\tilde{P} = \begin{pmatrix} \tilde{P}_1 & \tilde{P}_{12} \\ \tilde{P}_{12}^T & \tilde{P}_2 \end{pmatrix}. \tag{11}$$

$\partial L/\partial \Omega = 0$ and $\partial L/\partial A_m = 0$ yield

$$0 = 2M_c + 2\Omega M_o + 2(\tilde{P}_{12}^T\tilde{Q}_{12} + \tilde{P}_2\tilde{Q}_2), \qquad 0 = (A_m M_o)_{ii}, \quad 1 \le i \le n_m.$$

A straightforward calculation shows

$$\begin{aligned} \frac{\partial L}{\partial B_m} &= 2(\tilde{P}_{12}^T B + \tilde{P}_2 B_m)V + 2M_c B_m V, \\ \frac{\partial L}{\partial C_m} &= 2R(C_m\tilde{Q}_2 - C\tilde{Q}_{12}) + 2RC_m M_o. \end{aligned} \tag{12}$$

THEOREM 2 [3]. *The matrices M_c and M_o in (12) satisfy*

$$\begin{aligned} M_c &= -(\frac{1}{2}S + \Omega M_o), \\ (M_o)_{ii} &= -\frac{1}{(A_m)_{ii}} \sum_{\substack{j=1 \\ j\ne i}}^{n_m} (A_m)_{ij}(M_o)_{ji}, \\ (M_o)_{ij} &= \frac{(S)_{ij} - (S)_{ji}}{2(\omega_j - \omega_i)}, \quad \text{if } \omega_j \ne \omega_i, \end{aligned} \tag{13}$$

where

$$S = 2\big(\tilde{P}_{12}^T\tilde{Q}_{12} + \tilde{P}_2\tilde{Q}_2\big). \tag{14}$$

III. A HOMOTOPY APPROACH BASED ON THE INPUT NORMAL FORM.

A homotopy approach based on the input normal form is now described. Let A_f, B_f, C_f , R_f, and V_f denote A, B, C, R, and V in the above and define

$$\begin{aligned} A(\lambda) &= A_0 + \lambda(A_f - A_0), \\ B(\lambda) &= B_0 + \lambda(B_f - B_0), \\ C(\lambda) &= C_0 + \lambda(C_f - C_0), \end{aligned} \qquad \begin{aligned} R(\lambda) &= R_0 + \lambda(R_f - R_0), \\ V(\lambda) &= V_0 + \lambda(V_f - V_0). \end{aligned} \tag{15}$$

For brevity, $A(\lambda)$, $B(\lambda)$, $C(\lambda)$, $V(\lambda)$, and $R(\lambda)$ will be denoted by A, B, C, V, and R respectively in the following. Let

$$\begin{aligned} H_{B_m}(\theta,\lambda) &= \frac{\partial L}{\partial B_m} = 2\big(\tilde{P}_{12}^T B + \tilde{P}_2 B_m\big)V + 2M_c B_m V, \\ H_{C_m}(\theta,\lambda) &= \frac{\partial L}{\partial C_m} = 2R\big(C_m\tilde{Q}_2 - C\tilde{Q}_{12}\big) + 2RC_m M_o, \end{aligned}$$

where

$$\theta \equiv \begin{pmatrix} \text{Vec }(B_m) \\ \text{Vec }(C_m) \end{pmatrix}$$

denotes the independent variables B_m and C_m, M_o and M_c satisfy (13), and $\tilde{Q}$ and $\tilde{P}$ satisfy respectively (7) and (10) with partitioned forms (9) and (11). Vec(P) for a matrix $P \in \mathbf{R}^{p\times q}$ is the concatenation of its columns:

$$\text{Vec}(P) \equiv \begin{pmatrix} P_{\cdot 1} \\ P_{\cdot 2} \\ \vdots \\ P_{\cdot q} \end{pmatrix} \in \mathbf{R}^{p\times q}.$$

The probability-one homotopy map is defined as

$$\rho(\theta,\lambda) = \begin{pmatrix} \text{Vec }\big[H_{B_m}(\theta,\lambda)\big] \\ \text{Vec }\big[H_{C_m}(\theta,\lambda)\big] \end{pmatrix}, \tag{16}$$

and its Jacobian matrix is

$$D\rho(\theta,\lambda) = \big(D_\theta\rho(\theta,\lambda), D_\lambda\rho(\theta,\lambda)\big). \tag{17}$$

The vector $\big(\text{Vec }(A_0), \text{Vec }(B_0), \text{Vec }(C_0), \text{Vec }(V_0), \text{Vec }(R_0)\big)$ plays the role of the parameter vector in the probability-one homotopy theory [13]. Define

$$\begin{aligned} \hat{H}_{B_m}\big(\tilde{P}^{(j)}, M_c^{(j)}\big) &= 2\big(\tilde{P}_{12}^{T(j)} B + \tilde{P}_2^{(j)} B_m\big)V + 2M_c^{(j)} B_m V, \\ \hat{H}_{C_m}\big(\tilde{Q}^{(j)}, M_o^{(j)}\big) &= 2R\big(C_m\tilde{Q}_2^{(j)} - C\tilde{Q}_{12}^{(j)}\big) + 2RC_m M_o^{(j)}, \end{aligned}$$

where the superscript (j) means $\partial/\partial\theta_j$: $Y^{(j)} \equiv \frac{\partial Y}{\partial \theta_j}$. Using the above definitions, we have for $\theta_j = \left(B_m\right)_{kl}$,

$$\begin{aligned}
\frac{\partial H_{B_m}}{\partial (B_m)_{kl}} &= \hat{H}_{B_m}(\tilde{P}^{(j)}, M_c^{(j)}) + 2(\tilde{P}_2 + M_c)E^{(k,l)}V, \\
\frac{\partial H_{C_m}}{\partial (B_m)_{kl}} &= \hat{H}_{C_m}(\tilde{Q}^{(j)}, M_o^{(j)}),
\end{aligned} \tag{18}$$

and for $\theta_j = \left(C_m\right)_{kl}$,

$$\begin{aligned}
\frac{\partial H_{B_m}}{\partial (C_m)_{kl}} &= \hat{H}_{B_m}(\tilde{P}^{(j)}, M_c^{(j)}), \\
\frac{\partial H_{C_m}}{\partial (C_m)_{kl}} &= \hat{H}_{C_m}(\tilde{Q}^{(j)}, M_o^{(j)}) + 2RE^{(k,l)}(\tilde{Q}_2 + M_o),
\end{aligned} \tag{19}$$

where $E^{(k,l)}$ is a matrix of the appropriate dimension whose only nonzero element is $e_{kl} = 1$. $\tilde{P}^{(j)}$ and $\tilde{Q}^{(j)}$ can be obtained by solving the Lyapunov equations

$$\begin{aligned}
0 &= \tilde{A}^{(j)}\tilde{Q} + \tilde{A}\tilde{Q}^{(j)} + \tilde{Q}^{(j)}\tilde{A}^T + \tilde{Q}\tilde{A}^{T(j)} + \tilde{V}^{(j)}, \\
0 &= \tilde{A}^{T(j)}\tilde{P} + \tilde{A}^T\tilde{P}^{(j)} + \tilde{P}^{(j)}\tilde{A} + \tilde{P}\tilde{A}^{(j)} + \tilde{R}^{(j)}.
\end{aligned} \tag{20}$$

Similarly for λ, using a dot to denote $\partial/\partial\lambda$,

$$\begin{aligned}
\frac{\partial H_{B_m}}{\partial \lambda} &= \hat{H}_{B_m}(\dot{\tilde{P}}, \dot{M}_c) + 2\tilde{P}_{12}^T(\dot{B}V + B\dot{V}) + 2(\tilde{P}_2 + M_c)B_m\dot{V}, \\
\frac{\partial H_{C_m}}{\partial \lambda} &= \hat{H}_{C_m}(\dot{\tilde{Q}}, \dot{M}_o) + 2\dot{R}C_m(\tilde{Q}_2 + M_o) - 2(\dot{R}C + R\dot{C})\tilde{Q}_{12},
\end{aligned} \tag{21}$$

where $\dot{\tilde{P}}$ and $\dot{\tilde{Q}}$ are obtained by solving the Lyapunov equations

$$\begin{aligned}
0 &= \dot{\tilde{A}}\tilde{Q} + \tilde{A}\dot{\tilde{Q}} + \dot{\tilde{Q}}\tilde{A}^T + \tilde{Q}\dot{\tilde{A}}^T + \dot{\tilde{V}}, \\
0 &= \dot{\tilde{A}}^T\tilde{P} + \tilde{A}^T\dot{\tilde{P}} + \dot{\tilde{P}}\tilde{A} + \tilde{P}\dot{\tilde{A}} + \dot{\tilde{R}}.
\end{aligned}$$

IV. Numerical algorithm for input normal form homotopy.

The initial point $(\theta, \lambda) = (\theta_0, 0) = \left((B_m)_0, (C_m)_0, 0\right)$ is chosen so that the triple $\left((A_m)_0,\right.$ $(B_m)_0$, $\left.(C_m)_0\right)$ is in input normal form and satisfies $\rho(\theta_0, 0) = 0$. In the following algorithm $V_0 = V_f$ and $R_0 = R_f$ are used.

Theorem 3 [8]. *Suppose $\bar{A}$ is asymptotically stable. Then for every minimal $(\bar{A}, \bar{B}, \bar{C})$, i.e., $(\bar{A}, \bar{B})$ is controllable and $(\bar{A}, \bar{C})$ is observable, there exist a similarity transformation T and a positive definite matrix $\Lambda = \operatorname{diag}(d_1, d_2, \cdots, d_n)$ with $d_i \geq d_{i+1}$ such that $A = T^{-1}\bar{A}T$, $B = T^{-1}\bar{B}$, and $C = \bar{C}T$ satisfy*

$$\begin{aligned}
0 &= A\Lambda + \Lambda A^T + BVB^T, \\
0 &= A^T\Lambda + \Lambda A + C^T RC.
\end{aligned}$$

DEFINITION 2. The triple (A, B, C) in the above theorem is *balanced.*

According to Moore [8], under certain conditions, the leading principal $n_m \times n_m$ block of A, the leading principal $n_m \times m$ block of B, and the leading principal $l \times n_m$ block of C in balanced form are good approximations to the reduced order model. This suggests that the initial point $(\theta_0, 0)$ be chosen as follows:

1) Transform the given triple (A_f, B_f, C_f) to balanced form (A_b, B_b, C_b).
2) Partition (A_b, B_b, C_b) as

$$A_b = {}^{n_m}\{ \begin{pmatrix} \overbrace{A_{11}}^{n_m} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad B_b = {}^{n_m}\{ \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \quad C_b = \begin{pmatrix} \overbrace{C_1}^{n_m} & C_2 \end{pmatrix}.$$

3) (A_0, B_0, C_0) is chosen as

$$A_0 = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}, \quad B_0 = \begin{pmatrix} B_1 \\ 0 \end{pmatrix}, \quad C_0 = \begin{pmatrix} C_1 & 0 \end{pmatrix}.$$

4) The initial point for the reduced order model is chosen as

$$\bar{\theta}_0 = \begin{pmatrix} \text{Vec } (\bar{B}_m)_0 \\ \text{Vec } (\bar{C}_m)_0 \end{pmatrix} = \begin{pmatrix} \text{Vec } B_1 \\ \text{Vec } C_1 \end{pmatrix},$$

and $(\bar{A}_m)_0 = A_{11}$ by construction.

5) Transform the initial point $((\bar{A}_m)_0, (\bar{B}_m)_0, (\bar{C}_m)_0)$ to input normal form so that the initial reduced order model is

$$((A_m)_0, (B_m)_0, (C_m)_0) = (T^{-1}(\bar{A}_m)_0 T, \quad T^{-1}(\bar{B}_m)_0, \quad (\bar{C}_m)_0 T).$$

The initial point for the homotopy map is then $(\theta_0, 0)$, where

$$\theta_0 = \begin{pmatrix} \text{Vec } (B_m)_0 \\ \text{Vec } (C_m)_0 \end{pmatrix}.$$

(In general, the truncation to obtain the approximate reduced order model should be based on the component costs instead of on the sizes of the balanced gains d_i as done above [12]. This explains why in some cases the above algorithm for choosing the initial points did not lead to a reduced order model with a minimal cost.)

Once the initial point is chosen, the rest of the computation is as follows:

1) Set $\lambda := 0$, $\theta := \theta_0$.
2) Calculate A_m from (5), $\tilde{R}$, $\tilde{V}$, and compute $\tilde{Q}$ and $\tilde{P}$ according to (7) and (10).
3) Evaluate S from (14) and M_o and M_c according to (13).
4) Evaluate the homotopy map $\rho(\theta, \lambda)$ in (16) and $D\rho(\theta, \lambda)$ in (17).
5) Predict the next point $Z^{(0)} = (\theta^{(0)}, \lambda^{(0)})$ on the curve γ.
6) For $k := 0, 1, 2, \cdots$ until convergence do

$$Z^{(k+1)} = [D\rho(Z^{(k)})]^{\dagger} \rho(Z^{(k)}),$$

where $[D\rho(Z)]^\dagger$ is the Moore-Penrose inverse of $D\rho(Z)$. Let $(\theta_1, \lambda_1) = \lim_{k\to\infty} Z^{(k)}$.

7) If $\lambda_1 < 1$, then set $\theta := \theta_1$, $\lambda := \lambda_1$, and go to step 2).
8) If $\lambda_1 \geq 1$, compute the solution $\bar{\theta}$ at $\lambda = 1$. A_m is then obtained from (5).

An alternative strategy for choosing an initial point is as follows:

1) Modify A_f to $A'_f = c_1 I + c_2 A_f$, where $c_1 \leq 0$ and $c_2 \geq 0$.
1) Transform (A'_f, B_f, C_f) to balanced form and choose (A'_0, B'_0, C'_0) as before.
3) Compute the initial reduced order model $((A_m)_0, (B_m)_0, (C_m)_0)$ from the triple (A'_0, B'_0, C'_0) as before.

When $c_1 = 0$, $c_2 = 1$, this strategy reduces to the previous one. For some problems, our numerical experiments show that HOMPACK reaches $\lambda > 1$ in fewer steps with $c_1 \neq 0$ than with $c_1 = 0$. A modification to the homotopy map $\rho(\theta, \lambda)$ in (16) is

$$\rho_1(\theta, \lambda) = \lambda\rho(\theta, \lambda) + (1 - \lambda)(\theta - \theta_0),$$

where θ_0 denotes the initial value of θ at $\lambda = 0$. For some problems this homotopy map can be more efficient than the one in (16), while in other cases it can be less efficient.

V. Comparisons and Discussions.

The input normal form algorithm developed here was applied to Systems 1 through 9 in [15]. It successfully solved all of the problems except the following system:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -2 & -0.02 & 1 & 0.01 \\ 0 & 0 & 0 & 1 \\ 0.1 & 0.001 & -0.1 & -0.001 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix}.$$

For this problem, with the input normal form, when $n_m = 2, 3$, two of the initial ωs are approximately the same, which leads to a significant numerical error in computing M_o and the numerical failure of the homotopy algorithm. Therefore this technique for choosing initial points fails, and some modification to the algorithm is needed to avoid this kind of ill conditioning. However, it is not at all clear how to systematically avoid nearly equal ωs, and this remains an open question. It can be shown that the solutions, obtained by the optimal projection equation approach, also have close ωs, which implies that changing the strategy for choosing initial points will not suffice for this example.

For a given order n_m, the set

$$N = \{(A_m, B_m, C_m) : A_m \text{ is stable}, (A_m, B_m, C_m) \text{ is minimal and in input normal form}\}$$

is an *open* set. Therefore $J(A_m, B_m, C_m)$ may not *attain a minimum value* over this open set. e.g., if the optimal model of order n_m cannot be represented in input normal form, then J achieves its minimum on the *boundary* of N. The homotopy then, embodying the input normal form parametrization, must become more and more ill conditioned as the zero curve approaches the boundary of N. The starting point, form of homotopy map, and numerical algorithms used are irrelevant – the computation must eventually fail close enough to the boundary of N. Note that, contrary to what a minimal parametrization tacitly assumes, J *need not attain a minimum value with a particular structure for* (A_m, B_m, C_m).

Table 1 gives the comparison of the optimal projection equations approach and the input normal form formulation for System 8 (4th order) and System 9 (7th order) in [15]. System 8 [8] is given by

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -50 & -79 & -33 & -5 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad C = (50 \quad 15 \quad 1 \quad 0).$$

System 9 [4] is given by

$$A = \begin{pmatrix} -6.2036 & 15.054 & -9.8726 & -376.58 & 251.32 & -162.24 & 66.827 \\ 0.53 & -2.0176 & 1.4363 & 0 & 0 & 0 & 0 \\ 16.846 & 25.079 & -43.555 & 0 & 0 & 0 & 0 \\ 377.4 & -89.449 & -162.83 & 57.998 & -65.514 & 68.579 & 157.57 \\ 0 & 0 & 0 & 107.25 & -118.05 & 0 & 0 \\ 0.36992 & -0.14445 & -0.26303 & -0.64719 & 0.49947 & -0.21133 & 0 \\ 0 & 0 & 0 & 0 & 0 & 376.99 & 0 \end{pmatrix},$$

$$B = \begin{pmatrix} 89.353 & 0 \\ 376.99 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0.21133 \\ 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The data in Table 1 is typical of that for all systems tested. The time is CPU time on a DECstation 5000/200, and the steps are function evaluations along the homotopy zero curve, not the number of 2)–7) loop iterations, which is usually much smaller.

TABLE 1. COMPARISON OF METHODS.

	System 8			
	Optimal projection		input normal form	
n_m	# steps	time (sec)	# steps	time (sec)
1	31	0.6	10	0.20
2	59	2.7	18	0.50
3	89	14.	10	0.65
	System 9			
2	575	88	123	8.0
3	601	223	6	1.3
4	671	518	6	1.9

The optimal projection equations homotopy successfully solved *all* of the test problems, but Table 1, containing typical results, shows that the input normal form homotopy is much more efficient. However, when the input normal form is used, some restrictions are imposed on the structure of the triple (A_m, B_m, C_m), potentially resulting in ill conditioning. For the input normal form formulation, ill conditioning occurs if two diagonal elements of Ω in (4) are approximately the same. In other words, let Q_m and P_m be the controllability and observability Gramians of the system represented by (A_m, B_m, C_m), and let

$$Q_m = W\Sigma W^T, \quad P_m = W^{-T}\Sigma W^{-1},$$

where Σ is diagonal and is the controllability and observability Gramian in balanced form. If two diagonal elements of Σ are approximately the same, then ill conditioning occurs. For the example that input normal form fails, when $n_m = 2, 3$, both the initial point chosen using the given strategy and the solution obtained in [14]–[16] are ill conditioned, i.e., two diagonal elements of Ω are approximately the same. Hence the input normal form method will not be able to solve this problem; this has nothing to do with the initial points chosen or the particular homotopy maps used, but rather is an inherent failure of the input normal form parametrization.

REFERENCES

[1] D. S. Bernstein, private communication.

[2] A. E. Bryson and A. Carrier, "Second-order algorithm for optimal model order reduction," *AIAA Jour. Guidance, Control, & Dynamics,* vol. 13-5, pp. 887–892, 1990.

[3] L. D. Davis, E.G. Collins, Jr., and S. A. Hodel, "A parameterization of minimal plants," *Proc. 1992 American Contr. Conf.,* Chicago, IL, pp. 355–356, June 1992.

[4] J. Hickin and N. K. Sinha, "Model reduction for linear multivariable systems," *IEEE Trans. Autom. Contr.,* vol. AC-25, pp. 1121–1127, 1980.

[5] D. C. Hyland and D. S. Bernstein, "The optimal projection equations for fixed-order dynamic compensation," *IEEE Trans. Autom. Contr.,* vol. AC-29, pp. 1034–1037, 1984.

[6] ———, "The optimal projection equations for model reduction and the relationships among the methods of Wilson, Skelton, and Moore," *IEEE Trans. Autom. Contr.,* vol. AC-30, pp. 1201–1211, 1985.

[7] P. T. Kabamba, "Balanced forms: canonicity and parameterization," *IEEE Trans. Autom. Contr.,* vol. AC-30, pp. 1106–1109, 1985.

[8] B. C. Moore, "Principal component analysis in linear systems: controllability, observability, and model reduction," *IEEE Trans. Autom. Contr.,* vol. AC-26, pp. 17–32, 1981.

[9] S. Richter, "Reduced-order control design via the optimal projection approach: A homotopy algorithm for global optimality," *Proc. Sixth VPI&SU Symp. Dyn. Contr. Large Str.,* Blacksburg, VA, June 1987.

[10] ———, "A homotopy algorithm for solving the optimal projection equations for fixed-order dynamic compensation: Existence, convergence and global optimality," *Proc. American Contr. conf.,* Minneapolis, MN, pp. 1527-1531, June 1987.

[11] S. Richter and E. G. Collins, Jr., "A homotopy algorithm for reduced order compensator design using the optimal projection equations," *Proc. IEEE Conf. Dec. Contr.,* Tampa, FL, pp. 506–511, Dec. 1989.

[12] R. E. Skelton and P. Kabamba, "Comments on 'Balanced gains and their significance for L^2 model reduction'," *IEEE Trans. Autom. Contr.,* vol. AC–31, pp. 796–797, 1986.

[13] L. Watson, S. C. Billups and A. P. Morgan, "HOMPACK: a suite of codes for globally convergent homotopy algorithms," *ACM Trans. Math. Software,* vol. 13, pp. 281–310, 1987.

[14] D. Žigić, "Homotopy methods for solving the optimal projection equations for the reduced order model problem," M.S. thesis, Dept. of Computer Science, Virginia Polytechnic Institute and State Univ., Blacksburg, VA, June 1991.

[15] D. Žigić, L. Watson, E. G. Collins, Jr., and D. S. Bernstein, "Homotopy approaches to the H^2 reduced order model problem," *J. Math. Systems, Estimation, and Control,* to appear.

[16] ———, "Homotopy methods for solving the optimal projection equations for the H^2 reduced order model problem," TR 91-13, Dept. of Computer Science, Virginia Polytechnic Institute and State Univ., 1991.