beneficial to apply the proposed algorithm to obtain more accurate results (by the parsimony principle) and more efficient computations than is possible with the usual unconstrained models. In the white noise case, the algorithm becomes a recursive prediction error method, and as such its covariance attains the Cramér–Rao lower bound asymptotically when the true order is used (see, e.g., [3]). In the more general nonwhite noise case, the algorithm’s covariance can be evaluated using methods described, e.g., in [2, ch. 7]. This algorithm can be used for filter design and adaptive Nyquist rate estimation. The basic method used here can be applied for deriving system identification algorithms for other constrained transfer functions, such as band-pass and band-stop. Extension to adaptive parameter estimation of constrained ARMA signals with unknown inputs in the presence of noise is presented in [12].

APPENDIX

THE DERIVATIVE OF POLYNOMIAL COEFFICIENTS WITH RESPECT TO POLYNOMIAL ZEROS

In this Appendix we provide a simple proof of the formula (13). Let $C$ denote the following companion matrix:

\[ C = \begin{bmatrix}
-a_0 & 1 & 0 \\
-a_1 & 0 & 1 \\
\vdots & \vdots & \vdots \\
-a_n & 0 & 0 \\
\end{bmatrix} \]

associated with the polynomial $z^rA(z^{-1}) = z^n + a_1z^{n-1} + \cdots + a_n$. It is well known that the zeros $\lambda_k$ of $z^rA(z^{-1})$ are equal to the eigenvalues of $C$ (see, e.g., [11]). Let

\[ u_k = [u_{i,k} \cdots u_{n,k}]^T \neq 0 \]

denote a (nonzero) eigenvector corresponding to $\lambda_k$, thus,

\[ Cu_k = \lambda_k u_k \tag{A.1} \]

or, in a more detailed form

\[
\begin{aligned}
&u_{1,k} = a_1 u_{1,k} + \lambda_k u_{2,k} \\
&\vdots \\
&u_{n,k} = a_{n-1} u_{n-1,k} + \lambda_k u_{n,k} \\
\end{aligned}
\]

\[ 0 = a_n u_{n,k} + \lambda_k u_{n,k} \tag{A.2} \]

It readily follows by contradiction from (A.2) that $u_k \neq 0$ implies $u_{n,k} \neq 0$. Thus, we can set $u_{n,k} = 1$. In the following we assume that the eigenvector $u_k$ has been normalized such that $u_{n,k} = 1$.

Using forward substitution, we find from (A.2) that

\[ u_k = H v_k \tag{A.3} \]

where $H$ is the Hankel matrix

\[ H = \begin{bmatrix}
1 \\
0 & 1 \\
\vdots & \vdots & \vdots \\
0 & \cdots & 0 & 1
\end{bmatrix} \tag{A.4} \]

and

\[ u_k = [\lambda_k^{-1} \cdots \lambda_k]^T \tag{A.5} \]

Next, it can easily be verified that

\[ v_k^T C = \lambda_k v_k^T \tag{A.6} \]

which means that $u_k$ is a left eigenvector of $C$. Left and right eigenvectors associated with different eigenvalues must be orthogonal

\[ v_i^T u_k = 0 \quad \text{for} \; i \neq k \tag{A.7} \]

which can be seen as follows. From (A.1) and (A.6), we get

\[ v_i^T u_k - \lambda_k v_i^T u_k = v_i^T Cu_k - v_i^T Cu_k = 0 \]

which implies (A.7) since $\lambda_i \neq \lambda_k$. Writing out (A.7) for $i = 1, \ldots, n$, $i \neq k$, we obtain

\[ \lambda_i v_k^T u_i - \lambda_k v_k^T u_i = v_i^T Cu_k - v_i^T Cu_k = 0, \tag{A.8} \]

Since $\{\lambda_k\}$ are distinct by assumption, the vander Monde matrix appearing in (A.5) is nonsingular. Therefore, $u_k$ is uniquely determined by $\{\lambda_1, \ldots, \lambda_{k-1}, \lambda_{k+1}, \ldots, \lambda_n\}$ and, in particular, does not depend on $\lambda_k$. Using this property we get by differentiating (A.2) with respect to $\lambda_k$

\[ \frac{\partial u}{\partial \lambda_k} = -u_k \tag{A.9} \]

From (A.3) and (A.9) the Jacobian matrix $\frac{\partial u}{\partial \lambda}$ is equal to

\[ \frac{\partial u}{\partial \lambda_k} = -[u_1 \cdots u_n] = -HV \tag{A.10} \]

where $H$ was defined in (A.4) and $V$ is the vander Monde matrix $V = [v_1 \cdots v_n]$. Expression (15) is now proven immediately from the $i, k$ entry of the matrices in (A.10).

REFERENCES


Persistency of Excitation Results for Structured Nonminimal Models

MICHAEL HEYMANN

Abstract—It is frequently convenient to employ specially structured nonminimal models for parameter estimation such as in the case of direct

Manuscript received February 3, 1987. This work was supported in part by the Technion Fund for promotion of research.

The author is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa, Israel.

IEEE Log Number 8716170.
adaptive control where the model is parameterized in terms of the desired control law parameters. This is done, for example, in direct model reference adaptive control and in direct pole assignment adaptive control algorithms. It is shown that the parameters appearing in these nonminimal models can be uniquely estimated if and only if a certain design identity has a unique solution. The result is used to develop persistency of excitation results for these models.

I. INTRODUCTION

The problem of persistency of excitation for parameter convergence in estimation and adaptive control received a great deal of attention in the recent literature (see, e.g., [1]-[9]).

While in many cases minimal models (which have no redundancies through common factors) are used for estimation of parameters, there are situations where the employment of nonminimal models is necessary. This occurs typically in direct adaptive control applications where plant models are specially structured to be parameterized by the adaptive controller parameters. Examples of these instances are direct model reference adaptive control [10], direct pole assignment adaptive control [11], direct model reference adaptive pole assignment [12], as well as in multivariable systems where a simple left MFD is employed having a diagonal "denominator" matrix.

In these nonminimal models the question arises as to whether or not the parameters can be uniquely estimated from the plant input-output data. This often has important implications, e.g., in deciding whether an adaptive design is stable and convergent and in choosing suitable inputs that guarantee the desired stability and convergence.

In the present note we show that the parameters in structured (generally nonminimal) models can be uniquely determined if and only if a certain design identity has a unique solution. The latter is shown to be related to the output-reachability of an associated-signal system which in turn is used to develop persistency of excitation results.

The note focuses on single-input single-output systems but similar results can be developed for the multivariable multioutput case.

II. A GENERAL PARAMETERIZATION PROBLEM

We consider discrete time, time-invariant single-input single-output plants of the form

\[ p(D)y(t) = r(D)u(t) \]  
(2.1)

where \( p(D) \) and \( r(D) \) are real polynomials in the unit delay operator \( D \) [i.e., \( D^nx(t) = x(t - n) \)] of the form

\[ p(D) = 1 + \sum_{i=1}^{n} p_i D^i \]  
(2.2)

\[ r(D) = \sum_{i=1}^{n} r_i D^i \]  
(2.3)

We assume that the model (2.1) of the plant is minimal, that is, the polynomials \( p(D) \) and \( r(D) \) are coprime.

As a key step in setting up the parameter estimation problem we replace the minimal model (2.1) by a structured nonminimal model of the form

\[ g(D)y(t) = h(D)u(t) \]  
(2.4)

where the polynomials \( g(D) \) and \( h(D) \) are assumed to be parameterized as follows:

\[ g(D) = c(D) + \sum_{i=1}^{m_1} c_i d_i(D) \]  
(2.5)

\[ h(D) = d(D) + \sum_{i=1}^{m_2} \delta_i h_i(D) \]  
(2.6)

where \( m_1 \) and \( m_2 \) are positive integers, where \( a_i, \ldots, a_{m_1}, \beta_i, \ldots, \beta_{m_2} \) are real parameters and where

\[ a_i(D) = \sum_{k=1}^{i} a_k D^k, \quad i = 1, \ldots, m_1; \quad b_i(D) = \sum_{k=1}^{i} b_k D^k, \quad i = 1, \ldots, m_2; \quad c(D) = \sum_{i=1}^{\epsilon} c_i D^i, \quad d(D) = \sum_{i=1}^{\delta} d_i D^i. \]

It is assumed that the polynomials \( a_i(D), \quad i = 1, \ldots, m_1 \) are linearly independent over the reals, i.e., that there exists no nontrivial set of constants \( \gamma_1, \ldots, \gamma_{m_1} \) such that \( \sum_{i=1}^{\epsilon} \gamma_i c_i D^i = 0 \) (the zero polynomial).

It is similarly assumed that the \( b_j(D), \quad j = 1, \ldots, m_2 \) are linearly independent. These assumptions imply that \( m_1, m_2 \leq \epsilon \).

We note that for (2.4) to represent (in a generally nonminimal way) the plant (2.1) it is necessary (and sufficient) that \( g(D) = k(D)p(D) \) and \( h(D) = k(D)r(D) \) for some polynomial \( k(D) \).

We illustrate the parameterization (2.4) for two cases of interest in adaptive control.

Example 2.1 (Direct Model Reference Adaptive Control): We assume that in (2.3) \( \epsilon = 0 \) for \( i < d, \eta = 0 \) [if usually called the relative degree of (2.1)]. The reference model for the closed-loop plant is given by

\[ p^*(D)r^*(D)/D^q(t) \]  
(2.7)

where \( p^*(D) = 1 + \sum_{i=1}^{\epsilon} \rho_i D^i \) is a prespecified (stable) polynomial and where \( r^*(D) \) and \( u(t) \) are the reference model output and the command input, respectively. The control law is of the form

\[ u(t) = -F(D)y(t) + q(D)u(t) \]  
(2.8)

where \( g(D) = 1 + \sum_{i=1}^{\epsilon} \gamma_i D^i \) is a prespecified (stable) polynomial and where the polynomials \( E(D) = \sum_{i=1}^{\epsilon} e_i D^i \) and \( F(D) = \sum_{i=1}^{\epsilon} f_i D^i \) as well as the constant \( \gamma_i \) are to be determined so as to satisfy the closed-loop performance specification, i.e., (2.7).

Substitution of (2.8) in (2.1) to eliminate \( u(t) \) and equating the resultant expression with (2.7) yields the following nonminimal parameterization for (2.4):

\[ g(D) = p^*(D)q(D) - D^qF(D) \]  
(2.9)

\[ h(D) = D^q \left[ \frac{1}{\gamma_i} q(D) + E(D) \right] \]  
(2.10)

Equating (2.9) and (2.10) with (2.5) and (2.6), respectively, gives \( m_1 = n, m_2 = n + 1, a_i(D) = b_i(D) = D^{i+1}, i = 1, \ldots, n, b_i(D) = D^i q(D), c(D) = q(D)p^*(D) \) and \( d(D) = 0 \). The parameters to be estimated in the resultant model are \( \alpha_i = -\beta_i, \beta_i = e_i, i = 1, \ldots, n \) and \( \beta_{n+1} = 1/\gamma_i \). Note that the \( \alpha_i \) and \( \beta_i \) of the nonminimal model directly parameterize the model reference control law.

Example 2.2 (Direct Pole-Assignment Adaptive Control): Since \( p(D) \) and \( r(D) \) are coprime, there exist polynomials \( \gamma(D) = 1 + \sum_{i=1}^{\epsilon} \gamma_i D^i, b(D) = \sum_{i=1}^{\epsilon} b_i D^i \), \( r^*(D) = 1 + \sum_{i=1}^{\epsilon} \rho_i D^i \), \( \delta(D) = \sum_{i=1}^{\epsilon} \delta_i D^i \), \( \rho(D) = 1 + \sum_{i=1}^{\epsilon} \rho_i D^i \) and \( e(D) = \sum_{i=1}^{\epsilon} e_i D^i \) such that

\[ \gamma(D)p(D) + b(D)r(D) = p^*(D)q(D) \]  
(2.11)

\[ \rho(D)p(D) + e(D)r(D) = 1 \]  
(2.12)

where \( p^*(D) = 1 + \sum_{i=1}^{\epsilon} \rho_i D^i \) and \( q(D) = 1 + \sum_{i=1}^{\epsilon} \gamma_i D^i \) are prespecified stable polynomials.

Multiplying (2.1) by \( \gamma(D)y(t) \) and using (2.11) gives

\[ \gamma(D)p^*(D)q(D)y(t) = \gamma(D)(D)b(D)y(t) + \gamma(D)(D)u(t) \]  
(2.13)

Similarly, multiplying (2.1) by \( \rho(D) \cdot b(D) \) and using (2.11) gives

\[ \rho(D)p^*(D)q(D)u(t) = \rho(D)p(D)b(D)y(t) + \gamma(D)(D)u(t) \]  
(2.14)

Adding (2.13) and (2.14), and making use of (2.12) gives the following
nonminimal model of the plant:
\[ [(D) - p(D)q(D)\nu(D)]y(t) = [p(D)q(D)\rho(D) - \gamma(D)]u(t). \] (2.15)
This model is of the form (2.4) with the polynomials \( g(D) \) and \( h(D) \) parameterized as
\[ g(D) = b(D) - p(D)q(D)\nu(D) \] (2.16)
\[ h(D) = \rho(D)q(D)\rho(D) - \gamma(D). \] (2.17)
Comparing (2.16) and (2.17) to (2.5) and (2.6), respectively, gives
\[ c(D) = -p(D)q(D); \quad d(D) = \rho(D)q(D). \] (2.18)
\[ m_1 = m_2 = 2n, a_i(D) = b_i(D) = D; \quad i = 1, \ldots, n, a_i(D) = b_i(D) = d^{-i}p(D)q(D); \quad i = n + 1, \ldots, 2n, \quad a_i = b_i = -\gamma, \quad i = 1, \ldots, n, \quad \text{and} \quad a_i = b_i = \rho, \quad i = n + 1, \ldots, 2n. \]
Note that the common factor \( k(D) \) of \( g(D) \) and \( h(D) \) in this case is the polynomial \([\nu(D)\gamma(D) - \rho(D)\theta(D)]\). Following a procedure similar to the development in Example 2.1, it is readily verified that the nonminimal model (2.15) is obtained by using (2.1) the control law
\[ y(D)u(t) = -b(D)y(t) + q(D)u(t) \] (2.19)
which yields upon making use of (2.11) the closed-loop equation
\[ p(D)y(t) = r(D)u(t). \] (2.20)
Thus, the parameters \( \gamma \) and \( \delta \) directly parameterize the pole assignment control law.

III. CONDITIONS FOR UNIQUENESS OF PARAMETERS

Equation (2.4) with \( g(D) \) and \( h(D) \) given by (2.5) and (2.6) constitutes a model for the plant (2.1) if and only if the following design identity is satisfied
\[ \left[ c(D) + \sum_{i=1}^{n} a_i(D) \right] r(D) = \left[ d(D) + \sum_{i=1}^{n} b_i(D) \right] p(D). \] (3.1)

Since we wish to use the (possibly nonminimal) model (2.4) to uniquely estimate the parameters \( a_1, \ldots, a_n, b_1, \ldots, b_n \) from the plant input-output data, we conclude that a necessary and sufficient condition for existence and uniqueness of these parameters is that (3.1) is solvable by a unique vector
\[ \theta^* = [a_1, \ldots, a_n, b_1, \ldots, b_n]^T. \] (3.2)

For parameter estimation purposes it is convenient to represent the model (2.4) by a linear regression equation
\[ c(D)y(t) - d(D)u(t) = \phi(t)\theta^* \] (3.3)
where
\[ \phi(t) = [-a_1(D)y(t), \ldots, -a_n(D)y(t), b_1(D)u(t), \ldots, b_n(D)u(t)]. \] (3.4)

One can now use standard estimation procedures \([10]\) to estimate the parameter vector \( \theta^* \) using input-output data from the plant. For example, one could employ the recursive least squares (RLS) algorithm. Consider now a sequence of vectors \( \{\phi(t)\}_{t=1}^{\infty} \), \( \phi(t) \in \mathbb{R}^m \) and denote by \( \Phi_{n,k}(k) \) the \( m \times N \) matrix
\[ \Phi_{n,k}(k) := [\phi(k+1), \ldots, \phi(k+N)]. \] (3.5)
We say that \( \{\phi(t)\} \) is spanning (of order \( N \)) if there exists a number \( \epsilon > 0 \) and integers \( k \) and \( N \) such that
\[ \lambda_{\min}(\Phi_{n,k}(k)\Phi_{n,k}(k)^T) \geq \epsilon. \] (3.6)
The sequence will be called persistently spanning (of order \( N \)) if there exists a sequence of positive numbers \( \{\epsilon_i\}_{i=1}^{\infty} \) and positive integers \( k \) and \( N \) such that for all \( i \geq 0 \)
\[ \lambda_{\min}(\Phi_{n,k}(k+i\nu)\Phi_{n,k}(k+i\nu)^T) \geq \epsilon_i \] (3.7)
and it will be called uniformly persistently spanning (of order \( N \)) if for all \( i \geq 0 \), the \( \epsilon_i \) in (3.7) can be chosen so that \( \epsilon_i > \epsilon \) for some \( \epsilon > 0 \).
For a scalar sequence \( \{\epsilon_i\}_{i=1}^{\infty} \) we first construct an associated m-vector \( \delta_{\alpha}(t) \) as follows:
\[ \delta_{\alpha}(i) = [u(i+1), u(i+2), \ldots, u(i+m)]^T. \] (3.8)
We then denote by \( U_{n,k}(k) \) the \( m \times N \) matrix
\[ U_{n,k}(k) := [\delta_{\alpha}(k+1), \ldots, \delta_{\alpha}(k+N)]. \] (3.9)
The spanning properties of \( U_{n,k}(k) \) are defined just as for \( \Phi_{n,k}(k) \).

Now it is known \([10]\) that the RLS algorithm with covariance resetting yields a sequence of estimates that converges exponentially fast to \( \theta^* \), provided the sequence \( \{\phi(t)\} \) of regression vectors is uniformly persistently spanning. It is further known \([1], [3], [10]\) that the persistent spanning property of \( \{\phi(t)\} \) can be generated from the input \( u(t) \) provided a certain reachability property of \( \{\phi(t)\} \) holds.

In Section IV we shall exhibit a linear plant in state-space form that generates the vector \( \phi(t) \) as its output. We shall show that this plant, called the associated-signal system of (2.4), is output-reachable if and only if the design identity (3.1) has a unique solution vector \( \theta^* \).

Finally in Section V we will use this result to develop necessary and sufficient conditions on the input \( u(t) \) so that \( \{\phi(t)\} \) is uniformly persistently spanning.

IV. REACHABILITY OF THE REGRESSION VECTOR

We let the state of the associated-signal system be defined as
\[ x(t) = [y(t-1), \ldots, y(t-n), u(t-1), \ldots, u(t-n)]^T. \] (4.1)
It follows from (2.1) that \( x(t) \) satisfies the following discrete-time linear state equation
\[ x(t+1) = Ax(t) + bu(t) \] (4.2)
where
\[ A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \] (4.3)
with \( A_{i} \) \( l \times l \) matrices and \( b_1 \) \( l \times 1 \) matrices as follows:
\[ A_{11} = \begin{bmatrix} -p_1 & \cdots & -p_0 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} \] (4.4a)
\[ A_{12} = \begin{bmatrix} \tau_1 & \cdots & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} \] (4.4b)
\[ A_{21} = 0 \] (4.4c)
\[ A_{22} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} \] (4.4d)
The signal vector \( \phi(t) \) is obtained by passing \( x(t) \) through the following output equation:

\[
\phi(t) = Cx(t)
\]

where

\[
C = \begin{bmatrix}
C_{11} & 0 & \cdots & 0 \\
0 & C_{21} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & C_{m_2}
\end{bmatrix}
\]

with \( C_{11} \) an \( m_g \times 1 \) matrix and \( C_{21} \) an \( m_a \times 1 \) matrix as follows:

\[
C_{11} = \begin{bmatrix}
a_{11} & \cdots & a_{1r} \\
\vdots & \ddots & \vdots \\
a_{m_1,1} & \cdots & a_{m_1,m_g}
\end{bmatrix}, \quad C_{21} = \begin{bmatrix}
b_{11} & \cdots & b_{1r} \\
\vdots & \ddots & \vdots \\
b_{m_2,1} & \cdots & b_{m_2,m_g}
\end{bmatrix}
\]

We note that our assumption of Section II that the \( q_i(D) \) as well as the \( b_i(D) \) are linearly independent sets implies that rank \( C_{11} = m_g \) and rank \( C_{21} = m_a \), whence rank \( C = m_a + m_g \).

We recall that a linear system is output-reachable from the (input) vector and only if every vector in its output space can be generated (reached) using a suitable input sequence. Stated equivalently, the system is output-reachable if and only if the only vector orthogonal to all reachable outputs is the zero vector. Thus, the associated signal system (4.2)–(4.7) is output-reachable provided the (input) vector \( \mu = [\mu_{11}, \cdots, \mu_{m_a}, \mu_{21}, \cdots, \mu_{m_2}]^T \) satisfies

\[
\mu^T \phi(t) = 0
\]

for all \( \phi(t) = 0 \). Using the definition of \( \phi(t) \) in (3.4), (4.8) becomes

\[
\mu^T \phi(t) = \left[ \sum_{i=1}^{m_a} \mu_{ai}\alpha_i(D) \right] y(t) + \left[ \sum_{i=1}^{m_g} \mu_{bi}b_i(D) \right] u(t) = 0
\]

whence the associated signal system is output-reachable if and only if \( (4.8) \) has its only solution, for all signal sequences generated by (2.1), the trivial solution \( \mu = 0 \). Multiplying (4.9) by \( p(D) \) and employing (2.1), gives the (obvious) equivalent condition that the associated signal system is output-reachable if and only if the equation

\[
\left[ \sum_{i=1}^{m_a} \mu_{ai}\alpha_i(D) \right] r(D) + \left[ \sum_{i=1}^{m_g} \mu_{bi}b_i(D) \right] p(D) = 0
\]

has only the zero solution for \( \mu \). But (4.10) is the homogeneous equation of (3.1), whence the latter statement is equivalent to saying that (3.1) has a unique solution for \( (a_1, \cdots, a_{m_a}, b_1, \cdots, b_{m_g}) \). We thus have the following elementary but important:

**Theorem 4.1:** Assume that (3.1) is solvable. Then this solution is unique if and only if the associated signal system is output-reachable. □

We illustrate Theorem 4.1 by applying it to the second example in Section II.

**Example 2.2 (Continued).**

Equation (3.1) for this problem is

\[
[b(D) - p(D)q(D)e(D)]r(D) = [p(D)q(D)e(D) - \gamma(D)]p(D)
\]

and this equation is of course solvable in view of the co-prime assumption for \( r(D) \) and \( p(D) \). This equation has a unique solution if and only if the homogeneous equation

\[
[b(D) - p(D)q(D)e(D)]r(D) = [p(D)q(D)e(D) - \gamma(D)]p(D)
\]

has no nonzero solution for the parameters \( b_1, \cdots, b_a, \alpha_1, \cdots, \alpha_{m_g}, \beta_1, \cdots, \beta_g \). Writing (4.12) alternatively as

\[
D \left[ \sum_{i=1}^{m_a} \phi_i D^{-1} \right] r(D) + \left[ \sum_{i=1}^{m_g} \phi_i D^{-1} \right] p(D) = 0
\]

it is seen that (4.12) can have nontrivial solutions if degree \( q(D)p(D) \) is not greater than \( 2n \). If on the other hand, degree \( q(D)p(D) \) is at least \( 2n \) then every solution of (4.12) must satisfy

\[
\sum_{i=1}^{m_a} \phi_i D^{-1} r(D) + \sum_{i=1}^{m_g} \phi_i D^{-1} p(D) = 0
\]

Finally, for (4.14) and (4.15) to have no nontrivial solutions, it is necessary and sufficient for \( p(D) \) and \( r(D) \) to be coprime. (This is based on the well-known fact that two polynomials \( f(\lambda) \) and \( g(\lambda) \) of degree \( n \) are coprime if and only if the equation \( h(\lambda)f(\lambda) + k(\lambda)g(\lambda) = 0 \) has no nonzero solution polynomials \( h(\lambda) \) and \( g(\lambda) \) whose degrees are at most \( n - 1 \).) We conclude the example with the following summary: The associated signal system of the direct pole placement adaptive controller (as described, for example, in [11]) is output-reachable if and only if the following conditions both hold: (1) \( p(D), r(D) \) are coprime; and (2) degree \( (p(D)q(D)) \) is at least \( 2n \).

**V. UNIFORM PERSISTENT SPANNING**

We consider in this section the problem of achieving uniform persistent spanning of the output from the input of linear plants

\[
x(k+1) = Ax(k) + Bu(k)
\]

\[
\phi(k) = Cx(k)
\]

where \( u(k) \in \mathbb{R}^1 \), \( x(k) \in \mathbb{R}^n \), and \( \phi(k) \in \mathbb{R}^m \).

We first establish the following fact which extends a result of [3] to the case of output-reachable. We denote by \( \mu \) the dimension of the reachable subspace of (5.1), i.e.,

\[
\mu := \text{rank } \{b, Ab, \cdots, A^{\mu-1}b\}.
\]

**Lemma 5.1:** Assume that the plant (5.1) is output-reachable. Then for any nonzero \( m \)-vector \( \alpha \), there exist nonzero vectors \( \eta \) and \( \xi \) (dependent of \( \alpha \)) such that for any arbitrary \( x_0 \) and any arbitrary \( u_{k+1}, \cdots, u_{k+\mu} \),

\[
\eta^T Q_\alpha(k) = \alpha^T \Phi_{\alpha_\mu+1}(k) \xi
\]

where \( \Phi_{\alpha_\mu+1}(k) \) and \( Q_\alpha(k) \) are as defined in (3.4) and (3.8).

**Proof (Adapted from [3]):** Define \( z(k) = \alpha^T \phi(k) \). The sequence \( \{z_k\} \) can be regarded as an output of (5.1) evolving from \( x_0 \) via the input sequence \( u_{k+1}, \cdots, u_{k+\mu} \). Hence, it is related to \( \{u(k)\} \) by an equation of the form

\[
m(D)z(k) = (D)u(k)
\]

where

\[
b_k = \begin{bmatrix}
0 \\
\vdots \\
0 \\
1 \\
0 \\
\vdots \\
0
\end{bmatrix}, \quad b_2 = \begin{bmatrix}
0 \\
\vdots \\
0 \\
0 \\
\vdots \\
0 \\
0
\end{bmatrix}
\]

(4.4e)
where
\[ n(D) = n_1 D + \cdots + n_r D^r, \]
\[ m(D) = m_1 D + \cdots + m_s D^s. \]

The polynomial \( z(t) \) is a divisor of the characteristic polynomial of \( A \). Further, the polynomial \( n(D) \) is nonzero by virtue of the output-reachability of (5.1).

From (5.4) we then have
\[ z(k + 1) = m_1 z(k) + \cdots + m_s z(k + s) - n_1 u(k) + \cdots + n_r u(k + r). \]

whence (5.3) follows with
\[ \xi = [x_1, x_2, \ldots, x_r, x_{r+1}]^T \]
\[ \eta = [y_1, y_2, \ldots, y_r]^T. \]

With the aid of Lemma 1 the following result is easily established.

**Proposition 5.2:** Assume that the system (5.1) is output-reachable and let \( x_0 \) be an arbitrary initial state. Then \( \{\Phi(t)\} \) will be uniformly persistently spanning of order \( N \) provided \( \{\Phi(\ell)\} \) is uniformly persistently spanning of order \( N - \mu \).

**Proof (After [7]):** With the use of (5.3) and the Cauchy-Schwarz inequality, we obtain
\[ \|\Phi_0(k)\| \leq \sum_{i=0}^{\infty} \Phi_i(k) e^{-\alpha k}. \]

hence,
\[ \|U(n+N, -t) U^T(n+N, -t) \| \leq \sum_{i=0}^{\infty} \Phi_i(k) e^{-\alpha k}. \]

It follows that
\[ \min_{n=1} \|\alpha \| \|\Phi_{n-1}(k)\| e^{-\alpha (k + 1)} \]
for some fixed \( \alpha > 0 \), then the sequence \( \{\phi(k)\} \) of (5.12) is uniformly persistently spanning (of order \( N \)) for any reachable initial state \( x_0 \).

As a final remark it should be noted that Theorem 5.3 focuses attention on reachable initial states and (in general) not on all arbitrary initial states. For the case of arbitrary initial states the reader is referred to [13].

**ACKNOWLEDGMENT**

The author wishes to express his thanks to G. C. Goodwin for suggesting the structured-nonminimal-model concept and for his help in preparing this revised manuscript.

**REFERENCES**


