TECHNICAL REPORT 67-2

W. M. WONHAM

ON POLE ASSIGNMENT IN MULTI-INPUT CONTROLLABLE LINEAR SYSTEMS

FEBRUARY, 1967

CENTER FOR DYNAMICAL SYSTEMS

ON POLE ASSIGNMENT IN MULTI-INPUT CONTROLLABLE

LINEAR SYSTEMS*

by

W. M. Wonham

*This research was supported in part by the Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, under AFOSR Grant No. AF-AFOSR-693-66, in part by the National Science Foundation under Grant No. GK-967, and in part by National Aeronautics and Space Administration under Grant No. NGR-40-002-015.

ABSTRACT OF

ON POLE ASSIGNMENT IN MULTI-INPUT CONTROLLABLE LINEAR SYSTEMS

W. M. Wonham

It is shown that controllability of an open-loop system is equivalent to the possibility of assigning an arbitrary set of poles to the transfer matrix of the closed loop system, formed by means of suitable linear feedback of the state.

ON POLE ASSIGNMENT IN MULTI-INPUT CONTROLLABLE LINEAR SYSTEMS

W. M. Wonham

INTRODUCTION

Consider the system

$$dx(t)/dt = Ax(t) + Bu(t) \ . \tag{1}$$

Here and in the following, all vectors and matrices have real-valued elements and all matrices are constants. In (1), A, B are matrices of dimension respectively $n \times n$ and $n \times m$; x is the state, an n-vector; and u is an m-vector. As usual, u denotes an external input.

Let us "close the loop" by setting

$$u = Cx + v \ ,$$

for some $m \times n$ matrix C and new external input v. Then (1) becomes

$$dx(t)/dt = (A+BC)x(t) + Bv(t) \ . \tag{2}$$

In applications it is often desirable to choose C so that the matrix $A+BC$ has special properties: for example, stability. Intuitively it is clear that the possibility of such a choice depends on the controllability, in an appropriate sense, of the state x with respect to u. In this note we single out the property of "pole assignability", and show that it is equivalent to controllability of (1) in the usual sense.

To be precise, let

$$\Lambda = \{\lambda_1, \ldots, \lambda_n\}$$

be an arbitrary set of n complex numbers λ_i, such that any λ_i with $\mathrm{Im}\,\lambda_i \neq 0$ appears in Λ in a conjugate pair. We recall that the pair (A,B) is (completely) controllable if and only if the $n \times mn$ matrix

$$K = [B, AB, \ldots, A^{n-1}B]$$

is of full rank n. The result we wish to prove is the following.

THEOREM *The pair* (A,B) *is controllable if and only if, for every choice of the set* Λ, *there is a matrix* C *such that* $A + BC$ *has* Λ *for its set of eigenvalues.*

In other words, controllability is equivalent to the property that the closed-loop transfer matrix

$$[sI - (A+BC)]^{-1}B$$

can be assigned an arbitrary set of poles by a suitable choice of the feedback "gain" matrix C.

If B is an n-vector $(m=1)$ the result stated is well-known, and is obvious after a change of basis which transforms A to companion (rational canonical) form and B to the form

$$B = (0,0,\dots,0,1)'. \quad \text{(prime denotes transpose)}$$

Such a choice of basis is always possible if (A,B) is controllable; for the details see, for instance, [1], p. 113.

To prove the result in the general case we first transform (A,B) to a canonical form in which our multi-input system is exhibited as a triangular array of subsystems, each of the type just described. This transformation is due to Langenhop [2]. He proved a theorem similar to ours, in the case where $\dot{u} = Fx+Gu$, and the elements of the parameter matrices may be arbitrary complex numbers. The present restriction to real-valued elements apparently prevents the immediate application of Langenhop's result; however, basically we follow the same route.

PROOF OF THE THEOREM

1. Sufficiency.

The argument of [2] applies without change. In fact, let $\lambda_1, \ldots, \lambda_n$ be any distinct real numbers such that $\det(A-\lambda_i I) \neq 0$ $(i = 1,\ldots,n)$. By assumption there are n-vectors $x_i \neq 0$ $(i = 1,\ldots,n)$ and an $m \times n$ matrix C such that

$$(A + BC - \lambda_i I)x_i = 0 \qquad (i = 1,\ldots,n)$$

or

$$(A-\lambda_i I)^{-1}BCx_i = x_i \qquad (i = 1,\ldots,n). \tag{3}$$

Since

$$(A-\lambda I)^{-1} \equiv \sum_{j=1}^{n} \rho_j(\lambda)A^{j-1}$$

for suitable rational functions ρ_j, (3) implies

$$\sum_{j=1}^{n} \rho_j(\lambda_i)A^{j-1}BC\, x_i = x_i \qquad (i = 1,\ldots,n)\ . \tag{4}$$

Because the λ_i are distinct, the eigenvectors x_i of $A+BC$ are linearly independent; thus (4) states that the range of K is the whole space, that is, K has rank n.

2. Necessity.

The proof of necessity leans on the theory of cyclic subspaces ([3], Chapter 7). For completeness' sake we collect in 2.1 some definitions and preliminary results.

2.1 Denote the coordinate n-space by E and let the $n \times n$ matrix A be fixed. If $b \in E$ and $b \neq 0$ let $\nu(b)$ be the greatest integer such that the vectors

$$b, Ab, \ldots, A^{\nu(b)-1} b$$

are linearly independent, and let E_b denote their span. E_b is the <u>cyclic subspace generated by</u> b. E is <u>cyclic</u> if there exists $b \in E$ such that $E_b = E$, that is, such that (A,b) is controllable. The <u>minimal polynomial</u> (m.p.) of b is the (unique) monic polynomial $\beta(\lambda)$ of lowest degree such that $\beta(A)b = 0$. Thus b generates E if and only if the m.p. of b is of degree n. The <u>minimal polynomial</u> (m.p.) of E is, as usual, the (unique) monic polynomial $\alpha(\lambda)$ of least degree such that $\alpha(A) = 0$; $\alpha(\lambda)$ is the least common multiple (LCM) of the m.p. $\beta(\lambda)$ of the vectors $b \in E$. Evidently E is cyclic only if the degree of $\alpha(\lambda)$ is n; we state the converse as

LEMMA 1 ([3], p.180, Theorem 2)

<u>If the m.p.</u> $\alpha(\lambda)$ <u>of</u> E <u>is of degree</u> n <u>then</u> E <u>is cyclic</u>; <u>that is, there exists</u> $b \in E$ <u>such that the m.p. of</u> b <u>is</u> $\alpha(\lambda)$.

In particular the condition of Lemma 1 holds if the eigenvalues of A are distinct.

Two monic polynomials φ, ψ are <u>coprime</u> if their greatest common divisor (GCD) is 1. Then by ([4],p.75,Theorem 12) there are monic polynomials ρ, σ such that

$$\rho(\lambda)\varphi(\lambda) + \sigma(\lambda)\psi(\lambda) \equiv 1$$

LEMMA 2 (cf. [4],p.324,Theorem 19, Corollary).

Let E be cyclic, with m.p. α; and let c be a generator of E. If $b = \gamma(A)c$ for some polynomial γ and if α, γ are coprime, then b is also a generator of E.

The converse is also true, but will not be needed.

PROOF.

Let β be the m.p. of b. Since α, γ are coprime there exist ρ, σ such that $1 = \gamma\rho + \alpha\sigma$. Thus

$$c = \gamma(A)\rho(A)c + \alpha(A)\sigma(A)c = \rho(A)b \ .$$

If $x \in E$ is arbitrary then for some θ,

$$x = \theta(A)c = \theta(A)\rho(A)b \ ;$$

then $\beta(A)x = 0$, and so $\alpha | \beta$ (α divides β). Since α is the m.p. of E, $\beta | \alpha$. Hence $\beta = \alpha$, and so b generates E. The proof is complete.

Let B be an $n \times m$ matrix as before, and let $\{B\}$ denote the subspace of E spanned by the column vectors of B. Obviously if $\{B\}$ contains a generator of E then (A,B) is controllable. Since the m.p. of E may be of degree less than n, the converse statement is in general false. However, for the cyclic case we have

LEMMA 3

Let (A,B) _be controllable and let_ E _be cyclic_. _There exists an n-vector_ $b \in \{B\}$ _such that_ (A,b) _is controllable_.

PROOF.

For $i = 1,\ldots,m$ let β_i be the m.p. of b_i and E_i the subspace generated by b_i. Thus $E = E_1 + \cdots + E_m$ and, if α is the m.p. of E,

$$\alpha = \mathrm{LCM}(\beta_1,\ldots,\beta_m)$$

Let c be a generator of E. There are polynomials $\gamma_1,\ldots,\gamma_m$ such that $b_i = \gamma_i(A)c$, $i = 1,\ldots,m$. We shall prove the existence of real numbers $r_1,\ldots,r_m$ such that $r_1 b_1 + \cdots + r_m b_m$

generates E. By Lemma 2, it is enough to choose the r_i so that

$$\gamma(\lambda) = r_1\gamma_1(\lambda) + \cdots + r_m\gamma_m(\lambda)$$

is coprime with α.

For this, note first that

$$\text{GCD } \{\alpha(\lambda), \gamma_1(\lambda), \ldots, \gamma_m(\lambda)\} = 1 \tag{5}$$

In fact, if κ is the GCD on the left, then

$$\alpha = \kappa\hat{\alpha}, \quad \gamma_i = \kappa\hat{\gamma}_i \quad (i = 1, \ldots, m)$$

for suitable polynomials $\hat{\alpha}$, $\hat{\gamma}_i$. Then

$$\begin{aligned} \hat{\alpha}(A)b_i &= \hat{\alpha}(A)\kappa(A)\hat{\gamma}_i(A)c \\ &= \hat{\gamma}_i(A)\alpha(A)c \\ &= 0 \qquad (i = 1, \ldots, m). \end{aligned}$$

Thus $\beta_i | \hat{\alpha}$ $(i = 1, \ldots, m)$; hence (by definition of the LCM) $\alpha | \hat{\alpha}$; that is, $\kappa = 1$.

Finally, observe that α, γ are coprime if and only if $\gamma(\lambda_j) \neq 0$ $(j = 1, \ldots, n)$, where the λ_j are the (complex) zeros of $\alpha(\lambda)$. By (5), the quantities $\gamma_i(\lambda_j)$ $(i = 1,\ldots,m)$ cannot all vanish for any λ_j. Therefore numbers r_i with the required property exist.

The next observation will be useful; the simple proof is omitted.

LEMMA 4.

If (A,B) is controllable and C is any $m \times n$ matrix then $(A+BC,B)$ is controllable.

We shall also use the concept of congruence: for details see [3], Ch. 7, §3. Let E_1 be an invariant subspace of E, that is $Ax \in E_1$ whenever $x \in E_1$. A subspace E_2 is <u>invariant</u> (<u>mod</u> E_1) if $x \in E_2$ implies $Ax = y_1 + y_2$, where $y_1 \in E_1$ and $y_2 \in E_2$. We write $x \equiv 0 \pmod{E_1}$ if $x \in E_1$. Vectors $x_1,\ldots,x_k$ are <u>linearly independent</u> (<u>mod</u> E_1) if the relation $r_1x_1 + \ldots + r_kx_k \equiv 0 \pmod{E_1}$ implies $r_1 = \ldots = r_k = 0$, for all sets of scalars $r_1,\ldots,r_k$. For some $b \notin E_1$ let ν be the greatest integer such that the vectors

$$b,\ Ab,\ \ldots,\ A^{\nu-1}b$$

are linearly independent (mod E_1), and let E_2 be their span. Then E_2 is <u>cyclic</u> (<u>mod</u> E_1); clearly E_2 is invariant (mod E_1). The <u>relative m.p. of</u> E_2 (<u>mod</u> E_1) is the (unique) monic polynomial η of least degree such that $\eta(A)x \equiv 0 \pmod{E_1}$ for all $x \in E_2$.

2.2 We are now ready to prove necessity. Let b_i denote the i^{th} column of B:

$$B = [b_1, \ldots, b_m] \quad ;$$

let E_i be the cyclic subspace generated by b_i; and put $n_i = \dim(E_i)$. Since (A,B) is controllable,

$$E = E_1 + \ldots + E_m,$$

but in general the E_i are not independent. However, we can write E as a direct sum

$$E = \hat{E}_1 \oplus \hat{E}_2 \oplus \ldots \oplus \hat{E}_t \quad (t \leq m) \tag{6}$$

where the $\hat{E}_i$ are certain subspaces of the E_j. To see this, define $\hat{E}_1 = E_1$. If $b_2 \notin E_1$ let ν_2, $1 \leq \nu_2 \leq n_2$, be the greatest integer such that the vectors

$$b_1, Ab_1, \ldots, A^{n_1 - 1} b_1,\ b_2, Ab_2, \ldots, A^{\nu_2 - 1} b_2$$

are linearly independent, and let $\hat{E}_2$ be the span of the vectors $A^{j-1}b_2 (j = 1,\ldots,\nu_2)$. Thus $\hat{E}_2$ is cyclic (mod $\hat{E}_1$), and has relative m.p. of degree ν_2. Continuing in this way, for $i \geqq 2$ define ν_i, $1 \leqq \nu_i \leqq n_i$, to be the greatest integer such that the vectors

$$b_1,\ldots,A^{\nu_{i-1}-1}b_{i-1},\ b_i,Ab_i,\ldots,A^{\nu_i-1}b_i$$

are linearly independent, and let $\hat{E}_i$ be the span of $A^{j-1}b_i$ $(j = 1,\ldots,\nu_i)$. Then $\hat{E}_i$ is cyclic (mod $\hat{E}_1 \oplus \ldots \oplus \hat{E}_{i-1}$), with relative m.p. of degree ν_i. If at any stage $b_i \in \hat{E}_1 \oplus \ldots \oplus \hat{E}_{i-1}$ $(i \geqq 2)$, then b_i is skipped; by re-ordering the columns of B if necessary we can arrange that for $i = 1,\ldots,t$ the vectors b_i generate independent subspaces $\hat{E}_i$, where either $t = m$ or $b_s \in \hat{E}_1 \oplus \ldots \oplus \hat{E}_t$ $(s = t+1,\ldots,m)$. Since $A^{j-1}b_i \in \hat{E}_1 \oplus \ldots \oplus \hat{E}_t$ if $j \geqq \nu_i+1$ $(1 \leqq i \leqq m)$, it follows that (6) is true and thus $\nu_1 + \ldots + \nu_t = n$.

We shall now transform A to a convenient canonical form. Let

$$\lambda^{\nu_i} - \sum_{j=1}^{\nu_i} \alpha_{ij}\lambda^{j-1} \qquad (i = 1,\ldots,r) \qquad (7)$$

be the relative m.p. of b_i (mod $\hat{E}_1 + \ldots + \hat{E}_{i-1}$) or the absolute m.p. of b_1 in case $i = 1$. Following Langenhop [2] introduce vectors

$$e_{ik} = A^{\nu_i - k} b_i - \sum_{j=k+1}^{\nu_i} \alpha_{ij} A^{j-k-1} b_i \tag{8a}$$

($k = 1,2,\ldots,\nu_i$; $i = 1,2,\ldots,t$), where the summation does not appear in case $k = \nu_i$; that is

$$e_{i\nu_i} = b_i \qquad (i = 1,\ldots,t) \ . \tag{8b}$$

It is clear that for each i the vectors e_{ik} ($k = 1,\ldots,\nu_i$) are independent linear combinations of the vectors $A^{j-1}b_i$ ($j = 1,\ldots,\nu_i$), and thus the entire set of e_{ik} is a basis for E. Observe that for $2 \leqq k \leqq \nu_i$,

$$Ae_{ik} = e_{i,k-1} + \alpha_{ik} e_{i\nu_i} \quad (i = 1,\ldots,t). \tag{9a}$$

Furthermore, using the fact that the relative m.p. of b_i (mod $\hat{E}_1 + \ldots + \hat{E}_{i-1}$) is of degree ν_i, we have (cf. (7)).

$$Ae_{i1} = \alpha_{i1} e_{i\nu_i} + \sum_{k=1}^{i-1} \sum_{j=1}^{\nu_k} \gamma_{ijk} e_{kj} \tag{9b}$$

($i = 1,\ldots,t$) for certain scalars γ_{ijk}, and where the double summation does not appear if $i = 1$.

Using (9), we next compute the form of A, regarded as a linear transformation in E, relative to the basis $e_{11},\ldots,e_{1\nu_1}$, $e_{21},\ldots,e_{2\nu_2},\ldots,e_{t1},\ldots,e_{t\nu_t}$. That is, let H be the matrix with column vectors e_{ik} in the order just written and let $T = H^{-1}$. Then $TAT^{-1} = \tilde{A}$ has the block form

$$\tilde{A} = \begin{bmatrix} A_1 & A_{12} & \cdots & A_{1t} \\ & A_2 & \cdots & \vdots \\ & & \ddots & A_{t-1,t} \\ \bigcirc & & & A_t \end{bmatrix} \tag{10}$$

In (10) the matrices A_i which occur on the diagonal of $\tilde{A}$ are $\nu_i \times \nu_i$ companion matrices

$$A_i = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \\ 0 & 0 & \cdots & 0 & 1 \\ \alpha_{i1} & \alpha_{i2} & \cdots & & \alpha_{i\nu_i} \end{bmatrix} (i = 1,\ldots,t); \tag{11}$$

and the matrices A_{ij} have dimension $\nu_i \times \nu_j$.

Finally, $TB = \tilde{B}$, where

$$\tilde{B} = \left[\begin{array}{cccc|c} \tilde{b}_1 & & & & \\ & \tilde{b}_2 & & 0 & \\ & & \ddots & & \tilde{B}_o \\ 0 & & & \tilde{b}_t & \end{array}\right] \qquad (12a)$$

Here the ν_i-vectors $\tilde{b}_i = Tb_i$ have the form

$$\tilde{b}_i = (0,\ldots,0,1)', \qquad (12b)$$

and $\tilde{B}_o = T[b_{t+1},\ldots,b_m]$ is an $n \times (m-t)$ matrix (which does not appear if $t = m$).

To verify (10), (11), (12) in detail observe that Tx is the column of components of x in the basis $\{e_{ik}\}$. Thus the p'th column of $\tilde{A}$ is a list of the components, in this basis, of Ae_{ik}, where e_{ik} is the p'th basis vector. This list can be read off from (9). Similarly $\tilde{b}_i$ is the column of components of b_i, so that (12) follows at once from (8b).

With A,B in canonical form, we next show that the eigenvalues of $A+BC$ can be prescribed arbitrarily. Put $\tilde{C} = CT^{-1}$ and observe that the matrix $\tilde{A}+\tilde{B}\tilde{C}$ is similar under T to $A+BC$. From now on we drop the tilde on $\tilde{A},\tilde{B},\tilde{C}$. Next, take C in the form

$$C = \left[\begin{array}{ccccc} c_1' & & & & \\ & c_2' & & \bigcirc & \\ & & \cdot & & \\ & \bigcirc & & \cdot & \\ & & & & c_t' \\ \hdashline & & \bigcirc & & \end{array}\right] \tag{13}$$

where

$$c_1' = [c_{i1}, c_{i2}, \ldots, c_{i\nu_i}] \qquad (i = 1, \ldots, t)$$

In C the upper block is of dimension $t \times n$ and the lower block of dimension $(m-t) \times n$. From (10) through (13) we see that

$$A + BC - \lambda I = \begin{bmatrix} P_1(\lambda) & A_{12} & \cdot\ \cdot\ \cdot & A_{1t} \\ & P_2(\lambda) & & \\ & & \cdot & A_{t-1,t} \\ \bigcirc & & \cdot & \\ & & \cdot & P_t(\lambda) \end{bmatrix}$$

where, for $i = 1, \ldots, t$,

$$P_i(\lambda) = A_i + b_i c_i' - \lambda I = \begin{bmatrix} -\lambda & 1 & 0 & \cdots & \cdots & 0 \\ 0 & -\lambda & 1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \vdots \\ 0 & \cdots & \cdots & 0 & -\lambda & 1 \\ \alpha_{i1}+c_{i1}, & \alpha_{i2}+c_{i2}, & \cdots & \cdots & \cdots, & \alpha_{i\nu_i}+c_{i\nu_i}-\lambda \end{bmatrix}.$$

Let $p_i(\lambda)$ be the determinant of $P_i(\lambda)$; that is,

$$(-1)^{\nu_i} p_i(\lambda) = \lambda^{\nu_i} - \sum_{j=1}^{\nu_i} (\alpha_{ij}+c_{ij})\lambda^{j-1} \qquad (14)$$

From the triangular form of $A+BC-\lambda I$ it follows that

$$\det\,(A+BC-\lambda I) = p_1(\lambda) \ldots p_r(\lambda) .$$

Also, it is clear from (14) that the zeros of $p_i(\lambda)$ can be assigned arbitrarily (subject to conjugacy of complex zeros) by proper choice of the real coefficients c_{ij}. This shows that, in particular, <u>the system</u> (1) <u>can always be stabilized by appropriate choice of</u> C.

To show that an arbitrary set of eigenvalues Λ can be assigned, we must take account of the fact that the matrices $P_i(\lambda)$ are of fixed dimension: if for example $t=2$ and P_1, P_2 are each of dimension 1×1, it is impossible to assign an arbitrary complex pair of eigenvalues by independent adjustment of p_1, p_2. This is the added difficulty arising from our restriction to real parameters. To proceed, let $C = C_1 + C_2$, where C_1 is of the form (13), and is subject only to the requirement that all the eigenvalues of $A+BC_1$ be distinct; we have just seen that such a choice of C_1 is possible. Then, by Lemma 1, the space E is cyclic relative to the matrix $A+BC_1$. By Lemma 4, $(A+BC_1, B)$ is controllable. Applying Lemma 3, we find an n-vector $b \in \{B\}$ such that $(A+BC_1, b)$ is controllable. Now $b = Bg$ for a suitable m-vector g. Thus if we set

$$u = C_1 x + gv + w$$

our system (1) takes the form

$$\dot{x} = (A+BC_1)x + bv + Bw$$

Finally, because the theorem is true in the single-input case, we can choose

$$v = c'x$$

such that the matrix $A+BC_1+bc'$ has the desired set of eigenvalues Λ. That is, the feedback matrix

$$C = C_1 + gc'$$

has the required property. The theorem is proved.

COMMENTS

1. The construction in the proof of necessity can be summarized in terms of a block diagram (Figure). After choosing a suitable set of state variables (the canonical form (8)-(12)) the designer constructs an inner feedback loop (through the C_1 matrix) which renders the system cyclic: that is, controllable by a single input. The designer then picks a suitable input in the form gv $(Bg = b$ is a generator of $E)$, and completes the outer loop by setting $v = c'x$ to achieve the required disposition of poles. In this note we have not attempted to algorithmize the procedure.

2. In practice there may be many ways of choosing state variables and the quantities C_1, g, c to achieve a given assignment of poles. To exploit this freedom via suitable criteria of design is an interesting problem of current research. In this direction, see, for instance, [5].

3. It is worth noting that more than one 'canonical' form (10)-(12) may exist for a given (A,B) pair. That is, the matrix $\tilde{A}$ of (10) may not reveal the intrinsic structure of A, as do the usual rational canonical decompositions into block-diagonal forms ([3];Ch.7). The drawback of such block-diagonal representations here is that, in general, the cyclic subspaces corresponding to individual blocks need not have generators in the subspace $\{B\}$. If they do, the corresponding controls are completely non-interacting.

4. Although the proof has some intrinsic interest, it seems excessively clumsy beside the simplicity of the result.

REFERENCES

[1] Lefschetz, S., Stability of Nonlinear Control Systems, New York, Academic Press, 1965.

[2] Langenhop, C.E., On the stabilization of linear systems, Proc. Amer. Math. Soc., 15(5), 1964, pp. 735-742.

[3] Gantmacher, F.R., The Theory of Matrices, New York, Chelsea, Vol. 1, 1959.

[4] Birkhoff, G. and S. MacLane, A Survey of Modern Algebra, New York, MacMillan, 1953.

[5] Kalman, R.E., When is a linear control system optimal?, Trans. ASME, Ser. D, J. Basic Engrg., vol. 86, 1964, pp. 51-66.

Block diagram illustrating construction in proof, after choice of canonical state variables.