ON THE STABILIZATION OF LINEAR SYSTEMS

C. E. LANGENHOP

1. In a recent paper V. N. Romanenko [3] has given a necessary and sufficient condition that the system

\[ \frac{dx}{dt} = Ax + bu, \quad \frac{du}{dt} = px + qu \]

be "stabilizable." Here \( A \) is an \( n \) by \( n \) matrix, \( x \) and \( b \) are \( n \) by 1 column matrices (or vectors), \( p \) is a 1 by \( n \) row matrix and \( q \) and \( u \) are scalars. We shall assume that the elements of all these may be complex numbers. The vector \( x \) can be interpreted physically as the output of a linear system characterized by the matrix \( A \). The vector \( b \) corresponds to some feedback or control mechanism with \( u \) the controlling signal and \( p \) and \( q \) adjustable parameters in the controlling circuit. Romanenko calls the system \((A, b)\) stabilizable if for any nonempty set \( S \) of \( n + 1 \) or less complex numbers there exist \( p \) and \( q \) such that

\[ G = \begin{pmatrix} A & b \\ p & q \end{pmatrix} \]

has \( S \) as its set of characteristic values (spectrum). In particular then, if \((A, b)\) is stabilizable there exist \( p \) and \( q \) such that all characteristic values of \( G \) have negative real parts and every solution of (1.1) is such that \( x(t) \rightarrow 0 \) and \( u(t) \rightarrow 0 \) as \( t \rightarrow +\infty \).

In his paper Romanenko claims to be generalizing a known condition, which he attributes to Yu. M. Berezanskil, namely, that if the characteristic values of \( A \) are all distinct, then \((A, b)\) is stabilizable if and only if

\[ b, Ab, \ldots, A^{n-1}b \text{ are linearly independent.} \]

Romanenko's condition appears to be considerably more complicated than (1.3) while it is our intention here to show that in fact (1.3) is necessary and sufficient for \((A, b)\) to be stabilizable irrespective of any condition on the characteristic values of \( A \). Actually this is a corollary to our more general Theorem 1 given below.

As pointed out to the author by Dr. J. P. LaSalle the condition (1.3) and related ones have considerable significance in a seemingly...
different aspect of control theory, namely, the controllability of linear dynamical systems [2]. Undoubtedly there is a deeper connection between stabilizability and controllability. At any rate a certain canonical form developed in the study of the latter concept provides a simplification of our original proof and a generalization of the Berezanskiï-Romanenko result.

2. Rather than (1.2) we consider matrices of the form

\begin{equation}
G = \begin{pmatrix}
A & B \\
P & Q
\end{pmatrix}
\end{equation}

where \( A \) is \( n \) by \( n \), \( Q \) is \( m \) by \( m \) and \( B \) and \( P \) are correspondingly sized submatrices. We assume that the elements of all these may be complex. By \( S_r \) we denote a nonempty set of \( r \) or less complex numbers \( \mu_k \); that is, \( S_r = \{ \mu_k | k = 1, 2, \ldots, r \} \), where the \( \mu_k \) need not all be distinct. The rank \( h \) of the matrix \( H = (B, AB, \ldots, A^{n-1}B) \) is the significant feature of our results, the first of which is

**Theorem 1.** The condition \( h \geq r - m \) is necessary and sufficient that for each \( S_r \) there exist \( P \) and \( Q \) such that the spectrum of \( G \) contains \( S_r \).

**Proof of necessity.** Take any set \( S_r \) such that the \( \mu_k \) in \( S_r \) are distinct and different from any characteristic value of \( A \), and let \( P \) and \( Q \) be such that the spectrum of \( G \) contains this \( S_r \). Then there exist vectors

\[
\begin{pmatrix}
\xi_k \\
\eta_k
\end{pmatrix}, \quad k = 1, 2, \ldots, r,
\]

where \( \xi_k \) and \( \eta_k \) are \( n \) by \( 1 \) and \( m \) by \( 1 \) column matrices, respectively, such that

\[
\begin{pmatrix}
A & B \\
P & Q
\end{pmatrix}
\begin{pmatrix}
\xi_k \\
\eta_k
\end{pmatrix} = \mu_k \begin{pmatrix}
\xi_k \\
\eta_k
\end{pmatrix}
\]

or

\[
\begin{align}
A \xi_k + B \eta_k &= \mu_k \xi_k, \\
P \xi_k + Q \eta_k &= \mu_k \eta_k,
\end{align}
\]

From the first of (2.2) we may write

\[
\xi_k = (\mu_k I - A)^{-1} B \eta_k
\]

since \( A - \mu_k I \) is nonsingular. Using the characteristic equation of \( A \), one may write
for suitable scalar functions \( c_j(\mu) \). Substituting (2.4) into (2.3), we have

\[
(2.5) \quad \xi_k = \sum_{j=1}^{n} A^{j-1} B c_j(\mu_k) \eta_k = H \xi_k, \quad k = 1, 2, \ldots, r,
\]

where the transpose of the \( nm \) by 1 column matrix \( \xi_k \) is defined by

\[
\xi'_k = (c_1(\mu_k) \eta'_k, c_2(\mu_k) \eta'_k, \ldots, c_n(\mu_k) \eta'_k).
\]

Now by choice of \( S_r \) the vectors

\[
(\xi_k' \eta_k), \quad k = 1, 2, \ldots, r,
\]

are linearly independent and it is readily seen that the rank of the matrix \( (\xi_1, \xi_2, \ldots, \xi_r) \) must therefore be at least \( r - m \). From (2.5) it is then clear that the rank of \( H \) must likewise be at least \( r - m \).

**Proof of sufficiency.** The proof is accomplished by means of a similarity transformation on \( G \) to convert \( A \) and \( B \) to convenient canonical forms. We first dispose of the case \( B = 0 \), however. In this event \( h = 0 \) and the condition \( h \geq r - m \) becomes \( m \geq r \). It is clear then that we may choose \( P \) arbitrarily and \( Q \) such that the spectrum of \( Q \) and hence also that of \( G \) contains \( S_r \). Henceforth, then, we assume \( B \neq 0 \).

Consider now a matrix \( J \) of the form

\[
(2.6) \quad J = \begin{pmatrix} I & 0 \\ 0 & R \end{pmatrix}
\]

where \( I \) is an \( n \) by \( n \) identity and \( R \) is a nonsingular \( m \) by \( m \) matrix. Then

\[
(2.7) \quad J G J^{-1} = \begin{pmatrix} A & BR^{-1} \\RP & PQ R^{-1} \end{pmatrix}
\]

and it is evident that we may achieve any reordering of the columns of \( B \) with no essential change in the statement of the proposition to be proved. Thus if the columns of \( B \) are denoted by \( b_i, i = 1, 2, \ldots, m \), we may assume without loss of generality that the set of \( h \) columns of \( H \)

\[
(2.8) \quad b_1, Ab_1, \ldots, A^{h-1} b_1, \ldots, b_s, Ab_s, \ldots, A^{h-1} b_s
\]
is linearly independent. Here \(1 \leq s \leq m\), \(h_i \geq 1\), \(i = 1, 2, \cdots, s\), and \(\sum_{i=1}^{s} h_i = h\). We may further assume (see Chapter VII of [1]) that the sequence (2.8) is such that for \(i \geq 2\) the linear subspace \(V_i\) spanned by \(A^{j-1}b_i\), \(j = 1, 2, \cdots, k_i\), is invariant modulo \(V_1 + V_2 + \cdots + V_{i-1}\) under premultiplication by \(A\). That is,

\[
A^{h_i}b_i = \sum_{j=1}^{h_i} \alpha_{ij}A^{j-1}b_i + \sum_{k=1}^{i-1} \sum_{j=1}^{h_k} \beta_{ijk}A^{j-1}b_k,
\]

for some set of scalars \(\alpha_{ij}, \beta_{ijk}\) and where the double summation in (2.9) does not appear if \(i = 1\).

We now introduce a similarity transformation on \(G\) by means of a matrix \(K\) of the form

\[
K = \begin{pmatrix} T & 0 \\ 0 & I \end{pmatrix},
\]

where \(I\) is an \(m\) by \(m\) identity and \(T\) is a nonsingular \(n\) by \(n\) matrix. Then

\[
KGK^{-1} = \begin{pmatrix} TAT^{-1} & TB \\ PT^{-1} & Q \end{pmatrix}.
\]

The matrix \(T\) is defined by choosing certain combinations of the vectors in (2.8) as a new basis system for the space of \(n\) by 1 column matrices. Thus we introduce

\[
e_{ik} = A^{h_i-k}b_i - \sum_{j=k+1}^{h_i} \alpha_{ij}A^{j-1}b_i, \quad k = 1, 2, \cdots, h_i,
\]

and where the summation does not appear in case \(k = h_i\). That is

\[
e_{ih_i} = b_i, \quad i = 1, 2, \cdots, s.
\]

Note from (2.12) and (2.13) we have for \(k \geq 2\)

\[
Ae_{ik} = e_{i,k-1} + \alpha_{ik}e_{ih_i}, \quad i = 1, 2, \cdots, s,
\]

and, using (2.9) in addition, we have, since the \(A^{j-1}b_i\)'s may be expressed as linear combinations of the \(e_{ij}\)'s,

\[
Ae_{ii} = \alpha_{ii}e_{ih_i} + \sum_{k=1}^{i-1} \sum_{j=1}^{h_k} \gamma_{ijk}e_{kj}, \quad i = 1, 2, \cdots, s,
\]

for some set of scalars \(\gamma_{ijk}\) and where, again, the double summation is absent in the case \(i = 1\).

From the linear independence of the set (2.8) and the remark just preceding (2.15) it is evident that the set
is likewise linearly independent. If $h = \sum_{i=1}^{t} h_i < n$ we may adjoin to the set (2.16) $n - h$ additional vectors to get a set which spans the space of $n$ by 1 matrices. We now view the matrix $A$ as defining the linear transformation $x \rightarrow Ax$ and we let $T$ be defined so that for any $n$ by 1 matrix $x$ the matrix $Tx$ is the column of components of $x$ relative to the set (2.16) (augmented if required) as basis. With respect to this basis the matrix of the linear transformation $x \rightarrow Ax$ is the matrix $TAT^{-1}$ and, moreover, $Tb_i = u_{\sigma_i}$, $\sigma_i = h_1 + \cdots + h_i$ and $u_{\sigma}$ is an $n$ by 1 column all of whose components are zero except the $\sigma$th which is 1. From (2.14) and (2.15) we may thus infer that $TAT^{-1}$ has the following block form:

$$
TAT^{-1} = \begin{bmatrix}
C_1 & E_{12} & \cdots & E_{1s} & F_1 \\
0 & C_2 & \cdots & E_{2s} & F_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & C_s & F_s \\
0 & 0 & \cdots & 0 & L
\end{bmatrix},
$$

where $C_i$ has the canonical form

$$
C_i = \begin{bmatrix}
0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1 \\
\alpha_{i1} & \alpha_{i2} & \cdots & \alpha_{ih_i}
\end{bmatrix},
$$

Moreover,

$$
TB = \begin{bmatrix}
v_1 & 0 & \cdots & 0 & w_1 \\
v_2 & 0 & \cdots & 0 & w_2 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & v_s & w_s \\
0 & 0 & \cdots & 0 & 0
\end{bmatrix},
$$
where $v_i$ is an $h_i$ by 1 column all of whose components are zero except the last which is 1. The submatrices $w_i$ have $m - s$ columns.

We now choose $P$ and $Q$ in convenient form. Let

$$Q = \text{diag}(q_1, q_2, \ldots, q_m)$$

and let $P$ be such that

$$\begin{pmatrix}
    p_1 & 0 & \cdots & 0 & 0 \\
    0 & p_2 & \cdots & 0 & 0 \\
    \vdots & \vdots & \ddots & \vdots & \vdots \\
    0 & 0 & \cdots & p_s & 0 \\
    0 & 0 & \cdots & 0 & M
\end{pmatrix}
$$

(2.20)

where $p_i$ is a 1 by $h_i$ row and $M$ is $m - s$ by $n - h_i$. The determinant of $K G K^{-1} - \lambda I$ is now readily evaluated by means of Laplace expansions using minors from appropriate columns. Thus, if we expand by minors from the first $h_i$ columns along with the $(n + 1)$st column we find that only one of these is nonzero, namely, that using the first $h_i$ rows and the $(n + 1)$st row. The complementary minor will be a determinant array with exactly similar block structure to that of $K G K^{-1} - \lambda I$. It may thus be expanded similarly by using the corresponding columns, namely, those which appear in $K G K^{-1} - \lambda I$ as the second $h_2$ columns along with the $(n + 2)$nd. Continuing in this way we may write

$$\begin{pmatrix}
    L - \lambda I & 0 \\
    M & Q^* - \lambda I
\end{pmatrix}
\cdot
\begin{pmatrix}
    C_i - \lambda I & v_i \\
    p_i & q_i - \lambda
\end{pmatrix}
$$

(2.21)

where $Q^*$ = diag($q_{i+1}, \ldots, q_m$) and, of course, each appearance of $I$ denotes an identity of appropriate size.

First we observe that $Q^*$ may be chosen so that the spectrum of $G$ contains any prescribed set $S_{m-1}$. We next examine the determinants in the product factor in (2.21). Each of these has the same structure and may readily be evaluated by a Laplace expansion using minors from the last two rows. The result is

$$\begin{vmatrix}
    C_i - \lambda I & v_i \\
    p_i & q_i - \lambda
\end{vmatrix}
= (-\lambda)^{h+i+1} + (q_i - \alpha_{ih})(-\lambda)^{hi}$$

$$+ \sum_{j=2}^{h_i} (\alpha_{ij} + q_{ij} - p_{ij})(-\lambda)^{j-1} + (q_{ij} - \alpha_{ij} - p_{ij}),$$

where $\alpha_{ij}$ contains any prescribed set $S_{m-1}$. We next examine the determinants in the product factor in (2.21). Each of these has the same structure and may readily be evaluated by a Laplace expansion using minors from the last two rows. The result is

$$\begin{vmatrix}
    C_i - \lambda I & v_i \\
    p_i & q_i - \lambda
\end{vmatrix}
= (-\lambda)^{h+i+1} + (q_i - \alpha_{ih})(-\lambda)^{hi}$$

$$+ \sum_{j=2}^{h_i} (\alpha_{ij} + q_{ij} - p_{ij})(-\lambda)^{j-1} + (q_{ij} - \alpha_{ij} - p_{ij}),$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
where the $p_{ij}, j=1, 2, \ldots, h_i$, are the components of the row $p_i$. In this form it is evident that we may determine $q_i$ and the $p_{ij}, j=1, 2, \ldots, h_i$, so that the coefficients of this polynomial are any we desire. Thus $q_i$ and $p_i$ may be determined so that this factor of $\det(KGK^{-1} - \lambda I)$ has any given collection of $h_i+1$ roots. This is true for each $i$ so using this and the fact mentioned earlier regarding the choice of $Q^*$ it is clear that we may specify $P$ and $Q$ so that $G$ contains in its spectrum any given nonempty set of $m-s+\sum_{i=1}^{r} (h_i+1) = h+m$ or less complex numbers. Thus if $h \geq r-m$ we see that for any $S$, there exist $P$ and $Q$ such that $G$ contains $S_i$ in its spectrum.

**Corollary 1.** The condition $h=n$ is necessary and sufficient that for each $S_{n+m}$ there exist $P$ and $Q$ such that $G$ has $S_{n+m}$ as its spectrum.

**Proof.** In any case $h \leq n$, so if $r = n+m$, then the condition $h \geq r-m$ is equivalent to $h = n$.

**Corollary 2.** Condition (1.3) is necessary and sufficient that $(A, b)$ be stabilizable.

**Proof.** This is Corollary 1 in the case $m=1$.

**Remark.** Even in case $h < n$ there may exist $P$ and $Q$ such that all characteristic roots of $G$ have negative real parts. It is clear from (2.21) that this is the case provided the characteristic roots of $L$ have negative real parts which, in turn, is related to how the linear subspace complementary to that spanned by the columns of $H$ is associated with those characteristic roots of $A$ which have negative real parts. In any case this question does not appear to be directly answerable merely in terms of the rank of $H$.

3. In this section as an application of the results in §2 we point out the relevancy of condition (1.3) to the behavior of a more complicated system of differential equations than (1.1). Thus we consider the system

$$\begin{align*}
\frac{dx}{dt} &= Ax + bu, \\
\frac{d^{r+1}u}{dt^{r+1}} &= px + \sum_{k=0}^{r} q_k \frac{d^k u}{dt^k},
\end{align*}$$

(3.1)

where $A, b, x, u, p$ are as before and $q_k, k=0, 1, 2, \ldots, r$, are scalars.

**Theorem 2.** For any integer $r \geq 0$, if condition (1.3) holds, then there exist $p, q_k, k=0, 1, \ldots, r$, such that, for every solution of (3.1), $x(t) \to 0$ and $u(t) \to 0$ as $t \to +\infty$. 

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
PROOF. Introduce variables $u_k$ by the relations $u_0 = u$,

\begin{equation}
\tag{3.2}
\frac{du_k}{dt} = 1, \quad k = 1, 2, \ldots, r.
\end{equation}

Then (3.1) may be written in the equivalent matrix form

\begin{equation}
\tag{3.3}
\begin{pmatrix}
\frac{dx}{dt} \\
u_0 \\
\vdots \\
u_{r-1}
\end{pmatrix}
= \begin{pmatrix}
A & b & 0 & 0 & \cdots & 0 \\
0 & 0 & 1 & 0 & \cdots & 0 \\
\vdots \\
0 & 0 & 0 & 0 & \cdots & 1 \\
0 & 0 & 0 & 0 & \cdots & 0
\end{pmatrix}
\begin{pmatrix}
x \\
u_0 \\
\vdots \\
u_{r-1}
\end{pmatrix}
+ \begin{pmatrix}
0 \\
0 \\
\vdots \\
0 \\
0
\end{pmatrix}
\begin{pmatrix}
u_r \\
\vdots \\
\vdots \\
\vdots \\
\vdots
\end{pmatrix}
\begin{pmatrix}
u_{r-2} \\
u_{r-3} \\
\vdots \\
u_1 \\
u_0
\end{pmatrix}.
\end{equation}

\begin{equation}
\frac{du_r}{dt} = (p, q_0, \cdots, q_{r-1})(x', u_0, \cdots, u_{r-1})' + e_ru_r.
\end{equation}

This is the form of (1.1) with $(x', u_0, \cdots, u_{r-1})'$ playing the role of $x'$ there, $u_r$ the role of $u$, $(p, q_0, \cdots, q_{r-1})$ the role of $p$ and

\begin{equation}
\tag{3.4}
A^* = \begin{pmatrix}
A & b & 0 & 0 & \cdots & 0 \\
0 & 0 & 1 & 0 & \cdots & 0 \\
\vdots \\
0 & 0 & 0 & 0 & \cdots & 1 \\
0 & 0 & 0 & 0 & \cdots & 0
\end{pmatrix}\quad \text{and} \quad b^* = \begin{pmatrix}
0 \\
0 \\
\vdots \\
0 \\
0
\end{pmatrix}
\end{equation}

playing the roles of $A$ and $b$, respectively. From Corollary 2 $p, q_0, \cdots, q_r$ exist such that $x(t) \to 0$, $u_k(t) \to 0$, $k = 0, 1, \cdots, r$, as $t \to +\infty$ if $b^*, A^*b^*, A^{*2}b^*, \ldots, A^{*n+r-1}b^*$ are linearly independent.

From the form of $A^*$ and $b^*$ as given in (3.4) it is easy to verify that this is true if condition (1.3) holds.

Analogous applications of Corollary 1 may be made.

REFERENCES