Structural Controllability

CHING-TAI LIN, MEMBER, IEEE

Abstract—**The new concepts of "structure" and "structural controllability" for a linear time-invariant control system (described by a pair (A,b)) are defined and studied. The physical justification of these concepts and examples are also given.**

The graph of a pair (A,b) is also defined. This gives another way of describing the structure of this pair. The property of structural controllability is reduced to a property of the graph of the pair (A,b). To do this, the basic concept of a "cactus" and the related concept of a "precactus" are introduced. The main result of this paper states that the pair (A,b) is structurally controllable if an only if the graph of (A,b) is "spanned by a cactus." The result is also expressed in a more conventional way, in terms of some properties of the pair (A,b).

I. Introduction

CONSIDER a linear control system

$$\dot{x} = Ax + bu \tag{1.1}$$

with $x \in R^n$, $u \in R$. The matrices $A \in R^{n\times n}$ and $b \in R^n$ are assumed throughout to have compatible dimensions and to be time invariant. For convenience, the control system (1.1) is denoted throughout the paper by the pair (A,b).

It is known (see Lee and Markus [1]) that the set of all (completely) controllable pairs (1.1) is open and dense in the space of all pairs (A,b) (with the standard metric). Thus, if the pair (A_0,b_0) is not completely controllable, then for every $\epsilon > 0$, there exists a completely controllable pair (A,b) with $\|A - A_0\| < \epsilon$ and $\|b - b_0\| < \epsilon$ (where $\|\cdot\|$ denotes norm[1]). This result is of an obvious physical interest. Practically, for every pair (A,b), most of the entries of A and b are known only with the approximation of some errors of measurement. Only some of these entries are known with 100 percent precision. Normally, this happens for some entries which are equal to zeros. Thus we shall assume here that there are some entries of A and b which are precisely zero, while all the other entries are known only approximately.[2]

We shall say that the pair (A,b) has the same *structure* as another pair $(\tilde{A},\tilde{b})$, of the same dimensions, if for every fixed (zero) entry of the matrix (Ab), the corresponding entry of the matrix $(\tilde{A}\tilde{b})$ is fixed (zero) and, at the same time, for every fixed (zero) entry of $(\tilde{A}\tilde{b})$, the corresponding entry of (Ab) is also fixed (zero). Then one defines the pair (A_0,b_0) to be *structurally controllable* if and only if there exists a completely controllable pair (A,b) which has the same structure as (A_0,b_0).

The concept of "structural controllability" of a pair (A_0,b_0) makes the meaning of controllability (in the usual sense) more complete from the physical point of view. In fact, it is preferred whenever (A_0,b_0) represents an actual physical system (that involves parameters only approximately determined). Actually, the completely controllable pair (A,b) can be considered as "physically undistinguishable" from the pair (A_0,b_0) (see the following proposition). From this viewpoint, the concept of structural controllability has a physical justification. Indeed, the definition of structural controllability is equivalent to a stronger property as shown by the following proposition.

Proposition 1: The pair (A_0,b_0) is structurally controllable if and only if $\forall \epsilon > 0$, there exists a completely controllable pair (A_1,b_1) of the same structure as (A_0,b_0) such that $\|A_1 - A_0\| < \epsilon$ and $\|b_1 - b_0\| < \epsilon$.

Indeed, if the property in Proposition 1 is satisfied, then the pair (A_0,b_0) is obviously structurally controllable. Conversely, assume now that the pair (A_0,b_0) is structurally controllable, i.e., there exists a completely controllable pair (A_2,b_2) of the same structure as (A_0,b_0). Consider now the pairs, $A(\lambda) = (1-\lambda)A_0 + \lambda A_2$; $b(\lambda) = (1-\lambda)b_0 + \lambda b_2$, where $\lambda \epsilon [0,1]$. Then $\psi(\lambda) = \det(b(\lambda)\ A(\lambda)b(\lambda), A(\lambda)^{n-1}b(\lambda))$ is a polynomial in λ, and this polynomial is not identically zero (since it is different from zero for $\lambda = 1$). Clearly, given an arbitrary $\epsilon > 0$, one can find $\lambda_0 \epsilon [0,1]$ such that $\|A(\lambda) - A_0\| < \epsilon$ and $\|b(\lambda) - b_0\| < \epsilon, \forall \lambda \epsilon [0,\lambda_0]$. Further one can find $\lambda_1 \epsilon [0,\lambda_0]$ such that $\psi(\lambda_1) \neq 0$, i.e., the pair $(A(\lambda_1),b(\lambda_1))$ is completely controllable—and one has obtained the property in Proposition 1. Q.E.D.

It is immediate that every completely controllable pair (A,b) is also structurally controllable. Moreover, if no entry of (A,b) is fixed, then the pair (A,b) is structurally controllable (this also follows from the Lee–Markus result mentioned in the introduction). However, if some entries of (A,b) are fixed, the pair may not be structurally controllable. Consider, for instance, the pair (A,b) of the form

$$A = \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix}; \qquad b = \begin{pmatrix} 0 \\ b_2 \end{pmatrix} \tag{1.2}$$

where $A_{11} \in R^{k\times k}$, $A_{21} \in R^{(n-k)\times k}$, $A_{22} \in R^{(n-k)\times(n-k)}$, $b_2 \in R^{n-k} (1 \leq k \leq n)$, and by 0, one denotes a matrix or a vector containing only fixed (zero) entries. Then, it is easy to see that rank $(bAb \cdots A^{n-1}b)$ is less than n (independently of the values of A_{11}, A_{21}, A_{22}, and b_2). Thus the pair (1.2) is not structurally controllable.

Manuscript received November 15, 1972; revised August 16, 1973. Paper recommended by A. Stephen Morse, Chairman of IEEE S-CS Linear Systems Committee.

The author was with the Department of Electrical Engineering and Electronics Laboratory, University of Maryland, College Park, Md. 20742. He is now with the Bendix Field Engineering Corporation, Columbia, Md. 21043.

[1] For instance, if a_{ij} and b_i are the entries of A and b, respectively, then we may take $\|A\|_{i,j=1,\cdots,n} = \max |a_{ij}|$ and $\|b\|_{i=1,\cdots,n} = \max |b_i|$.

[2] Other possible assumptions are left for further research.

Consider also the pair (A,b) in which the $n \times (n+1)$ matrix (Ab) can be written as

$$(Ab) = \begin{pmatrix} P_1 \\ P_2 \end{pmatrix} \tag{1.3}$$

where P_2 is an $(n-k) \times (n+1)$ matrix, and P_1 is a $k \times (n+1)$ matrix $(k \geq 1)$ with no more than $k-1$ nonzero columns (all the other columns of P_1 having only fixed (zero) entries). Then one obtains directly that rank $(Ab) < n$, independently of the nonzero entries in (1.3), but rank $(Ab) < n$ implies that the pair (A,b) is not completely controllable. Thus the pair (1.3)—under the above assumptions—is not structurally controllable.

In this paper, we find the necessary and sufficient conditions of structural controllability. We shall prove the converse of the above results: *Every pair (A,b) which is not structurally controllable can be brought (after a suitable permutation of the coordinates) to one of the forms (1.2) and (1.3).*

II. The Graph of a Pair (A,b)

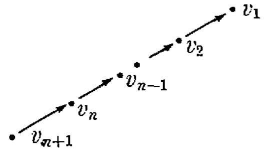
Given a pair (A,b), one defines its graph G as the graph which contains exactly $n+1$ nodes, $v_1, v_2, \cdots, v_{n+1}$, and all of whose edges are obtained as follows: For every nonfixed entry c_{ij} of the $n \times (n+1)$ matrix (Ab), the graph contains the *oriented* edge (v_j,v_i) (an arrow going from v_j to v_i). The node v_{n+1}, which corresponds to the $n+1$th column of (Ab), will be called the "*origin*" of G. (From the definition it follows that no arrow can point towards v_{n+1}.)

The set which contains all the nodes of G was called the *vertex set* of G [2]. For every oriented edge (v_j,v_i) in G, the node v_j will be called the "*origin*" of this edge; a node v in the vertex set of G will be called the "*final*" node, if v is not the origin of any oriented edge in G.

For example, the pair (A_1,b_1),

$$A_1 = \begin{bmatrix} 0 & \alpha_1 & 0 & \cdot & 0 \\ 0 & 0 & \alpha_2 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \alpha_{n-1} \\ 0 & 0 & 0 & \cdot & 0 \end{bmatrix}; \quad b_1 = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ 0 \\ \alpha_n \end{bmatrix} \tag{2.1}$$

(where all the entries denoted by zeros are fixed and all the other entries are not fixed) has the following graph:

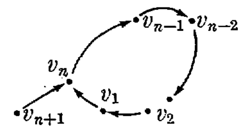


A graph of this form [i.e., the graph of a pair (A,b) of the form (2.1)] will be called a "*stem*."

As another basic example, consider the pair

$$A_2 = \begin{bmatrix} 0 & \alpha_1 & 0 & \cdot & 0 \\ 0 & 0 & \alpha_2 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \alpha_{n-1} \\ \alpha_{n+1} & 0 & 0 & \cdot & 0 \end{bmatrix}; \quad b_2 = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ 0 \\ \alpha_n \end{bmatrix} \tag{2.2}$$

where, as before, α_i are the only nonfixed entries. This pair has the following graph:



A graph of this form [i.e., the graph of a pair as in (2.2)] will be called a "*bud*." The node v_{n+1} is called the "*origin*" of the bud and the edge (v_{n+1},v_n) is called the "*distinguished edge*" of the bud. Both of the pairs (2.1) and (2.2) are easily seen to be structurally controllable.

Now let us give the graphs of the pairs (1.2) and (1.3) which are not structurally controllable. Consider, for instance, the following pair [of the form (1.2)]:

$$A_3 = \begin{pmatrix} \alpha_{11} & \alpha_{12} & 0 \\ \alpha_{21} & \alpha_{22} & 0 \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix}; \quad b_3 = \begin{pmatrix} 0 \\ 0 \\ \alpha_{34} \end{pmatrix} \tag{2.3}$$

where α_{ij} are the only nonfixed entries. The graph of this pair is shown in Fig. 1.

In Fig. 1, the nodes v_1 and v_2 are said to be "nonaccessible." In general, a node v_i (other than the origin) in the graph of a pair (A,b) is called *nonaccessible* if and only if there is no possibility of reaching the node v_i starting from the origin v_{n+1} and going to v_i only in the direction of the arrows, along a path in the graph of the pair (A,b).

It is easy to see that in general, the graph of a pair (A,b) of the form (1.2) contains at least one nonaccessible node. Moreover, the converse is also true: "*If the graph of a pair (A,b) is such that there exists at least one nonaccessible node, then (after a permutation of the coordinates) (A,b) can be brought to the form (1.2) and therefore (A,b) is not structurally controllable.*" Indeed, without loss of generality, one can assume that the nodes $v_1,v_2,\cdots,v_k$ are *all* the nonaccessible nodes, while $v_{k+1},\cdots,v_n$ are accessible (i.e., not nonaccessible). Then in the matrix (Ab), all the entries c_{ij}, $i = 1,2,\cdots,k$, $j = k+1, k+2, \cdots, n+1$, must be zero.

Consider now a pair of the form (1.3):

$$A_4 = \begin{pmatrix} 0 & \alpha_{12} & 0 \\ 0 & \alpha_{22} & 0 \\ 0 & \alpha_{32} & 0 \end{pmatrix}; \quad b_4 = \begin{pmatrix} \alpha_{14} \\ \alpha_{24} \\ \alpha_{34} \end{pmatrix} \tag{2.4}$$

where, as always, α_{ij} denote the only nonfixed entries. The graph of this pair is shown in Fig. 2. Consider here the set S formed by the nodes v_1, v_2, and v_3 ($S = \{v_1, v_2, v_3\}$). Determine the set $T(S)$ containing *all* the nodes v_j with the property that there is an oriented edge going from v_j to a node in S. Clearly, $T(S) = \{v_2,v_4\}$. Here, the $T(S)$ contains two elements while S contains three elements. One says that the graph contains a "*dilation*." More generally, the graph of a pair (A,b) contains a "*dilation*" if and only if there is a set S of k nodes in the vertex set of the graph—not containing the origin v_{n+1}—such that there are no more than $k-1$ nodes v_j in $T(S)$. (One denotes by $T(S)$ the set of all the nodes v_j with the property that there exists an oriented edge from v_j to a node in S. Note that the origin

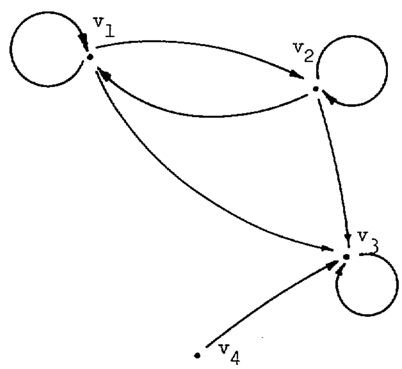


Fig. 1. A graph of the pair (A_3,b_3) as in (2.3).

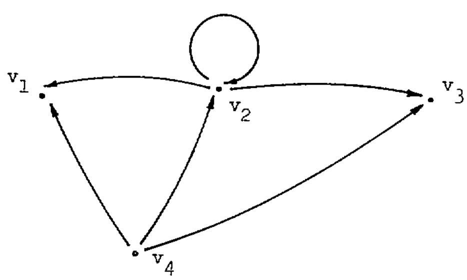


Fig. 2. A graph of the pair (A_4,b_4) as in (2.4).

v_{n+1} is not allowed to belong to S, but may belong to $T(S)$.)

One can easily see that if the pair (A,b) has the form (1.3), then its graph contains a dilation. *Conversely, if the graph of a pair (A,b) contains a dilation, then (after a permutation of the coordinates) the matrix (Ab) can be brought to the form of (1.3), and, therefore, the pair (A,b) is not structurally controllable.* Indeed, after a permutation of the nodes, one can always assume that $S = \{v_1,v_2,\cdots,v_k\}$. On the other hand, there are no more than $k-1$ nodes in $T(S)$. Each of these nodes corresponds to a different column in the matrix (Ab). From the definition of $T(S)$, it follows that for every other column p (different from the $k-1$ columns defined above), one must have $c_{ip} = 0$, $i = 1,2,\cdots,k$. Thus the pair (A,b) can be written as in (1.3).

The above discussion shows that if the graph of a pair (A,b) contains a *nonaccessible node* or a *dilation*, then the pair (A,b) is not structurally controllable. We shall prove later that the converse is true.

III. Cacti

We have seen in Section II that if the graph of a pair is a *stem* or a *bud*, then the pair is structurally controllable. Now we extend this conclusion to some special combinations of stems and buds which we call "cacti" (defined below).

First consider the following lemma:

Lemma 1: Suppose that G is a graph of a structurally controllable pair. Let B be a bud with the origin $\mathfrak{e}$, and suppose $\mathfrak{e}$ is the only node which belongs at the same time to the vertex set of G and to the vertex set of B. Then $G \cup B$ is the graph of a structurally controllable pair. (Here $\cup$ denotes union.)

Proof:

1) Consider first the case in which G and B have their origins as a common node. Then G is the graph of a structurally controllable pair (A,b), and B is the graph of a pair (A_1,b_1) of the form (2.2). Clearly $G \cup B$ is the graph of the pair

$$\begin{pmatrix} A & 0 \\ 0 & A_1 \end{pmatrix}; \quad \begin{pmatrix} b \\ b_1 \end{pmatrix}. \tag{3.1}$$

Since (A,b) is structurally controllable, there exists a completely controllable pair $(\tilde{A},\tilde{b})$ of the same structure as (A,b). Modify, if necessary, the nonfixed entries of A_1 such that A_1 and $\tilde{A}$ have no common eigenvalue. This can be done since the eigenvalues of A_1 are equal to the nth roots of the product of all the nonfixed entries in A_1. In order to prove that the pair (3.1) is structurally controllable, it will suffice to prove that the pair

$$\begin{pmatrix} \tilde{A} & 0 \\ 0 & A_1 \end{pmatrix}; \quad \begin{pmatrix} \tilde{b} \\ b_1 \end{pmatrix} \tag{3.2}$$

is completely controllable in the usual sense.

Recall that the pair (A,b) is completely controllable (in the usual sense) if and only if the relation $c^TA = \alpha c^T$, where α is a complex number, and $c \neq 0$ is an n-vector with complex entries, implies $c^Tb \neq 0$ (see W. Hahn or V. M. Popov [3] for reference). Thus, if (3.2) is not completely controllable, there exist c_1 and c_2, not both zeros, and α, a complex number, such that

$$(c_1^T \quad c_2^T)\begin{pmatrix} \tilde{A} & 0 \\ 0 & A_1 \end{pmatrix} = \alpha(c_1^Tc_2^T)$$

and

$$(c_1^Tc_2^T)\begin{pmatrix} \tilde{b} \\ b_1 \end{pmatrix} = 0.$$

That is, $c_1^T\tilde{A} = \alpha c_1^T$, $c_2^TA_1 = \alpha c_2^T$, and $c_1^T\tilde{b} + c_2^Tb_1 = 0$. This implies that either $c_1 = 0$ or $c_2 = 0$ (because A_1 and $\tilde{A}$ have no common eigenvalue).

Suppose $c_1 \neq 0$ and $c_2 = 0$. The, $c_1^T\tilde{A} = \alpha c_1^T$, and $c_1^T\tilde{b} = 0$. Therefore, $(\tilde{A},\tilde{b})$ is not completely controllable; a contradiction. Suppose $c_1 = 0$ and $c_2 \neq 0$. Then, $c_2^TA_1 = \alpha c_2^T$, and $c_2^Tb_1 = 0$, which means that the pair (A_1,b_1) is not completely controllable; a new contradiction. Thus the pair (3.1) is structurally controllable.

2) Consider now the case in which the origin of B is not the origin of G. Then, G is the graph of a structurally controllable pair (A,b), and B is the graph of a pair (A_1,b_1) of the form (2.2). Moreover $G \cup B$ is the graph of the pair

$$\begin{pmatrix} A & 0 \\ b_1q^T & A_1 \end{pmatrix}; \quad \begin{pmatrix} b \\ 0 \end{pmatrix} \tag{3.3}$$

where the *nonzero* vector q is chosen such that b_1q^T has one and only one nonzero entry, and its corresponding column corresponds to the only node belonging at the same time to the vertex set of B and to the vertex set of G.

Since (A,b) is structurally controllable, there exists a completely controllable pair $(\tilde{A},\tilde{b})$ of the same structure as (A,b). Choose A_1 of the form (2.2) such that A_1 and $\tilde{A}$

have no common eigenvalue. It remains to prove that there exists a pair of the form

$$\begin{pmatrix} \tilde{A} & 0 \\ b_1 q^T & A_1 \end{pmatrix}; \quad \begin{pmatrix} \tilde{b} \\ 0 \end{pmatrix} \tag{3.4}$$

which is completely controllable. Indeed, otherwise there exist c_1 and c_2, not both zeros, and α, a complex number, such that

$$c_1^T \tilde{A} + c_2^T b_1 q^T = \alpha c_1^T \tag{3.5}$$

$$c_2^T A_1 = \alpha c_2^T \tag{3.6}$$

and

$$c_1^T \tilde{b} = 0. \tag{3.7}$$

Moreover,

$$c_2^T b_1 \neq 0 \text{ and } \det(\alpha I - \tilde{A}) \neq 0.$$

Indeed, suppose $c_2^T b_1 = 0$. Then (3.5) becomes $c_1^T \tilde{A} = \alpha c_1^T$. This equation and (3.7) imply $c_1 = 0$, because $(\tilde{A}, \tilde{b})$ is completely controllable. On the other hand, $c_2^T b_1 = 0$ and (3.6) imply $c_2 = 0$, because (A_1, b_1) is completely controllable. Therefore, one obtains a contradiction since c_1 and c_2 are not both zeros. Suppose now that $\det(\alpha I - \tilde{A}) = 0$. Then, α is an eigenvalue of $\tilde{A}$. By (3.6), α is also the eigenvalue of A_1. Thus $\tilde{A}$ and A_1 have a common eigenvalue, contrary to the hypothesis.

From (3.5), one obtains $c_1^T(\alpha I - \tilde{A}) = c_2^T b_1 q^T$. Multiplying both sides of this equation by $(\alpha I - \tilde{A})^{-1} \tilde{b}$, it becomes $c_1^T \tilde{b} = c_2^T b_1 q^T (\alpha I - \tilde{A})^{-1} \tilde{b} = 0$. Consequently, $q^T(\alpha I - \tilde{A})^{-1} \tilde{b} = 0$ (since $c_2^T b_1 \neq 0$). If no pair of the form (3.4) is completely controllable, then the above conclusion $q^T(\alpha I - \tilde{A})^{-1} \tilde{b} = 0$ holds for every α such that $\det(\alpha I - \tilde{A}) \neq 0$ (see (3.6) and observe that, by modifying the nonfixed entries of A_1, one can always satisfy (3.6), for an arbitrary α). Recall now that if the pair $(\tilde{A}, \tilde{b})$ is completely controllable, then the identity $q^T(\alpha I - \tilde{A})^{-1} \tilde{b} = 0$ (for every α such that $\det(\alpha I - \tilde{A}) \neq 0$) implies $q = 0$ (see V. M. Popov [3]). Since in (3.3) we have $q \neq 0$, the contradiction proves that the pair (3.3) is structurally controllable. Lemma 1 is proved.

Lemma 1 gives a procedure of constructing structurally controllable pairs, by starting from a pair (A,b) whose graph is a *stem* and by applying Lemma 1 successively. The graph of a pair obtained in this way will be called a "*cactus*." More precisely, the graph of a pair (A,b) is a *cactus* P if it can be obtained by starting from a stem S and by constructing a sequence of graphs $G_0 \subset G_1 \subset \cdots \subset G_k \cdots \subset G_p$ as follows: The first graph G_0 in the sequence is the stem S, and the last graph G_p in the sequence is the cactus P. Every graph G_k, with $k = 1,2,\cdots,p$, in the sequence contains all the edges of G_{k-1}; all the edges of G_k which do not belong to G_{k-1} form a bud B_k. Moreover, the origin of the bud is also the origin of one oriented edge in G_{k-1}. There exists no other node which belongs at the same time to the vertex set of this bud and to the vertex set of G_{k-1}. (The origin of the bud may coincide with the origin of G_{k-1}.) Fig. 3 gives an illustration for the case $p = 3$.

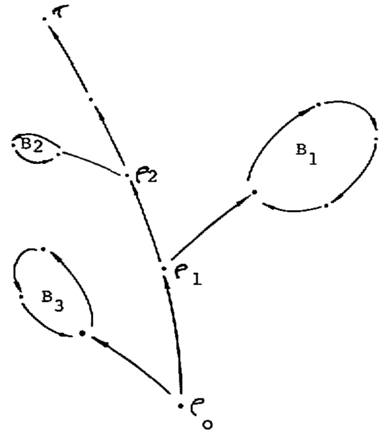


Fig. 3. A cactus.

The concept of a *cactus* can be also defined in the following descriptive form: The graph P of a pair (A,b) is a *cactus* if and only if one can write $P = S \cup B_1 \cup B_2 \cup \cdots \cup B_p$ where S is a stem and B_i are buds and, for every $i = 1,2,\cdots,p$, the origin e_i of B_i is also the origin of an (oriented) edge in the graph $S \cup B_1 \cup \cdots \cup B_{i-1}$. Moreover e_i is the only node which belongs at the same time to the vertex set B_i and to the vertex set of $S \cup B_1 \cup B_2 \cup \cdots \cup B_{i-1}$.

From the above definition and from Lemma 1, one obtains the following proposition.

Proposition 2: If the graph of a pair (A,b) is a *cactus*, then the pair is structurally controllable.

Let us say that the graph of a pair (A,b) is "*spanned by a cactus*" if it becomes a cactus after removing some or none of the edges from the graph. Then the following proposition is an easy consequence of Proposition 2 and of the definition.

Proposition 3: If the graph of a pair (A,b) is spanned by a *cactus*, then the pair (A,b) is structurally controllable.

We shall prove later that the inverse is also true.

IV. A Class of Graphs which are Cacti

In the following lemmas (2–8), we tacitly assume that G is the graph of a pair (A,b) and has the following properties.

1) There is no *nonaccessible* node in the vertext set of G.
2) There is no *dilation*.
3) G is *minimal* (i.e., after deleting any edge of the graph, one of the properties 1) and 2) is violated).

Lemma 2: Every node in G is *accessible* from the origin along one and only one *simple path*.[3]

Proof: Suppose a node α can be reached along two distinct paths and let β and $\tilde{\beta}$ be the last nodes met before α on these paths, respectively. One may assume $\beta \neq \tilde{\beta}$ since the two paths are distinct. After deleting the edge (β,α), one obtains a new graph which is denoted by G_1. According to condition 3), for G_1 one must have either a nonaccessible node or a dilation. But it is impossible to

[3] Recall that one is allowed to chase around the graph only in the direction of the arrows; a *simple path* does not go more than once through each node.

get in G_1 a nonaccessible node (since this node should necessarily be α and α can be anyhow reached along the path through $\tilde{\beta}$). Therefore, one must have a dilation in G_1. That is, (see the definition in Section II), there exists a set S of k nodes such that the set $T_1(S)$ (with respect to the graph G_1) contains no more than $k-1$ nodes. Let these sets be $S = \{\alpha_1,\alpha_2,\cdots,\alpha_k\}$ and $T_1(S) = \{\beta_1,\beta_2,\cdots,\beta_p\}$ with $p \leq k-1$. Then, 1) $\tilde{\beta} \in T_1(S)$ (by hypothesis); 2) $\alpha \in S$ (because, by replacing the edge (β,α), one must remove the dilation) and one may take $\alpha = \alpha_1$; 3) $\beta \notin T_1(S)$ (because, by replacing the edge (β,α), the set $T(S)$ (with respect to G) is $T(S) = T_1(S) \cup \{\beta\}$, and one must have $T_1(S) \cup \{\beta\} \neq T_1(S)$ since otherwise a dilation exists in the original graph).

Similarly in the graph G_2 obtained from the original graph after deleting the edge $(\tilde{\beta},\alpha)$, there exists a set $\tilde{S}$ of $\tilde{k}$ nodes such that $T_2(\tilde{S})$ (with respect to G_2) contains no more than $\tilde{k}-1$ nodes. Let these sets be $\tilde{S} = \{\tilde{\alpha}_1,\tilde{\alpha}_2,\cdots,\tilde{\alpha}_{\tilde{k}}\}$ and $T_2(\tilde{S}) = \{\tilde{\beta}_1,\tilde{\beta}_2,\cdots,\tilde{\beta}_{\tilde{p}}\}$ with $\tilde{p} \leq \tilde{k}-1$. Then as before, 1) $\beta \in T_2(\tilde{S})$; 2) $\alpha \in \tilde{S}$, and one takes $\tilde{\alpha}_1 = \alpha$; 3) $\tilde{\beta} \notin T_2(S)$ (and, after replacing the edge $(\tilde{\beta},\alpha)$, one obtains $T(\tilde{S}) = T_2(\tilde{S}) \cup \{\tilde{\beta}\}$).

Now define $\check{S} = S \cup \tilde{S}$ and consider $T(\check{S})$ (in the original graph G). One can see that

$$\begin{aligned} T(\check{S}) &= (T_1(S) \cup \{\beta\}) \cup (T_2(\tilde{S}) \cup \{\tilde{\beta}\}) \\ &= T_1(S) \cup \{\beta\} \cup T_2(\tilde{S}) \cup \{\tilde{\beta}\} \\ &= T_1(S) \cup T_2(\tilde{S}). \end{aligned}$$

Introduce the following notation: If M is any set of nodes, denote by $N(M)$ the number of distinct nodes in M. It has been seen that S and $\tilde{S}$ have always in common the node $\alpha = \alpha_1 = \tilde{\alpha}_1$. Consider first the case in which S and $\tilde{S}$ have in common only the node $\alpha = \alpha_1 = \tilde{\alpha}_1$. Then, $N(\check{S}) = N(S) + N(\tilde{S}) - 1 = k + \tilde{k} - 1$. Since $T(\check{S}) = T_1(S) \cup T_2(\tilde{S})$, one obtains

$$\begin{aligned} N(T(\check{S})) &= N(T_1(S) \cup T_2(\tilde{S})) \\ &\leq N(T_1(S)) + N(T_2(\tilde{S})) = p + \tilde{p} \\ &\leq (k-1) + (\tilde{k}-1) = k + \tilde{k} - 2. \end{aligned}$$

Therefore, one found a set $\check{S}$ of $k + \tilde{k} - 1$ nodes such that the set $T(\check{S})$ (in the original graph) contains no more than $k + \tilde{k} - 2$ nodes. Thus the original graph contains a dilation. The contradiction proves that the case in which $\tilde{S}$ and S have only the node α in common cannot take place.

Consider now the case in which S and $\tilde{S}$ have other nodes in common, besides α. Suppose that $\alpha_i = \tilde{\alpha}_i$, for $i = 1,\cdots,j$ where $1 < j \leq \min(k, \tilde{k})$, and that $\alpha_m \neq \tilde{\alpha}_l$, $\forall\, j < m \leq k$ and $\forall j < l \leq \tilde{k}$. Define $S_0 = \{\alpha_2,\alpha_3,\cdots,\alpha_j\}$ and consider the corresponding set $T(S_0)$ in the graph G. We claim that 1) $\beta \notin T(S_0)$; 2) $\tilde{\beta} \notin T(S_0)$; 3) $T(S_0) \subset T_1(S) \cap T_2(\tilde{S})$; and 4) $N(T(S_0)) \geq j-1$. To prove 1), assume $\beta \in T(S_0)$. Then, since $\alpha_1 \notin S_0$, one also obtains $\beta \in T_1(S)$ (in the graph G_1, with (β,α) removed). This is a contradiction (because $\beta \notin T_1(S)$). The same arguments prove (2). To prove (3), consider a node $v \in T(S_0)$. Then, $v \neq \beta$ and $v \neq \tilde{\beta}$ (because of (1) and (2)). One also has $v \in T_1(S) \cup \{\beta\}$ because $S_0 \subseteq S$, and, therefore, $v \in T(S_0)$ implies $v \in T(S)$. Moreover, from the definitions, it follows that $T(S) = T_1(S) \cup \{\beta\}$. Similarly, one obtains $v \in T_2(\tilde{S}) \cup \{\tilde{\beta}\}$, but $v \in T_1(S) \cup \{\beta\}$ and $v \neq \beta$ imply $v \in T_1(S)$. Similarly, $v \in T_2(\tilde{S}) \cup \{\tilde{\beta}\}$ and $v \neq \tilde{\beta}$ imply $v \in T_2(\tilde{S})$. Thus $v \in T_1(S) \cap T_2(\tilde{S})$ and (3) is proved. To prove (4), assume $N(T(S_0)) < j-1$. Since $N(S_0) = j-1$, one obtains a contradiction, namely that G has a dilation.

Using the above result and introducing, as before, the set $\check{S} = S \cup \tilde{S}$, one obtains

$$\begin{aligned} N(\boldsymbol{T}(\check{\boldsymbol{S}})) &= N(T_1(S) \cup T_2(\tilde{S})) \\ &= N(T_1(S)) + N(T_2(\tilde{S})) - N(T_1(S) \cap T_2(\tilde{S})) \\ &\leq N(T_1(S)) + N(T_2(\tilde{S})) - N(T(S_0)) \\ &\leq p + \tilde{p} - (j-1) \\ &\leq (k-1) + (\tilde{k}-1) - (j-1) = k \\ &\quad + \tilde{k} - j - 1. \end{aligned}$$

On the other hand, the number of distinct nodes in $\check{S}$ is $N(\check{S}) = j + (\tilde{k} - j) + (k - j) = k + \tilde{k} - j$. Hence, G has a dilation. The contradiction proves Lemma 2.

Let $\eta_1,\cdots,\eta_i,\cdots,\eta_r$ be the oriented edges in G with the property that their origins coincide with the origin of G. Define V_i to be the set of all the nodes which can be reached from the origin of G by passing through the edge η_i. Then Lemma 2 implies that the sets V_i are disjoint, i.e., $V_i \cap V_j = \phi$ for every couple of distinct integers i and j. Furthermore, from assumption 1), it follows that $V = V_1 \cup V_2 \cup \cdots \cup V_r \cup \{\mathfrak{e}\}$ where V is the vertex set of G and $\mathfrak{e}$ is the origin of G.

For every V_i as above, denote by G_i the subgraphs of G whose set of nodes is exactly $V_i \cup \{\mathfrak{e}\}$ and whose edges are all the edges from G of the form (α,β) with $\alpha \in V_i \cup \{\mathfrak{e}\}$ and $\beta \in V_i$. The subgraphs G_i, defined as above, will be called "*bunches.*" Clearly, one can write $G = G_1 \cup G_2 \cup . \cup G_r$ and G_i are (edge) disjoint (Lemma 2).

A subgraph G_i, defined as above, is called a "*terminal bunch*" if there exists a subset $S \subset V_i$ such that $N(T(S)) = N(S)$ and $T(S)$ contains the origin of G. Observe that from the definition, it follows that $T(S) \subset V_i \cup \{\mathfrak{e}\}$ and the set S contains one and only one *final* node v such that $v \in S$ and $v \notin T(S)$. Indeed, if there is no node with the mentioned properties in S, then $T(S) \supset S$ and $N(T(S)) > N(S)$ (because $\mathfrak{e} \in T(S)$ and $\mathfrak{e} \notin S$). This contradiction proves that S contains a final node v. If there are k such nodes, say $v_1, v_2,\cdots,v_k$ with $k > 1$, then $v_i \in S$ and $v_i \notin T(S)$ for $i = 1,2,\cdots,k$. Thus $N(T(S)) = N(S) + 1 - k < N(S)$ (because $k > 1$ by hypothesis). This contradiction proves that S contains only one terminal node. Note also the following lemmas.

Lemma 3: If G_i is not a *terminal bunch*, then for every set $S \subset V_i$ such that $T(S)$ contains the origin, one has $N(T(S)) - N(S) \geq 1$. (Indeed, the case $N(T(S)) < N(S)$ is contradictory since it implies the existence of a dilation in G. On the other hand, the case $N(T(S)) = N(S)$ is not possible since then G_i is a terminal bunch.)

Lemma 4: There exists at most one *terminal bunch* in G.

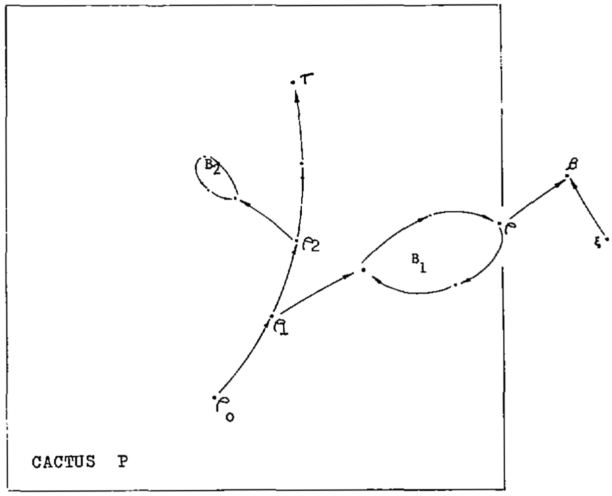


Fig. 4. An illustration of Remark 1.

Proof: Suppose there are more than one terminal bunches. Without loss of generality, let G_1 and G_2 be two terminal bunches in G. Then there exists $S_i \subset V_i$ such that $N(S_i) = N(T(S_i))$ and $T(S_i)$ contains the origin of G, with $i = 1,2$. Define $S = S_1 \cup S_2$ and consider the corresponding set $T(S)$ in G. Since V_1 and V_2 are disjoint, S_1 and S_2 are also disjoint, and $T(S_1)$ and $T(S_2)$ have in common only the origin of G. Therefore, one obtains

$$\begin{aligned} N(T(S)) &= N(T(S_1)) + N(T(S_2)) - 1 \\ &= N(S_1) + N(S_2) - 1 \\ &= N(S) - 1. \end{aligned}$$

The following remark states one of the properties of a terminal brunch G_i and is needed in the proof of Lemma 8.

Remark 1: With reference to Fig. 4, suppose that a cactus P, whose vertex set contains the origin of G_i, is a subgraph of the terminal bunch G_i. Let $(\mathfrak{e},\beta) \not\subseteq P$ be an oriented edge in G_i such that the node $\mathfrak{e}$ of $(\mathfrak{e},\beta)$ has in common with the vertex set of P one and only one node (excepting the final node of P). Then there exists "another" edge (ξ,β) entering β. (Indeed, let R be the set of all nodes which can be reached from $\mathfrak{e}$ "against" the oriented edge in G_i and all the nodes which belong to the vertex set of the stem, (excepting the origin of P). Define $S = \{\beta\} \cup R$. Suppose there is no other edge entering β, excepting $(\mathfrak{e},\beta)$. Clearly, $T(S) = R \cup \{\mathfrak{e}_0\} / \{\tau\}$ and $N(T(S)) = N(R)$, where $\mathfrak{e}_0$ and τ indicate the origin and the final node of the stem, respectively. Therefore, $N(T(S)) < N(S)$ because $N(S) = N(R) + 1$. Thus one obtains a dilation; a contradiction.) Now the concept of a "*precactus*" is defined as follows.

A graph $\textcircled{H}$ is a "*precactus*" if and only if one can write $\textcircled{H} = B_1 \cup B_2 \cup \cdots \cup B_p$, where B_i are the buds, such that for every $i = 2,3,\cdots,p$, the origin $\mathfrak{e}_i$ of B_i is also the origin of one oriented edge in the graph of $B_1 \cup B_2 \cup \cdots \cup B_{i-1}$. Moreover, $\mathfrak{e}_i$ is the only node which belongs at the same time to the vertex set of B_i and to the vertex set of $B_1 \cup B_2 \cup \cdots \cup B_{i-1}$.

Lemma 5: Every *precactus* becomes a *cactus* after eliminating one or more suitable edges.

Proof: From the above definition, a *precactus* can be written in the form $\textcircled{H} = B_1 \cup B_2 \cup \cdots \cup B_p$. Let $\mathfrak{e}_1$ be the origin of B_1 and $(\mathfrak{e}_1,\beta_1)$ be the distinguished edge of B_1. Then besides the edge $(\mathfrak{e}_1,\beta_1)$, the bud contains one and only one edge of the form $(\mathfrak{e}_2,\beta_1)$ where $\mathfrak{e}_2$ is another node belonging to the vertex set of B_1. (One has $\mathfrak{e}_2 \neq \mathfrak{e}_1$; one might have $\mathfrak{e}_2 = \beta_1$.) Remove the edge $(\mathfrak{e}_2,\beta_1)$. Then one can write $B_1 = S_1 \cup (\mathfrak{e}_2, \beta_1)$, where S_1 is a stem whose vertex set is identical to the vertex set of B_1. If $\mathfrak{e}_2$ is not the origin of any other bud, $B_2, B_3, \cdots, B_p$, then the proof is complete. If not, then $\mathfrak{e}_2$ is the origin of another bud (one can assume this bud is B_2), and the same procedure can be repeated obtaining a new node $\mathfrak{e}_3$ and a new relation $B_1 \cup B_2 = S_2 \cup (\mathfrak{e}_2,\beta_1) \cup (\mathfrak{e}_3,\beta_2)$ where S_2 is a stem whose vertex set is identical to the vertex set of $B_1 \cup B_2$. After applying the procedure the maximal number of times, one obtains a node $\mathfrak{e}_{n+1} (n \leq p)$ which is not the origin of any bud, $B_{n+1}, B_{n+2}, \cdots, B_p$. Therefore, one has $B_1 \cup B_2 \cup \cdots \cup B_n = S_n \cup (\mathfrak{e}_2,\beta_1) \cup (\mathfrak{e}_3,\beta_2) \cup \cdots \cup (\mathfrak{e}_{n+1},\beta_n) = S_n \cup G_0$ where $G_0 = (\mathfrak{e}_2,\beta_1) \cup (\mathfrak{e}_3,\beta_2) \cup \cdots \cup (\mathfrak{e}_{n+1},\beta_n)$ and S_n is a stem whose vertex set is identical to the vertex set of $B_1 \cup B_2 \cup \cdots \cup B_n$. Thus $B_1 \cup B_2 \cup \cdots \cup B_n \cup B_{n+1} \cup \cdots \cup B_p = S_n \cup G_0 \cup B_{n+1} \cup \cdots \cup B_p = P \cup G_0$, where $P = S_n \cup B_{n+1} \cup \cdots \cup B_p$. Clearly, the graph P satisfies the conditions of a cactus. Q.E.D.

Lemma 6: Any nonterminal bunch becomes a precactus, possibly after eliminating some edges of the bunch. (In other words, any nonterminal bunch is "spanned" by a precactus.)

Proof: With η_i, V_i, and G_i defined as before, let G_1 be a nonterminal bunch and $\beta_1 \in V_1$ be the only successor of the origin along the (oriented) edge η_1 (issuing from the origin). Then there must exist another edge entering β_1 from a node ξ_1. (Otherwise one can define $S = \{\beta_1\}$ and $T(S) = \{\mathfrak{e}\}$ = the origin. This is contrary to Lemma 3.) One might have $\xi_1 = \beta_1$. Assume first that $\xi_1 \neq \beta_1$. There is one and only one simple path which connects ξ_1 to the origin (Lemma 2). This path goes through β_1 since G_1 is

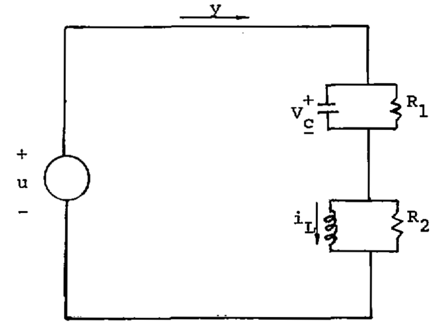


Fig. 5. An RLC circuit.

disjoint of the other bunches of G. Therefore, one can find a loop whose set of nodes contains the node β_1. This loop together with the (distinguished) edge $(\mathfrak{e}_1,\beta_1)$ entering β_1 forms a bud B_1 (where $\mathfrak{e}_1$ is the origin). If $\beta_1 = \xi_1$, one still has a bud.

If the vertex set of B_1 is the same as the vertex set of G_1, then G_1 is "spanned" by B_1, which, by definition, is a precactus. If not, one takes a new node q. There is one and only one simple path π connecting the origin to q. Let $\mathfrak{e}_2$ be the last node in the vertex set of B_1 which is met along this path. Let β_2 be the first node which is met after $\mathfrak{e}_2$ on this path. Applying to the node β_2 the same arguments which were used before connecting the node β_1, one obtains another bud B_2. Moreover $\mathfrak{e}_2$ is the origin of B_2 and is the only node which belongs at the same time to the vertex set of B_1 and to the vertex set of B_2. The same procedure can be applied successively until all nodes of V_1 belong to the vertex set of some buds $B_1,B_2,\cdots,B_p$ (because V_1 contains only a finite number of nodes). Clearly, the origin $\mathfrak{e}_i$ of the bud B_i $(1 < i \leq p)$ is the only node which belongs at the same time to the vertex set of B_i and to the vertex set of $B_1 \cup B_2 \cup \cdot \cup B_{i-1}$. Therefore, Ⓗ $= B_1 \cup B_2 \cup \cdots \cup B_p$ is a precactus according to the previous definition. Lemma 6 is proved.

Lemma 7: There always exists a *terminal bunch* in G.

Proof: Suppose there are no terminal bunches in G. One saw that all these nonterminal bunches in G are disjoint from one another, but every nonterminal bunch is spanned by a precactus (Lemma 6). Thus G is spanned by a precactus. But then G does not satisfy assumption 3) at the beginning of this section.

Indeed, Lemma 5 shows that, after eliminating one or several suitable edges from the precactus, one obtains a cactus P. By Proposition 2 in Section III, P is the graph of a structurally controllable pair (A,b). As shown at the end of Section II, this implies that P satisfies the assumptions 1) and 2) at the beginning of this section. Thus assumption 3) is contradicted and Lemma 7 is proved.

Lemma 8: Any *terminal bunch* G_1 is spanned by a *cactus.*

Proof: The reasonings are similar to those in the proof of Lemma 6. Since G_1 is a terminal bunch, one can find a final node v. As proved before, there exists one and only one simple path from the origin to v (Lemma 2). Define the stem S as this path. If all the nodes of V_1 belong to the vertex set of this stem, then Lemma 8 is proved. Otherwise let q be a node which does not belong to the vertex set of the stem. Moreover, q is not a final node of G_1, because the final node v is unique.

There is one and only one simple path π connecting the origin to q. Let $\mathfrak{e}_1$ be the last node in the vertex set of S which is met along this path. Let β_1 be the first node which is met after $\mathfrak{e}_1$ along this path. Then, as before, there must exist "another" edge entering β_1 from a node ξ_1 (see Remark 1). Now the same arguments and procedure as in Lemma 6 can be applied here to obtain the conclusion that G_1 is spanned by a graph $S \cup B_1 \cup B_2 \cup \cdots \cup B_p$ where B_i are buds. Moreover, the origin $\mathfrak{e}_i$ of the bud $B_i (1 \leq i \leq p)$ is the only node which belongs at the same time to the vertex set of B_i and to the vertex set of $S \cup B_1 \cup B_2 \cup \cdots \cup B_{i-1}$. Thus $S \cup B_1 \cup B_2 \cup \cdots \cup B_p$ is a cactus and Lemma 8 is proved.

Now we can state the conclusion of this section.

Proposition 4: If the graph of a pair (A,b) satisfies the properties 1) through 3) at the beginning of this section, then G is a cactus.

Proof: Indeed, we proved that if 1)–3) are satisfied, then G is the graph of a finite number of bunches G_i. Moreover, one and only one of these bunches is a terminal bunch (see Lemmas 4 and 7), and this terminal bunch is spanned by a cactus (Lemma 8) while all the other bunches are spanned by precacti. This implies that G itself is spanned by a cactus. As in the proof of Lemma 7, condition 3) implies that G is a cactus.

V. Conclusions

The results of this paper are summarized by the following theorem.

Theorem: The following properties are equivalent.

1° The pair (A,b) is *structurally controllable.*

2° There is no permutation of coordinates, bringing the pair (A,b) to one of the forms (1.2) and (1.3).

3° The graph of (A,b) contains no *nonaccessible* node and no *dilation.*

4° The graph of (A,b) is spanned by a *cactus.*

Indeed, in Section I we proved that $1° \Rightarrow 2°$. In Section II, we have also shown that $2° \Leftrightarrow 3°$. In Section IV, we proved that $3° \Rightarrow 4°$. Proposition 3 of Section III states that $4° \Rightarrow 1°$. All these results prove the theorem.

The concepts obtained above imply that the set of all the noncompletely controllable pairs (A,b) of given dimension can be partitioned into two sets: The set of the pairs (A,b), which are not structurally controllable, and the set of the pairs (A,b), which are structurally controllable. We also gave several necessary and sufficient conditions of structural controllability. The introduction of the structurally controllable set is justified by the remark that, taking into account the presence of the errors of measurements, we cannot decide, by physical measurements, whether or not a pair (A,b) is completely controllable or only structurally controllable. For instance, consider the circuit shown in Fig. 5, in which u is the input voltage and y is the output current. It is a single-input single-output system. Let the voltage across the capacitor and the current flowing through the inductor be the state-variables x_1 and x_2, respectively [4]. Then the system can be easily described as $\dot{x} = Ax + bu$, (denoted by the pair (A,b)), where

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \; A = \begin{bmatrix} -\frac{1}{C}\left(\frac{1}{R_1} + \frac{1}{R_2}\right)\frac{1}{C} \\ -\frac{1}{L} & 0 \end{bmatrix} \text{ and } b = \begin{bmatrix} \frac{1}{R_2C} \\ \frac{1}{L} \end{bmatrix}.$$

Suppose the components of the circuit, R_1, R_2, C, and L, have unit values, and let (A_0,b_0) be the pair with these specified values in the entries of (A,b). Then the pair (A_0,b_0) is of the form $A_0 = \begin{pmatrix} -2 & 1 \\ -1 & 0 \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Clearly, the pair (A_0,b_0) is not completely controllable since rank $(b_0 \; A_0b_0) < 2$. However, by the definition of structural controllability or the theorem in this section, one sees that the pairs (A,b) and (A_0,b_0), having the same structure, are both structurally controllable.

The study of the structure and the graph of a pair (A,b) (1.1) in this work results in a more precise controllability property in the control system. Furthermore, it will also be possible to derive some basic properties in matrix and graph theories. For instance, by examining the theorem in this section, structural controllability equivalence conditions could be related to the irreducibility properties in matrix theory, ergodic properties of Markov chains, and the idea of the shortest length in graph theory. Moreover, the control system (1.1) studied here has a single input and is time invariant. It is left for further research to establish results like the theorem in Section V for the multi-input and time-variant cases. It would be also interesting to study the dual property of structural observability of a linear observed system. Thus the concepts of "structure" and "structural controllability" and the results obtained here are only the first step in what is believed to be a promising new branch in control system theory.

Acknowledgment

The author wishes to express his gratitude to Prof. V. M. Popov of the University of Maryland for his suggestion and guidance in the course of this research.

References

[1] E. B. Lee and L. Markus, *Foundations of Optimal Control Theory*. New York: Wiley, 1967.

[2] O. Ore, "Theory of Graphs," *Amer. Math. Soc. Colloquium Publ.*, vol. 38, 1962.

[3] V. M. Popov, *Hyperstavilitates sistemelor automate* (in Romanian). New York: Springer, to be published.

[4] L. A. Zadeh and C. A. Desoer, *Linear System Theory*. New York: McGraw-Hill, 1963.

[5] F. R. Gantmacher, *The Theory of Matrices*. New York: vol. 1 and 2, Chelsea Publishing Company, 1960.

[6] R. E. Kalman, "Mathematical description of linear dynamical systems," *SIAM J. Contr.*, vol. 1, 1963.

[7] R. W. Brockett, *Finite Dimensional Linear Systems*. New York: Wiley, 1970.

[8] R. S. Varga, *Matrix Iterative Analysis*. Englewood Cliffs, N.J.: Prentice-Hall, 1962.

[9] J. G. Kemeny and J. L. Snell, *Finite Markov Chains*. Princeton, N. J.: Van Nostrand, 1960.

[10] D. G. Luenberger, "Canonical forms for linear multivariable systems," *IEEE Trans. Auto. Contr.*, June 1967.

[11] Dénes König, *Theories der Endlichen und Unendlichen Graphen*. New York: Chelsea Publishing Company, 1950.

[12] C. T. Lin, "Structural controllability," Ph.D. dissertation, Dep. Elec. Eng., Univ. of Maryland, College Park, Md., 1972.

Ching-Tai Lin (S'73–M'73) was born in Taipei, China, on October 14, 1944. He received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, China in 1967, and the M.S. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, Md. in 1970 and 1972, respectively.

From 1967 to 1968, he was a Second Lieutenant Electronics Officer in the Chinese Air Force, Taiwan, China. He was a Research Assistant and a Teaching Assistant at the Electronics Research Laboratory, University of Maryland during 1968–1969 and 1969–1972, respectively. In 1973, he, as an Instructor, taught at the Electrical Engineering Department, Howard University, Washington, D.C. Since June 1973, he has been with Bendix Field Engineering Corp., Columbia, Md. where he is now a Senior System Engineer. His current research interests include linear system theory and process control, and stability theory.