Manuscript submitted: 28 July 1966

The research reported herein was supported in part by the National Science Foundation under Grant No. GK-569/716 and the U. S. Army Research Office, Durham, N. C., Grant No. DA-ARO-D-31-124-G576.

INVERTIBILITY, EQUIVALENCE, AND THE DECOMPOSITION PROPERTY OF ABSTRACT SYSTEMS

by

K. Doty and H. Frank

Memorandum No. ERL-M181

16 September 1966

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

ABSTRACT

The concepts of inverse and left inverse systems, weakly equivalent and strongly equivalent systems, and discrete time and finite memory systems are examined. Various relations between these concepts are found, and the role of the decomposition property of the system input-output-state relation is investigated. Among the results obtained are: (a) under suitable controllability assumptions, left inverse systems are inverse systems, (b) weakly equivalent finite state systems with the decomposition property are strongly equivalent, and (c) finite state systems with the decomposition property have finite memory.

INTRODUCTION

An abstract system, or simply a system, $\mathcal{A}$ is defined by Zadeh and Desoer[1] as a "partially interconnected set of abstract objects $\mathcal{A}_1, \mathcal{A}_2, \ldots,$ termed the components of $\mathcal{A}$. . . ." The components of $\mathcal{A}$, as well as $\mathcal{A}$ itself, may be defined by specifying an input-output relation, or equivalently, an admissible set of input-output pairs.

Let $\mathcal{T}$ be the set of points $(-\infty, \infty)$ on the real line. If $T \subset \mathcal{T}$ let (u_T, y_T) be an ordered pair of time functions defined on T such that for $t \in T$, $u(t)$ and $y(t)$ are the values of u_T and y_T at time t, respectively. Now, let T be a semiclosed interval $[t_0, t)$ on the real line where $t_0 < t$, and let $(u_{[t_0, t)}, y_{[t_0, t)})$ be the corresponding ordered pair of time functions defined on this interval. An abstract system $\mathcal{A}$ will be specified as a set of ordered pairs $(u_{[t_0, t)}, y_{[t_0, t)})$ $\forall t_0, t \in \mathcal{T} \ni t_0 < t\}$; this relation is denoted as $\mathcal{A} =$ $\{(u_{[t_0, t)}, y_{[t_0, t)})\}$ and the pair $(u_{[t_0, t)}, y_{[t_0, t)}) \in \mathcal{A}$ is known as an input-output pair of $\mathcal{A}$.

The time-function segments $u_{[t_0, t)}$ and $y_{[t_0, t)}$ are elements of the function spaces $R_a[u_{[t_0, t)}]$ and $R_a[y_{[t_0, t)}]$, respectively, where

$$R_a[u_{[t_0, t)}] = \{u_{[t_0, t)} \mid (u_{[t_0, t)}, y_{[t_0, t)}) \in \mathcal{A} \text{ for some } y_{[t_0, t)}\}, \tag{1}$$

and

$$R_a[y_{[t_0, t)}] = \{y_{[t_0, t)} \mid (u_{[t_0, t)}, y_{[t_0, t)}) \in \mathcal{A} \text{ for some } u_{[t_0, t)}\}. \tag{2}$$

Then we define the input function space $\mathcal{R}_a[u]$ and the output function space $\mathcal{R}_a[y]$ as, respectively,

$$\mathcal{R}_a[u] = \{R_a[u_{[t_0, t)}] \ \forall t_0, t \in \mathcal{I} \ni t_0 < t\}, \tag{3}$$

and

$$\mathcal{R}_a[y] = \{R_a[y_{[t_0, t)}] \ \forall t_0, t \in \mathcal{I} \ni t_0 < t\} \tag{4}$$

In general, for each input $u^*_{[t_0, t)} \in R_a[u_{[t_0, t)}]$, the set of all $(u^*_{[t_0, t)}, y_{[t_0, t)}) \in \mathcal{A}$ contains more than one element. One way of associating a unique $y_{[t_0, t)}$ with each $u_{[t_0, t)}$ is by parameterization of the space of input-output pairs. We will assume that a parameterization can be found which obeys

<u>Assumption 1.</u> There exists: (1) a set $\sum_a$; (2) a causal function $\bar{A}$ on the product space $\sum_a \times \mathcal{R}_a[u]$ into $\mathcal{R}_a[y]$ such that for any $y_{[t_0, t)} \in \mathcal{R}_a[y]$, there exists an $\alpha \in \sum_a$ and a $u_{[t_0, t)} \in R_a[u_{[t_0, t)}]$ where

$$y_{[t_0, t)} = \bar{A}(\alpha; u_{[t_0, t)}); \tag{5}$$

and (3) a causal function S_a on the product space $\sum_a \times \mathcal{R}_a[u]$ into $\sum_a$ such that for any α^*, $u^*_{[t_0, t)}$ and $y^*_{[t_0, t)}$ which satisfy (5),

$$y^*_{[\tau, t)} = \bar{A}(S_a(\alpha^*; u^*_{[t_0, \tau)}); u^*_{[\tau, t)}) \tag{6}$$

for all τ such that $t_0 < \tau < t$.

For simplicity, we will denote $u_{[t_0, t)}$ by u, and $y_{[t_0, t)}$ by y. Equation (5) then becomes

$$y = \overline{A}(\alpha ; u) . \tag{7}$$

The set $\sum_a$ is known as the state space of $\mathcal{A}$ and the element α in the right-hand side of Eq. (5) is called the state of $\mathcal{A}$ at time t_0. The function $S(\alpha ; u)$ is the state transition function of $\mathcal{A}$, while Eq. (5) is the input-output-state equation of $\mathcal{A}$.

The "state space" approach to system theory is conceptually related to the theory of discrete time systems and automata.[2, 3] Briefly, in a discrete time system, the input u and the output y are sequences of the form $u_0u_1u_2 \ldots$ and $y_0y_1y_2 \ldots$, where u_λ is the value of u at time t_λ, with λ ranging over the integers. The state equation for a discrete time system is of the form

$$y_\lambda = \overline{A}(\alpha ; u_\lambda). \tag{8}$$

A discrete time system $\mathcal{A}$ is said to be a finite state system if the state space $\sum_a$ of $\mathcal{A}$ has a finite number of elements. In this paper, we investigate some concepts of system theory and their equivalents in the theory of discrete time and finite state systems. Specifically, we investigate some problems concerning the existence of inverse systems, equivalence in weakly equivalent systems, and finite memory in discrete time systems. In many of our results, the decomposition property[1] plays an important role.

INVERSE SYSTEMS

Informally, any two systems $\mathcal{A}$ and $\mathcal{B}$ are inverses if $\mathcal{A}$ is capable of "undoing" whatever $\mathcal{B}$ is capable of doing. The idea of inverses and system invertibility is closely related to the concept of the "information lossless machines" of the theory of automata.[2, 4, 5] If a finite state machine is information lossless, then given the machine in any known initial state, knowledge of the response to an unknown input sequence is sufficient[1] to identify that excitation. We will see below that if $\mathcal{A}$ has a <u>left inverse</u>, then $\mathcal{A}$ is information lossless. However, at this point we begin with the following definition

<u>Definition.</u> Let $\mathcal{A}$ and $\mathcal{B}$ be characterized by input-output state relations of the form

$$\mathcal{A}: \quad y = \bar{A}(\alpha\ ;\ u), \alpha \in \sum_a ,\ u \in \mathcal{R}_a[u],\ y \in \mathcal{R}_a[y], \tag{9}$$

$$\mathcal{B}: \quad w = \bar{B}(\beta\ ;\ v), \beta \in \sum_b ,\ v \in \mathcal{R}_b[v],\ w \in \mathcal{R}_b[w], \tag{10}$$

where u and v denote input segments to $\mathcal{A}$ and $\mathcal{B}$, respectively, y and w are the corresponding output segments and $\mathcal{R}_a[u] \equiv \mathcal{R}_b[w]$, $\mathcal{R}_a[y] \equiv \mathcal{R}_b[v]$. (The input function space of $\mathcal{A}$ is the output function space of $\mathcal{B}$, and vice versa.) Then $\mathcal{A}$ and $\mathcal{B}$ are <u>inverse systems</u> if and only if to every state α of $\mathcal{A}$, there is an <u>inverse state</u> β_α of $\mathcal{B}$, such that

$$\bar{B}(\beta_\alpha\ ;\ \bar{A}(\alpha\ ;\ u)\ ) = u \qquad \forall\ u \in \mathcal{R}_a[u], \tag{11}$$

and conversely, to every state β of $\mathcal{B}$, there is an inverse state α_β of $\mathcal{A}$ such that

$$\bar{A}(\alpha_\beta\ ;\ \bar{B}(\beta\ ;\ v)\ ) = v \qquad \forall\ v \in \mathcal{R}_b[v]. \tag{12}$$

Remark. The assumption that $\mathcal{R}_a[u] \equiv \mathcal{R}_b[w]$ and $\mathcal{R}_a[y] \equiv \mathcal{R}_b[v]$ is vital since, otherwise, the operations described by Eqs. (11) and (12) are meaningless.

The above definition is illustrated in Fig. 1. If $\mathcal{A}$ and $\mathcal{B}$ are inverses (written $\mathcal{A} = \mathcal{B}^{-1}$ or $\mathcal{B} = \mathcal{A}^{-1}$), given the system $\mathcal{A}$ in state α, we can construct the tandem system $\mathcal{B}\mathcal{A}$ such that if $\mathcal{B}$ is in state β_α, the tandem system behaves like an identity operator or "unitor" on any input u. A corresponding statement can be made about the tandem system $\mathcal{A}\mathcal{B}$. To denote the equivalence of $\mathcal{B}\mathcal{A}$ (and $\mathcal{A}\mathcal{B}$) to a unitor, we write $\mathcal{B}\mathcal{A} \overset{*}{=} I$ (and $\mathcal{A}\mathcal{B} \overset{*}{=} I$). The asterisk above the equal sign indicates that the equivalence is conditional[1] and only holds if $\mathcal{A}$ and $\mathcal{B}$ are in their proper inverse states.

The definition of inverse systems is symmetrical with respect to an interchange of the names of systems $\mathcal{A}$ and $\mathcal{B}$ ($\mathcal{A}$ is inverse to $\mathcal{B}$ and $\mathcal{B}$ is inverse to $\mathcal{A}$). In many cases, only one of the relations given by Eqs. (11) and (12) will hold. Such a situation is shown in Fig. 2. It is easily seen that the tandem combination $\mathcal{B}\mathcal{A}$ in Fig. 2(a) is conditionally equivalent to a unitor but the system $\mathcal{A}\mathcal{B}$ shown in Fig. 2(b) is not (unless v_1 and v_2 are specially constrained). This example leads us to the following definition:

Definition. Let $\mathcal{A}$ and $\mathcal{B}$ be characterized by input-output state relations of the form

$$\mathcal{A}: \; y = \bar{A}(\alpha; u), \; \alpha \in \Sigma_a, \; u \in \mathcal{R}_a[u], \; y \in \mathcal{R}_a[y], \tag{13}$$

$$\mathcal{B}: \; w = \bar{B}(\beta; v), \; \beta \in \Sigma_b, \; v \in \mathcal{R}_b[v], \; w \in \mathcal{R}_b[w], \tag{14}$$

where $\mathcal{R}_a[y] \subseteq \mathcal{R}_b[v]$ and $\mathcal{R}_a[u] \subseteq \mathcal{R}_b[w]$. Then $\mathcal{B}$ is a left inverse of $\mathcal{A}$ if for every state α of $\mathcal{A}$, there exists a left inverse state β_α of $\mathcal{B}$ such that

$$\bar{B}(\beta_\alpha ; \bar{A}(\alpha ; u)) = u \qquad \forall\ u \in \mathcal{R}_a[u]. \tag{15}$$

An immediate consequence of this definition is that if $\mathcal{B}$ is a left inverse of $\mathcal{A}$ (i.e., $\mathcal{B}\mathcal{A} \overset{*}{=} I$) and $\mathcal{A}$ is a left inverse of $\mathcal{B}$ ($\mathcal{A}\mathcal{B} \overset{*}{=} I$), then $\mathcal{A}$ and $\mathcal{B}$ are inverse systems. It should be clear that finding a left inverse of a finite state system is equivalent to finding a finite set of rules with which, given the initial state and observed response of the system, the unknown excitation can always be determined. Thus, systems with left inverses are information lossless systems.[2] The converse is not true. Not all information lossless systems have left inverses. An example of such a system is shown in Fig. 3.

We can also define a right inverse. Briefly, $\mathcal{B}$ is a right inverse of $\mathcal{A}$ if $\mathcal{R}_b[w] \subseteq \mathcal{R}_a[u]$, $\mathcal{R}_b[v] \subseteq \mathcal{R}_a[y]$ and for every state α of $\mathcal{A}$, there exists a right inverse state β_α of $\mathcal{B}$ such that

$$\bar{A}(\alpha ; \bar{B}(\beta_\alpha ; v)) = v \qquad \forall v \in \mathcal{R}_b[v]. \tag{16}$$

If a system $\mathcal{A}$ has a left or right inverse, it is natural to ask whether the system is invertible. Clearly, from the example in Fig. 2, in general the system is not invertible. However, if systems $\mathcal{A}$ and $\mathcal{B}$ have additional structure, a more positive conclusion can be reached.

Definition. A system $\mathcal{A}$ is said to be completely state controllable[6] if given any $\alpha_0, \alpha_1 \in \sum_a$, there exists a $u_{[t_0, t)} \in \mathcal{R}_a[u]$ for any $t_0 \in \mathcal{T}$ such that $\alpha_1 = S_a(\alpha_0 ; u_{[t_0, t)})$. A system $\mathcal{A}$ is said to be functionally reproducible[7] if, given any $\alpha_0 \in \sum_a$ and any $y_{[t_0, t)} \in \mathcal{R}_a[y]$, there exists a $u_{[t_0, t)} \in R_a[u_{[t_0, t)}]$

such that $y_{[t_0, t)} = \bar{A}(\alpha_0 ; u_{[t_0, t)})$.

Theorem 1. If $\mathcal{B}$ is a completely state controllable left inverse of a functionally reproducible system $\mathcal{A}$ such that $\mathcal{R}_b[w] \equiv \mathcal{R}_a[u]$, then $\mathcal{A}$ and $\mathcal{B}$ are inverse systems.

Proof. We first note that if, in the tandem connection $\mathcal{B}\mathcal{A}$, the state of $\mathcal{A}$ is α and the state of $\mathcal{B}$ is β_α (the left inverse state of α), then for any $u \in \mathcal{R}_a[u]$, the state $S_b(\beta_\alpha ; \bar{A}(\alpha ; u))$ is a left inverse state of $S_a(\alpha ; u)$. In other words, the input of $\mathcal{A}$ takes $\mathcal{A}$ to state $\alpha^* = S_a(\alpha ; u)$ and the output of $\mathcal{A}$ takes $\mathcal{B}$ to the left inverse state $\beta_{a^*} \triangleq S_b(\beta_\alpha ; \bar{A}(\alpha ; u))$ of α^*. Furthermore, if $\mathcal{A}$ is functionally reproducible, then β_{α^*} is the unique[3] left inverse state of α^*. This follows since otherwise, suppose β' is also a left inverse state of α^*, then

$$\bar{B}(\beta_{\alpha^*} ; \bar{A}(\alpha^* ; u)) = u = \bar{B}(\beta' ; \bar{A}(\alpha^* ; u)). \tag{17}$$

But, since $\mathcal{A}$ is functionally reproducible and $\mathcal{R}_a[y] \equiv \mathcal{R}_b[v]$,

$$\mathcal{R}_a[y] \equiv \{ \bar{A}(\alpha ; u) \mid u \in \mathcal{R}_a[u] \} \equiv \mathcal{R}_b[v]. \tag{18}$$

Hence,

$$\bar{B}(\beta_{\alpha^*} ; v) = \bar{B}(\beta' ; v) \qquad \forall v \in \mathcal{R}_b[v], \tag{19}$$

and so $\beta_{\alpha^*} \cong \beta'$.

Consider the tandem connection $\mathcal{A}\mathcal{B}\mathcal{A}$, where each system $\mathcal{A}$ is in state α and $\mathcal{B}$ is in the unique left inverse state β_α. Apply the input u_1 to the compound system. Now,

$$\bar{B}(\beta_\alpha ; \bar{A}(\alpha ; u_1)) = u_1 \tag{20}$$

and so, if we let $v_1 = \bar{A}(\alpha\ ;\ u_1)$, we have

$$v_1 = \bar{A}(\alpha\ ;\ \bar{B}(\beta_\alpha\ ;\ v_1)) . \tag{21}$$

From the functional reproducibility of $\mathcal{A}$, Eq. (21) is true for all $v_1 \in \mathcal{R}_b[v_1]$. This statement is illustrated in Fig. 4.

Finally, since $\mathcal{B}$ is completely state controllable, every state $\beta \in \sum_b$ is the inverse state of some state of $\mathcal{A}$. Therefore, for each state β of $\mathcal{B}$, we can find a state α_β of $\mathcal{A}$ such that

$$v = \bar{A}(\alpha_\beta\ ;\ \bar{B}(\beta\ ;\ v)) \qquad \forall v \in \mathcal{R}_b[v] . \tag{22}$$

This implies that $\mathcal{A}$ is a left inverse of $\mathcal{B}$ and the theorem is proved.

Remark. We have actually proved the somewhat stronger result that the systems $\mathcal{A}$ and $\mathcal{B}$ are inverse systems, such that each state α of $\mathcal{A}$ has a unique inverse state β_α of $\mathcal{B}$ and, conversely, each state β of $\mathcal{B}$ has a unique inverse state α_β of $\mathcal{A}$. The first part of this statement follows from the first paragraph in the proof of the theorem. The second part follows from a theorem of Zadeh's and Desoer's (Ref. [1], p. 119) which states, "If $\mathcal{B}$ is invertible and $\bar{B}(\beta\ ;\ v_1) = \bar{B}(\beta\ ;\ v_2)$ then $v_1 \equiv v_2$." Thus, if α_β and α'_β are inverse states of β,

$$u = \bar{B}(\beta\ ;\ \bar{A}(\alpha_\beta\ ;\ u)) = \bar{B}(\beta\ ;\ \bar{A}(\alpha'_\beta\ ;\ u)) \ \forall u \in \mathcal{R}[u], \tag{23}$$

and consequently,

$$\bar{A}(\alpha_\beta\ ;\ u) = \bar{A}(\alpha'_\beta\ ;\ u) \quad \forall u \in \mathcal{R}_a[u] . \tag{24}$$

Therefore, $\alpha_\beta \simeq \alpha'_\beta$, as we have claimed.

An analogous statement for right inverse systems is possible. Since the proof is similar to the proof of Theorem 1, we simply state the result as a theorem.

Theorem 2. If $\mathcal{B}$ is a completely state controllable, functionally reproducible, right inverse of the system $\mathcal{A}$ such that $\mathcal{R}_a[y] \equiv \mathcal{R}_b[v]$, then $\mathcal{A}$ and $\mathcal{B}$ are inverse systems.

The concepts of left and right inverses are connected by the following theorem:

Theorem 3. Let $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ be systems characterized by the input-output-state relations

$$\mathcal{A}: \quad y = \bar{A}(\alpha ; u) \qquad \alpha \in \sum_a, \quad u \in \mathcal{R}_a[u], \quad y \in \mathcal{R}_a[y], \tag{25}$$

$$\mathcal{B}: \quad w = \bar{B}(\beta ; v) \qquad \beta \in \sum_b, \quad v \in \mathcal{R}_b[v], \quad w \in \mathcal{R}_b[w], \tag{26}$$

$$\mathcal{C}: \quad z = \bar{C}(\gamma ; x) \qquad \gamma \in \sum_c, \quad x \in \mathcal{R}_c[x], \quad z \in \mathcal{R}_c[z]. \tag{27}$$

Then, if $\mathcal{B}$ is a left inverse of $\mathcal{A}$ and $\mathcal{C}$ is a right inverse of $\mathcal{A}$, to every right inverse state of $\mathcal{C}$ there is an equivalent left inverse state of $\mathcal{B}$ and vice versa, provided that $\mathcal{R}_b[v] \equiv \mathcal{R}_c[x]$.

Proof. Construct the tandem system $\mathcal{B}\mathcal{A}\mathcal{C}$ as shown in Fig. 5. Let $\mathcal{A}$ be in state α and set $\mathcal{B}$ and $\mathcal{C}$ in the corresponding left and right inverse states β_α and γ_α, respectively. Since $\mathcal{C}$ is a right inverse state of $\mathcal{A}$,

$$\bar{A}(\alpha ; \bar{C}(\gamma_\alpha ; x)) = x \qquad \forall x \in \mathcal{R}_c[x] . \tag{28}$$

The input to the system $\mathcal{B}$ in the tandem connection $\mathcal{B}\mathcal{A}\mathcal{C}$ is also x. Furthermore, since β_α is a left inverse state of α,

$$\bar{B}(\beta_\alpha ; \bar{A}(\alpha ; \bar{C}(\gamma_\alpha ; x))) = \bar{C}(\gamma_\alpha ; x) \quad \forall x \in \mathcal{R}_c[x] . \tag{29}$$

Substituting Eq. (28) into Eq. (29), we obtain

$$\bar{B}(\beta_\alpha ; x) = \bar{C}(\gamma_\alpha ; x) \quad \forall x \in \mathcal{R}_c[x] , \tag{30}$$

and since $\mathcal{R}_c[x] \equiv \mathcal{R}_b[v]$, β_α and γ_α are equivalent states. This is the desired result; hence, the theorem.

The above theorem implies that if either $\mathcal{B}$ or $\mathcal{C}$ is completely state controllable, then $\mathcal{B}$ is a right inverse system or $\mathcal{C}$ is a left inverse system. Another result of this nature is given in Theorem 4.

Theorem 4. If $\mathcal{B}$ is a left inverse of a functionally reproducible system $\mathcal{A}$ such that $\mathcal{R}_a[y] \equiv \mathcal{R}_b[v]$, then $\mathcal{B}$ is a right inverse of $\mathcal{A}$.

Proof. Consider the system $\mathcal{B}\mathcal{A}$ with $\mathcal{A}$ in state α and $\mathcal{B}$ in the left inverse state β_α of α. For any $u \in \mathcal{R}_a[u]$

$$\bar{B}(\beta_\alpha ; \bar{A}(\alpha ; u)) = u, \tag{31}$$

Since $y = \bar{A}(\alpha ; u)$ and $\mathcal{R}_a[y] \equiv \mathcal{R}_b[v]$, functional reproducibility of $\mathcal{A}$ implies that

$$\{ v \mid v = \bar{A}(\alpha ; u) \} \equiv \mathcal{R}_b[v]. \tag{32}$$

If we substitute the expression for u given by

$$u = \bar{B}(\beta_\alpha ; y) \tag{33}$$

into the equation $y = \bar{A}(\alpha ; u)$, we obtain

$$y = \bar{A}(\alpha ; \bar{B}(\beta_\alpha ; y)) \quad \forall y \in \mathcal{R}_a[y]. \tag{34}$$

Therefore,

$$v = \bar{A}(\alpha ; \bar{B}(\beta_{\alpha} ; v)) \qquad \forall v \in \mathcal{R}_b[v], \tag{35}$$

and $\mathcal{B}$ is a right inverse of $\mathcal{A}$.

In the remark following the proof of Theorem 1, we stated that if $\mathcal{A}$ is invertible, and the response of $\mathcal{A}$ to an input u_1 in any state α is identical with the response of $\mathcal{A}$ to an input u_2 starting in α, then u_1 must be identical with u_2. This statement is also true if $\mathcal{A}$ has only a left inverse $\mathcal{B}$, since if

$$\bar{A}(\alpha ; u_1) = \bar{A}(\alpha ; u_2), \tag{36}$$

there exists a state β_{α} of $\mathcal{B}$ such that

$$\bar{B}(\beta_{\alpha} ; \bar{A}(\alpha ; u_1)) = u_1 \text{ and } \bar{B}(\beta_{\alpha} ; \bar{A}(\alpha ; u_2)) = u_2, \tag{37}$$

and hence $u_1 = u_2$.

This means that for each α in the state space of $\mathcal{A}$, $\bar{A}(\alpha ; u)$ defines a one-to-one correspondence between the input segment space and the output segment space of $\mathcal{A}$. Suppose that now we are given a system $\mathcal{A}$ with this property. Then, we claim that $\mathcal{A}$ has a left inverse.

<u>Theorem 5.</u> The system $\mathcal{A}$ with input-output-state relation $y = \bar{A}(\alpha ; u)$ and state function $S_a(\alpha ; u)$ has a left inverse if and only if $\bar{A}(\alpha ; u)$ defines a one-to-one correspondence between the input segment space and the output segment space of $\mathcal{A}$.

<u>Proof.</u> We have already shown that any left invertible system has this property. We will show by construction that any system with this property has a left inverse. Let α be an arbitrary state of $\mathcal{A}$ and let $(u, y(\alpha, u))$ be an input-output pair for the system in this state so that

$$y(\alpha,\ u) = \bar{A}(\alpha\ ;\ u), \tag{38}$$

and

$$\alpha^*(\alpha,\ u) = S_a(\alpha\ ;\ u). \tag{39}$$

Now we define a system $\mathcal{B}$ with state space $\sum_b$, input space $\mathcal{R}_b[v]$, output space $\mathcal{R}_b[w]$, input-output-state relation $w = \bar{B}(\beta\ ;\ v)$, and state function $S_b(\beta\ ;\ v)$ with the properties:

(1) there is a one-to-one correspondence between states in $\sum_a$ and states in $\sum_b$. If a state in $\sum_a$ is denoted by α, the corresponding state in $\sum_b$ is denoted by β_α.

(2) $\mathcal{R}_b[v] \equiv \mathcal{R}_a[y]$ and $\mathcal{R}_b[w] \equiv \mathcal{R}_a[u]$,

(3) If $y(\alpha,\ u) = \bar{A}(\alpha\ ;\ u)$ then $\bar{B}(\beta_\alpha\ ;\ y(\alpha,\ u)) = u$, (40)

(4) If $S_a(\alpha\ ;\ u) = \alpha^*$ then $S_b(\beta_\alpha\ ;\ y(\alpha,\ u)) = \beta_{\alpha^*}$. (41)

Clearly, the system $\mathcal{B}$ is not yet completely defined since if $\mathcal{B}$ is in state β_α and we apply any input $y(\bar{\alpha},\ u) = \bar{A}(\bar{\alpha}\ ;\ u)$ where $\bar{\alpha} \neq \alpha$, the behavior of the system is not specified. However, given the excitation $y(\alpha,\ u)$ to $\mathcal{B}$ in state β_α, the response and final state are uniquely defined because of the one-to-one correspondence between inputs and outputs of the system $\mathcal{A}$. Hence, the tandem system $\mathcal{B}\mathcal{A}$ with $\mathcal{A}$ in state α and $\mathcal{B}$ in state β_α will behave as a unitor because

$$\bar{B}(\beta_\alpha\ ;\ \bar{A}(\alpha\ ;\ u)) = u \quad \forall u \in \mathcal{R}_a[u]. \tag{42}$$

To complete the definition of $\mathcal{B}$, we can let[4]

$$\bar{B}(\beta_\alpha\ ;\ y(\bar{\alpha}\ ;\ u)) = 0 \quad \forall \beta_\alpha \in \textstyle\sum_b,\ \forall\ \bar{\alpha} \neq \alpha,\ \forall u \in \mathcal{R}_a[u], \tag{43}$$

and

$$S_b(\beta_\alpha\ ;\ y(\bar{\alpha},\ u)) = \beta_\alpha \quad \forall \beta_\alpha \in \textstyle\sum_b,\ \forall \bar{\alpha} \neq \alpha,\ \forall u \in \mathcal{R}_a[u]. \tag{44}$$

The definition of $\mathcal{B}$ is consistent with Assumption 1, and $\mathcal{B}$ exhibits the desired left inverse properties. Hence, the theorem is proved.

As an example of the application of the above theorem, consider the finite state system $\mathcal{A}$ shown in Fig. 6(a). The left inverse system $\mathcal{B}$ is shown in Fig. 6(b) where transitions indicated by solid lines are specified according to Eqs. (40)and (41) while transitions indicated by dashed lines are specified according to Eqs. (43) and (44).

WEAK EQUIVALENCE, EQUIVALENCE, AND SOME IMPLICATIONS OF THE DECOMPOSITION PROPERTY

One of the most important concepts in system theory is the concept of "equivalence." Roughly speaking, two systems $\mathcal{A}$ and $\mathcal{B}$ are equivalent if they exhibit the same behavior when subject to the same excitation. If an experimenter is given a "black box" and asked to determine whether the system in the box is either $\mathcal{A}$ or $\mathcal{B}$, it is necessary to specify the type of experiment he is allowed to perform. To this end, we can define two basic kinds of equivalence.

Definition. $\mathcal{A}$ and $\mathcal{B}$ are weakly equivalent, or equivalent under a single experiment, if every input-output pair of $\mathcal{A}$ is an input-output pair of $\mathcal{B}$, and vice versa. Weak equivalence is denoted by $\mathcal{A} = \mathcal{B}$ and implies that the input and output spaces of $\mathcal{A}$ and $\mathcal{B}$ are identical.

If $\mathcal{A} = \mathcal{B}$ and the experimenter is allowed to apply only a single input u, he cannot distinguish between $\mathcal{A}$ and $\mathcal{B}$ by the observed output y since the input-output pair (u, y) belongs to both $\mathcal{A}$ and $\mathcal{B}$. However, the experimenter may have a number of copies of the unknown system, each of which is in the same (unknown) state. If he has as many copies as he needs, he may, by applying a different input to each copy, be able to distinguish between the weakly equivalent systems $\mathcal{A}$ and $\mathcal{B}$. If not, $\mathcal{A}$ and $\mathcal{B}$ are equivalent in a stronger sense, namely,

$\mathcal{A}$ and $\mathcal{B}$ are equivalent under a multiple experiment, or more briefly, $\mathcal{A}$ and $\mathcal{B}$ are "strongly equivalent." Strong equivalence is denoted by $\mathcal{A} \equiv \mathcal{B}$.

<u>Definition.</u> Let $\mathcal{A}$ and $\mathcal{B}$ be characterized by input-output-state relations of the form

$$\mathcal{A}: \quad y = \bar{A}(\alpha\,;\,u), \ \alpha \in \textstyle\sum_a, \ u \in \mathcal{R}_a[u], \ y \in \mathcal{R}_a[y], \tag{45}$$

$$\mathcal{B}: \quad w = \bar{B}(\beta\,;\,v), \ \beta \in \textstyle\sum_b, \ v \in \mathcal{R}_b[v], \ w \in \mathcal{R}_b[w], \tag{46}$$

where $\mathcal{R}_a[u] \equiv \mathcal{R}_b[v]$, and $\mathcal{R}_a[y] \equiv \mathcal{R}_b[w]$. $\mathcal{A}$ is said to be <u>strongly equivalent</u> to $\mathcal{B}$ or <u>equivalent under a multiple experiment,</u> if and only if to every state α of $\mathcal{A}$, there corresponds at least one equivalent state β of $\mathcal{B}$, and vice versa. In symbols,

$$\mathcal{A} \equiv \mathcal{B} \quad \text{iff} \quad \{\forall \alpha \, \exists \beta \, \forall u \mid \bar{A}(\alpha\,;\,u) = \bar{B}(\beta\,;\,u)\},$$

$$\{\forall \beta \, \exists \alpha \, \forall u \mid \bar{A}(\alpha\,;\,u) = \bar{B}(\beta\,;\,u)\}. \tag{47}$$

It should be clear that strong equivalence always implies weak equivalence. However, the converse is not always true. Zadeh[1] has shown that if the systems $\mathcal{A}$ and $\mathcal{B}$ are linear differential systems of finite order, weak equivalence always implies (strong) equivalence. It is reasonable to ask whether this statement can be made for more general systems. To this end, we define the term "decomposition property."

<u>Definition.</u> A system $\mathcal{A}$ with a unique zero state[5] θ_a has the <u>decomposition property</u> if its response to any input u with $\mathcal{A}$ in any state α, can be expressed as the sum of the zero input response of $\mathcal{A}$ starting in state α, and the zero state response of $\mathcal{A}$ to u. That is,

$$\bar{A}(\alpha \,;\, u) = \bar{A}(\alpha \,;\, 0) + \bar{A}(\theta_a \,;\, u) \quad \forall \alpha \in \Sigma_a, \ \forall u \in \mathcal{R}_a[u] \ . \tag{48}$$

It can be shown that all "linear" systems have the decomposition property.[1] On the other hand, a large number of nonlinear systems also have the decomposition property. For example, the system characterized by the input-output-state relation

$$y(t) = e^{-(t-t_0)}\alpha + \int_{t_0}^{t} e^{-(t-\xi)} u^2(\xi)\, d\xi \ , \tag{49}$$

where α (on the real line) represents the state of the system (at t_0), has this property but is nonlinear. We now prove Theorem 6.

Theorem 6. Let $\mathcal{A}$ and $\mathcal{B}$ be weakly equivalent systems with the decomposition property. Then, if $\mathcal{A}$ and $\mathcal{B}$ have a pair of equivalent states, $\alpha_1 \in \Sigma_a$ and $\beta_1 \in \Sigma_b$, $\mathcal{A}$ and $\mathcal{B}$ are strongly equivalent.

Proof. Since $\alpha_1 \cong \beta_1$,

$$\bar{A}(\alpha_1 \,;\, u) = \bar{B}(\beta_1 \,;\, u) \quad \forall u \in \mathcal{R}_a[u] . \tag{50}$$

By the decomposition property

$$\bar{A}(\alpha_1 \,;\, 0) + \bar{A}(\theta_a ;\, u) = \bar{B}(\beta_1 \,;\, 0) + \bar{B}(\theta_b \,;\, u) \ , \tag{51}$$

where θ_a and θ_b are the zero states of $\mathcal{A}$ and $\mathcal{B}$, respectively. By the equivalence of α_1 and β_1

$$\bar{A}(\alpha_1 \,;\, 0) = \bar{A}(\beta_1 \,;\, 0) \ , \tag{52}$$

so that

$$\bar{A}(\theta_a ,\ u) = \bar{B}(\theta_b ,\ u) \quad \forall u \in \mathcal{R}_a[u] , \tag{53}$$

Consequently, the zero states of $\mathcal{A}$ and $\mathcal{B}$ are equivalent.

We are given that $\mathcal{A}$ and $\mathcal{B}$ are weakly equivalent. Choose any $\alpha \in \sum_a$ and an arbitrary $u^* \in \mathcal{R}_a[u]$. By weak equivalence there exists a state $\beta \in \sum_b$ such that

$$\bar{A}(\alpha ; u^*) = \bar{B}(\beta ; u^*) \tag{54}$$

Applying the decomposition property, we obtain

$$\bar{A}(\alpha ; 0) + \bar{A}(\theta_a ; u^*) = \bar{B}(\beta, 0) + \bar{B}(\theta_b ; u^*). \tag{55}$$

Using Eq. (53),

$$\bar{A}(\alpha ; 0) = \bar{B}(\beta ; 0) . \tag{56}$$

Then, for any $u \in \mathcal{R}_a[u]$, by adding Eqs. (53) and (56) together, we obtain

$$\begin{aligned} \bar{A}(\alpha ; u) = \bar{A}(\alpha ; 0) + \bar{A}(\theta_a ; u) &= \bar{B}(\beta ; 0) + \bar{B}(\theta_b ; u) \\ &= \bar{B}(\beta ; u) \quad \forall u \in \mathcal{R}_a[u], \end{aligned} \tag{57}$$

and therefore, $\alpha \cong \beta$. Consequently, for every state $\alpha \in \sum_a$, there is an equivalent state $\beta \in \sum_b$. The converse is shown in the same manner, and thus, $\mathcal{A} \equiv \mathcal{B}$.

By the above theorem, weakly equivalent systems with the decomposition property are equivalent if they have a pair of equivalent states. Thus, a complete proof of the equivalence of weakly equivalent <u>linear</u> systems require only the demonstration that their impulse responses (and hence zero state responses) are equal. For finite state systems, a stronger statement can be made. We first state a theorem due to Zadeh.[8] The proof of this theorem is given in the Appendix.

Theorem 7. (Zadeh) Weakly equivalent finite state systems have at least one pair of equivalent states.

If we combine the last two theorems, we obtain:

Theorem 8. Weakly equivalent finite state systems with the decomposition property are strongly equivalent.

We can derive an interesting result from our theorem on weak equivalence. First, we introduce a concept closely related to invertibility. We say that $\mathcal{A}$ and $\mathcal{B}$ are converse systems if every input-output pair (u, y) of $\mathcal{A}$ has the property that (y, u) (with y as input and u as output) is an input-output pair of $\mathcal{B}$ and vice versa.

All inverse systems are converse systems; however, all converse systems are not inverse systems. The essential difference between converse and inverse systems is that for converse systems $\mathcal{A}$ and $\mathcal{B}$, for each state α of $\mathcal{A}$ and input u, there exists a converse state $\beta_{\alpha, u}$ of $\mathcal{B}$ such that

$$\bar{A}(\alpha ; \bar{B}(\beta_{\alpha, u} ; u)) = u \tag{58}$$

and for each state β of $\mathcal{B}$ and each input v, there exists a state $\alpha_{\beta, v}$ of $\mathcal{A}$ such that

$$\bar{B}(\beta ; \bar{A}(\alpha_{\beta, v} ; v)) = v , \tag{59}$$

while for inverse systems, the selection of the states $\beta_{\alpha, u}$ and $\alpha_{\beta, v}$ is independent of u and v.

As a corollary to Theorem 7, we can prove (see Appendix) that if $\mathcal{A}$ and $\mathcal{B}$ are finite state converse systems, then $\mathcal{A}$ and $\mathcal{B}$ have a pair of inverse states. Furthermore, it should be obvious that all systems converse to a given system are weakly equivalent. Then, we can conclude:

Theorem 9. Let $\mathcal{A}$ be an invertible finite state system with inverse $\mathcal{B}$. Suppose that $\mathcal{B}$ has the decomposition property. Then any converse system $\mathcal{C}$ with the decomposition property is strongly equivalent to the inverse system $\mathcal{B}$.

Remark. It may be shown ([1], pp. 151-152) that if the system $\mathcal{A}$ is invertible and $\mathcal{A}$ has the decomposition property, then the inverse of $\mathcal{A}$ has the decomposition property if $\mathcal{A}$ is zero state additive; i.e., $\bar{A}(\theta_a, u_1 - u_2) = \bar{A}(\theta_a, u_1) - \bar{A}(\theta_a, u_2)$ for all $u_1, u_2 \in R_a[u_{[t_0, t)}]$ and for all $t_0, t\ (t_0 < t)$. Another important property of systems with decomposition is that the cascade system $\mathcal{B}\mathcal{A}$ has the decomposition property, if $\mathcal{A}$ and $\mathcal{B}$ have the decomposition property and $\mathcal{B}$ is zero state additive.

Finally, we will give a necessary and sufficient condition for two weakly equivalent finite state systems $\mathcal{A}$ and $\mathcal{B}$ to be strongly equivalent. A state α is said to be primitive transient if it is not reachable from any other state. A primitive transient machine is either a single primitive transient state or a strongly connected[6] set of two or more states, none of which is reachable from any state not in that set.

Theorem 10. Weakly equivalent finite state systems $\mathcal{A}$ and $\mathcal{B}$ are strongly equivalent if and only if every primitive transient machine in $\mathcal{A}$ contains a state equivalent to some state in $\mathcal{B}$ and every primitive transient machine in $\mathcal{B}$ contains a state equivalent to some state in $\mathcal{A}$.

Proof. The necessity of the condition is trivial. To prove sufficiency, we first note that if we form[9] the nonoriented maximally connected subgraphs of $\mathcal{A}$ and $\mathcal{B}$, we obtain a number of disjoint components, each of which is either weakly or strongly connected. In addition, each weakly connected component must contain at least one primitive transient machine and no strongly connected component can contain a primitive transient machine. Furthermore, every state in a weakly connected (but not strongly connected) component is reachable from at least one state in a primitive transient machine.

Now, by the last sentence of the last paragraph every state in a weakly connected component of $\mathcal{A}$ is equivalent to some state of $\mathcal{B}$. This follows since equivalent states transit to equivalent states. If we examine the states in weakly connected components of $\mathcal{B}$, we can arrive at a similar conclusion. Let us pick a state α in a strongly connected component of $\mathcal{A}$, and apply an arbitrary input u_1. By weak equivalence there is a state β of B such that $\bar{A}(\alpha ; u_1) = \bar{B}(\beta ; u_1)$. We now repeat the argument of the Appendix used to prove Theorem 7 to conclude that there is some state in this strongly connected component of $\mathcal{A}$ which is equivalent to a state of $\mathcal{B}$. By strong connectivity, every state in this component is equivalent to some state of $\mathcal{B}$. Treating each strongly connected component of $\mathcal{A}$ in the same manner we are lead to the same conclusion, and reversing the argument proves that every state in a strongly connected component of $\mathcal{B}$ is equivalent to some state of $\mathcal{A}$. Hence, the theorem.

FINITE STATE SYSTEMS AND SOME IMPLICATIONS OF THE DECOMPOSITION PROPERTY

A discrete time system $\mathcal{A}$ is a <u>finite memory</u> system[2] if the response at any sampling interval depends on only a finite number of past and present excitations and past responses. The output at time k of a finite memory system $\mathcal{A}$ is of the form

$$y_k = \bar{A}(\alpha(k) ; u_k) = f(u_k , \ldots , u_{k-\mu_1}, y_{k-1}, \ldots , y_{k-\mu_2-1}), \quad (60)$$

where $u_i = u_{[i, i+1)}$, $y_i = y_{[i, i+1)}$, and $\alpha(i)$ is the state of $\mathcal{A}$ at time i $(i = 1, 2, \ldots)$. The maximum of μ_1 and μ_2 is said to be the maximal memory of the system. The major feature of a finite memory system is that every input-output sequence of at least length $\mu = \max(\mu_1, \mu_2)$ uniquely determines the final state of the system. That is, if α_1 and α_2 are any two states of $\mathcal{A}$, then for any input sequence $u_1^* u_2^* \ldots u_\mu^*$ such that $\bar{A}(\alpha_1 ; u_1^* \ldots u_\mu^*) = \bar{A}(\alpha_2 ; u_1^* \ldots u_\mu^*)$,

$$S_a(\alpha_1 ; u_1^* u_2^* \ldots u_\mu^*) \simeq S_a(\alpha_2 ; u_1^* u_2^* \ldots u_\mu^*). \tag{61}$$

Hence, given the known system $\mathcal{A}$ in an unknown initial state, we need apply at most μ arbitrary inputs to the system, and by observing the outputs, we can uniquely (to within equivalent states) determine the final state of the system.

Arbitrary discrete time systems, even finite state systems are not finite memory. In fact, if we form the tandem connection of two finite memory systems, the resulting system may not be finite memory. An example[7] of such a system is shown in Fig. 7. In this example, let the last k inputs to $\mathcal{B}\mathcal{A}$ be $a_0, a_0, \ldots, a_0$ where k is an arbitrary number. Suppose that the last k outputs of the system were $c_0, c_0, \ldots, c_0$. This is consistent with the specifications of $\mathcal{A}$ and $\mathcal{B}$. At this point, apply the input a_1 to $\mathcal{B}\mathcal{A}$. Then the output can either be c_2 or c_5. Since k is an arbitrary integer, there cannot exist an input-output relation for $\mathcal{B}\mathcal{A}$ of the form of Eq. (60). Hence, the system does not have finite memory.

Theorem 11. If a discrete time system $\mathcal{A}$ has the decomposition property, every pair of distinct states (α_1, α_2) is distinguishable by a finite length sequence of inputs. Furthermore, there exists a number K_{α_1, α_2} such that any input sequence of length K or greater is a distinguishing sequence.

Proof. Suppose the theorem is false. Then there exist two nonequivalent states, say α_1 and α_2, which have an identical input-output sequence of infinite length. Let $u_1^* \; u_2^* \; . \; . \; . \; u_i^* \; . \; . \; .$ be such a sequence. Then

$$\bar{A}(\alpha_1 \; ; u_1^* \; u_2^* \; . \; . \; . \; u_i^*) = \bar{A}(\alpha_2 \; ; u_1^* \; u_2^* \; . \; . \; . \; u_i^*) \quad i = 1, \; 2, \; 3, \; . \; . \; . \; . \tag{62}$$

Since the system has the decomposition property

$$\bar{A}(\alpha_j \; ; u_1^* \; u_2^* \; . \; . \; . \; u_i^*) = \bar{A}(\alpha_j \; ; 0_1 \, 0_2 \; . \; . \; . \; 0_i) + \bar{A}(\theta_a \; ; u_1^* u_2^* \; . \; . \; . \; u_i^*)$$

$$j = 1, \; 2 \; ; \; i = 1, \; 2, \; 3, \; . \; . \; . \; . \; . \; . \; . \tag{63}$$

and

$$\bar{A}(\alpha_1 \; ; 0_1 \, 0_2 \; . \; . \; . \; 0_i) = \bar{A}(\alpha_2 \; ; 0_1 \, 0_2 \; . \; . \; . \; 0_i) \quad i = 1, \; 2, \; 3, \; . \; . \; . \; . \tag{64}$$

However, Eq. (64) implies that since the system has the decomposition property, for any input sequence $u_1 \, u_2 \; . \; . \; . \; u_i$,

$$\bar{A}(\alpha_1 \; ; u_1 \, u_2 \; . \; . \; . \; u_i) = \bar{A}(\alpha_2 \; ; u_1 \, u_2 \; . \; . \; . \; u_i) \; . \tag{65}$$

Therefore, $\alpha_1 \simeq \alpha_2$, which is a contradiction. Consequently, the theorem is true.

Now that we have shown that every input sequence of length K_{α_1, α_2} or greater is a distinguishing sequence for the <u>initial</u> state, it is obvious that a <u>finite</u> state system with the decomposition property is finite memory since on the basis of the initial state, the last K inputs and the last K outputs, we can easily find the final state and final output. (where $K = \max_{i, j} K_{\alpha_i, \alpha_j}$). Thus, we have:

<u>Theorem 12.</u> If a <u>finite state</u> system $\mathcal{A}$ has the decomposition property, it has finite memory.

Once a finite state system is shown to possess finite memory, it becomes important to find the maximal memory of the system, or a priori, to find bounds on the memory. For example, if $\mathcal{A}$ has η states, none of which is equivalent, $\mu \leq \eta(\eta - 1) / 2$ ([2] , p. 162). Moreover, Gill[10] has shown that this bound is the best possible upper bound in general. If $\mathcal{A}$ has the decomposition property, this upper bound can be considerably strengthened.

Theorem 13. If a minimal state system of η states has the decomposition property, then its maximal memory is less than η.

Proof. Since there are η states, the zero input response of the system must repeat after at most η inputs. That is, if the arbitrary state $\alpha \in \sum_a$ is taken into the state α' by the input $0_1 \, 0_2 \, . \, . \, . \, 0_\eta$, then

$$\bar{A}(\alpha \; ; \; 0_1) = \bar{A}(\alpha' \; ; \; 0_{\eta+1}). \tag{66}$$

Suppose two states α_1 and α_2 have identical input-output sequences of length η .

$$\bar{A}(\alpha_1 \; ; \; u_1^* \; u_2^* \; . \, . \, . \; u_\eta^*) = \bar{A}(\alpha_2 \; ; \; u_1^* \; u_2^* \; . \, . \, . \; u_\eta^*). \tag{67}$$

By the decomposition property

$$\bar{A}(\alpha_1 \; ; \; 0_1 \, 0_2 \, . \, . \, . \, 0_\eta) = \bar{A}(\alpha_2 \; ; \; 0_1 \, 0_2 \, . \, . \, . \, 0_\eta) , \tag{68}$$

and Eq. (66) implies

$$\bar{A}(\alpha_1 \; ; \; 0_1 \, 0_2 \, . \, . \, . \, 0_i) = \bar{A}(\alpha_2 \; ; \; 0_1 \, 0_2 \, . \, . \, . \, 0_i) \quad i = 1, 2, 3, \, . \, . \, . \, . \tag{69}$$

In other words, there is an infinite input sequence which will not distinguish between α_1 and α_2 . By Theorem 11, this is impossible; thus, we have reached the desired conclusion.

Finally, we give a result related to the last two theorems.

Theorem 14. Let $\mathcal{A}$ and $\mathcal{B}$ be weakly equivalent discrete time systems. If $\mathcal{A}$ has finite maximal memory μ, then $\mathcal{B}$ has finite maximal memory μ.

Proof. Let $\mathcal{A}$ be in state α_1 and apply an input sequence $u_1^* u_2^* \ldots u_\mu^*$ to $\mathcal{A}$. Suppose that α_1 transits to the state α^*. Since $\mathcal{A}$ has finite memory μ, for all $\alpha \in \sum_a$ such that $\bar{A}(\alpha ; u_1^* u_2^*, \ldots, u_\mu^*) = \bar{A}(\alpha_1 ; u_1^* u_2^*, \ldots u_\mu^*)$, then

$$S_a(\alpha ; u_1^* \ldots u_\mu^*) \cong S_a(\alpha_1 ; u_1^* \ldots u_\mu^*) \triangleq \alpha^* . \tag{70}$$

Since $\mathcal{A}$ and $\mathcal{B}$ are weakly equivalent, there exists a state β_1 of $\mathcal{B}$ such that

$$\bar{B}(\beta_1 ; u_1^* \ldots u_\mu^*) = \bar{A}(\alpha_1 ; u_1^* \ldots u_\mu^*). \tag{71}$$

To show that $\mathcal{B}$ has finite memory μ or less, we must show that for any state β of $\mathcal{B}$ such that $\bar{B}(\beta ; u_1^* \ldots u_\mu^*) = \bar{B}(\beta_1 ; u_1^* \ldots u_\mu^*)$,

$$S_b(\beta ; u_1^* \ldots u_\mu^*) \cong S_b(\beta_1 ; u_1^* \ldots u_\mu^*) . \tag{72}$$

Let $\underline{\Upsilon}$ be an arbitrary input sequence of arbitrary length, and let us apply the input sequence $u_1^* \ldots u_\mu^* \underline{\Upsilon}$ to $\mathcal{B}$ in any state β with $\bar{B}(\beta ; u_1^* \ldots u_\mu^*) = \bar{B}(\beta_1 ; u_1^* \ldots u_\mu^*)$. By weak equivalence, there exists a state α of $\mathcal{A}$ such that

$$\bar{A}(\alpha ; u_1^* \ldots u_\mu^* \underline{\Upsilon}) = \bar{B}(\beta ; u_1^* \ldots u_\mu^* \underline{\Upsilon}). \tag{73}$$

But, since $\mathcal{A}$ has finite memory

$$S_a(\alpha ; u_1^* \ldots u_\mu^*) \cong \alpha^* . \tag{74}$$

So that if we define β^* by

$$\beta^* = S_b(\beta\, ;\, u_1^* \, . \, . \, . \, u_\mu^*), \tag{75}$$

we have

$$\bar{B}(\beta^*\, ;\, \underline{\Upsilon}) = \bar{A}(\alpha^*\, ;\, \Upsilon) \tag{76}$$

for any input sequence $\underline{\Upsilon}$. Therefore $\beta^* \cong \alpha^*$, and since β was any arbitrary state such that $\bar{B}(\beta\, ;\, u_1^* \, . \, . \, . \, u_\mu^*) = \bar{B}(\beta_1\, ;\, u_1^* \, . \, . \, . \, u_\mu^*)$,

$$S_b(\beta\, ;\, u_1^* \, . \, . \, . \, u_\mu^*) \cong \alpha^*, \ \forall \beta \ni \bar{B}(\beta\, ;\, u_1^* \, . \, . \, u_\mu^*) = \bar{B}(\beta_1\, ;\, u_1^* \, . \, . \, . \, u_\mu^*). \tag{77}$$

Therefore, $\mathcal{B}$ has finite maximal memory μ or less. To show that $\mathcal{B}$ has maximal memory exactly μ, reverse the roles of $\mathcal{A}$ and $\mathcal{B}$. If $\mathcal{B}$ has maximal memory ν, using the same argument as above, we can conclude that $\mu \leq \nu$. Therefore $\mu = \nu$ and the theorem is proved.

In our proof of Theorem 14, we showed that the weakly equivalent finite memory systems $\mathcal{A}$ and $\mathcal{B}$ have a pair of equivalent states. Thus, we have actually derived a more important result which we state below.

Theorem 15. Weakly equivalent finite memory systems with the decomposition property are strongly equivalent.

CONCLUSIONS AND FURTHER REMARKS

We have seen that the decomposition property has many strong implications for abstract systems. Many properties, usually proved only for the restricted class of "linear" systems, also hold for the more general class of systems with the decomposition property. There are several areas for further research. For example, the authors believe that all weakly equivalent systems with the decomposition property are strongly equivalent. This statement is supported by the conclusions reached in the present paper and also by the fact that continuous systems with the decomposition property have finite memory in the following sense: Any two states with an identical input-output pair $(u_{[t_0, \infty)}, y_{[t_0, \infty)})$ must be equivalent.

The concepts of inverse and left inverse systems seem to have potential, and besides abstract interest, appear to have a more concrete impact in terms of feedback control systems.[11] The idea of an inverse system could be generalized in several ways. For instance, we might be willing to allow a finite period of time to elapse before recovering the original inputs. The results reported here appear to be generalizable to this case, and the authors are presently working on this and other extensions.

APPENDIX

<u>Proof of Theorem 7:</u> Weakly equivalent finite state systems have a pair of equivalent states.

Let $\mathcal{A}$ and $\mathcal{B}$ be weakly equivalent finite state systems. Choose an arbitrary state α_0 in $\sum_a$. Apply an arbitrary input u_0 to $\mathcal{A}$ in state α_0. By the weak equivalence of $\mathcal{A}$ and $\mathcal{B}$, there exists a non-empty set of states in $\sum_b$, which yields the same output $\bar{A}(\alpha_0 ; u_0)$. Let $K_0(\alpha_0, u_0) \triangleq \{\beta_0 \mid \bar{B}(\beta_0 ; u_0) = \bar{A}(\alpha_0 ; u_0)\}$ be this set of states.

Suppose $\alpha_1 \triangleq S_a(\alpha_0 ; u_0)$. Let $\mathcal{T}_0(\beta_0, u_0)$ be the set of states reachable from $K_0(\alpha_0, u_0)$ via the input u_0. In other words, $\mathcal{T}_0(\beta_0, u_0) \triangleq \{\beta \mid \beta = S_b(\beta_0; u_0), \beta_0 \in K_0(\alpha_0, u_0)\}$. Either all the states in $\mathcal{T}_0(\beta_0, u_0)$ are equivalent to α_1, or there is at least one state β_1

in $\mathcal{I}_0(\beta_0, u_0)$ not equivalent to α_1. If the latter possibility is true, there exists an input u_1 such that

$$\overline{B}(\beta_1 ; u_1) \neq \overline{A}(\alpha_1 ; u_1).$$

The input u_1 cannot distinguish α_1 from every state in $\mathcal{I}_0(\beta_0, u_0)$ since by weak equivalence

$$\overline{A}(\alpha_0 ; u_0\, u_1) = \overline{B}(\gamma ; u_0\, u_1) \text{ for } \gamma \in \sum_b .$$

This implies

$$\overline{A}(\alpha_0 ; u_0) = \overline{B}(\gamma; u_0), \quad \gamma \in K_0 [\alpha_0, u_0],$$

and

$$\overline{A}(\alpha_1 ; u_1) = \overline{B}(\hat{\gamma}, u_1), \quad \hat{\gamma} = S_b(\gamma ; u_0) \in \mathcal{I}_0(\beta_0, u_0).$$

Therefore, the set $K_1(\alpha_0, u_0 u_1) \triangleq \{\beta_{01} \in \mathcal{I}_0(\beta_0, u_0) \mid \overline{B}(\beta_{01}; u_1) = \overline{A}(\alpha_1 ; u_1)\}$ is not empty.

Let $N(P)$ be the number of distinct states in an arbitrary set P of states. $N(K_0) > N(K_1)$, since the input u_0 can take each state of K_0 into at most one state; therefore, $\mathcal{I}_0$ has at most $N(K_0)$ states, and since K_1 has one less state than $\mathcal{I}_0$, the inequality follows.

Let $\alpha_2 \triangleq S_a(\alpha_1 ; u_1)$ and let $\mathcal{I}_1(\beta_0, u_0 u_1)$ be the set of states reachable from $K_1(\alpha_0, u_0 u_1)$ via the input u_1. Either all states in $\mathcal{I}_1$ are equivalent to α_2, (proving the theorem), or there exists at least one state β_2 in $\mathcal{I}_1$ such that there is an input u_2 and

$$\overline{B}(\beta_2 ; u_2) \neq \overline{A}(\alpha_2 ; u_2)$$

As before, the set $K_2(\alpha_0, u_0 u_1 u_2) = \{\beta_{02} \in \mathcal{I}_1(\beta_0, u_0 u_1) \mid \overline{B}(\beta_{02} ; u_2) = \overline{A}(\alpha_2 ; u_2)\}$ is not empty since by the weak equivalence of $\mathcal{A}$ and $\mathcal{B}$, there is a state γ in K_0 such that

$$\bar{A}(\alpha_0 ; u_0u_1u_2) = \bar{B}(\gamma ; u_0u_1u_2)$$

Then, using the same reasoning as above, we can conclude that $N(K_1) \geq N(\mathcal{T}_1) > N(K_2)$.

Either we continue the preceding process, or for some j, all the states in $\mathcal{T}_j$ are equivalent to the state α_{j+1} where $\alpha_{j+1} = S_a(\alpha_0 ; u_0u_1 \ldots u_j)$, $\mathcal{T}_j(\beta_0, u_0u_1 \ldots u_j) \triangleq \{\beta_{0j} \mid \beta_{0j}$

$= S_b(\beta_{0j-1} ; u_j)$, $\beta_{0j} \in K_j(\alpha_0, u_0u_1 \ldots u_j)\}$ and

$K_j(\alpha_0, u_0u_1 \ldots u_j) \triangleq \{\beta_{0j} \mid \beta_{0j} \in \mathcal{T}_{j-1}$ and $\bar{B}(\beta_{0j}, u_j) = \bar{A}(\alpha_j, u_j)\}$.

The former possibility cannot continue indefinitely since $N(K_0)$, $N(K_1)$, $\ldots$, $N(K_j)$ is a strictly decreasing sequence of numbers with $N(K_0)$ finite. If the process does continue indefinitely, there exists an integer m such that $N(K_m) = 0$. This violates the assumption that $\mathcal{A} = \mathcal{B}$, which requires that each K_j be nonempty. Therefore, the process must terminate at some $\mathcal{T}_j$ for which all states in $\mathcal{T}_j$ are equivalent to α_{j+1}. This proves the theorem.

<u>Corollary.</u> Converse finite state systems have a pair of inverse states.

<u>Proof.</u> We proceed in the same manner as in the proof of the theorem except that we now define $K_0(\alpha_0, u_0) \triangleq \{\beta_0 \mid \bar{B}(\beta_0 ; \bar{A}(\alpha_0 ; u_0)) = u_0\}$, $\mathcal{T}_0(\beta_0, \bar{A}(\alpha_0 ; u_0))$ is the set of states reached from $K_0(\alpha_0, u_0)$ via $\bar{A}(\alpha_0 ; u_0)$, and in general

$$K_j(\alpha_0, u_0u_1 \ldots u_j) = \{\beta_{0j} \in \mathcal{T}_{j-1}(\beta_0, \bar{A}(\alpha_0 ; u_0u_1 \ldots u_j) \mid$$

$$\bar{B}(\beta_{0j} ; \bar{A}(\alpha_j ; u_j)) = u_j\}.$$

The converse property implies that each K_j is not empty and by construction $N(K_{j-1}) > N(K_j)$. Thus, the process must terminate for a finite state converse system and consequently finite state converse systems contain a pair of inverse states.

REFERENCES

1. L. A. Zadeh and C. A. Desoer, Linear System Theory - The State Space Approach, McGraw-Hill Book Co., Inc., New York; 1963.
2. A. Gill, Introduction to the Theory of Finite State Machines, McGraw-Hill Book Co., Inc., New York; 1962.
3. S. Ginsberg, An Introduction to Mathematical Machine Theory, Addison-Wesley Pub. Co., Inc., Reading, Mass. ; 1962.
4. D. A. Huffman, "Canonical Forms for Information-Lossless Finite State Machines," IRE Trans. on Circuit Theory, Vol. CT-6, special supplement, pp. 41-59; 1959.
5. S. Even, "On Information Lossless Automata of Finite Order," IEEE Trans. on Electronic Computers, Vol. EC-14 No. 4, pp. 561-570; August, 1965.
6. E. Kreindler and P. E. Sarachik, "On the Concepts of Controllability and Observability of Linear Systems," IEEE Trans. on Automatic Control, Vol. AC-9, pp. 129-136; April, 1964.
7. R. W. Brockett and M. D. Mesarovic, "The Reproducibility of Multivariable Systems," Math. Analysis and Applications, Vol. 11, pp. 548-564; 1965.
8. L. A. Zadeh, unpublished notes on system theory, University of California, Berkeley.
9. S. Seshu and M. B. Reed, Linear Graphs and Electrical Networks, Addison-Wesley Pub. Co., Inc., Reading Mass.; 1961.
10. A. Gill, "On the Bound to the Memory of a Sequential Machine," IEEE Trans. on Electronic Computers, Vol. EC-14, No. 3, pp. 464-466; June, 1965.
11. R. W. Brockett, "Poles, Zeros, and Feedback: State Space Interpretation," IEEE Trans. on Automatic Control, Vol. AC-10, No. 2, pp. 129-135; April, 1965.

FOOTNOTES

1. Here, we assume that we can perform an experiment with the system after the unknown excitation has been applied.

2. A necessary and sufficient condition for a finite state system to be information lossless is that it contain no state which can be reached from any initial state via two or more distinct input sequences which yield identical output sequences.[2]

3. Naturally, we mean unique to within equivalent states. Two states α_1 and α_2 are equivalent if for all $u \in \overline{\mathcal{R}_a[u]}$, $\bar{A}(\alpha_1 ; u) = \bar{A}(\alpha_2 ; u)$. If α_1 and α_2 are equivalent, we write $\alpha_1 \cong \alpha_2$.

4. If $0 \notin \mathcal{R}_b[w]$, we can suitably expand $\mathcal{R}_b[w]$ to include this element, or define $\bar{B}(\beta_\alpha ; \bar{A}(\alpha ; u))$ in some other convenient way.

5. We assume that $\mathcal{R}_a[u]$ and $\mathcal{R}_a[y]$ are sufficiently well defined to have an addition operation, and an (additive) zero input $0_{[t_0, t]}$ and zero output $0_{[t_0, t)}$ for all t_0, t with $t_0 < t$. A zero state θ_a is a state such that $0_{[t_0, t)} = \bar{A}(\theta_a ; 0_{[t_0, t)}) \; \forall \, t_0, t, (t > t_0)$.

6. Completely state controllable finite state systems are said to be strongly connected.

7. Suggested by Prof. L. Haines, Dept. of Electrical Engineering, University of California, Berkeley.

LIST OF FIGURES

Figure No.	
1	Inverse systems $\mathcal{A}$ and $\mathcal{B}$.
2	A left invertible system $\mathcal{A}$ which is not invertible.
3	A two-state information lossless machine which is not left invertible.
4.	The tandem system $\mathcal{A}\mathcal{B}\mathcal{A}$ used in proof of Theorem 1.
5	The tandem connection $\mathcal{B}\mathcal{A}\mathcal{C}$ where $\mathcal{B}$ and $\mathcal{C}$ are left and right inverses of $\mathcal{A}$.
6	A three-state system $\mathcal{A}$ and its left inverse $\mathcal{B}$.
7	A finite state, nonfinite memory system.

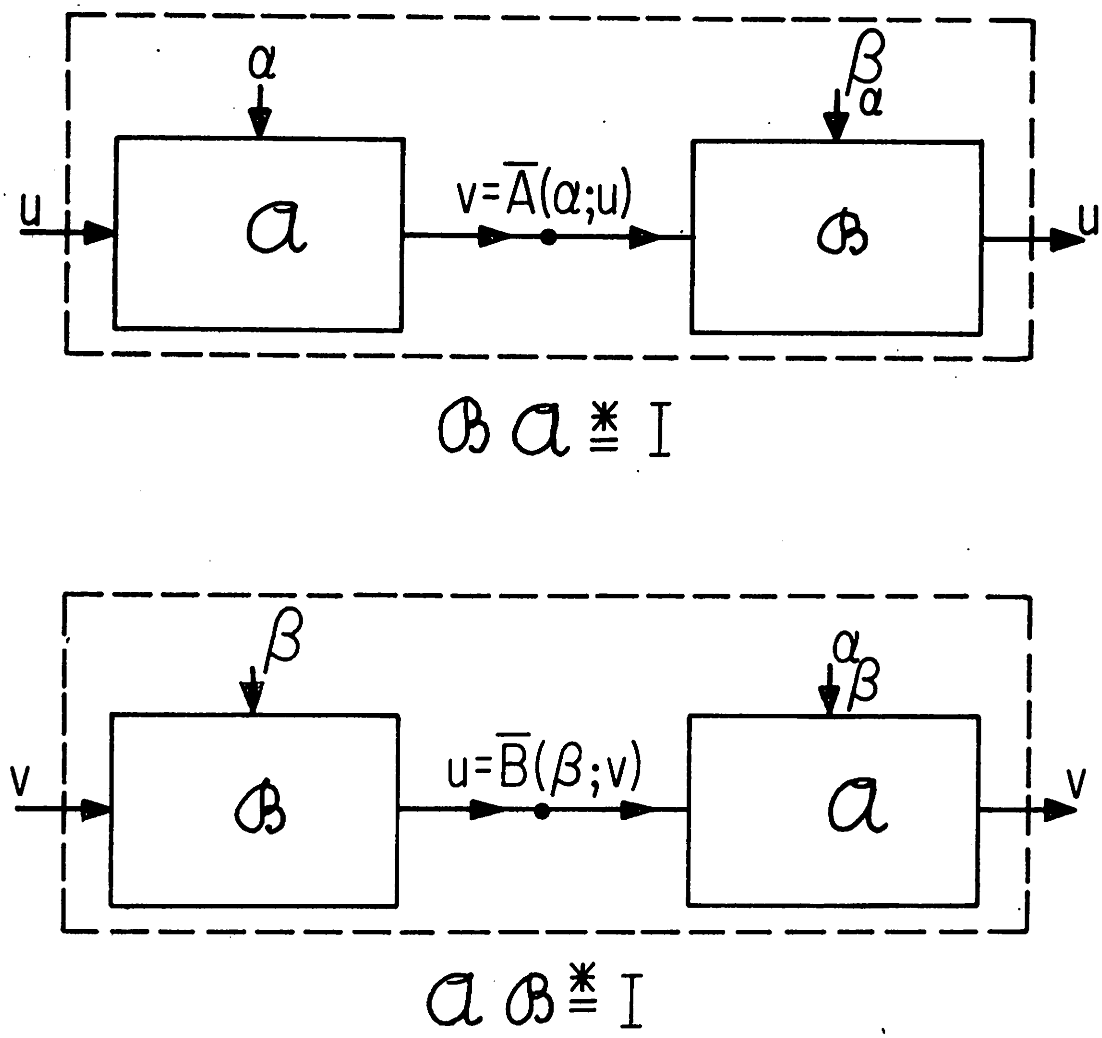


Fig. 1. Inverse systems $\mathcal{A}$ and $\mathcal{B}$.

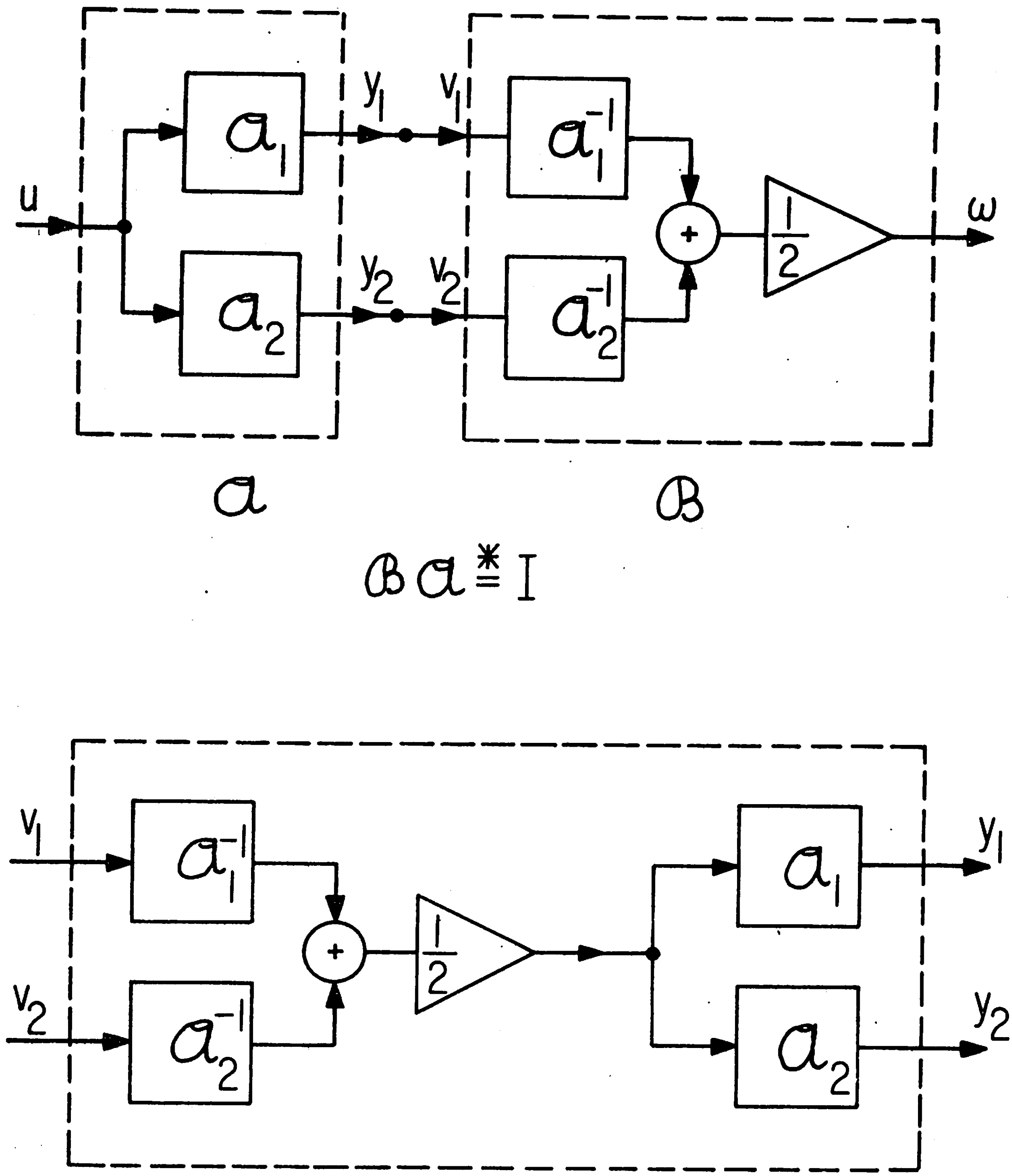


Fig. 2. A left invertible system $\mathcal{A}$ which is not invertible.

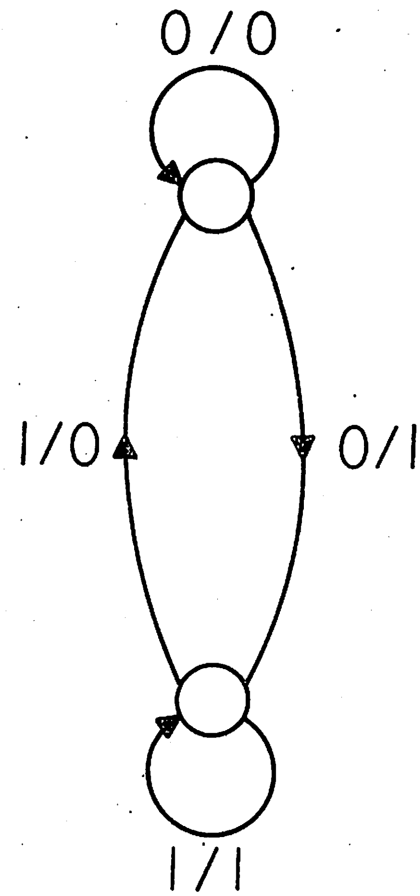


Fig. 3. A two-state information lossless machine which is not left invertible.

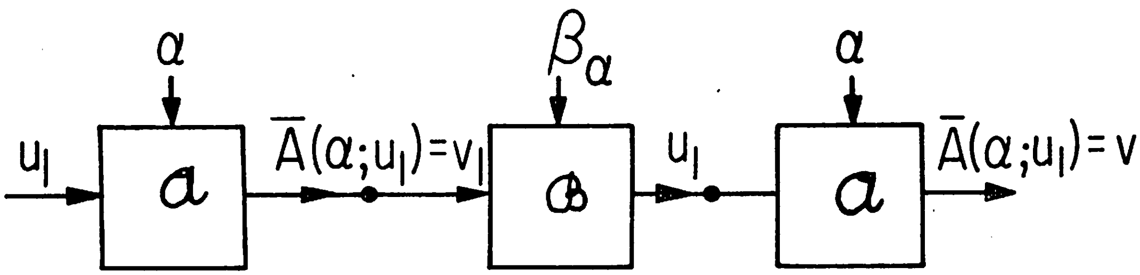


Fig. 4. The tandem system $\mathcal{A}\mathcal{B}\mathcal{A}$ used in proof of Theorem 1.

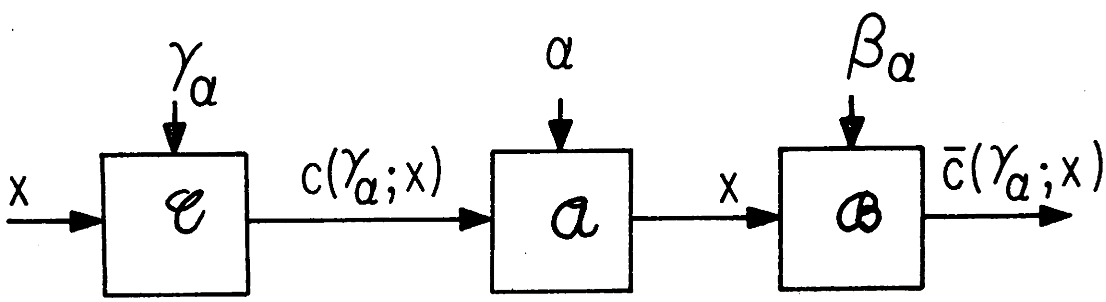


Fig. 5. The tandem connection $\mathcal{B}\mathcal{A}\mathcal{C}$ where $\mathcal{B}$ and $\mathcal{C}$ are left and right inverses of $\mathcal{A}$.

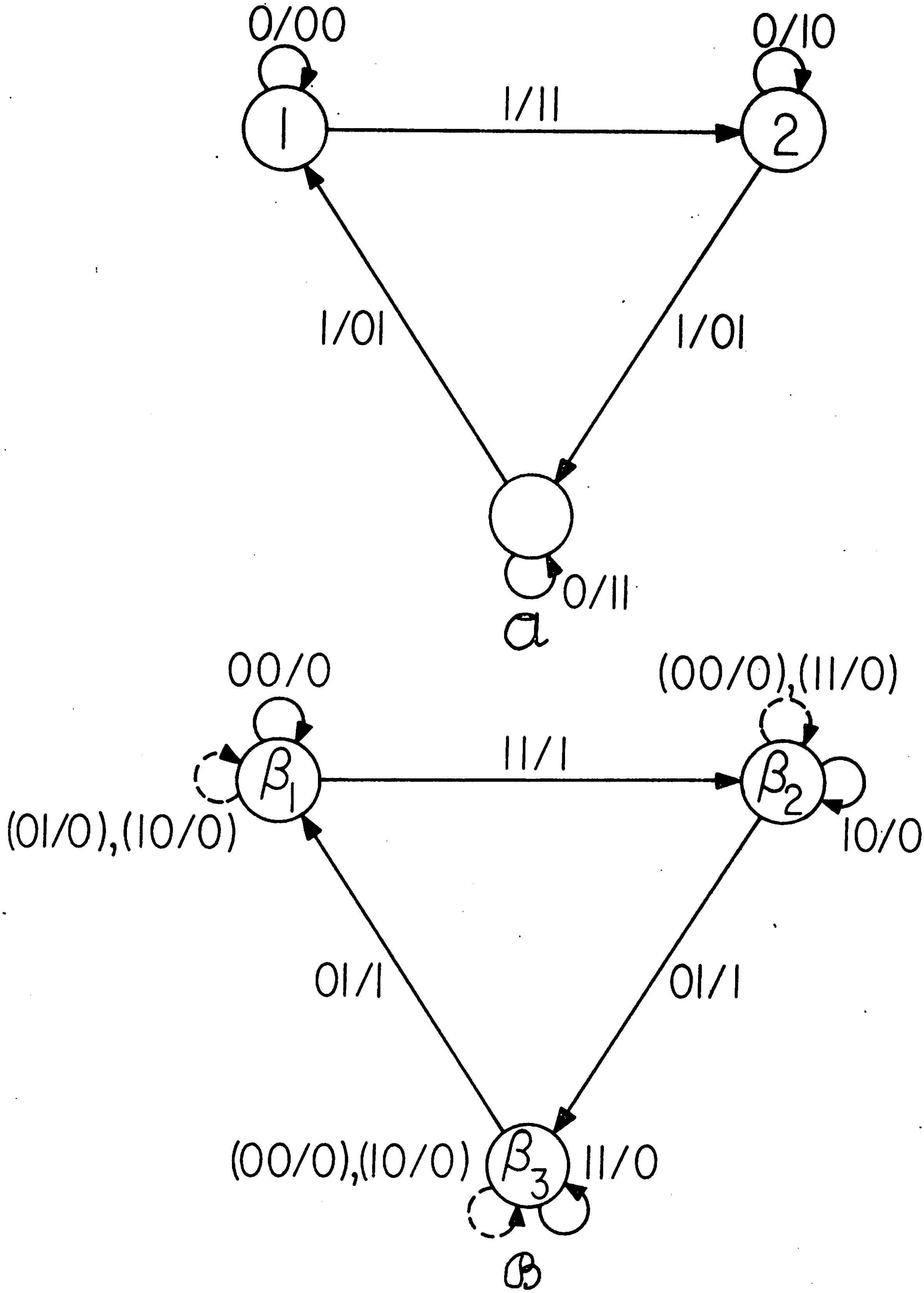
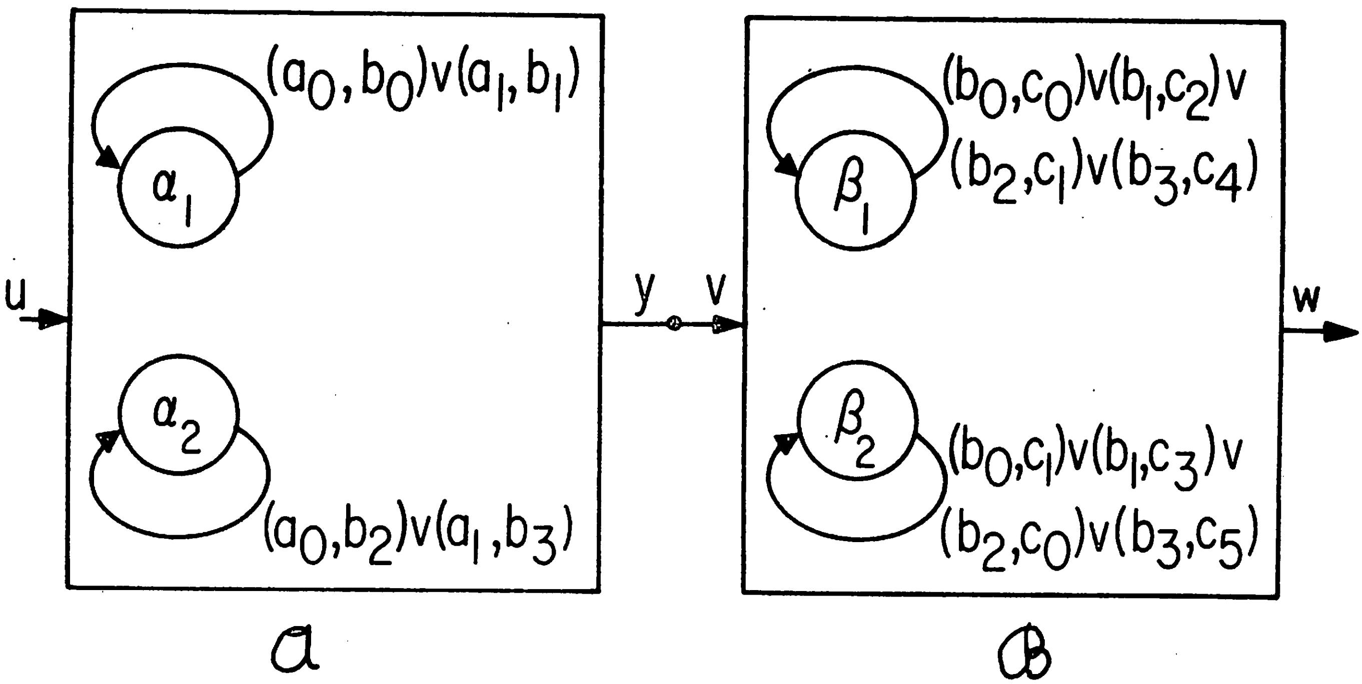


Fig. 6. A three-state system $\mathcal{A}$ and its left inverse $\mathcal{B}$.



$$\mathcal{R}_a[u] = \{a_0, a_1\}$$

$$\mathcal{R}_a[y] = \{b_0, b_1, b_2, b_3\}$$

$$\mathcal{R}_b[v] = \{b_0, b_1, b_2, b_3\}$$

$$\mathcal{R}_b[w] = \{c_0, c_1, c_2, c_3, c_4, c_5\}$$

Fig. 7. A finite state, nonfinite memory system.