SIAM J. CONTROL
Vol. 10, No. 2, May 1972

CONTROLLABILITY, REALIZATION AND STABILITY OF DISCRETE-TIME SYSTEMS*

LEONARD WEISS†

Abstract. The following problems are discussed and solved in this paper: finding computable necessary and sufficient conditions for complete reachability and complete observability of a linear, time-varying, discrete-time system; finding sufficient conditions for local controllability of nonlinear discrete-time systems; relating reachability to the concept of discrete Pfaffian systems; obtaining a minimal-dimension difference equation (with possibly variable coefficients) from a given input/output function of a system; finding necessary and sufficient conditions for Lyapunov stability and finite-time stability of nonlinear difference equations; and, giving an algorithm for determining whether a linear difference equation is stable in the finite-time sense.

1. Introduction. In a certain sense, the theory of discrete-time systems dates back at least to the time of De Moivre and Laplace, who were the first to use the concept of generating functions in connection with the study of discrete random variables [1]. In the modern engineering literature, these functions are called Z-transforms [2], about which we shall say nothing further in this paper. The systematic study of difference equations (about which we shall say a good deal) began much later, with major landmarks in the development of the theory being provided by the treatises of Boole [3] and Milne-Thomson [4]. The interest in these equations in both mathematics and engineering stems from their usefulness in various applications. Numerical analysts deal with them in designing and analyzing algorithms for the numerical solution of differential and other equations [5], while engineers are often confronted with physical systems whose description by difference equations is quite natural [6]. In addition, any system whose internal structure may be unknown but whose input/output behavior can be (at least partially) obtained through experiment is a candidate for a difference equation model, and this accounts, in part, for the popularity of such models with economists, psychologists and statisticians (see [7] for many examples from these areas).

With the advent of digital computers, such models have become of compelling interest, and an accelerating growth of literature on various aspects of discrete-time systems has been the result [8].

In this paper, we study some selected problems in the mathematical theory of discrete-time systems, and we derive various results for both linear and nonlinear deterministic systems in the areas mentioned in the paper's title. (See [40] for a discussion of stochastic discrete-time systems.)

2. Preliminaries. Consider the difference equation

$$x(k+1) = f(k, x(k), u(k)), \qquad k \in \mathfrak{Z}, \tag{1}$$

* Received by the editors September 21, 1971, and in revised form October 28, 1971. Presented at the NSF Regional Conference on Control Theory, held at the University of Maryland Baltimore County, August 23–27, 1971.

† Department of Electrical Engineering and Institute for Fluid Dynamics and Applied Mathematics, University of Maryland, College Park, Maryland 20742, and Mathematics Research Center, Naval Research Laboratory, Washington, D.C. This work was supported in part by the Air Force Office of Scientific Research under Grant AFOSR-69-1646.

where $\mathfrak{Z}$ is the set of all integers, $x(k) \in \mathfrak{R}^n$, $u(k) \in \mathfrak{R}^p$, $p \leqq n$. Then, as long as f is a well-defined function on $\mathfrak{Z} \times \mathfrak{R} \times \mathfrak{R}^p$, there is no problem regarding existence and uniqueness of solutions to (1) starting from any given initial condition.

A linear discrete-time system is a system of the form (1) in which f is a linear function of x and u, i.e., (1) becomes

$$x(k + 1) = A(k)x(k) + B(k)u(k), \qquad k \in \mathfrak{Z}, \tag{2}$$

where $A(k)$ is $n \times n$ and $B(k)$ is $n \times p$, $p \leqq n$.

In connection with (2), we define a real $n \times n$ matrix-valued function Ω on $\mathfrak{Z} \times \mathfrak{Z}$ by the formulas

$$\begin{aligned} &\Omega(k, j) = A(k)A(k - 1) \cdots A(j + 1)A(j), && j, k \in \mathfrak{Z}, \quad k \geqq j, \\ &\Omega(k, k + 1) \triangleq I \text{ (the identity)}, && k \in \mathfrak{Z}, \\ &\Omega(k, j) \quad \text{undefined for } j > k + 1. && \end{aligned} \tag{3}$$

By iteration, the solution of (2) at the kth instant starting from initial time k_0 and state x_0 is

$$x(k; k_0, x_0, u) = \Omega(k - 1, k_0)x_0 + \sum_{j=k_0}^{k-1} \Omega(k - 1, j + 1)B(j)u(j). \tag{4}$$

It should be noted that, unlike ordinary linear differential equations, it is possible for the set of all solutions of (2), with $u(\cdot) \equiv 0$, to be located in a proper subspace of $\mathfrak{R}^n$ (for example, take $A(\cdot) \equiv 0$). That is, discrete-time systems can be *pointwise degenerate* [9]. (If the set of all solutions to (2), with $u = 0$, spans $\mathfrak{R}^n$, the system (2) is said to be *pointwise complete* [10].) This property plays an important role in seeking necessary conditions for controllability, as will be shown later on. The possibility of pointwise degeneracy adds interest to the study of linear discrete-time systems, for it means that theories developed for linear ordinary continuous-time systems (where pointwise completeness always holds [11]) may not have 1–1 correspondence to their analogues in the discrete-time case. It is therefore of interest to develop results for discrete-time systems which are independent of the properties of pointwise completeness or degeneracy (and to point out circumstances under which considerations of these properties cannot be avoided).

Finally, we denote the set of positive integers by $\mathfrak{Z}^+$, and we define a discrete interval $[\alpha, \beta]$, where $\alpha, \beta \in \mathfrak{Z}$, $\alpha < \beta$, as the set of integers $\{\alpha, \alpha + 1, \cdots, \beta - 1, \beta\}$.

3. The concepts of controllability and reachability. One of the most fundamental contributions to the mathematical theory of systems over the past fifteen years has been the formulation and characterization of the property of controllability, which was first done by Kalman [12] for linear, time-invariant, discrete-time systems. Since then, much development of the theory of controllability (and its conceptual partner, the theory of reachability) has occurred for differential equations (see [13], [14]). In the sequel, we present some results for reachability and controllability of time-varying linear and nonlinear discrete-time systems using a variety of techniques.

In the following definition, the *phase space* $\mathfrak{P} = \mathfrak{R}^n \times \mathfrak{Z}$.

DEFINITION 1. (a) For the system (1), the phase $(x, v) \in \mathfrak{P}$ is *reachable* (or, *N-step reachable*) if there exists $N \in \mathfrak{Z}^+$, and a control sequence $\mathcal{U} = \{u(v - N + 1), u(v - N + 2), \cdots, u(v)\}$ such that the phase $(0, v - N)$ is transferred to (x, v) under the action of $\mathcal{U}$ (written $(0, v - N) \xrightarrow{\mathcal{U}} (x, v)$).

(b) If, for all $x \in \mathfrak{R}^n$, (x, v) is reachable (or N-step reachable), then the system (1) is *completely* (*N-step*) *reachable* at time v.

(c) Complete (N-step) reachability with no time designated implies that (1) is *completely* (*N-step*) *reachable* at all times.

DEFINITION 2. The phase (x, v) is *controllable* (or *N-step controllable*) if there exists $N \in \mathfrak{Z}^+$ and $\mathcal{U} = \{u(v), u(v + 1), \cdots, u(v + N - 1)\}$ such that $(x, v) \xrightarrow{\mathcal{U}} (0, v + N)$.

Definitions of complete (N-step) controllability follow in an entirely analogous fashion from Definition 1.

THEOREM 1. *A necessary and sufficient condition for* (2) *to be completely M-step reachable at time v is that*

$$\begin{aligned}&\operatorname{rank}[B(v-1), \Omega(v-1, v-1)B(v-2),\\ &\qquad \cdots, \Omega(v-1, v-M+1)B(v-M)] = n. \end{aligned} \tag{5}$$

Proof. Sufficiency. Let

$$\begin{aligned}\mathcal{R}_k(v-1) = {}&[B(v-1), \Omega(v-1, v-1)B(v-2),\\ &\cdots, \Omega(v-1, v-k+1)B(v-k)]\end{aligned} \tag{6}$$

and suppose rank $\mathcal{R}_M(v - 1) = n$. The solution to (2) at time v starting from the zero state at time $v - M$ is

$$x(v; v - M, 0, u) = \mathcal{R}_M(v-1)\begin{bmatrix} u(v-1) \\ u(v-2) \\ \vdots \\ u(v-M)\end{bmatrix}, \tag{7}$$

or, simplifying the notation,

$$x_M(v) = \mathcal{R}_M(v-1)U_M(v), \tag{8}$$

where

$$U_M(v) = \begin{bmatrix} u(v-1) \\ \vdots \\ u(v-M)\end{bmatrix}.$$

Now, define an n-vector $V_M(v)$ by the relation

$$U_M(v) = \mathcal{R}'_M(v-1)V_M(v), \tag{9}$$

where the "prime" indicates transpose. Then, from (8) and (9),

$$V_M(v) = [\mathscr{R}_M(v - 1)\mathscr{R}'_M(v - 1)]^{-1}x_M(v), \tag{10}$$

and so, we can solve for $V_M(v)$ and thus obtain, from (9), the appropriate sequence of controls needed to reach any given $x_M(v)$.

Necessity. Suppose rank $\mathscr{R}_M(v - 1) < n$ but the system (2) is completely M-step reachable at time v. Then there exists a nonzero vector $\eta \in \mathfrak{R}^n$ such that $\eta'\mathscr{R}_M(v - 1) = 0$. Hence, premultiplying both sides of (7) with η' yields $\eta'x(v; v - M, 0, u) = 0$ regardless of u. Since the system is completely M-step reachable at time v, choose $\{u(v - M), \cdots, u(v - 1)\}$ such that $x(v; v - M, 0, u) = \eta$. Then $\eta'\eta = 0$ which contradicts the assumption that $\eta \neq 0$.

COROLLARY 1. *The system* (2) *is completely M-step reachable at time v if and only if the rows of* $\Omega(v - 1, v - k + 1)B(v - k)$, *considered as functions of k, are linearly independent over the discrete k-interval* $[1, M]$.

Proof. Let

$$\Theta(v - 1, v - k) = \Omega(v - 1, v - k + 1)B(v - k).$$

If the rows of $\Theta(v - 1, v - k)$ are linearly dependent on $[1, M]$, then there exists an n-vector $\eta \neq 0$ such that

$$\eta'\Theta(v - 1, v - k) = 0 \quad \text{for all integers } k \in [1, M]. \tag{11}$$

But (11) implies $\eta'\mathscr{R}_M(v - 1) = 0$, where$\mathscr{R}_M$ is given by (6). Hence rank$\mathscr{R}_M(v - 1) < n$, and, by Theorem 1, the system is not completely reachable at time v. Reversing the argument proves the converse.

Remark 1. The criterion (5) is also a sufficient condition for complete (M-step) controllability of (2) at time $v - M$. It is *not*, however, a *necessary* condition for controllability as defined in Definition 2 (with $N = M$) unless $A(\cdot)$ is invertible on the discrete interval $[v - M + 1, v - 1]$ (the pointwise completeness condition for linear discrete-time systems).

Remark 2. The proof of Theorem 1 shows that complete M-step reachability at time v implies the ability to reach any fixed state at time v from *any* given state (not just the origin) at time $v - M$.

Remark 3. Notice that complete M-step reachability at time v implies complete N-step reachability at time v for all integers $N \geqq M$. This statement is false if reachability is replaced by controllability unless $A(\cdot)$ is invertible for all integers $\geqq v + M$.

Remark 4. In the time-invariant case, (5) reduces to the standard condition

$$\operatorname{rank}[B, AB, \cdots, A^{M-1}B] = n \quad \text{for some } M \leqq n. \tag{12}$$

It therefore follows that in an n-dimensional time-invariant system, complete reachability (or controllability) implies complete M-step reachability (or controllability), with M any integer satisfying (12). The minimum possible value of M is $n - p + 1$, where p is the number of control variables.

4. Controllability for nonlinear discrete-time systems.

Using Theorem 1 plus a technique developed by Lee and Markus [15], one can obtain a result for controllability of systems of the form (1). In this case, it is expedient to assume that

f is a differentiable function of its arguments although only the values of f and its derivatives at discrete instants are of interest. (Hence this type of result is applicable to sample-data systems.)

Consider the system (1) with $f(k, 0, 0) \equiv 0$ and f differentiable in all its arguments.

DEFINITION 3. The system (1) is *locally controllable* at time v if there exists a neighborhood, $\mathcal{N}_0$, of the origin in $\mathfrak{R}^n$ such that for every state $x \in \mathcal{N}_0$, there exists $M \in \mathfrak{Z}^+$ and a sequence, $\{u(v), u(v+1), \cdots, u(v+M-1)\}$, such that $x(v+M; v, x, u) = 0$.

Now consider the system (2), where

$$A(k) = \frac{\partial f}{\partial x}(k, 0, 0), \qquad B(k) = \frac{\partial f}{\partial u}(k, 0, 0).$$

THEOREM 2. *The system* (1) *is locally controllable at time v if the system* (2), *with $A(\cdot)$ and $B(\cdot)$ as given above, is completely controllable at time v.*

Proof. For simplicity, let $v = 0$. For some fixed integer λ, and n-dimensional vector parameter ξ, let

$$u(k, \xi) = B'(k)[\Omega(\lambda - 1, k + 1)]'\xi, \qquad k = 0, 1, \cdots, \lambda - 1.$$

Define

$$x(k + 1, \xi) = f(k, x(k, \xi), u(k, \xi)), \qquad k = 0, 1, 2, \cdots,$$

with $x(0, \xi) = 0$, and let

$$J(k) = \frac{\partial x}{\partial \xi}(k, \xi)|_{\xi=0}.$$

Notice that $u(k, 0) = 0$ for all k so that $x(k, 0) = 0$ for all k. In that case, it is easy to show that J satisfies the difference equation

$$\begin{aligned} J(k+1) &= A(k)J(k) + B(k)\frac{\partial u}{\partial \xi}(k, \xi)|_{\xi=0} \\ &= A(k)J(k) + B(k)B'(k)[\Omega(\lambda - 1, k + 1)]', \end{aligned} \tag{13}$$

where $J(0) = 0$ and $k = 0, 1, \cdots, \lambda - 1$. Iterating (13), we obtain

$$J(\lambda) = \mathscr{R}_\lambda(\lambda - 1)\mathscr{R}'_\lambda(\lambda - 1), \tag{14}$$

where $\mathscr{R}_\lambda(\lambda - 1)$ is given by (6). By hypothesis, there exists an integer M such that rank $\mathscr{R}_M(M - 1) = n$. Then $J(M)$ is nonsingular and the implicit function theorem allows one to obtain a solution to the equation $x(M; 0, x_0, \xi) = 0$ in terms of a mapping $\Pi: \mathcal{N}_0 \to \mathfrak{R}^n$ such that $\xi = \Pi(x_0)$. Hence, (1) is locally controllable at time 0.

5. An alternative approach to reachability. In this section we show that a discrete version of Pfaffian systems can be used to generate results on reachability for linear discrete-time systems.

Consider the system (2) and let $G(k)$ be an $(n - p) \times n$ matrix such that rank $G(k) = n -$ rank $B(k)$ and $G(k)B(k) = 0$ for all integers $k \in [v - M, v - 1]$, for some $M \in \mathfrak{Z}^+$, $M > 1$. Then, from (2), we have that

$$G(k)x(k+1) - G(k)A(k)x(k) = 0 \tag{15}$$

for all k on the discrete interval $[v - M, v - 1]$. We shall call (15) the *discrete Pfaffian* associated with (2). Let g_i' denote the ith row of G and let $g'(k)$ be an arbitrary nonzero linear combination of the rows of $G(k)$. That is,

$$g'(k) = \sum_i \alpha_i(k) g_i(k).$$

DEFINITION 4. The discrete Pfaffian system (15) is *summable* on $[v - M, v - 1]$ if and only if there exists some nonzero $g'(k)$ such that the expression

$$g'(k)x(k+1) - g'(k)A(k)x(k) \tag{16}$$

is an exact forward difference on $[v - M, v - 2]$.

More precisely, summability of (15) on $[v - M, v - 1]$ means that there exists a scalar-valued function $\varphi(k, x)$, where $(k, x) \in [v - M, v - 2] \times \mathfrak{R}^n$, such that

$$\begin{aligned} \Delta_x \varphi(k, x) &= g'(k), \\ \Delta_k \varphi(k, x) &= g'(k)[A(k) - I]x, \end{aligned} \tag{17}$$

where Δ_k represents the forward difference operator with respect to the variable k and

$$\Delta_x \varphi(k, x) = (\Delta_{x_1} \varphi(k, x), \cdots, \Delta_{x_n} \varphi(k, x)).$$

Clearly, this type of definition will also work for nonlinear systems of the form (1), provided the control $u(k)$ appears linearly. In the case at hand, the choice of φ is obvious, since, by inspection of (16), we have the following.

PROPOSITION 1. *A necessary and sufficient condition for* (16) *to be an exact forward difference on* $[v - M, v - 2]$ *is that*

$$g'(k-1) = g'(k)A(k)$$

for all k on the discrete interval $[v - M + 1, v - 1]$.

The function φ in (17) can therefore be taken as $\varphi(k, x) = g'(k)x$.

The main result we wish to prove in this section is as follows.

THEOREM 3. *The following statements are equivalent.*

(i) *The system* (2) *is completely M-step reachable at time v.*

(ii) *The discrete Pfaffian* (15) *is nonsummable on some discrete subinterval of* $[v - M, v - 1]$.

(iii) $\operatorname{Rank} \mathscr{R}_M(v - 1) = n$, *where* $\mathscr{R}_M$ *is given by* (16).

Proof. (i) ⇔ (iii): This was proved as Theorem 1.

(i) ⇒ (ii): Suppose the Pfaffian is summable on the entire discrete interval $[v - M, v - 1]$. Then there exists a nonzero row vector $g'(k)$ such that $g'(k)B(k) = 0$ for all integers $k \in [v - M, v - 1]$, and by Proposition 1, $g'(k - 1) = g'(k)A(k)$

for all $k \in [v - M + 1, v - 1]$. Hence,

$$\begin{aligned} g'(v-1)\mathscr{R}_M(v-1) &= g'(v-1)[B(v-1), \Omega(v-1, v-1)B(v-2), \\ &\qquad \cdots, \Omega(v-1, v-M+1)B(v-M)] \\ &= [g'(v-1)B(v-1), g'(v-2)B(v-2), \cdots, g'(v-M)B(v-M)] \\ &= 0. \end{aligned}$$

But this implies that rank $\mathscr{R}_M(v-1) < n$, so that the system (2) is not completely M-step reachable at time v. Taking the contrapositive establishes that (i) ⇒ (ii).

(ii) ⇒ (iii): Suppose rank $\mathscr{R}_M(v-1) = q < n$. Then there exists an $n \times n$ nonsingular matrix T such that

$$T\mathscr{R}_M(v-1) = \begin{bmatrix} \hat{\mathscr{R}}_M(v-1) \\ 0 \end{bmatrix}, \tag{18}$$

where $\mathscr{R}_M$ has q rows and rank $\hat{\mathscr{R}}_M(v-1) = q$. Let η' be a fixed nonzero row vector whose first q entries are 0. Define

$$g'(k) = \eta' T\Omega(v-1, k+1), \qquad k \in [v-M, v-1]. \tag{19}$$

Then

$$g'(v-1) = \eta' T$$

and

$$g'(k-1) = g'(k)A(k), \qquad k \in [v-M+1, v-1]. \tag{20}$$

It then follows from (18) and (19) that $g'(k)B(k) = 0$ for all integers $k \in [v - M, v - 1]$, and (20) implies that the discrete Pfaffian associated with the system (2) is summable on the entire discrete interval $[v - M, v - 1]$. Taking the contrapositive proves (ii) ⇒ (iii), which proves the theorem.

For a discussion of the Pfaffian technique applied to continuous-time systems, the reader is referred to Hermes [16] and Weiss [9].

6. Observability. The duality principle discovered by Kalman [12] is a statement of the fact that the mathematical structure of the optimal control, quadratic cost problem in control theory, and the optimal estimation, minimum variance problem in filtering theory, are identical for linear ordinary differential equations. The role of observability in filtering theory is completely dual to that of reachability in control theory.

The appropriate model for our study of observability in the discrete-time case is as follows:

$$\begin{aligned} x(k+1) &= A(k)x(k) + B(k)u(k), \\ y(k) &= C(k)x(k), \end{aligned} \tag{21}$$

where $k \in \mathfrak{Z}$, $y(k) \in \Re^m$ and represents the output of the system. $A(\cdot)$, $B(\cdot)$, $x(\cdot)$, $u(\cdot)$ are as in (2).

DEFINITION 5. (a) The system (21) is *completely* (*N-step*) *observable* at time μ if and only if there exists $N \in \mathfrak{Z}^+$ such that knowledge of $y(\mu), y(\mu + 1), \cdots, y(\mu + N - 1)$ and $u(\mu), u(\mu + 1), \cdots, u(\mu + N - 2)$ is sufficient to determine $x(\mu)$.

Exactly Y-observable

(b) The system (21) is *completely* (*N-step*) *determinable* at time μ if and only if there exists $N \in \mathfrak{Z}^+$ such that any state at time μ can be determined from knowledge of $y(\mu - N + 1), \cdots, y(\mu)$ and $u(\mu - N + 1), \cdots, u(\mu - 1)$.

(c) Complete observability (or determinability) without any time designation denotes *complete observability* (or *determinability*) at all times.

The definition of determinability differs from that of observability in that in the former case, we determine the "present" state from "past" measurements, while in the latter case, we determine a "past" state from "future" measurements.

THEOREM 4. *The system* (21) *is completely N-step observable at time μ if and only if*

$$\text{(22)}\quad \operatorname{rank}\,[C'(\mu), [\Omega(\mu, \mu)]'C'(\mu + 1), \cdots, [\Omega(\mu + N - 2, \mu)]'C'(\mu + N - 1)] = n.$$

Proof. Sufficiency. The solution to (21) at the mth instant starting from initial time μ and initial state $x(\mu)$ is

$$\text{(23)}\quad y(m; \mu, x(\mu), u) = C(m)\Omega(m - 1, \mu)x(\mu) + \sum_{k=\mu}^{m-1} C(m)\Omega(m - 1, k + 1)B(k)u(k).$$

Let

$$\tilde{y}(m, \mu) = y(m; \mu, x(\mu), u) - \sum_{k=\mu}^{m-1} C(m)\Omega(m - 1, k + 1)B(k)u(k).$$

Then

$$\text{(24)}\qquad \tilde{y}(m, \mu) = C(m)\Omega(m - 1, \mu)x(u), \qquad m = \mu, \mu + 1, \cdots.$$

Let

$$\text{(25)}\qquad \mathcal{Y}_N(\mu) = \begin{bmatrix} \tilde{y}(\mu, \mu) \\ \tilde{y}(\mu + 1, \mu) \\ \vdots \\ \tilde{y}(\mu + N - 1, \mu) \end{bmatrix}$$

and let

$$\text{(26)}\quad \mathcal{O}_N(\mu) = [C'(\mu), [\Omega(\mu, \mu)]'C'(\mu + 1), \cdots, [\Omega(\mu + N - 2, \mu)]'C'(\mu + N - 1)].$$

Then it follows from (24)–(26) that

$$\text{(27)}\qquad x(\mu) = [\mathcal{O}_N(\mu)\mathcal{O}_N'(\mu)]^{-1}\mathcal{O}_N(\mu)\mathcal{Y}_N(\mu)$$

so that $x(\mu)$ can be computed as long as rank $\mathcal{O}_N(\mu) = n$.

Necessity. Suppose rank $\mathcal{O}_N(\mu) < n$ but the system (21) is completely N-step observable at time μ. Then there exists a nonzero vector $\xi \in \Re^n$ such that $\xi'\mathcal{O}_N(\mu) = 0$. From (24) and (25) we have

$$\text{(28)}\qquad \mathcal{Y}_N(\mu) = \mathcal{O}_N'(\mu)x(\mu).$$

Setting $x(\mu) = \xi$ implies $\mathcal{Y}_N(\mu) = 0$ which violates complete N-step observability (the "output" is identically zero over $[\mu, \mu + N - 1]$ although the state at time μ is not zero).

Remark 5. The duality between Theorems 1 and 4 is obvious. The system (21) is completely N-step reachable (or completely N-step observable) at time ν if and only if the system

$$(29)\qquad \begin{aligned} z(k-1) &= A'(k)z(k) + C'(k)\tilde{u}(k),\\ \tilde{y}(k) &= B'(k)z(k), \end{aligned}$$

with the time scale reversed about ν, is completely N-step observable (or completely N-step reachable) at time ν.

Remark 6. One would expect the criterion for determinability (see Definition 5(c)) to be similar to that for observability (see (22)), but there is an important difference. The criterion (22) is only *sufficient* for complete N-step determinability at μ unless the matrix $A(\cdot)$ is nonsingular over $[\mu, \mu + N - 1]$. That is, pointwise degeneracy could force the "present" state to be zero regardless of "past" values of y. Since knowledge of the homogeneous system equation is presumed to be available to the observer, one could then, under the aforementioned circumstances, determine the "present" state regardless of the rank of the determinability matrix.

Remark 7. From condition (22) it is evident that complete N-step observability at time μ implies complete M-step observability at time μ for any integer $M \geqq N$. This is not the case for determinability, however, unless the pointwise completeness condition holds at all integral times $\leqq \mu - N$.

Remark 8. The discussion in Remarks 1, 5, 6 and 7 indicates that the pairing of reachability–observability and controllability–determinability as dual variables is most natural. This will become still clearer in the next section. For remarks on this problem within the continuous-time framework, where the issue is slightly less transparent, see Weiss [17] and Kalman [18].

Finally, the dual result of Corollary 1 is given.

COROLLARY 2. *The system* (21) *is completely N-step observable at time μ if and only if the columns of $C(\mu + k)\Omega(\mu + k - 1, \mu)$, considered as functions of k, are linearly independent over the discrete k-interval* $[0, N - 1]$.

7. Realization of input/output functions for linear discrete-time systems.

As a result of different analysis or design considerations, dynamical systems are represented in various ways, e.g., by means of transfer functions, impulse responses and weighting patterns, input/output operator equations, state-variable equations, etc. For any kind of system with such different representations, it is desirable to be able to move easily from one representation to another. Exactly how one does this in the case of finite-dimensional systems has been the subject of many papers in recent years (see Kalman [19], Weiss and Kalman [20], and Youla [21]). A complete solution to the problem of constructing a state-variable differential or difference equation from input/output functions of "smooth" linear, time-invariant, finite-dimensional systems was given by B. L. Ho [22] via a now well-known algorithm.

Our objective in this section is to develop the background for a (partial) solution to the following problem: Given a graph of a matrix function of two discrete variables, $W(k, l)$, find (if possible) a linear, discrete-time system of the form of (21), having minimal state dimension, which generates the given data as the graph of the system's "unit pulse response" (see Definition 6 below).

We begin by considering the linear system (21) with its concomitant solution (23).

DEFINITION 6. The *unit pulse response matrix* for the system (21) is the $m \times p$ matrix function $W(k, l)$ given by

$$W(k,l) = \begin{cases} C(k)\Omega(k-1, l+1)B(l), & k > l, \\ 0, & k \leqq l. \end{cases} \tag{30}$$

Remark 9. The name "unit pulse response" is suggested by the fact that, in (21), if $u = \operatorname{col}(u_1, \cdots, u_p)$, $y = (y_1, \cdots, y_m)$, and $W = (w_{ij})$, then the ith column, W_i, of W can be expressed as

$$W_i(n, r) = y(n; r, 0, u),$$

where $u(k) = \operatorname{col}(0, \cdots, 0, \delta_{kr}, 0, \cdots, 0)$, and the Kronecker delta δ_{kr} appears in the ith entry.

Remark 10. In ordinary linear differential systems, the kernel matrix $W(t, \tau)$ (referred to as the "weighting pattern" [23] in system theory) is defined for all t, τ while the "causal impulse response" $W_C(t, \tau)$ is defined as

$$W_C(t,\tau) = \begin{cases} W(t,\tau), & t \geqq \tau, \\ 0, & t < \tau. \end{cases}$$

Except for the case where the system (21) is pointwise complete for all $k \in \mathfrak{Z}$ (i.e., the "A" matrix is invertible for all k), the unit pulse response $W(k, l)$ is not naturally well-defined for $l > k$, and is arbitrarily set to 0 in the latter region. In this sense, the unit pulse response and the causal impulse response are analogous.

PROPOSITION 2. *The unit pulse response* (30) *of a system* (21) *is invariant under coordinate transformations.*

Proof. Consider an arbitrary coordinate transformation $\hat{x}(k) = T(k)x(k)$. Then $\{A(\cdot), B(\cdot), C(\cdot)\}$ is transformed into $\{\hat{A}(\cdot), \hat{B}(\cdot), \hat{C}(\cdot)\}$ according to the relations

$$\begin{aligned} \hat{A}(k) &= T(k+1)A(k)T^{-1}(k), \\ \hat{B}(k) &= T(k+1)B(k), \\ \hat{C}(k) &= C(k)T^{-1}(k). \end{aligned}$$

Then

$$\begin{aligned} \hat{W}(k,l) &= \hat{C}(k)\tilde{\Omega}(k-1, l+1)\hat{B}(l) \\ &= C(k)T^{-1}(k)T(k)\Omega(k-1, l+1)T^{-1}(l+1)T(l+1)B(l) \\ &= C(k)\Omega(k-1, l+1)B(l) = W(k,l). \end{aligned}$$

Q.E.D.

Now consider (30) and, for some fixed integer λ, write it as

$$W(k,l) = C(k)\Omega(k-1, \lambda+1)\Omega(\lambda, l+1)B(l), \qquad l \leqq \lambda < k.$$

Let

$$\Psi(k,\lambda) \triangleq C(k)\Omega(k-1, \lambda+1), \qquad k > \lambda, \tag{31}$$

$$\Theta(\lambda, l) \triangleq \Omega(\lambda, l+1)B(l). \qquad l \leqq \lambda. \tag{32}$$

Then

$$W(k,l) = \Psi(k,\lambda)\Theta(\lambda,l), \qquad l \leqq \lambda < k. \tag{33}$$

Let

$$\begin{aligned} n_\theta(\lambda) &\triangleq \text{number of rows of } \Theta(\lambda,\cdot\,) \\ &\quad\ \text{which are linearly independent on } (-\infty,\lambda], \\ n_\psi(\lambda) &\triangleq \text{number of columns of } \Psi(\,\cdot\,,\lambda) \\ &\quad\ \text{which are linearly independent on } [\lambda + 1,\infty), \end{aligned}$$

$$n_0 \triangleq \max_{\lambda} \min\,\{n_\theta(\lambda), n_\psi(\lambda)\}. \tag{34}$$

(Note that n_0 is uniquely determined by (34).)

Define $\hat{\lambda}$ by the relation

$$n_0 = \min\,\{n_\theta(\hat{\lambda}), n_\psi(\hat{\lambda})\}. \tag{35}$$

DEFINITION 7. The unit pulse response $W(k,l)$ in (30) is *globally reduced* if and only if $n_0 = n_\theta(\hat{\lambda}) = n_\psi(\hat{\lambda})$.

LEMMA 1. *Every nonzero unit pulse response has a globally reduced form.*

Proof. Let $W(k,l)$ be a unit pulse response given by (30) and suppose it is not globally reduced. Assume, without loss of generality, that $n_\psi(\hat{\lambda}) < n_\theta(\hat{\lambda}) < n$ $=$ number of rows (columns) of Ω. Then there exists an $n \times n$ nonsingular constant matrix T_1 such that

$$T_1\Theta(\hat{\lambda},\cdot\,) = \begin{bmatrix} \Theta_1(\hat{\lambda},\cdot\,) \\ 0 \end{bmatrix},$$

where the $n_\theta(\hat{\lambda})$ rows of Θ_1 are linearly independent on $(-\infty,\hat{\lambda}]$. Partitioning Ψ conformably with Θ_1 and multiplying Ψ on the right by T_1^{-1}, we get

$$\begin{aligned} W(k,l) &= [\Psi_{11}(k,\hat{\lambda}) \quad \Psi_{12}(k,\hat{\lambda})]\begin{bmatrix} \Theta_1(\hat{\lambda},l) \\ 0 \end{bmatrix} \\ &= \Psi_{11}(k,\hat{\lambda})\Theta_1(\hat{\lambda},l). \end{aligned}$$

Since the $n_\theta(\hat{\lambda})$ columns of Ψ_{11} are not linearly independent on $[\hat{\lambda} + 1,\infty)$, there exists an $n_\theta(\hat{\lambda}) \times n_\theta(\hat{\lambda})$ nonsingular matrix T_2 such that

$$W(k,l) = \Psi_{11}(k,\hat{\lambda})T_2T_2^{-1}\Theta_1(\hat{\lambda},l),$$

where

$$\Psi_{11}(k,\hat{\lambda})T_2 = [\Psi(k,\hat{\lambda}) \quad 0]$$

and the $n_\psi(\hat{\lambda})$ columns of $\tilde{\Psi}$ are linearly independent over $[\hat{\lambda} + 1,\infty)$. Writing

$$\tilde{\Theta}(\hat{\lambda},l) = T_2^{-1}\tilde{\Theta}_1(\hat{\lambda},l)$$

we then have that

$$W(k,l) = \tilde{\Psi}(k,\lambda)\tilde{\Theta}(\lambda,l) = \tilde{C}(k)\hat{\Omega}(k-1,l+1)\tilde{B}(l)$$

which is globally reduced with $\tilde{C}$ of dimension $m \times n_0$ and $\tilde{B}$ of dimension $n_0 \times p$.

DEFINITION 8. A *global realization* of a unit pulse response $W(k, l)$ is a linear finite-dimensional discrete-time system, defined on $\mathfrak{J}$, whose unit pulse response coincides with $W(k, l)$. A realization on a smaller time set is a *local realization.*

DEFINITION 9. A *minimal realization* of a globally reduced unit pulse response is one which has the lowest state-space dimension of all global realizations.

Now consider the following lemma.

LEMMA 2. (a) *A minimal realization of a globally reduced unit pulse response is completely reachable and completely observable at some time* v.

(b) *Conversely, any realization which is completely reachable and completely observable at some time* v *is minimal.*

Proof. (a) Let $W(k, l)$ be given by (33). Suppose there is no time such that the realization $\{A(\cdot), B(\cdot), C(\cdot)\}$ is completely reachable and completely observable. Then, by Corollaries 1 and 2, and for any integer λ, either the columns of $\Psi(\cdot, \lambda)$ or the rows of $\Theta(\lambda, \cdot)$ are linearly dependent on the infinite discrete interval $[\lambda + 1, \infty)$ or $(-\infty, \lambda]$, respectively. But this contradicts the assumption that W is globally reduced.

(b) Suppose $\{A(\cdot), B(\cdot), C(\cdot)\}$ is a nonminimal completely reachable and completely observable realization of a unit pulse response. Let $\{\hat{A}(\cdot), \hat{B}(\cdot), \hat{C}(\cdot)\}$ be a minimal realization. Then

$$C(k)\Omega(k-1, l+1)B(l) = \hat{C}(k)\hat{\Omega}(k-1, l+1)\hat{B}(l)$$

and both expressions represent the same globally reduced unit pulse response. But $\dim \Omega(\cdot) > \dim \hat{\Omega}(\cdot)$ and this contradicts the uniqueness of n_0 in (34).

COROLLARY 3. *The dimension of any minimal realization of a unit pulse response* (30) *is the number* n_0 *in* (34).

DEFINITION 10. Two realizations of a given unit pulse response are *algebraically equivalent* if and only if they are related by a coordinate transformation.

Although a unit pulse response is invariant under coordinate transformations, and all minimal realizations have the same dimension, it does not follow that all minimal realizations are algebraically equivalent. This is due to the existence of "extra degrees of freedom" in the choice of $C(k)$ or $B(l)$ provided by the arbitrary setting of $W(k, l) = 0$ for $l \geqq k$. To illustrate this, consider the scalar systems $\{a(\cdot), b(\cdot), c(\cdot)\}$ and $\{a(\cdot), b(\cdot), \hat{c}(\cdot)\}$ in which

$$a \equiv 1, \qquad b(k) = \begin{cases} 1, & k = 3, 4, \\ 0, & \text{otherwise}, \end{cases}$$

$$c(k) = \begin{cases} 1, & k = 5, 6, \\ 0, & \text{otherwise}, \end{cases} \qquad \hat{c}(k) = \begin{cases} 1, & k = 1, 2, 5, 6, \\ 0, & \text{otherwise}. \end{cases}$$

Then the unit pulse response for both systems is

$$W(k, l) = \begin{cases} 1, & k = 5, 6, \quad l = 3, 4, \\ 0, & \text{otherwise}, \end{cases} \tag{36}$$

but the two systems, which are both minimal realizations of the unit pulse response (36), are not algebraically equivalent.

We now delineate a class of systems of the form (21) in which all minimal realizations of the associated unit pulse responses are algebraically equivalent.

The proof of the following theorem is analogous to one given by Youla [21] (see also Kalman [13]) for a result on algebraic equivalence of continuous-time systems announced in [23] (see also [20]).

THEOREM 5. *Two realizations of a given unit pulse response are algebraically equivalent if they are completely N-step reachable and completely M-step observable for some $N, M \in \mathfrak{Z}^+$.*

Proof. Let $\{A_i(\cdot), B_i(\cdot), C_i(\cdot)\}$, $i = 1, 2$, be the two realizations. Let

$$\Psi_i(k, \lambda) = C_i(k)\Omega_i(k - 1, \lambda + 1), \qquad i = 1, 2,$$

$$\Theta_i(\lambda, l) = \Omega_i(\lambda, l + 1)B_i(l), \qquad i = 1, 2.$$

Then, for any $\lambda \in \mathfrak{Z}$, the rows of $\Theta_i(\lambda, \cdot)$ and the columns of $\Psi_i(\cdot, \lambda)$ are linearly independent on the discrete intervals $I_N = [\lambda - N + 1, \lambda]$ and $J_M = [\lambda + 1, \lambda + M - 1]$, respectively. Let

$$K_i = \sum_{k \in J_M} \Psi_i'(k, \lambda)\Psi_i(k, \lambda), \qquad i = 1, 2,$$

$$L_i = \sum_{k \in I_N} \Theta_i(\lambda, k)\Theta_i'(\lambda, k), \qquad i = 1, 2.$$

Then, since

$$W(k, l) = \Psi_i(k, \lambda)\Theta_i(\lambda, l), \qquad l \leqq \lambda < k, \quad i = 1, 2,$$

we have

$$\Theta_2(\lambda, l) = K_2^{-1}\left[\sum_{k \in J_M} \Psi_2'(k, \lambda)\Psi_1(k, \lambda)\right]\Theta_1(\lambda, l) \tag{37}$$
$$= U\Theta_1(\lambda, l),$$

where

$$U = K_2^{-1} \sum_{k \in J_M} \Psi_2'(k, \lambda)\Psi_1(k, \lambda).$$

Similarly,

$$\Psi_2(k, \lambda) = \Psi_1(k, \lambda)\left[\sum_{l \in I_N} \Theta_1(\lambda, l)\Theta_2'(\lambda, l)\right]L_2^{-1} \tag{38}$$
$$= \Psi_1(k, \lambda)V,$$

where

$$V = \left[\sum_{l \in I_N} \Theta_1(\lambda, l)\Theta_2'(\lambda, l)\right]L_2^{-1}.$$

Then

$$\Psi_1(k, \lambda)\Theta_1(\lambda, l) = \Psi_2(k, \lambda)\Theta_2(\lambda, l) = \Psi_1(k, \lambda)VU\Theta_1(\lambda, l).$$

Hence $VU = I$ or $U = V^{-1}$. It then follows from (37), (38), and the definitions of Θ_i and Ψ_i, that $\{A_1(\cdot), B_1(\cdot), C_1(\cdot)\}$ is algebraically equivalent to $\{A_2(\cdot), B_2(\cdot), C_2(\cdot)\}$.

8. Construction of minimal realizations. We now present an algorithm for constructing a minimal realization of a unit pulse response given in the form of numerical data. The algorithm, which will work whenever the given unit pulse response possesses a realization which is completely N-step reachable and completely N-step observable for some $N \in \mathfrak{Z}^+$, is structurally analogous to one given by Silverman [24] for certain special classes of continuous-time systems. Important computational advantages are gained in the present context, however, and the result can be viewed as part of a procedure for synthesizing optimal linear digital filters [25].

We begin by defining the generalized Hankel matrix (see [26]) for the unit pulse response $W(k, l)$ given by (30).

DEFINITION 11. The *generalized Hankel matrix* of a unit pulse response $W(k, l)$ is given by

$$\mathscr{W}_N(k, l) =$$

(39)
$$\begin{bmatrix} W(k,l) & W(k,l-1) & \cdots & W(k,l-N+1) \\ W(k+1,l) & W(k+1,l-1) & \cdots & W(k+1,l-N+1) \\ \vdots & & & \\ W(k+N-1,l) & W(k+N-1,l-1) & \cdots & W(k+N-1,l-N+1) \end{bmatrix}.$$

From (30), (6), and (26), we have

$$\mathscr{W}_N(k, k-1) = \mathscr{O}'_N(k)\mathscr{R}_N(k-1). \tag{40}$$

THEOREM 6. *Let $W(k, l)$ be an $m \times p$ matrix function of two discrete variables. Suppose W is a unit pulse response with an n-dimensional realization $\{A(\cdot), B(\cdot), C(\cdot)\}$ as in* (30), *and, for some $N \in \mathfrak{Z}^+$, $\mathscr{W}_N(k, k-1)$ contains a fixed $n \times n$ submatrix $\Gamma(k)$ which is nonsingular for all $k \in \mathfrak{Z}$. Then there exists an n-dimensional, completely N-step reachable and completely N-step observable realization of W, given by*

$$\{\Gamma_1\Gamma^{-1}, \Lambda, \Phi\Gamma^{-1}\}, \tag{41}$$

where

$\Lambda(k)$ = *submatrix of $\mathscr{W}_N(k, k-1)$ consisting of those rows of the first m-column block whose indices match those of the rows of* Γ;

$\Phi(k)$ = *submatrix of $\mathscr{W}_N(k, k-1)$ consisting of those columns of the first p-row block whose indices match those of the columns of* Γ;

$\Gamma_1(k)$ = *submatrix of $\mathscr{W}_N(k+1, k-1)$ consisting of those elements whose indices match those of the elements of* Γ.

Proof. Since we postulate existence of an n-dimensional realization and rank $\mathscr{W}_N(k, k-1) \geqq n$ for all k, it follows from (40) that rank $\mathscr{W}_N(k, k-1)$ = rank $\mathscr{O}_N(k)$ = rank $\mathscr{R}_N(k-1) = n$ for all k. But, by Theorems 1 and 2, this implies that the realization $\{A(\cdot), B(\cdot), C(\cdot)\}$ is completely N-step reachable and completely N-step observable, and by Lemma 2, this realization is minimal. Now, it follows from (40) that

$$\Gamma(k) = \Gamma_o(k)\Gamma_r(k-1), \tag{42}$$

where $\Gamma_o(k)$ consists of those rows of $\mathcal{O}'_N(k)$ whose indices match those of the rows of Γ, and $\Gamma_r(k-1)$ consists of those columns of $\mathcal{R}_N(k-1)$ whose indices match those of the columns of Γ. Since Γ is nonsingular, it follows that Γ_o and Γ_r are nonsingular. Defining $\Gamma_1(k)$ as the matrix contained in

$$\mathcal{W}_N(k+1, k-1) = \begin{bmatrix} C(k+1)A(k) \\ C(k+2)C(k+1)A(k) \\ \cdot \\ \cdot \\ C(k+N-1)\cdots C(k+1)A(k) \end{bmatrix} [B(k-1), A(k-1)B(k-2), \cdots, A(k-1)\cdots A(k+N-1)B(k-N)],$$

the indices of whose elements match those of the elements of Γ, it is clear that

$$\Gamma_1(k) = \Gamma_o(k+1)A(k)\Gamma_r(k-1).$$

But since $\Gamma_1(k)$ is a submatrix of $\mathcal{W}_N(k, k-1)$ and rank $\mathcal{W}_N(k, k-1) =$ rank $\mathcal{W}_{N+1} \cdot (k, k-1)$ it follows that there exists an $n \times n$ matrix $\hat{A}(k)$ such that

$$\Gamma_1(k) = \hat{A}(k)\Gamma(k). \tag{43}$$

Then

$$\Gamma_o(k+1)A(k)\Gamma_r(k-1) = \hat{A}(k)\Gamma_o(k)\Gamma_r(k-1)$$

which yields

$$\hat{A}(k) = \Gamma_o(k+1)A(k)\Gamma_o^{-1}(k) = \Gamma_1(k)\Gamma^{-1}(k). \tag{44}$$

In like manner, it follows from (40) plus the definitions of Λ and Φ that

$$\Lambda(k) = \Gamma_o(k+1)B(k) \tag{45}$$

and

$$\Phi(k) = C(k)\Gamma_r(k-1) = C(k)\Gamma_o^{-1}(k)\Gamma(k). \tag{46}$$

It is now a simple matter to note that $\{A(\cdot), B(\cdot), C(\cdot)\}$ is algebraically equivalent to $\{\Gamma_1\Gamma^{-1}, \Lambda, \Phi\Gamma^{-1}\}$, where the associated coordinate transformation matrix is Γ_o. The realization (41) is completely N-step reachable and completely N-step observable (since these properties are invariant under algebraic equivalence), and the realization is obviously minimal.

Remark 11. Although the algorithm proceeds from the assumption that a realization with the appropriate properties exists, one need not know what this realization is in order to obtain (41).

Remark 12. The realization (41) will be time-invariant if a time-invariant realization exists, since $\mathcal{W}_N(k, k-1)$ will be a constant matrix under those circumstances.

Remark 13. The algorithm will yield a realization defined on any discrete k-interval on which the function $\mathcal{W}_N(k, k-1)$ is defined.

Remark 14. The "partial realization" problem discussed by Kalman [27] and Tether [28] has an analogue in the present framework, but we shall not consider that problem here.

9. Lyapunov stability for discrete-time systems. Two of the most fundamental questions arising in control system analysis and design are:

(a) What are the possibilities for identification and alteration of system behavior?

(b) Is the system stable (in some sense)?

Up to this point, we have been concerned only with question (a) and its ramifications with respect to the realization problem. We now turn to question (b) and discuss some aspects of the stability problem for discrete-time systems.

We first consider the nonlinear, homogeneous, n-dimensional system

$$(47) \qquad x(k+1) = f(k, x(k)), \qquad k \in \mathfrak{S} = \{k_0, k_0 + 1, k_0 + 2, \cdots\},$$

defined within a region $\mathfrak{J} = \{x \in R^n \colon \|x\| \leqq H\}$. It is assumed that f is finite for all finite values of its arguments and that $f(k, 0, 0) = 0$ for all $k \in \mathfrak{S}$.

DEFINITION 12. The zero solution of the system (47) is *stable* if, for any $\varepsilon > 0$, there exists $\delta(\varepsilon, k_0) > 0$ such that $\|x(k_0)\| < \delta$ implies $\|x(k; k_0, x(k_0))\| < \varepsilon$ for all $k \in \mathfrak{S}$.

Sufficient conditions for stability of (47) in terms of existence of Lyapunov functions have been known for some time (see [29]). The converse problem, however, has been treated only sparsely in the literature, and usually only with very restrictive assumptions. For example, Hahn [30] has proved a converse theorem within the context of "general motions" of dynamical systems, but his assumptions, when applied to (47), essentially include the requirement that f be "invertible" (i.e., that $x(k)$ can be expressed as a function of $x(k + 1)$).

We now show that this assumption is unnecessary for the concrete model (47) by proving the following Lyapunov-type stability theorem and its converse.

THEOREM 7 (Weiss and Lam [31]). *The system* (47) *is stable if and only if there exists a real-valued function* $V(k, x)$ *defined on* $(k, x) \in \mathfrak{S} \times \mathfrak{J}$ *and continuous in* x *at* $x = 0$, *such that*:

(i) $V(k, 0) = 0$ *for all* $k \in \mathfrak{S}$.

(ii) *There exists a real-valued, continuous, monotone-increasing positive definite function* $a(\cdot)$ *such that* $V(k, x) \geqq a(\|x\|)$ *for all* $k \in \mathfrak{S}$, *all* $x \in \mathfrak{J}$.

(iii) $\Delta V(k; x(k; k_0, x)) \leqq 0$ *for all* $k \in \mathfrak{S}, x \in \mathfrak{J}$, *where* Δ *is the forward difference operator.*

Proof. Sufficiency. By the assumptions on V, corresponding to any given $\varepsilon > 0$, $(\varepsilon < H)$, we can choose $\delta > 0$ such that $\|x_0\| < \delta$ implies $\|x(k_1; k_0, x_0)\| = \varepsilon' \geqq \varepsilon$. By hypothesis (iii),

$$V(k_1, x(k; k_0, x_0)) \leqq V(k_0, x_0).$$

Hence, we have

$$a(\varepsilon) \leqq a(\varepsilon') \leqq V(k_1, x(k_1; k_0, x_0)) \leqq V(k_0, x_0) < a(\varepsilon)$$

which is a contradiction.

Necessity. Let

$$V(k, x) = \min \{H; \sup_{\substack{j \geqq k \\ j, k \in \mathfrak{S}}} \{\|x(j; k, x)\|\}\},$$

$$a(\|x\|) = \|x\|.$$

Then $V(k, 0) = 0$ for all $k \in \mathfrak{S}$, and since (47) is assumed to be stable V is continuous in x at $x = 0$. Moreover, it is simple to check that hypothesis (ii) is satisfied. Now, for any $k \in \mathfrak{S}$ and any $x \in \mathfrak{J}$, we have

$$\Delta V(k, x(k; k_0, x)) = \min \{H; \sup_{\substack{j \geqq k+1 \\ j,k+1 \in \mathfrak{S}}} \{\|x(j; k+1, x(k+1; k_0, x))\|\}\}$$
$$- \min \{H; \sup_{\substack{j \geqq k \\ j,k \in \mathfrak{S}}} \{\|x(j; k, x(k; k_0, x))\|\}\}.$$

Since

$$\sup_{\substack{j \geqq k+1 \\ j,k+1 \in \mathfrak{S}}} \|\cdot\| \leqq \sup_{\substack{j \geqq k \\ j,k \in \mathfrak{S}}} \|\cdot\|$$

we have that $\Delta V(k; x(k; k_0, x)) \leqq 0$ for all $k \in \mathfrak{S}$, all $x \in \mathfrak{J}$.

It should be noted that without further information about the behavior of trajectories of (47), the "sup" function alone cannot suffice as an appropriate V-function in order to prove the necessity of the hypotheses of Theorem 7. If one is interested in boundedness rather than stability, however, such a V-function is admissible [31].

10. Finite-time stability of nonlinear discrete-time systems. In certain practical situations, stability in the sense of Definition 12 may be irrelevant. It may be more pertinent to require some function of the state to remain bounded in a particular way over a fixed finite number of discrete instants rather than to consider asymptotic properties as $k \to \infty$. The theory of finite-time stability has been created in response to such types of problems within a continuous-time context.

We now develop some results in this theory (see Weiss and Lam [27]) for nonlinear discrete-time systems.

Let $\mathfrak{S}_N = \{k_0, k_0 + 1, \cdots, k_0 + N\}$.

DEFINITION 13. The system (47) is *stable with respect to* $(\alpha, \beta, \mathfrak{S}_N)$, $\alpha < \beta$, if $\|x_0\| \leqq \alpha$ implies $\|x(k; k_0, x_0)\| < \beta$ for all $k \in \mathfrak{S}_N$.

DEFINITION 14. The system (47) is *uniformly stable with respect to* $(\alpha, \beta, \mathfrak{S}_N)$, $\alpha < \beta$, if $\|x\| \leqq \alpha$ implies $\|x(j; k, x)\| < \beta$ for $j \in \{k, k+1, \cdots, k_0 + N\}$, for all $k \in \mathfrak{S}_N$.

DEFINITION 15. The system (47) is *contractively stable with respect to* $(\alpha, \beta, \gamma, \mathfrak{S}_N)$, $\beta < \alpha < \gamma$, if it is stable with respect to $(\alpha, \gamma, \mathfrak{S}_N)$ and $\|x_0\| \leqq \alpha$ implies $\|x(k_0 + N; k_0, x_0)\| < \beta$.

We use the following notation:

$$V^a_m(k) = \min_{\|x\| = a} V(k, x), \qquad V^a_{\underline{m}}(k) = \min_{\|x\| \geqq a} V(k, x),$$

$$V^a_M(k) = \max_{\|x\| = a} V(k, x), \qquad V^a_{\overline{M}}(k) = \max_{\|x\| \leqq a} V(k, x),$$

$$\mathfrak{B}(a) = \{x \in R^n : \|x\| < a\}, \qquad \bar{\mathfrak{B}}(a) = \{x \in R^n : \|x\| \leqq a\}.$$

THEOREM 8. *The system* (47) *is stable with respect to* $(\alpha, \beta, \mathfrak{S}_N)$, $\alpha < \beta$, *if there exists a real-valued function* $V(k, x)$, *defined for all* $k \in \mathfrak{S}_N$, *and continuous in* $x \in \mathfrak{R}^n$, *and a real-valued function* $\varphi(k)$ *defined on* $\mathfrak{S}_{N-1}$, *such that*

(i) $\Delta V(k, x(k; k_0, x_0)) \leqq \varphi(k)$ *for all* $k \in \mathfrak{S}_{N-1}$, *all* $x \in \bar{\mathfrak{B}}(\beta)$,

(ii) $\sum_{\substack{j \in \mathfrak{S}_{N-1} \\ j < k}} \varphi(j) < V^\beta_{\underline{m}}(k) - V^\alpha_{\overline{M}}(k_0)$ *for all* $k \in \mathfrak{S}_N$.

Furthermore, if the function f in (47) *is "invertible", then the converse holds.*

Proof. It follows from (i) that for all $k \in \mathfrak{S}_N$,

$$V(k, x(k; k_0, x_0)) \leqq V(k_0, x_0) + \sum_{\substack{j \in \mathfrak{S}_{N-1} \\ j < k}} \varphi(j).$$

Suppose $\|x_0\| \leqq \alpha$ and there exists $l \in \mathfrak{S}_N$ (the first such point) such that $\|x(l; k_0, x_0)\| \geqq \beta$, while $\|x(l-1; k_0, x_0)\| < \beta$. Then

$$\begin{aligned} V_{\underline{m}}^{\beta}(l) \leqq V(l; x(l; k_0, x_0)) &\leqq V(k_0, x_0) + \sum_{\substack{j \in \mathfrak{S}_{N-1} \\ j < l}} \varphi(j) \\ &\leqq V_{\overline{M}}^{\alpha}(k_0) + \sum_{\substack{j \in \mathfrak{S}_{N-1} \\ j < l}} \varphi(j) \end{aligned}$$

which contradicts (ii). This proves the first part.

To prove the converse under the additional hypothesis on f, we take $\varphi(j) \equiv 0$ and define

$$V(k, x) = \|x(k_0; k, x)\|, \qquad k \in \mathfrak{S}_N, \quad x \in \mathfrak{R}^n.$$

Then V is continuous in x and

$$V(k, x(k; k_0, x_0)) = \|x(k_0; k, x(k; k_0, x_0))\| = \|x_0\|$$

for all $k \in \mathfrak{S}_N$.

Hence,

$$\Delta V(k, x) = 0 \quad \text{for all } k \in \mathfrak{S}_N, \quad \text{all } x \in \mathfrak{R}^n,$$

so (i) is satisfied. Now,

$$V(k_0, x) = \|x(k_0; k_0, x)\| = \|x\|$$

so that

$$V_{\overline{M}}^{\alpha}(k_0) = \alpha. \tag{48}$$

Since the system is stable with respect to $(\alpha, \beta, \mathfrak{S}_N)$, it follows that for any pair (k_1, x_1), with $k_1 \in \mathfrak{S}_N$, $\|x_1\| \geqq \beta$, we have

$$V(k_1, x_1) = \|x(k_0; k_1, x_1)\| > \alpha.$$

By continuity of V in x,

$$V_{\underline{m}}^{\beta}(k) > \alpha \quad \text{for all } k \in \mathfrak{S}_N. \tag{49}$$

Combining (48) and (49) yields condition (ii), and the theorem is proved.

For application of this theorem to finite-time stability of *linear* systems, see Weiss and Lee [32].

We now state, without proof, the corresponding theorem for uniform stability.

THEOREM 9. *The system* (47) *is uniformly stable with respect to* $(\alpha, \beta, \mathfrak{S}_N)$, $\alpha < \beta$, *if there exists a real-valued function* $V(k, x)$, *defined for all* $k \in \mathfrak{S}_N$, *and continuous in* $x \in \mathfrak{R}^n$, *and a real-valued function* $\phi(k)$ *defined on* $\mathfrak{S}_{N-1}$, *such that*:

(i) $\Delta V(k, x) \leqq \phi(k)$ *for all* $k \in \mathfrak{S}_{N-1}$, *all* $x \in \overline{\mathfrak{B}}(\beta)$,

(ii) $\sum_{j=k_1}^{k_2-1} \phi(j) < V_{\underline{m}}^{\beta}(k_2) - V_{\overline{M}}^{\alpha}(k_1)$ *for all* $k_1, k_2 \in \mathfrak{S}_N, k_2 > k_1$.

Furthermore, if f in (47) *is "invertible," the converse holds.*

Finally, we present a theorem on contractive stability for systems of the form (47). The result is analogous to that of Kayande [33] for differential equations, and is in the spirit of a result given by Hurt [34]. The latter gives some interesting applications of results of this type to error analysis in numerical computation. Since the proof requires only one extra step beyond that of Theorem 8, it is omitted.

THEOREM 10. *The system* (47) *is contractively stable with respect to* $(\alpha, \beta, \gamma, \mathfrak{S}_N)$, $\beta < \alpha < \gamma$, *if there exist a real-valued function* $V(k, x)$ *defined for all* $k \in \mathfrak{S}_N$, *all* $x \in \mathfrak{R}^n$, *and a real-valued function* $\phi(k)$ *defined for all* $k \in \mathfrak{S}_N$, *such that hypotheses* (i) *and* (ii) *in Theorem* 8 *are satisfied, and in addition*

$$\sum_{j \in \mathfrak{S}_{N-1}} \varphi(j) < \min_{\beta \leqq \|x\| \leqq \alpha} V(k_0 + N, x) - V^{\alpha}_{\underline{M}}(k_0).$$

Furthermore, if f in (47) *is "invertible," the converse holds.*

Remark 15. Sufficiency conditions for finite-time stability of (47) were first obtained by Michel and Wu [35].

Remark 16. Necessary and sufficient conditions for finite-time stability of (47) have also been derived by Heinen [36], but in a form different from that presented here.

11. Finite-time stability of linear discrete-time systems. The basic linear theory for finite-time stability of discrete-time systems has been developed in [32], which contains general sufficient conditions, general necessary conditions, as well as results on mean-square finite-time stability under white noise sequence perturbations.

Our objective in this section is to indicate how the Hermite–Fujiwara form of the Schur–Cohn criterion for asymptotic stability of linear constant-coefficient difference equations can be used to obtain a computationally simple test for finite-time stability of linear time-invariant systems. Our exposition of the classical result follows that of Kalman [37].

First we characterize finite-time stability for linear discrete-time systems.

Let F be a real $n \times n$ matrix. $\{\lambda(F)\}$ is the set of eigenvalues of F. If the latter are real, $\tilde{\lambda} = \max\{\lambda(F)\}$. Define the *spectral norm* of F as

$$\|F\|^* = \sqrt{\tilde{\lambda}(F'F)}.$$

Now consider the system of linear equations

$$x(k + 1) = A(k)x(k), \quad k = k_0, k_0 + 1, k_0 + 2, \cdots, \tag{50}$$

where $A(k)$ is $n \times n$.

The solution at the lth instant starting with initial state x_0 at time k_0 is

$$x(l; k_0, x_0) = \Omega(l - 1, k_0)x_0, \tag{51}$$

where Ω is defined by (3). From (51) and Definition 13 we obtain the following result.

THEOREM 11 (Weiss and Lee [33]).[1] *The system* (51) *is stable* (*Definition* 13) *if and only if*

$$\|\Omega(k, k_0)\|^* < \beta/\alpha, \qquad k = 1, \cdots, N. \tag{52}$$

[1] In [33], $\|x_0\|$ cannot be equal to α in the definition of finite-time stability. Hence (52) and (53) differ from the corresponding conditions in [33] by being strict inequalities.

COROLLARY 4. *If $A(\cdot)$ in* (50) *is constant with time, then* (51) *is stable* (*Definition* 13) *if and only if*

$$\|A^k\|^* < \beta/\alpha, \qquad k = 1, \cdots, N. \tag{53}$$

Now, let A in (50) be a constant matrix, and denote the characteristic polynomial of A by

$$P(z) = z^n + \alpha_1 z^{n-1} + \cdots + \alpha_n. \tag{54}$$

Then the constant system (50) is (classically) asymptotically stable if and only if all the zeros of $P(z)$ lie inside the unit circle on the complex plane.

A criterion for $P(z)$ to have all its zeros inside the unit circle was given by Fujiwara [38], using a classical technique of Hermite, as follows.

Let P^* be the polynomial defined by

$$P^*(z) = z^n P(z^{-1}), \tag{55}$$

and define

$$Q(z, w) = \left[\frac{P(z)P^*(w) - P(w)P^*(z)}{z - w}\right] = \sum_{i,j=1}^{n} z^{i-1}\psi_{ij}w^{j-1}. \tag{56}$$

By (55), Q also has the representation

$$Q(z, w) = w^{n-1}\left[\frac{P(z)P(w^{-1}) - P^*(w^{-1})P^*(z)}{zw^{-1} - 1}\right] = w^{n-1} \sum_{i,j=1}^{n} z^{i-1}\varphi_{ij}w^{1-j}. \tag{57}$$

The matrices $\Phi = (\varphi_{ij})$ and $\Psi = (\psi_{ij})$ are $n \times n$ symmetric matrices and are related by $\psi_{ij} = \varphi_{i,n-j}$. In fact, from (54), (55), and (56), one can easily compute φ_{ij} as

$$\varphi_{ij} = \sum_{k=1}^{\min(i,j)} (\alpha_{i-k}\alpha_{j-k} - \alpha_{n-i+k}\alpha_{n-j+k}), \qquad i,j = 1, \cdots, n. \tag{58}$$

THEOREM 12 (Fujiwara). *The zeros of the polynomial P in* (54) *lie inside the unit circle on the complex plane if and only if the matrix Φ defined by* (58) *is positive definite.*

To apply this to the finite-time stability problem, we need only consider the simple proposition below.

PROPOSITION 3. *If z_0 is an eigenvalue of a matrix F, then cz_0 is an eigenvalue of cF.*

Now, consider the system (50) with A a constant $n \times n$ matrix, and let the characteristic polynomial of $(\alpha^2/\beta^2)(A'^k A^k)$ be given by

$$P_k(z) = z^n + \alpha_{1k}z^{n-1} + \cdots + \alpha_{nk}. \tag{59}$$

Let $\Phi_k = (\varphi_{ij}^{(k)})$ be the $n \times n$ matrix defined by

$$\varphi_{ij}^{(k)} = \sum_{l=1}^{\min(i,j)} (\alpha_{i-l,k} \quad \alpha_{j-l,k} - \alpha_{n-i+l,k} \quad \alpha_{n-j+l,k}), \qquad i,j = 1, \cdots, n. \tag{60}$$

Then, from Corollary 4, Theorem 12 and Proposition 3, we have the main result of the section.

THEOREM 13. *The system* (50), *with A constant, is stable* (*Definition* 13) *if and only if the matrices Φ_k defined by* (60) *are positive definite for $k = 1, 2, \cdots, N$.*

Using a result of Parks [39], an alternative form of Theorem 13 can be obtained. Let P_k be the characteristic polynomial of $(\alpha^2/\beta^2)(A'^k A^k)$ as in (59). Let $P_k^*(x) = z^n P_k(z^{-1})$. Define the n-vector $q_k = \operatorname{col}(q_{k1}, \cdots, q_{kn})$ by the expression

$$\alpha_{kn} P_k(z) - P_k^*(z) = q_{kn} z^{n-1} + \cdots + q_{k1}.$$

Let Ψ_k denote the companion matrix

$$\Psi_k = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & 1 \\ -\alpha_{kn} & \cdot & \cdots & -\alpha_{k1} \end{bmatrix}.$$

Then we have the following theorem.

THEOREM 14. *The system* (50), *with A constant, is stable* (*Definition* 13) *if and only if there exists a sequence $\{S_k\}$ of symmetric, positive definite matrices which satisfy the equations*

$$\Psi_k' S_k \Psi_k - S_k = -q_k' q_k, \qquad k = 1, \cdots, N. \tag{61}$$

REFERENCES

[1] W. FELLER, *An Introduction to Probability Theory and its Applications*, vol. I, John Wiley, New York, 1950.

[2] E. I. JURY, *Theory and Application of the Z-Transform Method*, John Wiley, New York, 1964.

[3] G. BOOLE, *Calculus of Finite Differences*, 4th ed., Chelsea, New York, 1958.

[4] L. M. MILNE-THOMSON, *The Calculus of Finite Differences*, Macmillan, London, 1933.

[5] P. HENRICI, *Discrete Variable Methods in Ordinary Differential Equations*, John Wiley, New York, 1962.

[6] J. R. RAGAZZINI AND G. FRANKLIN, *Sampled-Data Control Systems*, McGraw-Hill, New York, 1958.

[7] S. GOLDBERG, *Introduction to Difference Equations*, John Wiley, New York, 1958.

[8] E. I. JURY AND YA. Z. TSYPKIN, *Theory of Discrete Automatic Systems* (*Review*), Avtomat. i Telemekh., (1970), pp. 57–81.

[9] L. WEISS, *Controllability for various linear and nonlinear system models*, Seminar on Differential Equations and Dynamical Systems. II, Lecture Notes in Mathematics, Springer-Verlag, New York, 1970, pp. 250–261.

[10] ———, *On the controllability of delay-differential systems*, this Journal, 5 (1967), pp. 575–587.

[11] E. A. CODDINGTON AND N. LEVINSON, *Theory of Ordinary Differential Equations*, John Wiley, New York, 1955.

[12] R. E. KALMAN, *On the general theory of control systems*, Proc. First IFAC, Butterworth's, London, 1960, pp. 481–492.

[13] R. E. KALMAN, P. L. FALB AND M. A. ARBIB, *Topics in Mathematical System Theory*, McGraw-Hill, New York, 1969.

[14] L. WEISS, *Lectures on controllability and observability*, C.I.M.E. Seminar on Controllability and Observability, Edizioni Cremonese, Rome, 1969, pp. 205–289.

[15] E. B. LEE AND L. MARKUS, *Optimal control for nonlinear processes*, Arch. Rational Mech. Anal., 8 (1961), pp. 36–38.

[16] H. HERMES, *Controllability and the singular problem*, this Journal, 3 (1965), pp. 241–260.

[17] L. WEISS, *On the structure theory of linear differential systems*, this Journal, 6 (1968), pp. 659–680.

[18] R. E. KALMAN, *Realization theory for nonconstant linear systems*, unpublished notes.

[19] R. E. KALMAN, *Mathematical description of linear dynamical systems*, this Journal, 1 (1963), pp. 152–192.

[20] L. WEISS AND R. E. KALMAN, *Contributions to Linear System Theory*, Internat. J. Engrg. Sci., 3 (1964), pp. 141–171.

[21] D. C. YOULA, *The synthesis of linear dynamical systems from prescribed weighting patterns*, SIAM J. Appl. Math., 14 (1966), pp. 527–549.

[22] B. L. HO AND R. E. KALMAN, *Effective construction of state-variable models from input/output functions*, Regelungstechnik, 12 (1966), pp. 545–548.

[23] R. E. KALMAN, *Canonical structure of linear dynamical systems*, Proc. Nat. Acad. Sci., 48 (1962), pp. 596–600.

[24] L. SILVERMAN AND H. MEADOWS, *Equivalent realizations of linear systems*, SIAM J. Appl. Math., 17 (1969), pp. 393–408.

[25] W. BELFIELD AND L. WEISS, *Realization theory with applications to the design of digital filters*, to appear.

[26] F. R. GANTMACHER, *The Theory of Matrices. I, II*, Chelsea, New York, 1959.

[27] R. E. KALMAN, *On partial realizations of a linear input/output map*, Aspects of Network and System Theory, Holt, Rinehart and Winston, New York, 1970.

[28] A. J. TETHER, *Construction of minimal linear state-variable models from finite input/output data*, IEEE Trans. Automatic Control, AC-15 (1970), pp. 427–436.

[29] W. HAHN, *Über die Anwendung der Methode von Ljapunov auf Differenzengleichungen*, Math. Ann., 136 (1958), pp. 430–441.

[30] ———, *Stability of Motion*, Springer-Verlag, New York, 1967.

[31] L. WEISS AND L. LAM, *Stability of nonlinear discrete-time systems*, Internat. J. Control, to appear.

[32] L. WEISS AND J. S. LEE, *Finite-time stability of linear discrete-time systems*, Avtomat. i Telemekh., to appear.

[33] A. A. KAYANDE, *A theorem on contractive stability*, to appear.

[34] J. HURT, *Some stability theorems for difference equations*, SIAM J. Numer. Anal., 4 (1967), pp. 582–596.

[35] A. N. MICHEL AND S. H. WU, *Stability on discrete systems over a finite interval of time*, Internat. J. Control, 9 (1969), pp. 679–693.

[36] J. A. HEINEN, *Quantitative stability of disqrete systems*, Michigan Math. J., 17 (1970), pp. 211–216.

[37] R. E. KALMAN, *On the Hermite-Fujiwara theorem in stability theory*, Quart. Appl. Math., 23 (1965), pp. 279–282.

[38] M. FUJIWARA, *Über die Algebraischen Gleichungen, deren Wurzeln in einem Kreise oder in einer Halbene Liegen*, Math. Z., 24 (1926), pp. 160–169.

[39] P. C. PARKS, *Lyapunov and the Schur-Cohn stability criterion*, IEEE Trans. Automatic Control, AC-8 (1964), p. 121.

[40] H. W. SORENSON, *Controllability and observability of linear, stochastic, time-discrete control systems*, Advances in Control Systems, C. T. Leondes, ed., vol. 6, Academic Press, New York, 1968.