Some Fundamental Control Theory I: Controllability, Observability, and Duality

William J. Terrell

1. INTRODUCTION. It is well known that a single n-th order nonhomogeneous linear differential equation is equivalent to a system of n first order linear differential equations. Specifically, an n-th order linear equation

$$y^{(n)} + k_1 y^{(n-1)} + k_2 y^{(n-2)} + \cdots + k_n y = u(t), \tag{1}$$

with real constant coefficients k_i, is equivalent, via the standard definition of the vector variable $z = [y\ y'\ y'' \ldots y^{(n-1)}]^T$, to the linear system

$$z' = Pz + du(t), \tag{2}$$

where

$$P = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & 1 \\ -k_n & -k_{n-1} & -k_{n-2} & \cdots & -k_1 \end{bmatrix} \tag{3}$$

is a companion matrix, the k_i are as in (1), and

$$d = [0\ 0 \ldots 1]^T \tag{4}$$

is the n-th standard basis vector.

What about the converse? When can a constant coefficient linear system

$$x' = Ax + bu(t), \tag{5}$$

where A is $n \times n$ and b is $n \times 1$, be transformed to (2) by a nonsingular linear transformation of the *state variable*, $z = Tx$, where T is a constant matrix? Since $z' = Tx' = (TAT^{-1})(Tx) + Tbu = (TAT^{-1})z + (Tb)u$, we are led to ask: When is there a nonsingular T such that TAT^{-1} is a companion matrix and Tb is the n-th standard basis vector?

The answer to this question is known [**8**, Chapter 2], although it seems not to be common knowledge outside the mathematical control community. A linear transformation of the state x that transforms (5) to (2) is not always possible, as can be seen by considering the diagonal system

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}. \tag{6}$$

In advanced courses in dynamics the subject of normal forms achieved by coordinate transformations is an important topic. And in the literature of mathematical control theory, questions concerning alternative system representations have always been important. However, we know of no elementary differential equations text outside the control-theoretic literature that systematically addresses the question of a transformation from (5) to (2).

The primary purpose of this article is to introduce a circle of ideas in mathematical control theory. The approach is via the question of "equivalence" between n-dimensional first order linear systems like (5) and n-th order linear equations like (1). The full answer to the equivalence question introduces some of the central concepts of modern control theory. We derive some classical results concerning the important control-theoretic concepts of *controllability* and *observability*. We also consider the relationships of these concepts with other important topics in control, such as stabilization of equilibria, and linearization of nonlinear systems using coordinate change and state feedback.

- In Sections 2 and 3 we clarify the relationship between the system (5) and the equation (1) and derive a necessary and sufficient condition for equivalence.
- In Section 4 we explore the equivalence condition of Section 3 by motivating and explaining its meaning as a *controllability* condition. We then rephrase our original equivalence problem and introduce the concept of *observability*. An easy step in Section 5 then shows the algebraic duality of controllability and observability.
- In Section 6 we indicate briefly the importance of these developments to questions of asymptotic behavior such as stability.
- Finally, Section 7 briefly discusses some extensions of Sections 4–6 to the case of linear systems with multivariable input and multivariable output.

2. A SIMPLE EXAMPLE. Let us begin with a naive approach to transforming a simple example and then consider a precise definition of *linear equivalence* of systems.

Example 1. Consider the system,

$$x_1' = -2x_1 + 2x_2 + u(t) \tag{7a}$$

$$x_2' = x_1 - x_2, \tag{7b}$$

which has the form (5) with

$$A = \begin{bmatrix} -2 & 2 \\ 1 & -1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Differentiate (7b) and substitute from (7a) to obtain $x_2'' + 3x_2' = u(t)$. This second order equation for x_2 has the form (1) for $n = 2$, and a solution of it for $x_2(t)$ determines a function $x_1(t)$ (using $x_1 = x_2' + x_2$ from (7b)) so that system (7) is solved. Thus, system (7) can reasonably be said to be equivalent to the second order equation $y'' + 3y' = u(t)$.

Is there another second order equation of the form $y'' + k_1 y' + k_0 y = u(t)$ that is also equivalent to (7)? For example, we might try to get a second order equation for x_1. This question is handled using a precise definition of equivalence. Note that the equation $y'' + 3y' = u(t)$ has the linear system form

$$z' = \begin{bmatrix} 0 & 1 \\ 0 & -3 \end{bmatrix} z + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t). \tag{8}$$

We expect that there is a transformation from R^2 to itself that transforms our original system (7) to the form (8). Since the differential equations are linear, we expect that the transformation is linear, say $z = Tx$. Differentiation then gives: $z' = TAT^{-1}z + Tbu(t)$.

Definition 1. The system $x' = Ax + bu(t)$ is *linearly equivalent* to the system $z' = Ez + fu(t)$ if there exists a nonsingular matrix T such that

$$TAT^{-1} = E, \quad Tb = f. \tag{9}$$

Thus, (5) is equivalent to (2) if and only if there is a nonsingular T such that $TAT^{-1} = P$ and $Tb = d$, where P and d are given in (3) and (4).

In Example 1 we obtained a second order equation for the variable x_2. If we set $z_1 = x_2$, $z_2 = x_2'$, then $z_2 = x_2' = x_1 - x_2$, so a transformation demonstrating the equivalence of (7) and (8) is given by

$$T = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}.$$

The natural questions concerning existence, uniqueness, and computation of T arise. Before proceeding to answer these questions it is instructive to try to transform the following example to the form (2).

Example 2. Let A be as in Example 1, but let $b = [1\ \ 1]^T$. Show that this system cannot be transformed to the form (2). Hint: In Example 1 we knew P once we had the second order equation, but, in fact, we know P anyway because by similarity we know P's characteristic polynomial.

3. EQUIVALENCE AND THE COMPANION MATRIX P. System (2) is very special; we call it a *companion system* because P is a *companion matrix* defined by the characteristic polynomial $\lambda^n + k_1\lambda^{n-1} + k_2\lambda^{n-2} + \cdots + k_{n-1}\lambda + k_n$, which is the same as the characteristic polynomial of any matrix that is similar to P. Example 1 shows that system (5) may be equivalent to a companion system, and we have seen two examples of (5) that are not equivalent to a companion system, namely, Example 2 and a two-dimensional system with diagonal A having a repeated eigenvalue.

3.1 A Similarity Invariant. It is convenient to make the following definition.

Definition 2. The vector x is a *cyclic vector* for the square matrix A if the n vectors $x, Ax, \ldots, A^{n-1}x$ are linearly independent.

In (2), d is a cyclic vector for P; one way to see this is by direct calculation of

$$[dPd\ P^2d\ \cdots\ P^{n-1}d] = \begin{bmatrix} 0 & \cdot & \cdot & \cdot & \cdot & 0 & 1 \\ 0 & \cdot & \cdot & \cdot & 0 & 1 & * \\ 0 & \cdot & \cdot & 0 & 1 & * & * \\ & & & \vdots & & & \\ 0 & 1 & * & * & * & * & * \\ 1 & * & * & * & * & * & * \end{bmatrix},$$

which is nonsingular. Existence of a cyclic vector for a matrix is a similarity invariant. If A is similar to P and $TAT^{-1} = P$, then A has a cyclic vector given by $T^{-1}d$.

The next proposition gives a useful condition that is equivalent to similarity between A and P.

Proposition 1. [**6**, Theorem 3.3.15] *The matrix A is similar to the companion matrix P of its characteristic polynomial if and only if the minimal and characteristic polynomials of A are identical.*

Proof: Similar matrices have the same characteristic polynomial and minimal polynomial, and the minimal polynomial of a companion matrix P is the same as its characteristic polynomial [**6**, pp. 146–147]. Thus, if A is similar to P, then the minimal polynomial and the characteristic polynomial of A must be identical.

On the other hand, if the minimal polynomial and characteristic polynomial of A are identical, then the Jordan canonical form of A must contain exactly one Jordan block for each distinct eigenvalue; the size of each Jordan block is equal to the multiplicity of the corresponding eigenvalue as a zero of the characteristic (minimal) polynomial of A. In this case, the Jordan canonical form of the companion matrix P has the same Jordan block structure as A, and hence it must be similar to A. Thus, A must be similar to P. ∎

Proposition 1 makes it easy to construct examples of matrices that have (or do not have) cyclic vectors. The similarity condition holds in Examples 1 and 2, where the characteristic (and minimal) polynomial for A is $\lambda(\lambda + 3)$. We conclude that there is some other obstruction in Example 2 to an equivalence with system (2), and the obstruction must involve the b vector. Thus, the problem with transforming Example 2 is related to the way the forcing function u enters the equations. We pursue this observation in Section 4.

3.2 Uniqueness of the Transformation T. Assume that we have a nonsingular T such that $TAT^{-1} = P$ and $Tb = d$. Then $TAT^{-1}d = TAb$, and $TA^kb = TA^kT^{-1}d = (TAT^{-1})^kd = P^kd$ for all $k \geq 0$. Nonsingularity of T implies that

$$n = rank\left[d \ Pd \ P^2d \ \cdots \ P^{n-1}d\right] = rank\left[b \ Ab \ A^2b \ \cdots \ A^{n-1}b\right].$$

Moreover, T is uniquely determined by its action on the basis defined by the vectors $b, Ab, \dots, A^{n-1}b$. Thus, we have the following uniqueness result and necessary condition.

Proposition 2. *There is at most one nonsingular linear transformation*, $z = Tx$, *taking* (5) *to the companion form* (2). *Such a* T *exists only if*

$$rank\left[b \ Ab \dots A^{n-1}b\right] = n. \tag{10}$$

Example 2 is explained by this result, because in that example we have

$$\left[b \ Ab\right] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

We return to Example 2 later for additional insight. Proposition 2 also explains why we cannot get a second order equation (1) for the variable x_1 in Example 1: the second order equation for $y = z_1$ must be a *unique* linear combination of the components of x.

We now show that the criterion (10) is sufficient for there to be a nonsingular linear transformation T from (5) to (2). We also show how to construct T by a simple direct method.

3.3 The Rank Condition (10) is Sufficient for Equivalence. Referring back to Example 1, the key in transforming (5) to (2) is to identify the variable z_1 that satisfies an equivalent n-th order equation (1). Note that $z_1 =$ (first row of T) $\cdot x$. Let τ denote the first row of T. Since $Tb = d = [0 \ 0 \dots 0 \ 1]^T$ and $TA^kb = P^kd$, we must have $\tau \cdot b = 0$, $\tau \cdot Ab = 0, \dots, \tau \cdot A^{n-2}b = 0$, and $\tau \cdot A^{n-1}b = 1$. Write this as

$$\tau\left[b \ Ab \dots A^{n-1}b\right] = [0 \dots 0 \ 1] = d^T. \tag{11}$$

Now if we assume $rank[b \ Ab \dots A^{n-1}b] = n$, there is a unique solution for τ in (11). (Again, the crucial z_1 variable must be a *unique* linear combination of the x components.) What about the rest of T? The form of the companion system requires that

$$z_2 = (\text{second row of } T)x = z_1' = \tau \cdot x' = \tau \cdot (Ax + bu) = \tau Ax + \tau bu = \tau Ax;$$

therefore the second row of T is τA. Continuing in this way, the equations defining τ and the form of the z system imply that

$$T = \begin{bmatrix} \tau \\ \tau A \\ \tau A^2 \\ \vdots \\ \tau A^{n-1} \end{bmatrix}. \tag{12}$$

We combine this argument with Proposition 2 as follows.

Theorem 1. *The system $x' = Ax + bu(t)$ with $x \in R^n$ can be transformed to the companion system, $z' = Pz + du(t)$, by a nonsingular linear transformation, $z = Tx$, if and only if $rank[b \ Ab \dots A^{n-1}b] = n$; in this case, T is uniquely defined by* (12), *where τ is the unique solution of* (11).

Theorem 1 answers our original question. If the basic algebraic fact concerning the existence of a cyclic vector for the companion matrix of A is understood, then the situation regarding equivalence between (5) and (2) becomes transparent.

Our original question got us to this point. But there is much more involved here, if we re-examine things. Think about varying the nonhomogeneous term in (5). What if we apply different input functions $u(t)$? To what extent can this affect the solutions of the system?

We consider the question of varying the input $u(t)$ in the next section. By doing so, we obtain an analytic, control-theoretic meaning of the rank condition in Theorem 1.

4. CONTROLLABILITY. System (5) is often called a *single-input* system because the input function u is scalar-valued rather than vector-valued. We show in this section that a natural concept of *controllability* for the single-input system (5) coincides with b being a cyclic vector for A.

In an elementary differential equations course the nonhomogeneous term in (1) is considered to be a fixed, specified function of t. But we now ask: What happens with the system dynamics as we change u? More specifically, to what extent can the motion of the state vector $x(t)$ be influenced, starting from an initial state x_0 and using fairly arbitrary inputs, $u(t)$? The next definition describes a concept of complete controllability of the state. Before stating this definition, we should specify a set $\mathscr{U}$ of admissible input functions. The solutions of linear constant coefficient systems of differential equations are defined on the entire real line, and generally we want the same property for the inputs. However, for some questions, the inputs are restricted to an interval $[t_0, t_f]$. Thus, with an appropriate restriction of domain when necessary, we could consider several vector spaces of functions for the set $\mathscr{U}$, including piecewise constant, continuous, or locally integrable inputs. A real-valued function $u(t)$ is locally integrable if

$$\int_{t_1}^{t_2} |u(s)| \, ds < \infty$$

for each $t_1 < t_2$. The set of locally integrable functions is the largest vector space of inputs for which (13) makes sense; therefore we assume our inputs are locally integrable.

Definition 3. The linear system (5) is *completely controllable* if, given any x_0, $x_f \in R^n$, there exists a $t_f > 0$ and a control function $u(t)$, defined for $0 \le t \le t_f$, such that the solution to (5) with initial condition $x(0) = x_0$ satisfies $x(t_f) = x_f$.

The solution of (5) with $x(0) = x_0$ is

$$x(t) = e^{tA}\left(x_0 + \int_0^t e^{-sA} bu(s)\, ds\right), \tag{13}$$

where

$$e^{tA} = I + tA + \frac{t^2}{2!}A^2 + \cdots + \frac{t^k}{k!}A^k + \cdots.$$

By the Weierstrass M-test, this series is absolutely and uniformly convergent for $|t| \le t_f$ for any finite t_f [**10**, pp. 134–135]. The linear system (5) is completely controllable if for any given x_0, x_f there is some t_f and some locally integrable function u on $0 \le t \le t_f$ such that

$$x_f = e^{t_f A}\left(x_0 + \int_0^{t_f} e^{-sA} bu(s)\, ds\right). \tag{14}$$

It may be surprising that the solvability of (14) for arbitrary x_0, x_f is determined by a purely algebraic criterion; the explanation lies with the Cayley-Hamilton Theorem: the matrix A satisfies $p(A) = 0$, where $p(\lambda)$ is the characteristic polynomial of A. The rank condition (10) is known as the *controllability rank condition*, and the matrix $[b\ Ab \dots\ A^{n-1}b]$ is called the *controllability matrix*, because of the next theorem.

Theorem 2. *The linear system $x' = Ax + bu(t)$ in* (5) *is completely controllable if and only if rank* $[b\ Ab \dots A^{n-1}b] = n$.

Proof: By the Cayley-Hamilton theorem, for each $k \ge n$, A^k can be expressed as a linear combination of the powers $A, A^2, \dots, A^{n-1}$. Let $\mathscr{R}$ denote the column space (range) of $[b\ Ab \dots A^{n-1}b]$. From the definition of the matrix exponential and the fact that $\mathscr{R}$ is a closed subspace of R^n, we can conclude that the range of $e^{-sA}b$ must lie in $\mathscr{R}$ for every s. Thus, the integral on the right side of (13) must lie in $\mathscr{R}$ for all t. Take $x_0 = 0$, so the states that are reachable from the origin in finite time, by means of some input $u(t)$, must all lie within $\mathscr{R}$. Thus, if the rank condition does *not* hold, then the system is *not* completely controllable because there are states that cannot be reached from x_0. This establishes the implication: complete controllability $\Rightarrow$ *rank* $[b\ Ab \dots A^{n-1}b] = n$.

Conversely, suppose *rank* $[b\ Ab \dots A^{n-1}b] = n$. We must now show that (5) is completely controllable. Choose any finite time $t_f > 0$, and consider the symmetric $n \times n$ matrix

$$M = \int_0^{t_f} e^{-sA} bb^T e^{-sA^T}\, ds.$$

We first show that M is nonsingular, and then we show that nonsingularity of M implies complete controllability. So suppose that $Mv = 0$ for some v; then also

$v^T M v = 0$, and this implies that

$$0 = v^T M v = \int_0^{t_f} v^T e^{-sA} b b^T e^{-sA^T} v\, ds = \int_0^{t_f} (\psi(s))^2\, ds,$$

where $\psi(s) = v^T e^{-sA} b$. Since $(\psi(s))^2$ is continuous and nonnegative, we conclude that $\psi(s) \equiv 0$. It follows that

$$\psi(0) = v^T b = 0, \quad \psi'(0) = -v^T A b = 0, \dots, \psi^{(n-1)}(0) = \pm v^T A^{n-1} b = 0.$$

Therefore v is perpendicular to $\mathscr{R}$. By the rank assumption, we must have $v = 0$, and therefore M is nonsingular. Now take any two points x_0, x_f in R^n, and define the control $u(s) = b^T e^{-sA^T} x$ for $0 \le s \le t_f$, where x remains to be chosen. The solution $x(t)$ with input u and initial condition x_0 has final point x_f at time t_f provided that x can be chosen so that

$$x_f = e^{t_f A}\left(x_0 + \left(\int_0^{t_f} e^{-sA} b b^T e^{-sA^T}\, ds \right) x \right) = e^{t_f A}(x_0 + Mx).$$

But $e^{t_f A}$ is nonsingular because $(e^{t_f A})^{-1} = e^{-t_f A}$, and M is nonsingular, so $x = M^{-1}(e^{-t_f A} x_f - x_0)$. Thus, any x_0 can be steered to any x_f in time t_f, so the system is completely controllable. ∎

Our proof that the controllability rank condition is sufficient for complete controllability follows an argument in [**9**, pp. 167–168]

Let us illustrate both Theorem 2 and the idea of controllability by re-examining Example 2.

Example 3. (*Example* 2 *continued*) The system is

$$x' = \begin{bmatrix} -2 & 2 \\ 1 & -1 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u.$$

Note that $\lambda = 0$ is an eigenvalue of A, and $b = [1\ 1]^T$ is a corresponding eigenvector, so the controllability rank condition does not hold. However, A is similar to its companion matrix. Using the T computed before and $z = Tx$ we have the system

$$z' = \begin{bmatrix} 0 & 1 \\ 0 & -3 \end{bmatrix} z + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u.$$

Differentiation of the z_1 equation and substitution produces a second order equation for z_1:

$$z_1'' + 3z_1' = 3u + u',$$

which does not match (1) due to the u' term. One integration produces a first order equation

$$z_1' + 3z_1 = 3\int u\, ds + u,$$

which shows that the action of arbitrary inputs u affects the dynamics in only a one-dimensional space. The original x equations might lead us to think that u can fully affect both x_1 and x_2, but notice that the z_2 equation says that u has no affect on the dynamics of the difference $x_1 - x_2 = z_2$. Only when the initial condition for z involves $z_2(0) = 0$ can u be used to control a trajectory. That is, the inputs completely control only the states that lie in the subspace $span[b\ Ab] = span\{b\} = span[1\ 1]^T$. Solutions starting with $x_1(0) = x_2(0)$ satisfy $x_1(t) = x_2(t) = \int_0^t u\, ds + x_1(0)$. One can steer along the line $x_1 = x_2$ from any initial point to any final point $x_1(t_f) = x_2(t_f)$ at any finite time t_f by appropriate choice of $u(t)$. On the other hand, if the initial condition lies off the line $x_1 = x_2$, then the difference $z_2 = x_1 - x_2$ decays exponentially so there is no chance of steering to an arbitrarily given final state in finite time.

When a system is completely controllable, there are generally many input functions that can implement the transfer from x_0 to x_f. This flexibility can be exploited in some applications to optimize the behavior of the system in some way, for example by minimizing a measure of the cost of carrying out the transfer. In particular, if the cost of the control action is measured by the integral

$$\int_{t_0}^{t_f} |u(s)|^2 \, ds,$$

then a control that minimizes this cost can be determined. For additional details on this problem for linear systems, see [**1**, pp. 102–105]. *Optimal control theory* is concerned with optimizing various performance indices of systems such as (5).

5. OBSERVABILITY AND DUALITY. Suppose we have a system (5) for which a certain linear combination of the state components x_i is directly measured, perhaps by some combination of instruments. We write the system and its measured output as

$$x' = Ax + bu \tag{15a}$$

$$y = c^T x, \tag{15b}$$

where c is a constant vector. The function $y(t)$ is our known output.

We now ask, when is c^T the first row of a transformation T to the companion system (2) where y is the dependent variable in (1)? If such a T exists, *then T must have the form* (12) *with $\tau = c^T$*. We must also have $Tb = d = [0 \ldots 0 \;\; 1]^T$. Thus,

$$\mathit{rank}\, T = \mathit{rank} \begin{bmatrix} c^T \\ c^T A \\ \vdots \\ c^T A^{n-1} \end{bmatrix} = n. \tag{16}$$

In addition, since $Tb = d$, we have

$$\begin{bmatrix} c^T \\ c^T A \\ \vdots \\ c^T A^{n-1} \end{bmatrix} b = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} b^T \\ b^T A^T \\ \vdots \\ b^T (A^{n-1})^T \end{bmatrix} c. \tag{17}$$

You can check that the differential equation for $z = [y \; y' \ldots y^{(n-1)}]^T$ really is the companion form (2), by remembering that A satisfies its own characteristic polynomial: $A^n + k_1 A^{n-1} + \cdots + k_{n-1} A + k_n I = 0$.

Proposition 3. *There exists a nonsingular T transforming* (15a) *to companion form* (2) *with $z_1 = y = c^T x$, if and only if the rank condition* (16) *holds and* (17) *is satisfied. In this case, T is uniquely determined and is the matrix in* (16).

Note that the matrix in (16) has the same rank as the matrix $[c \, A^T c \ldots (A^T)^{n-1} c]$, so that $y = c^T x$ satisfies (1) if and only if the system

$$x' = A^T x + cu \tag{18}$$

is completely controllable, by Theorem 2. Moreover, (17) shows that b^T is the first row of the transformation that takes (18) to companion form.

The connection of Proposition 3 with the system (18) leads to a fundamental duality between complete controllability and the concept of *complete observability.* The rank condition (16) is known as the *observability rank condition* and the matrix

$$\begin{bmatrix} c^T \\ c^TA \\ \vdots \\ c^TA^{n-1} \end{bmatrix} \tag{19}$$

is called the *observability matrix* for the system (15). The rank condition implies that the system state x can be reconstructed from knowledge of y, u, and their derivatives. Here is a basic definition.

Definition 4. The system (15) is *completely observable* if, for any $x_0 = x(0)$, there is a finite time $t_f > 0$ such that knowledge of the input $u(t)$ and output $y(t)$ on $[0, t_f]$ suffices to determine x_0 uniquely.

Definition 4 could be restated using only the zero input, $u \equiv 0$. To see how a determination of x_0 is made when the observability rank condition holds, differentiate the output equation (15b) $n - 1$ times and set $t = 0$ to get

$$\begin{bmatrix} y(0) \\ y'(0) \\ \vdots \\ y^{(n-1)}(0) \end{bmatrix} = \begin{bmatrix} c^T \\ c^TA \\ \vdots \\ c^TA^{n-1} \end{bmatrix} x_0 + \text{terms dependent on } u. \tag{20}$$

Under the observability rank condition, the coefficient of x_0 is nonsingular, and we can solve for x_0 in terms of y and u and their derivatives. To illustrate, consider the system of Example 2 with output $y = [1\,0]x$. In this case, (20) gives

$$\begin{aligned} \begin{bmatrix} y(0) \\ y'(0) \end{bmatrix} &= \begin{bmatrix} c^T \\ c^TA \end{bmatrix} x_0 + \begin{bmatrix} 0 \\ c^Tb \end{bmatrix} u \\ &= \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} x_0 + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u. \end{aligned}$$

Therefore the system is completely observable.

Theorem 3. *The system* (15) *is completely observable if and only if the observability rank condition* (16) *holds.*

Proof: We have already shown the sufficiency of the rank condition (16). Now assume that complete observability holds; we must show that (16) holds. For the purpose of contradiction, suppose also that the observability matrix has deficient rank; then, there is a nonzero vector v such that

$$\begin{bmatrix} c^T \\ c^TA \\ \vdots \\ c^TA^{n-1} \end{bmatrix} v = 0. \tag{21}$$

Now take $x_0 = v$ and consider the output $y = c^T e^{tA} v$ using input $u \equiv 0$. Using (21) and the definition of the matrix exponential, the series expansion for y must have all coefficients equal to zero. Thus, $y \equiv 0$; but this is also the output when $x_0 = 0$ under zero input. This contradicts the complete observability assumption. ∎

Motivated by the comments following Proposition 3, we define the *dual system* of (15) to be

$$x' = A^T x + cu(t) \tag{22a}$$

$$y = b^T x. \tag{22b}$$

Then the dual of the dual of a system is the original system. With this definition we can encapsulate the discussion thus far with the following classical duality statement.

Theorem 4. *The system* (15) *is completely observable if and only if the system* (22a) *is completely controllable. The system* (15a) *is completely controllable if and only if the dual system* (22) *is completely observable.*

6. FEEDBACK, STABILIZATION, OBSERVERS, AND DUALITY. A major theme in control theory is the use of *feedback* to modify the system dynamics to achieve some desired behavior, for example to stabilize an otherwise unstable equilibrium point. In this section we indicate some advantages of an equivalence with the companion system (2) with regard to these issues. We also present one additional consequence of duality. The considerations in this section help to indicate that much can be accomplished with the control of linear systems, and thus it is desirable to have an extension of the solution of the equivalence problem involving systems (2) and (5) to the case where (5) is replaced by a single-input *nonlinear* system.

Definition 5. The linear system $x' = Ax$ is *stable* if all eigenvalues of A lie in the open left half-plane.

From the theory of linear differential equations, it is known that all solutions $x(t) \to 0$ as $t \to \infty$ if all the eigenvalues of A have negative real part. In this case, the equilibrium at the origin is asymptotically stable.

Definition 6. In system (5), *linear state feedback* is specified by $u = Kx$ where K is a real $1 \times n$ matrix. The corresponding *closed loop system* is $x' = (A + bK)x$.

Consider the companion form (2). Using linear state feedback, $u = Kx$, it is possible to assign eigenvalues arbitrarily to the resulting closed loop system, provided that the complex eigenvalues of $A + bK$ occur in complex conjugate pairs. Specifically, by setting $u = Kx = [-\alpha_n - \alpha_{n-1} \dots -\alpha_1]x$ in (2) we get the closed loop system $z' = \tilde{P}z$, where $\tilde{P}$ has the same form as P in (3) except the last row is now $[-(k_n + \alpha_n) - (k_{n-1} + \alpha_{n-1}) \dots -(k_1 + \alpha_1)]$. Suppose that $m_1, m_2, \dots, m_n$ are the desired coefficients of the characteristic polynomial of the closed loop, $z' = \tilde{P}z$. With the k_i known and the m_i specified, then $\alpha_i = m_i - k_i$. Thus, the coefficients of the characteristic polynomial of $A + bK$ may be chosen so that all its roots lie in the open left half plane. And the exponential rate of convergence of $z(t)$ to the origin can be increased, for example, by shifting the roots to the left in the complex plane.

A system that is not stable might be made stable if modified by appropriate linear feedback.

Definition 7. The linear system $x' = Ax + bu(t)$ is *stabilizable* if there exists a $1 \times n$ matrix K such that the linear system $x' = (A + bK)x$ is stable.

Theorem 5. *If $x' = Ax + bu(t)$ is completely controllable then it is stabilizable and the eigenvalues of $x' = (A + bK)x$ can be assigned arbitrarily (provided that complex eigenvalues occur in conjugate pairs) by appropriate choice of K.*

Proof: We have discussed the proof only for the special case of a companion system. By complete controllability, there is a nonsingular T with $z = Tx$ such that $z' = TAT^{-1}z + Tbu$ is a companion system. Therefore the eigenvalues of $TAT^{-1} + Tb\tilde{K}$ can be assigned as described, where $u = \tilde{K}z$ represents linear feedback for the companion system. Now the similarity

$$TAT^{-1} + Tb\tilde{K} = T(A + b\tilde{K}T)T^{-1} = T(A + bK)T^{-1}; \quad K \equiv \tilde{K}T$$

shows that the eigenvalues of $A + bK$ can be assigned by appropriate choice of feedback $u = Kx$. ■

There is a concept dual to stabilizability that involves the state-to-output interaction of system (15). We give only a very brief discussion.

Definition 8. System (15) is *detectable* if there exists an $n \times 1$ matrix L such that the system $x' = (A + Lc^T)x$ is stable.

Forming the matrix $A + Lc^T$ corresponds to output feedback given by $u = Ly = Lc^Tx$. The eigenvalues for $A + Lc^T$ are the same as those for $A^T + cL^T$, which corresponds to state feedback $u = L^T x$ in the dual system (22a). Thus *a system is detectable if and only if the dual system is stabilizable*. These are purely algebraic statements. An analytic interpretation of detectability derives from its implication that linear output feedback can be used to "detect" system trajectories asymptotically through a construction known as an *observer* system. Specifically, consider the system

$$\xi' = A\xi + Bu - L(y - c^T\xi) \tag{23}$$

where ξ is an auxiliary state that can be initialized at any vector $\xi(0)$. The auxiliary state ξ is intended to approximate the true state x, and L, a so-called "output error" gain matrix, is to be chosen so that ξ approximates x. Define the error by

$$e = x - \xi.$$

The objective is to choose L so that $e \to 0$ as $t \to \infty$. Now, subtraction of (23) from (15a) gives

$$e' = (A + Lc^T)e.$$

Theorem 6. *If the system* (15) *is detectable then L can be chosen in system* (23) *so that $e = x - \xi \to 0$ as $t \to \infty$, independently of the initial condition* $\xi(0)$.

Some comments on this construction are in order. The observer system (23) is an alternative to computing the solutions of the system (15) with a direct numerical method. By using the known data provided by A, b, and c, together with y and u, system (23) can be simulated numerically with a guarantee that the estimated state *asymptotically* reconstructs the true system state x for (15), *independently of the initial* $\xi(0)$. If we were so lucky as to have $\xi(0) = x(0)$, then the observer equation (23) implies that $\xi(t) = x(t)$ for all t, a perfect estimate. You can think of (23) as a system with inputs u and y, and with output ξ, the desired approximation. The estimate ξ itself can be fed back to (15a), via $u = K\xi$, in place of the true state for purposes of stabilization of (15a), provided that (15a) is stabilizable. In other words, the eigenvalues of the closed loop system can be placed somewhere within the left half-plane even though only the output y is measured. Moreover, an important feature of this construction is that the controller (that is, the matrix K) and the observer (essentially the matrix L) can be designed independently while ensuring that the overall, interconnected observer/controller system is stable. To see this, use (15a) together with (23) to write the combined system for (x, ξ) as

$$\begin{bmatrix} x' \\ \xi' \end{bmatrix} = \begin{bmatrix} A & bK \\ -Lc^T & A + Lc^T + bK \end{bmatrix} \begin{bmatrix} x \\ \xi \end{bmatrix}. \tag{24}$$

We can obtain the characteristic polynomial for this system by using the following similarity transformation:

$$\begin{bmatrix} I & 0 \\ -I & I \end{bmatrix} \begin{bmatrix} A & bK \\ -Lc^T & A + Lc^T + bK \end{bmatrix} \begin{bmatrix} I & 0 \\ I & I \end{bmatrix} = \begin{bmatrix} A + bK & bK \\ 0 & A + Lc^T \end{bmatrix}.$$

Thus, the characteristic polynomial of the coefficient matrix in (24) is the product of the characteristic polynomials of $A + bK$ and $A + Lc^T$. This means that K can be designed without regard to the fact that only state estimates will be fed back, and the observer error gain L can be designed without reference to the fact that the resulting state estimates are fed back for stabilization purposes. This independent design feature is often called the *Separation Principle*.

Let us consider two examples illustrating stabilizability and detectability.

Example 4. We return to Example 2 once more, and adjoin the output equation $y = x_1$. Then the system coefficients are

$$A = \begin{bmatrix} -2 & 2 \\ 1 & -1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad c^T = [1 \quad 0].$$

This system is both stabilizable and detectable, using the feedback matrix K and observer matrix L given by

$$K = [-2 \quad 0], \quad L = \begin{bmatrix} 0 \\ -1 \end{bmatrix},$$

because $A + bK$ then has eigenvalues -2, -3, and $A + Lc^T$ has eigenvalues -2, -1. Other choices for K and L are also possible.

Example 5. Stabilizability and detectability are not guaranteed. Consider the linear system with coefficients

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad c^T = [0 \quad 1].$$

In this case, any 1×2 feedback matrix K produces a closed loop matrix $A + bK$ with zero as an eigenvalue; therefore, the system is not stabilizable. Also, any 2×1 matrix L yields a matrix $A + Lc^T$ with zero as an eigenvalue; therefore the system is not detectable.

7. A BRIEF NOTE ON EXTENSIONS. Let us briefly describe an extension of our discussion of the single-input systems in Sections 4–6 to the case of linear systems with multivariable input and output,

$$x' = Ax + Bu(t) \tag{25a}$$

$$y = Cx, \tag{25b}$$

where $u \in R^m$, $y \in R^p$, and thus B is $n \times m$ and C is $p \times n$. As noted before, the rank tests for controllability and observability allow for a statement of algebraic duality between these concepts, once an appropriate dual system is identified. The same principles extend to (25).

Definition 3 (complete controllability) makes sense for the m-input case; the admissible control functions are R^m-valued functions $u(t)$ such that every entry of $u(t)$ is locally integrable.

The characterization of complete controllability in Theorem 2 directly carries over to (25a) with no change in the statement. In this case, the controllability matrix $[B \; AB \dots A^{n-1}B]$ has size $n \times nm$, and the proof proceeds as before from (13). With careful attention to the dimensions involved, the same proof carries through; the M matrix is still $n \times \; n$ while ψ is $1 \times m$.

Complete observability of the system (25) is defined exactly as in Definition 4, and the system (25) is completely observable if and only if the observability matrix,

$$\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix},$$

which is now $pn \times n$, has rank n. One checks that the proof of Theorem 3 carries through as before.

The dual system of (25) is defined by

$$x' = A^T x + C^T u(t) \tag{26a}$$

$$y = B^T x, \tag{26b}$$

with matrix dimensions determined, of course, by (25). Theorem 4, which documents the algebraic duality of complete observability and complete controllability, is also valid when applied to (25) and its dual system.

The extension of Theorem 5 to the case of m-input controllable systems can be based on the single-input result: see [**13**, pp. 49 – 51] for an accessible proof that essentially reduces the m-input case to the single-input case.

Definition 7 and Definition 8 have straightforward extensions to the m-input and p-output cases. Theorem 6 on observer construction is easily seen to extend to (25); the extension is essentially notational.

Once we move to linear systems with time-dependent coefficient matrices, additional technical issues arise in any extension of observability, controllability, and their duality, although several extensions have been accomplished. For example, several useful definitions of controllability for time-varying systems are possible, all of which coalesce in the linear constant coefficient case to describe the same concept. These definitions may involve the initial time t_0, the particular initial state x_0 considered, and the time interval over which control action is to take place. The reader interested in these issues is invited to explore the references. However, let us give one further example to illustrate that time-varying systems require alternative approaches in order to describe controllability properties.

Example 6. [**11**] Consider the system

$$x' = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} x + \begin{bmatrix} b_1(t) \\ b_2(t) \end{bmatrix} u.$$

The general solution is

$$x(t) = \begin{bmatrix} x_1(t) \\ \\ x_2(t) \end{bmatrix} = \begin{bmatrix} e^t\left(x_1(0) + \int_0^t e^{-s} b_1(s) u(s)\, ds\right) \\ \\ e^{2t}\left(x_2(0) + \int_0^t e^{-2s} b_2(s) u(s)\, ds\right) \end{bmatrix}.$$

If b_1 and b_2 are constant, then Theorem 2 ensures that the system is completely controllable. Suppose now that $b_1(t) = e^t$ and $b_2(t) = e^{2t}$; the solution is then

$$x(t) = \begin{bmatrix} x_1(t) \\ \\ x_2(t) \end{bmatrix} = \begin{bmatrix} e^t\left(x_1(0) + \int_0^t u(s)\, ds\right) \\ \\ e^{2t}\left(x_2(0) + \int_0^t u(s)\, ds\right) \end{bmatrix}.$$

Thus, solutions that start on the line $x_2 = x_1$ at $t = 0$ always satisfy the condition $x_2(t) = e^t x_1(t)$, and from this condition we can conclude that the system is not completely controllable according to Definition 3, for if $x_1(0) = x_2(0)$, then the motion is confined to the first or third quadrant since x_1 and x_2 must have the same sign. In particular, the set of points reachable from the origin lies within those two quadrants. If we consider the controllability rank condition in a pointwise manner, that is, if we consider the following matrix for each time instant t,

$$[B(t)\ AB(t)] = \begin{bmatrix} e^t & e^t \\ e^{2t} & 2e^{2t} \end{bmatrix},$$

we obtain a nonsingular matrix. This example shows that a pointwise interpretation of the controllability rank condition of Theorem 2 does not lead to a satisfactory criterion for complete controllability of a time-varying linear system.

8. FURTHER READING. Three major themes in control theory (and in this article) involve (i) the input-to-state interaction: *controllability*, (ii) the state-to-output interaction: *observability*, and (iii) transitions between different representations of a dynamical system. We have tried to illustrate those themes in a discussion of an equivalence problem for single-input linear systems.

Two comprehensive texts that focus on time-invariant linear systems are [**7**] and [**9**]. For more on multivariable input and output, and time-varying linear systems, see [**1**], [**2**], [**3**], [**11**], [**12**], and [**13**]. The linear algebra text [**4**] provides a mathematician's view of some fundamental results of control-theoretic interest. The presentation of linear control theory in [**13**] is nicely unified around the concept of invariant subspace. Additional perspective on linear systems theory from the mathematical point of view can be obtained from [**5**].

ACKNOWLEDGMENT. Thanks to Dr. Robert E. Terrell at Cornell University for a helpful reading of an early version of this article.

REFERENCES

1. S. Barnett and R. G. Cameron, *Introduction to Mathematical Control Theory*, Oxford Clarendon Press, Second Edition, 1985.
2. R. W. Brockett, *Finite Dimensional Linear Systems*, John Wiley and Sons, Inc., 1970.
3. C.-T. Chen, *Linear System Theory and Design*, Holt, Rinehart and Winston, 1984.
4. P. A. Fuhrmann, *A Polynomial Approach to Linear Algebra*, Springer-Verlag, 1996.
5. I. Gohberg, P. Lancaster, and L. Rodman, *Invariant Subspaces of Matrices with Applications*, Canadian Mathematical Society Monographs, John Wiley and Sons, 1986.
6. R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.
7. T. Kailath, *Linear Systems*, Prentice-Hall, Inc., 1980.
8. R. E. Kalman, P. L. Falb, and M. A. Arbib, *Topics in Mathematical Control Theory*, McGraw-Hill, 1969.
9. J. W. Polderman and J. C. Willems, *Introduction to Mathematical Systems Theory: A Behavioral Approach*, Texts in Applied Mathematics **26**, Springer, 1998.
10. W. Rudin, *Principles of Mathematical Analysis*, McGraw-Hill, Inc., Second Edition, 1964.
11. W. J. Rugh, *Linear System Theory*, Prentice-Hall, Inc., Second Edition, 1996.
12. E. D. Sontag, *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, Second Edition, Springer-Verlag, 1998.
13. W. M. Wonham, *Linear Multivariable Control*, Third Edition, Springer-Verlag, 1985.

WILLIAM J. TERRELL received his Ph.D. in Applied Mathematics from North Carolina State University in 1990. Since 1991 he has been a member of the Department of Mathematical Sciences at Virginia Commonwealth University, Richmond, Virginia. His research interests are in the area of differential equations with special focus on systems and control theory.
Department of Mathematical Sciences, Virginia, Commonwealth University, Richmond, VA 23284-2014
wterrell@atlas.vcu.edu