Linear Output-Reversible Systems

Dennis S. Bernstein[1]
Department of Aerospace Engineering
The University of Michigan
Ann Arbor, MI 48109-2140
dsbaero@umich.edu

Sanjay P. Bhat
Department of Aerospace Engineering
Indian Institute of Technology
Powai, Mumbai 400076, India
bhat@aero.iitb.ernet.in

1 Introduction

The arrow of time remains one of physics' most puzzling questions. The most widely accepted explanation depends on dimensionality and randomness, which give rise to a monotonically increasing quantity known as entropy. Additional candidate mechanisms that are responsible for the arrow of time are discussed in [1].

In the present paper we consider the free response of a linear system involving dynamics and an output map. A second system is an *output reversal* of the original system if it can produce a time-reversed image of every output of the original system. As a special case, a system is *output reversible* if it can produce a time-reversed image of every one of its free output responses.

Our main result is a spectral symmetry condition that provides a complete characterization of single-input, single-output, output-reversible systems. In particular, we show that a linear system is output-reversible if and only if its non-imaginary spectrum is symmetric with respect to the imaginary axis. As special cases, the class of output-reversible systems includes rigid body and Hamiltonian systems. This result suggests that stability and instability play a key role in the arrow of time, independently of dimensionality, nonlinearity, and sensitivity.

The present paper is directed toward the goal of placing thermodynamics on a system-theoretic foundation. For related work, see [2,3].

2 Output-Reversible Systems

We begin by considering the dynamical system

$$\dot{x}(t) = f(x(t)), \quad t \geq 0, \quad x(0) = x_0, \tag{2.1}$$

with output

$$y(t) = g(x(t)), \tag{2.2}$$

where $x(t) \in \mathbb{R}^n$, $y(t) \in \mathbb{R}$, and $f : \mathbb{R}^n \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}$ are continuous. We assume that solutions of (2.1) exist and are unique on all finite intervals $[0, T)$. For clarity we write the solution of (2.1) as $x(t, x_0)$ with the output given by $y(t) = y(t, x_0) = g(x(t, x_0))$.

Definition 2.1. The system (2.1), (2.2) is *output reversible* if, for all $x_0 \in \mathbb{R}^n$ and $t_1 > 0$, there exists $\hat{x}_0 \in \mathbb{R}^n$ such that

$$y(t, \hat{x}_0) = y(t_1 - t, x_0), \quad t \in [0, t_1]. \tag{2.3}$$

We wish to determine whether a given system (2.1), (2.2) is output reversible.

Next, we consider the linear system

$$\dot{x}(t) = Ax(t), \quad t \geq 0, \quad x(0) = x_0, \tag{2.4}$$

with output

$$y(t) = Cx(t), \tag{2.5}$$

where $A \in \mathbb{R}^{n \times n}$ and $C \in \mathbb{R}^{1 \times n}$. We assume that the pair (A, C) is observable. It then follows from Definition 2.1 that (2.4), (2.5) is output reversible if and only if, for all $x_0 \in \mathbb{R}^n$ and $t_1 > 0$, there exists $\hat{x}_0 \in \mathbb{R}^n$ such that

$$Ce^{At}\hat{x}_0 = Ce^{A(t_1 - t)}x_0, \quad t \in [0, t_1]. \tag{2.6}$$

Note that output reversibility is a basis-independent property.

Proposition 2.2. If (2.4), (2.5) is output reversible, then $\hat{x}_0$ in (2.6) is given uniquely by

$$\hat{x}_0 = \mathcal{O}^{-1} S \mathcal{O} e^{At_1} x_0, \tag{2.7}$$

where

$$\mathcal{O} \triangleq \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}, \quad S \triangleq \begin{bmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & -1 \\ & & & & \ddots \end{bmatrix}.$$

Next, we write the matrix exponential e^{At} as a polynomial in A of the form

$$e^{At} = \sum_{i=0}^{n-1} \phi_i(t) A^i. \tag{2.8}$$

The coefficients $\phi_0(t), \ldots, \phi_{n-1}(t)$ are real linear combinations of terms of the form $t^r \mathrm{Re}\, e^{\lambda t}$ and $t^r \mathrm{Im}\, e^{\lambda t}$, where λ is an eigenvalue of A and r is a nonnegative integer. Since (A, C) is observable, the matrix A is cyclic (nonderogatory) and thus its minimal polynomial coincides with its characteristic polynomial. (Recall that A

is cyclic if and only if A has exactly one Jordan block associated with each distinct eigenvalue.) Consequently, the coefficients $\phi_i(t)$ satisfying (2.8) are unique.
For the following result, define

$$\phi(t) \triangleq \begin{bmatrix} \phi_0(t) \\ \vdots \\ \phi_{n-1}(t) \end{bmatrix}. \tag{2.9}$$

Substituting (2.8) into (2.6) yields

$$\phi^{\mathrm{T}}(t)\mathcal{O}\hat{x}_0 = \phi^{\mathrm{T}}(-t)\mathcal{O}e^{At_1}x_0, \quad t \geq 0. \tag{2.10}$$

Note that (2.6) and (2.10) are equivalent.

Proposition 2.3. The linear system (2.4), (2.5) is output reversible if and only if

$$\phi(-t) = S\phi(t), \quad t \geq 0. \tag{2.11}$$

Proposition 2.3 shows that the output reversibility of (2.4), (2.5) is independent of C so long as (A, C) is observable.

Example 2.4. Let $A \in \mathbb{R}$ be scalar so that $\phi_0(t) = e^{At}$. Hence $\phi_0(-t) = e^{-2At}\phi_0(t)$ and thus (2.11) is satisfied if and only if $A = 0$. Hence, for $n = 1$, (2.4), (2.5) is output reversible if and only if $A = 0$.

Example 2.5. Let $A = \left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right]$ so that $e^{At} = \begin{bmatrix} e^t & 0 \\ 0 & e^{-t} \end{bmatrix} = \frac{1}{2}(e^t + e^{-t})I + \frac{1}{2}(e^t - e^{-t})A$. Hence $\phi(-t) = S\phi(t)$ so that (2.11) is satisfied. Therefore, (2.4), (2.5) is output reversible. Furthermore, it can be seen that if $A = \left[\begin{smallmatrix} a & 0 \\ 0 & b \end{smallmatrix}\right]$, where $a \neq b$, then (2.4), (2.5) is output reversible if and only if $b = -a$.

Example 2.6. Let $A = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$, which represents rigid body motion. Then $e^{At} = \left[\begin{smallmatrix} 1 & t \\ 0 & 1 \end{smallmatrix}\right] = I + tA$. Hence $\phi(t) = \left[\begin{smallmatrix} 1 & t \end{smallmatrix}\right]^{\mathrm{T}}$ so that $\phi(-t) = S\phi(t)$ and thus (2.11) is satisfied. Therefore, (2.4), (2.5) is output reversible. Next, let $C = \left[\begin{smallmatrix} 1 & 0 \end{smallmatrix}\right]$ so that $y(t)$ represents the particle's position. Then $\mathcal{O} = I$ and (2.7) implies that

$$\hat{x}_0 = \begin{bmatrix} t_1 x_{02} \\ -x_{02} \end{bmatrix}. \tag{2.12}$$

Hence the reversed output is a consequence of the state trajectory that arises from an initial position given by the endpoint position $t_1 x_{02}$ of the forward trajectory as well as an initial velocity $-x_{02}$ given by the sign-reversed endpoint velocity of the forward trajectory. This is, of course, exactly what we would intuitively expect.

Example 2.7. As an extension of the previous example, let $A \in \mathbb{R}^{n \times n}$ be nilpotent and cyclic so that rank $A = n - 1$. Hence $e^{At} = I + tA + \cdots + \frac{1}{(n-1)!}t^{n-1}A^{n-1}$ and thus $\phi(t) = \left[\begin{smallmatrix} 1 & t & \cdots & \frac{1}{(n-1)!}t^{n-1} \end{smallmatrix}\right]^{\mathrm{T}}$. Hence $\phi(-t) = S\phi(t)$ so that (2.4), (2.5) is output reversible.

The following result shows that a linear system is output reversible if and only if its spectrum is symmetric with respect to the imaginary axis.

Theorem 2.8. The system (2.4), (2.5) is output reversible if and only if $p(-s) = (-1)^n p(s)$.

The following observation is valid whether or not A is cyclic.

Proposition 2.9. Suppose that $p(-s) = (-1)^n p(s)$. If n is even, then p is even and the algebraic multiplicity of the zero eigenvalue of A is even. If n is odd, then p is odd and the algebraic multiplicity of the zero eigenvalue of A is odd.

The following results depend on the fact that A is cyclic.

Proposition 2.10. $p(-s) = (-1)^n p(s)$ if and only if A has the following property: if λ is an eigenvalue of A, then so is $-\lambda$, and λ and $-\lambda$ have the same algebraic multiplicity.

Note that Proposition 2.10 places no restriction on eigenvalues of A whose real part is zero. Since A is cyclic, the condition specified in Proposition 2.10 implies that A and $-A$ have the same similarity invariants. This observation yields the following result.

Proposition 2.11. $p(-s) = (-1)^n p(s)$ if and only if A and $-A$ are similar.

Since A and A^{T} are similar (whether or not A is cyclic), we have the following variation of the previous result.

Proposition 2.12. $p(-s) = (-1)^n p(s)$ if and only if A and $-A^{\mathrm{T}}$ are similar.

Recall that the matrix $A \in \mathbb{R}^{2r \times 2r}$ is Hamiltonian if $A = -J^{-1}A^{\mathrm{T}}J$, where $J = \left[\begin{smallmatrix} 0 & I_r \\ -I_r & 0 \end{smallmatrix}\right]$.

Corollary 2.13. Assume that n is even and A is Hamiltonian. Then (2.4), (2.5) is output reversible.

[1] B. Bernstein and T. Erber, "Reversibility, Irreversibility: Restorability, Non-Restorability," *J. Phys. A: Math. Gen.*, V. 32, 7581-7602, 1999.

[2] D. S. Bernstein and D. C. Hyland, "Compartmental Modeling and Second-Moment Analysis of State Space Systems," *SIAM J. Matrix Anal. Appl.*, V. 14, 880-901, 1993.

[3] D. S. Bernstein and S. P. Bhat, "Energy Equipartition and the Emergence of Damping in Lossless Systems," CDC, 2913-2918, 2002.