Proceedings of the 7th Mediterranean Conference on Control and Automation (MED99) Haifa, Israel - June 28-30, 1999

From Physical Realizations to Nonlinear Stability, Passivity and Optimality

Michael Margaliot and Gideon Langholz*
Department of Electrical Engineering - Systems
Tel Aviv University, Israel

Abstract

In this paper we consider the following problem: Given a dynamical system, described by a differential equation, determine an appropriate measure for the energy stored in the system. This problem is of great importance because it is related directly to fundamental concepts such as stability, passivity, and optimality.

First, we solve the problem for linear dynamical systems. We realize the linear differential equation using a linear electrical circuit, so that the energy stored in the system is just the energy stored in the circuit's components. This leads to an intuitive proof of the Routh-Hurwitz stability criterion. In addition, we use the energy balance in the circuit to derive passivity and optimality relations.

Then, we extend the results to a class of nonlinear systems by replacing the linear components in the circuit with more general nonlinear ones. We derive explicit storage functions for passivity analysis of these nonlinear systems and for the formulation and explicit solution of a novel nonlinear optimal control problem.

Keywords: Nonlinear systems, stability, passivity, optimal control.

1 Introduction

Energy is a fundamental concept in the study of dynamical systems. The energy stored in the system and its rate of change are intimately related to stability, passivity, and optimality.

Unfortunately, even if we are given a mathematical model of the system, it is generally very difficult to derive a meaningful energy function. Consequently, it is often very difficult to find a suitable Lyapunov function for stability analysis, or a suitable storage function for passivity analysis in the case of a system with inputs and outputs.

In this paper, we utilize the following age-old[1] idea–*given a mathematical model of the system, realize it as an electrical circuit and then define the system's energy as the energy stored in the circuit's components*–to derive some new results:

- An intuitive physically-based proof of the Routh-Hurwitz (RH) stability criterion;
- A simple *explicit* storage function for a class of passive nonlinear systems;

*Corresponding author: Prof. Gideon Langholz, Department of Electrical Engineering - Systems, Tel Aviv University, Tel Aviv 69978. Tel. 972-3-6408738, Fax. 972-3-6407221, Email: langholz@eng.tau.ac.il.

[1] See for example, (Hirsch & Smale, 1974, Ch. 10).

- An extension of H_2 optimal control to a large class of nonlinear systems. In particular, we show that if the solution of a certain Riccati equation is positive-definite and *diagonal*, then a large class of nonlinear optimal control problems has a simple *explicit* closed-form solution.

We begin with the relatively simple case of linear dynamical systems. The mathematical model of the system is a characteristic polynomial $D(s)$. We show how to realize $D(s)$ using an autonomous electrical circuit comprising of capacitors, inductors and resistors. If these components are all positive, then they are passive and the energy stored in the circuit can never increase. On the other hand, if there is a sign change among the components, then at least one component is gaining energy, and the total energy of the circuit increases. This sign-change condition immediately brings to mind the RH stability criterion. Indeed, we will see that our approach yields a very intuitive proof for this criterion.

Adding inputs and outputs to our autonomous linear electrical circuit, we use the energy stored in the circuit as a storage function for passivity analysis. Because passivity and optimality are fundamentally related (Sepulchre et al., 1997), this leads to a simple and intuitive derivation of the optimal H_2 infinite-time regulator.

Furthermore, since stability, passivity, and optimality are directly related to the energy stored in the electrical circuit, we can extend our results to the *nonlinear* case simply by replacing the linear components in the circuit with the more general nonlinear ones and modifying the energy function appropriately.

This leads to the derivation of explicit storage functions and passivity relations for a class of nonlinear systems. Also, based again on the relationship between passivity and optimality, we can formulate and *explicitly* solve a nonlinear H_2 optimal control for these nonlinear systems.

The rest of the paper is organized as follows. In Section 2 we examine linear dynamical systems by realizing them using a linear electrical circuit. We relate stability, passivity, and optimality of the system to the electrical circuit, leading to a simple proof of the RH stability criterion and to a simple derivation of the optimal H_2 regulator. In Section 3 we extend these results to a class of nonlinear systems and show how the expression for the stored energy must be modified. Using the energy we derive new results regarding stability, passivity, and optimality of these nonlinear systems. The final section concludes.

2 Linear Systems

In this section we show that a linear dynamical system, described by a characteristic polynomial $D(s)$, can be realized as a linear electrical circuit by using the information contained in the RH array of $D(s)$. Furthermore, we show how the energy stored in the circuit's components can be used to study stability, passivity, and optimality. As a by-product we obtain an intuitive proof of the RH stability criterion.

2.1 Realization and the Routh-Hurwitz array

Consider a linear electrical circuit consisting of $n+1$ linear components (capacitors, inductors, and resistors). Each linear component is defined by a voltage-current relation and a parameter (namely, capacitance, inductance, and resistance). Denote these parameters by $\mathbf{K} = (K_1, K_2, ..., K_{n+1})^T$ and define the state-variables x_i, $i = 1, 2, ...$, as the voltages on the capacitors and the currents through the inductors. Using Kirchhoff's laws we can derive the circuit's state-space description $\dot{\mathbf{x}} = A\mathbf{x}$, and the characteristic polynomial $P(s) = det(sI - A) = p_n s^n + p_{n-1}s^{n-1} + ... + p_0$.

This process can be viewed as a transformation: $T : (K_1, K_2, ..., K_{n+1}) \to (p_0, p_1, ..., p_n)$. Hence, if we can find the inverse transformation T^{-1}, then, given the coefficients of a characteristic polynomial, we would be able to determine the parameters of the components in the linear circuit that implements it. The purpose of this section is to show that this inverse transformation is nothing but (a slightly modified version of) the RH array.

We define a modified version of the RH array as follows:

- Given a characteristic polynomial $D(s)$, calculate the RH array as usual.
- Divide each element in the first column of the RH array by the element right above it (the first element is left unchanged).

For example, the conventional RH array for $D(s) = s^3 + 3s^2 + 16s + 30$ is:

$$\begin{array}{ll} 1 & 16 \\ 3 & 30 \\ 6 & 0 \\ 30 & \end{array}$$

and the modified one is:

$$\begin{array}{ll} 1 & 16 \\ 3 & 30 \\ 2 & 0 \\ 5 & \end{array} \tag{1}$$

To explain the usefulness of (1) consider the electrical circuit depicted in Fig. 1. It consists of four linear components whose voltage-current relationships[2] are shown in Fig. 1, where K_i, $i = 1, ..., 4$, are constants.

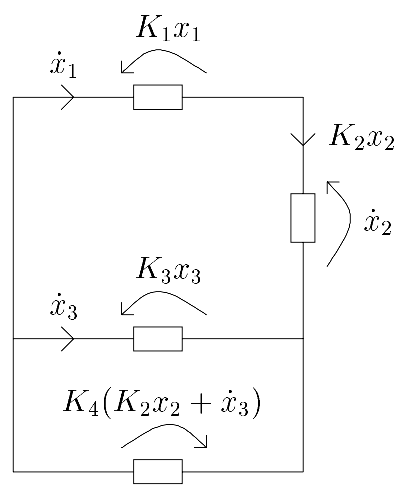

Figure 1: A third-order system

Using Kirchhoff's laws, we obtain the state-space equation: $\dot{\mathbf{x}} = A\mathbf{x}$, where $\mathbf{x} = (x_1, x_2, x_3)^T$ and

$$A = \begin{pmatrix} 0 & K_2 & 0 \\ -K_1 & 0 & K_3 \\ 0 & -K_2 & -\frac{K_3}{K_4} \end{pmatrix} \tag{2}$$

[2]These components are actually two capacitors, an inductor, and a resistor. However, we use the "black-box" notation because we are going to generalize them to nonlinear components in the sequel.

and the characteristic polynomial is $P(s) = det(sI - A) = s^3 + \frac{K_3}{K_4}s^2 + K_2(K_1 + K_3)s + \frac{1}{K_4}K_1K_2K_3$. It is easy to verify that for the particular choice: $(K_4, K_3, K_2, K_1) = (1, 3, 2, 5)$ we obtain:

$$P(s) = s^3 + 3s^2 + 16s + 30 = D(s)$$

Notice though, that $(1, 3, 2, 5)$ is exactly the first column of the (modified) RH array we obtained for $D(s)$ (see (1)). In other words, the values in the first column are just the parameters of the components required to implement $D(s)$ using our electrical circuit. More generally, we have:

Theorem 1 *Given a characteristic polynomial $D(s) = s^n + d_{n-1}s^{n-1} + ... + d_0$, let $\mathbf{z}$ denote the vector of elements in the first column of the modified RH array of $D(s)$. Then, $D(s)$ can be realized as an electrical circuit containing $n+1$ linear components (capacitors, inductors, and resistors). Furthermore, the parameters K_i of these components are the entries of the vector $\mathbf{z}$.*

Proof. See the Appendix.

Special cases arise when a zero entry appears in the first column of the RH array because, then, $\mathbf{z}$ cannot be constructed. However, it can be shown that our approach can be easily extended to handle these cases as well.

Once $D(s)$ is realized as an electrical circuit we can associate with it the energy stored in this circuit. For example, for the circuit in Fig. 1:

$$\begin{aligned} V(\mathbf{x}) &= \int K_1x_1(t)\dot{x}_1(t)dt + \int K_2x_2(t)\dot{x}_2(t)dt + \int K_3x_3(t)\dot{x}_3(t)dt \\ &= \frac{1}{2}(K_1x_1^2 + K_2x_2^2 + K_3x_3^2) \\ &= \mathbf{x}^TP\mathbf{x} \end{aligned} \tag{3}$$

where $P = diag(\frac{K_1}{2}, \frac{K_2}{2}, \frac{K_3}{2})$. The rate of change of the energy is:

$$\dot{V} = 2\mathbf{x}^TPA\mathbf{x} = -\frac{1}{K_4}K_3^2x_3^2$$

which is just the power dissipated by the resistor.

We can use V for stability analysis. Indeed, it is easy to verify that, if all the K_i's are positive, then V is a Lyapunov function guaranteeing asymptotic stability[3]; if all the K_i's are negative, then the same holds with $-V$. On the other hand, if there is a sign-change among the K_i's, then we can use V (or $-V$) to prove that the system is unstable. Thus, a necessary and sufficient condition for asymptotic stability is that there will be no sign-changes among the K_i's.

Actually, we can derive this conclusion in a more intuitive manner. If all the K_i's are positive, then all the circuits components are passive and any initial energy will dissipate and lead to $\lim_{t\to\infty} \mathbf{x}(t) = 0$. The same holds, symmetrically, if all the K_i's are negative. On the other hand, if there is a sign change among the K_i's, then at least one component is gaining energy and the circuit is unstable.

This is related to the (modified) RH array because the K_i's constitute its first column, hence, we immediately obtain:

Theorem 2 *Given a characteristic polynomial $D(s) = s^n + d_{n_1}s^{n-1} + ... + d_0$, let $\mathbf{z}$ denote the vector of elements in the first column of the modified RH array of $D(s)$. Then, $D(s)$ is asymptotically stable if and only if there are no sign changes among the elements of $\mathbf{z}$.*

[3]Note that Eq. (2) implies that the the only trajectory contained in the set $\{\mathbf{x}|x_3 = 0\}$ is $\mathbf{x} = \mathbf{0}$.

The rational behind the RH stability criterion can now be explained as follows. Given a linear differential equation (in the form of a characteristic polynomial), realize it as an electrical circuit and examine the behavior of the circuit's energy, which depends only on the signs of the circuit's components.

A natural generalization comes to mind: Given a nonlinear differential equation, realize it as an electrical circuit and use the energy stored in the circuit as a Lyapunov function to analyze its stability. Thus, the idea behind the RH stability criterion can be extended to the nonlinear case. We demonstrate this approach in Section 3.1.

Note that other proofs of the RH stability criterion (e.g., Chapellat et al., 1990; Parks, 1962) are in general less intuitive than ours because they are based on mathematical abstractions rather than on physical entities.

In the following two subsections we obtain well-known results on passivity and optimality of linear systems directly from the analysis of an electrical circuit. This will allow the derivation of new results for the nonlinear case in Section 3.

2.2 Passivity

In this section we use the energy balance in the electrical circuit to derive some passivity relations. Consider, for example, the circuit in Fig. 2 which consists of three linear components and two

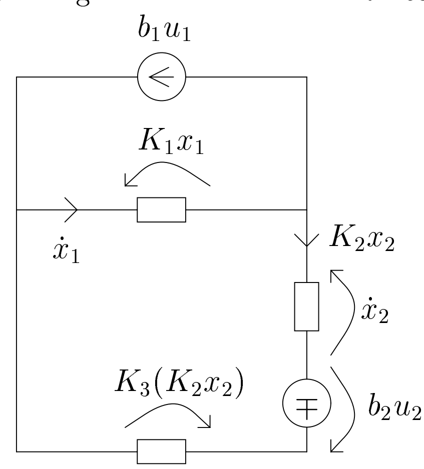

Figure 2: A linear circuit with sources

sources: a current source in parallel with the capacitor and a voltage source in series with the inductor. Kirchhoff's laws now yield:

$$\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u} \tag{4}$$

where $\mathbf{x} = (x_1,\ x_2)^T$, $\mathbf{u} = (u_1,\ u_2)^T$,

$$A = \begin{pmatrix} 0 & K_2 \\ -K_1 & -K_3K_2 \end{pmatrix}, \quad \text{and } B = \begin{pmatrix} b_1 & 0 \\ 0 & b_2 \end{pmatrix} \tag{5}$$

The energy stored in the circuit is the energy stored in the capacitor and inductor, namely:

$$V(\mathbf{x}) = \mathbf{x}^T\overline{P}\mathbf{x}, \qquad \overline{P} = diag(\frac{K_1}{2}, \frac{K_2}{2}) \tag{6}$$

The power supplied by the two sources is:

$$b_1u_1K_1x_1 + b_2u_2K_2x_2 = 2\mathbf{x}^T\overline{P}B\mathbf{u} = 2\mathbf{y}^T\mathbf{u}$$

where we define the circuit's outputs as:

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = B^T \overline{P} \mathbf{x} \tag{7}$$

If all the circuit's components are passive, then the increase in the energy stored in the system cannot be greater than the power supplied by the sources, that is:

$$\dot{V} \leq 2\mathbf{y}^T \mathbf{u}$$

Indeed, this is easily verified by calculating:

$$\begin{aligned} \dot{V} &= 2\mathbf{x}^T \overline{P}(A\mathbf{x} + B\mathbf{u}) \\ &= \mathbf{x}^T(\overline{P}A + A^T\overline{P})\mathbf{x} + 2\mathbf{x}^T\overline{P}B\mathbf{u} \\ &\leq 2\mathbf{y}^T\mathbf{u} \end{aligned} \tag{8}$$

where we used the fact that $\overline{P}A + A^T\overline{P} \leq 0$. In other words, the system defined by (4) and (7) is passive with respect to the storage function $S(\mathbf{x}) = \frac{1}{2}V(\mathbf{x})$.

In general, however, the components in the circuit are not necessarily passive, hence, $V(\mathbf{x})$ in (6) is no longer a valid storage function because the K_i's might be negative. Instead, let $V(\mathbf{x}) = \mathbf{x}^T P \mathbf{x}$, where P is some non-negative definite matrix. To derive passivity relations for $\dot{V}$ in this case, we rewrite (8) as:

$$\begin{aligned} \dot{V} &= \mathbf{x}^T(PA + A^TP - PBB^TP + PBB^TP)\mathbf{x} + 2\mathbf{x}^TPB\mathbf{u} \\ &= -\mathbf{x}^TQ\mathbf{x} + \mathbf{y}^T\mathbf{y} + 2\mathbf{y}^T\mathbf{u} \end{aligned} \tag{9}$$

where we denote:

$$Q = PBB^TP - PA - A^TP \tag{10}$$

Hence, if we can find matrices $P \geq 0$ and $Q \geq 0$ so that (10) holds, then:

$$\dot{V} \leq \mathbf{y}^T\mathbf{y} + 2\mathbf{y}^T\mathbf{u}$$

that is, the system is dissipative with respect to the storage function $S(\mathbf{x}) = \frac{1}{2}V(\mathbf{x})$ and the supply rate $\frac{1}{2}\mathbf{y}^T\mathbf{y} + \mathbf{y}^T\mathbf{u}$ (or, in the terminology of (Sepulchre et al., 1997), the system is Output Feedback Passive $(-\frac{1}{2})$).

Note that Eq. (10) is just the famous Riccati equation arising in optimal control theory. Indeed, we can easily relate the above derivation to the linear-quadratic regulator problem.

2.3 Optimality

In this section we use the fundamental relation between passivity and optimality (Sepulchre et al., 1997, Ch. 3) to derive a simple exposition of the linear-quadratic optimal regulator (Anderson & Moore, 1990).

We rewrite (9) as:

$$\mathbf{x}^TQ\mathbf{x} + \mathbf{u}^T\mathbf{u} = -\dot{V} + (\mathbf{u} + \mathbf{y})^T(\mathbf{u} + \mathbf{y}) \tag{11}$$

Defining:

$$J(\mathbf{x}(t_0), \mathbf{u}) = \int_{t_0}^{\infty} [\mathbf{x}(t)^T Q\mathbf{x}(t) + \mathbf{u}^T(t)\mathbf{u}(t)]dt$$

we obtain from (11):

$$J(\mathbf{x}(t_0), \mathbf{u}) = V(\mathbf{x}(t_0)) - \lim_{t \to \infty} V(\mathbf{x}(t)) + \int_{t_0}^{\infty} (\mathbf{u} + \mathbf{y})^T (\mathbf{u} + \mathbf{y}) dt \tag{12}$$

Now, let U be the set of controllers that yield, in the closed-loop system, $\lim_{t \to \infty} \mathbf{x}(t) = 0$ and consider the problem of finding $\overline{\mathbf{u}} \in U$ that minimizes J. Using (12) and the fact that $V(\mathbf{0}) = 0$, we immediately see that this optimal controller is just $\overline{\mathbf{u}} = -\mathbf{y} = -B^T P \mathbf{x}$ and that the minimal value of J is $V(\mathbf{x}(t_0))$. Note that if there exists a matrix $Q > 0$ that solves (10), then $\overline{\mathbf{u}} = -B^T P \mathbf{x}$ indeed belongs to U.

Since the power supplied by the sources is $2\mathbf{y}^T \mathbf{u}$, then setting $\mathbf{u} = \overline{\mathbf{u}} = -\mathbf{y}$ corresponds to maximizing the power that the sources extract from the circuit. This yields a very intuitive explanation of the optimal regulator.

Note that we have related passivity and optimality directly to the energy stored in the circuit. This will allow us to derive nonlinear versions of our results by using nonlinear circuits.

3 Nonlinear Systems

In this section we extend the previous results to nonlinear systems. Consider now the circuit depicted in Fig. 3. Note that it is a modification of the linear circuit in Fig. 2 (without the sources), namely, we have replaced the linear components of Fig. 2 with the more general nonlinear ones. In particular, we can obtain the linear circuit by setting $f_i(x) = K_i x$, $i = 1, 2, 3$.

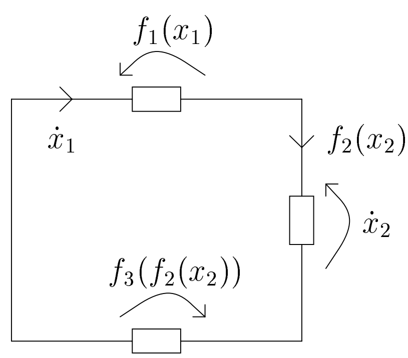

Figure 3: Nonlinear electrical circuit

The state-space equations[4] are now:

$$\begin{aligned} \dot{x}_1 &= f_2(x_2) \\ \dot{x}_2 &= -f_1(x_1) - f_3(f_2(x_2)) \end{aligned} \tag{13}$$

and the energy stored in the circuit at time t is:

$$V(\mathbf{x}(t)) \quad = \quad V(\mathbf{x}(t_0)) + \int_{t_0}^{t} f_1(x_1(\tau)) \dot{x}_1(\tau) d\tau + \int_{t_0}^{t} f_2(x_2(\tau)) \dot{x}_2(\tau) d\tau$$

[4]Throughout this paper we implicitly assume that the nonlinear functions satisfy conditions that guarantee existence and uniqueness of solutions.

$$\begin{aligned} &= V(\mathbf{x}(t_0)) + \int_{x_1(t_0)}^{x_1(t)} f_1(y)dy + \int_{x_2(t_0)}^{x_2(t)} f_2(y)dy \\ &= \int_0^{x_1(t)} f_1(y)dy + \int_0^{x_2(t)} f_2(y)dy \end{aligned} \tag{14}$$

where we assume that the initial state at time t_0 is $\mathbf{x}(t_0) = \mathbf{0}$.

The derivative of V, along the trajectories of (13), is simply:

$$\begin{aligned} \dot{V} &= f_1(x_1)\dot{x}_1 + f_2(x_2)\dot{x}_2 \\ &= f_1(x_1)f_2(x_2) + f_2(x_2)(-f_1(x_1) - f_3(f_2(x_2))) \\ &= -f_2(x_2)f_3(f_2(x_2)) \end{aligned} \tag{15}$$

which is just the power dissipated by the nonlinear resistor.

A simple yet important observation follows. Given a nonlinear system in the form (13) it can be realized immediately as the circuit in Fig. 3, pointing to a natural candidate for the system's energy, namely, the energy $V(\mathbf{x})$ stored in the circuit's components.

3.1 Nonlinear stability analysis

If we can realize a nonlinear system by an electrical circuit, we can use the associated energy V for stability analysis. We demonstrate this in the following example.

Example 1 *(Lotka-Volterra systems)*

Consider the following second-order Lotka-Volterra system (Hubbard & West, 1995):

$$\begin{aligned} \dot{x} &= (2-y)x \\ \dot{y} &= (3x-1)y \end{aligned} \tag{16}$$

where $x(t)$ and $y(t)$ can attain only non-negative values (e.g., they represent sizes of populations). Hence, we are interested in the behavior of the system only in the open first quadrant: $M = \{(x,y) \ : \ x > 0, \ y > 0\}$.

Defining the coordinate transformation: $x_1 = \ln(x)$ and $x_2 = \ln(y)$ yields:

$$\begin{aligned} \dot{x}_1 &= 2 - e^{x_2} \\ \dot{x}_2 &= 3e^{x_1} - 1 \end{aligned} \tag{17}$$

Compared with (13), we see that (17) can be realized using our nonlinear electrical circuit (Fig. 3) with:

$$f_1(x) = 1 - 3e^x, \quad f_2(x) = 2 - e^x, \quad f_3(x) = 0$$

In this case, (14) and (15) become:

$$V(x_1, x_2) = x_1 - 3e^{x_1} + 2x_2 - e^{x_2}$$

and

$$\dot{V}(x_1, x_2) = 0$$

respectively. Or, in terms of our original variables:

$$V(x,y) = \ln(x) - 3x + 2\ln(y) - y \tag{18}$$

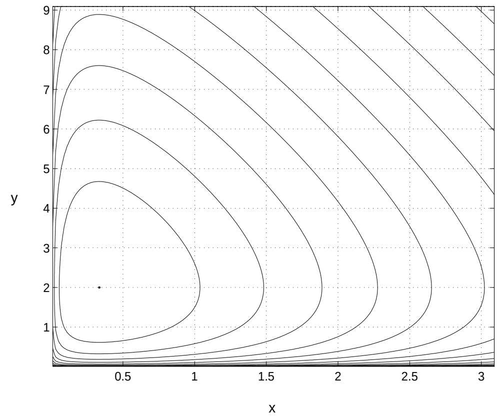

Figure 4: Contours of $V(x, y)$

(note that V is well-defined on M) and

$$\dot{V}(x, y) = 0$$

Note that we have solved the problem of associating an appropriate energy function with the system (16) by realizing it as a nonlinear electrical circuit and using the energy stored in the components of the circuit. In view of Theorem 1, what we have done can clearly be considered an extension of the idea behind the RH stability criterion.

Fig. 4 depicts contours of the function $V(x, y)$ in (18) (i.e., curves that satisfy $V(x, y) = const.$). Since $\dot{V}(x, y) = 0$ along the trajectories of (16), we immediately conclude that these contours are just the trajectories of the dynamical system (16). Note that the oscillatory behavior depicted in Fig. 4 agrees with our interpretation of the system as the combination of a (nonlinear) capacitor and a (nonlinear) inductor with no resistor. Hence, our approach led, in a very simple manner, to a complete analysis of the behavior of (16) in M.

3.2 Passivity

In Section 2.2 we used the energy stored in our linear circuit to derive passivity relations. We now extend these results to the nonlinear case. Consider the nonlinear circuit depicted in Fig. 5, which is a generalization of the circuit in Fig. 2. The state-space equations are now:

$$\begin{aligned} \dot{x}_1 &= f_2(x_2) + b_1 u_1 \\ \dot{x}_2 &= -f_1(x_1) - f_3(f_2(x_2)) + b_2 u_2 \end{aligned} \tag{19}$$

First, we assume that all the circuit's components are passive. For example, let $f_1(x) = K_1 \tanh(x)$, $f_2(x) = K_2 \tanh(x)$, and $f_3(x) = K_3 x$, $K_i > 0$ $(i = 1, 2, 3)$. In this case, (19) becomes:

$$\dot{\mathbf{x}} = A \tanh(\mathbf{x}) + B\mathbf{u} \tag{20}$$

where $\tanh(\mathbf{x}) = (\tanh(x_1), \tanh(x_2))^T$, and A and B are given by Eq. (5).

The energy stored in the circuit is:

$$V(\mathbf{x}) = K_1 \ln(\cosh(x_1)) + K_2 \ln(\cosh(x_2)) = 2\overline{p}_1 \ln(\cosh(x_1)) + 2\overline{p}_2 \ln(\cosh(x_2))$$

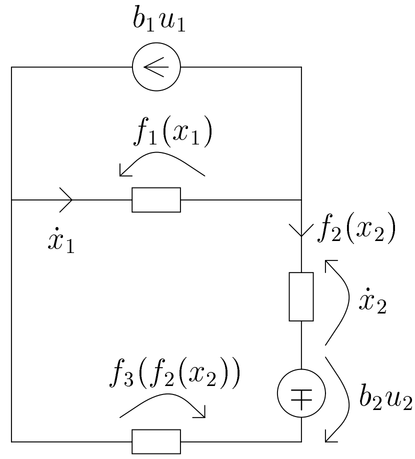

Figure 5: Nonlinear circuit with inputs

where the $\overline{p}_i$'s are diagonal elements of the matrix $\overline{P}$ (recall that $\overline{P} = diag(\frac{K_1}{2}, \frac{K_2}{2})$). The power supplied by the sources is:

$$b_1u_1K_1\tanh(x_1) + b_2u_2K_2\tanh(x_2) = 2\mathbf{u}^TB^T\overline{P}\tanh(\mathbf{x})$$

Hence, if we define the circuit's outputs as:

$$\mathbf{y} = B^T\overline{P}\tanh(\mathbf{x}) \tag{21}$$

then we must have, just as in the linear case, $\dot{V} \leq 2\mathbf{y}^T\mathbf{u}$. Indeed:

$$\begin{aligned} \dot{V} &= 2\overline{p}_1\tanh(x_1)\dot{x}_1 + 2\overline{p}_2\tanh(x_2)\dot{x}_2 \\ &= 2\tanh^T(\mathbf{x})\overline{P}\dot{\mathbf{x}} \\ &= 2\tanh^T(\mathbf{x})\overline{P}(A\tanh(\mathbf{x}) + B\mathbf{u}) \\ &= \tanh^T(\mathbf{x})(\overline{P}A + A^T\overline{P})\tanh(\mathbf{x}) + 2\mathbf{y}^T\mathbf{u} \\ &\leq 2\mathbf{y}^T\mathbf{u} \end{aligned} \tag{22}$$

Note that the transition from the first to the second line in (22) is possible only because P is diagonal.

Equation (22) implies that the system defined by (20) and (21) is passive with respect to the storage function $S(\mathbf{x}) = \frac{1}{2}V(\mathbf{x})$.

Now, let us consider the more general case in which the nonlinear components of the circuit are not necessarily passive. Let $V(\mathbf{x}) = \mathbf{x}^TP\mathbf{x}$ where P is a non-negative definite matrix. Rewriting (22) yields:

$$\begin{aligned} \dot{V} &= \tanh^T(\mathbf{x})(PA + A^TP - PBB^TP + PBB^TP)\tanh(\mathbf{x}) + 2\mathbf{y}^T\mathbf{u} \\ &= -\tanh^T(\mathbf{x})Q\tanh(\mathbf{x}) + \mathbf{y}^T\mathbf{y} + 2\mathbf{y}^T\mathbf{u} \end{aligned} \tag{23}$$

where Q is defined in (10). Hence, if we can find a *diagonal* matrix $P \geq 0$ and a matrix $Q \geq 0$ so that (10) holds, then the system is dissipative with respect to the storage function $S(\mathbf{x}) = \frac{1}{2}V(\mathbf{x})$ and the supply rate $\frac{1}{2}\mathbf{y}^T\mathbf{y} + \mathbf{y}^T\mathbf{u}$.

Note that we are not restricted to using components defined by the function $\tanh(\cdot)$, but we can follow (Kaszkurewicz & Bhaya, 1993) and define the more general set of possible functions:

Definition 1 *(The set S_c)*
Let S_c be the set of all functions $f(\cdot): R \to R$ satisfying:

- *$f(\cdot)$ is continuous*
- *$f(0) = 0$, and for all other $x \in R$: $xf(x) > 0$*
- *$\int_0^x f(y)dy \to \infty$ as $|x| \to \infty$*

Note that the last two properties imply that if $V(x) = \int_0^x f(y)dy$, where $f \in S_c$, then $V(x)$ is positive-definite and radially unbounded.

We can now generalize our previous result:

Theorem 3 *Consider the nonlinear system:*

$$\dot{\mathbf{x}} = A\mathbf{f}(\mathbf{x}) + B\mathbf{u}$$

where $\mathbf{f}(\mathbf{x}) = (f_1(x_1), ..., f_n(x_n))^T$, with $f_i(\cdot) \in S_c$, $i = 1, ..., n$. Suppose that there exist $Q \geq 0$ and $P \geq 0$, with $P = diag(p_1, ..., p_n)$, that solve (10) and define the systems outputs as:

$$\mathbf{y} = B^T P \mathbf{f}(\mathbf{x})$$

Then, the system is dissipative with respect to the storage function

$$S(\mathbf{x}) = \sum_{i=1}^{n} p_i \int_0^{x_i(t)} f_i(y)dy$$

and the supply rate $\frac{1}{2}\mathbf{y}^T\mathbf{y} + \mathbf{y}^T\mathbf{u}$.

3.3 Optimality

We can use the relation between passivity and optimality to formulate and solve a nonlinear optimal control problem. We begin by rewriting (23) as:

$$\tanh^T(\mathbf{x}) Q \tanh(\mathbf{x}) + \mathbf{u}^T\mathbf{u} = -\dot{V} + (\mathbf{u} + \mathbf{y})^T(\mathbf{u} + \mathbf{y}) \tag{24}$$

Defining:

$$J(\mathbf{x}(t_0), \mathbf{u}) = \int_{t_0}^{\infty} [\tanh^T(\mathbf{x}(t)) Q \tanh(\mathbf{x}(t)) + \mathbf{u}^T(t)\mathbf{u}(t)]dt$$

we obtain from (24):

$$J(\mathbf{x}(t_0), \mathbf{u}) = V(\mathbf{x}(t_0)) - \lim_{t\to\infty} V(\mathbf{x}(t)) + \int_{t_0}^{\infty} (\mathbf{u} + \mathbf{y})^T(\mathbf{u} + \mathbf{y})dt \tag{25}$$

Now, let U be the set of controllers that yield, in the closed-loop system, $\lim_{t\to\infty} \mathbf{x}(t) = 0$ and consider the problem of finding $\overline{\mathbf{u}} \in U$ that minimizes J. Using (25) and the fact that $V(\mathbf{0}) = 0$, we immediately see that this optimal controller is just $\overline{\mathbf{u}} = -\mathbf{y} = -B^T P \tanh(\mathbf{x})$ and that the minimal value of J is $V(\mathbf{x}(t_0))$. Note that if there exists a matrix $Q > 0$ that solves (10), then $\overline{\mathbf{u}} = -B^T P \tanh(\mathbf{x})$ indeed belongs to U.

More generally, we obtain the following extension of the linear-quadratic H_2 optimal regulation problem to the nonlinear case.

Theorem 4 *Consider the nonlinear system*[5]*:*

$$\begin{aligned} \dot{\mathbf{x}} &= A\mathbf{f}(\mathbf{x}) + B\mathbf{u} \\ \mathbf{x}(t_0) &= \mathbf{x_0} \end{aligned} \tag{26}$$

where $\mathbf{f}(\mathbf{x}) = (f_1(x_1), ..., f_n(x_n))^T$, *with* $f_i \in S_c$ *for* $i = 1, ..., n$. *Define the cost functional:*

$$J(\mathbf{x_0}, t_0, \mathbf{u}) = \int_{t_0}^{\infty} [\mathbf{f}^T(\mathbf{x}(t))Q\mathbf{f}(\mathbf{x}(t)) + \mathbf{u}^T(t)\mathbf{u}(t)]dt \tag{27}$$

where $Q > 0$ *is a symmetric matrix, and denote the minimal cost:*

$$\overline{J}(\mathbf{x_0}, t_0) = \min_{\mathbf{u} \in U} J(\mathbf{x_0}, t_0, \mathbf{u}) \tag{28}$$

If there exists a positive-definite and diagonal *matrix* $P = diag(p_1, ..., p_n)$ *that solves (10), then:*

$$\overline{J}(\mathbf{x_0}) = 2\sum_{i=1}^{n} p_i \int_0^{x_i(t_0)} f_i(y)dy \tag{29}$$

and the optimal controller, yielding $\overline{J}$, *is given by:*

$$\overline{\mathbf{u}}(t) = -B^T P\mathbf{f}(\mathbf{x}(t)) \tag{30}$$

Theorem 4 first appeared in (Margaliot & Langholz, 1999a) for the particular case $f_i(x) = \tanh(k_i x)$, $k_i > 0$. It turns out that in this case, the optimal controller is in fact a *fuzzy controller* and this result supplies a theoretical explanation for the well-demonstrated success of fuzzy controllers. The extension to the H_∞ framework is quite straight forward (see (Margaliot & Langholz, 1999b)).

4 Conclusions

In this paper we studied the following problem: Given the mathematical model of a dynamical system, find a meaningful measure of the energy stored in the system.

The problem was first considered in the framework of linear systems. Here, the mathematical model is a characteristic polynomial. We showed that it can be realized by an electrical circuit composed of three types of linear components. The energy of the system is then just the energy stored in the circuit's components. This led to a very intuitive proof of the famous Routh-Hurwitz stability criterion.

For the more general case of linear systems with inputs, we used the energy stored in the circuit as a storage function and derived passivity relations based on the energy balance in the circuit. Using the relation between passivity and optimality we also formulated and solved, in a very intuitive manner, the infinite-time H_2 optimal regulation problem.

As the concepts of stability, passivity, and optimality are related directly to the electrical circuit, we were able to extend these results to a class of *nonlinear* systems by replacing the linear components in the circuit with more general nonlinear components. We showed how to define the energy that is stored in the circuit and how to use this energy for nonlinear stability, passivity, and optimality analysis and synthesis. In particular, we were able to give an intuitive derivation of a class of nonlinear optimal control problems enjoying a simple *explicit* solution.

[5]Again, we assume that the equations have a unique solution and, furthermore, that the $f_i(\cdot)$'s are sufficiently smooth.

References

Anderson, B. D. O. and J. B. Moore (1990). *Optimal Control: Linear Quadratic Methods*, Prentice-Hall.

Chapellat, H., M. Mansour, and S. P. Bhattacharyya (1990). "Elementary proofs of some classical stability criteria," *IEEE Trans. Education*, **33**, no. 3, pp. 232–239.

Hirsch, M. W. and S. Smale (1974). *Differential Equations, Dynamical Systems, and Linear Algebra*, Academic Press.

Hubbard, J. H. and B. H. West (1995). *Differential Equations: A Dynamical Systems Approach, Part II: Higher-Dimensional Systems*, Springer-Verlag.

Kaszkurewicz, E. and A. Bhaya (1993). "Robust stability and diagonal Liapunov functions," *SIAM J. Matrix Analysis Applications*, **14**, no. 2, pp. 508–520.

Margaliot, M. and G. Langholz (1999a). "Hyperbolic optimal control and fuzzy control," *IEEE Trans. Systems, Man, Cybernetics: Part A*, **29**, no. 1, pp. 1–10.

Margaliot, M. and G. Langholz (1999b). "Hyperbolic approach to fuzzy modeling and control," *Submitted to FUZZ-IEEE'99.*

Ogata, K. (1967). *State Space Analysis of Control Systems*, Prentice-Hall.

Parks, P. C. (1962). "A new proof of the Routh-Hurwitz stability criterion using the second method of Liapunov," *Proc. Cambridge Philosophical Society (Mathematical and Physical Sciences)*, **58, Part 4**, pp. 694–702.

Sepulchre, R., M. Janković, and P. V. Kokotović (1997). *Constructive Nonlinear Control*, Springer-Verlag.

Appendix: Proof of Theorem 1

We begin by defining recursively a linear electrical circuit $C(n)$, $n = 1, 2, \ldots$. The circuit $C(1)$, shown in Fig. 6(a), consists of the parallel connection of a capacitor and 1Ω resistor. The circuit $C(2)$, shown in Fig. 6(b), is obtained by replacing the resistor in $C(1)$ with the serial connection of an inductor K_2 and a 1Ω resistor. For any $n > 2$, $C(n)$ is defined recursively:

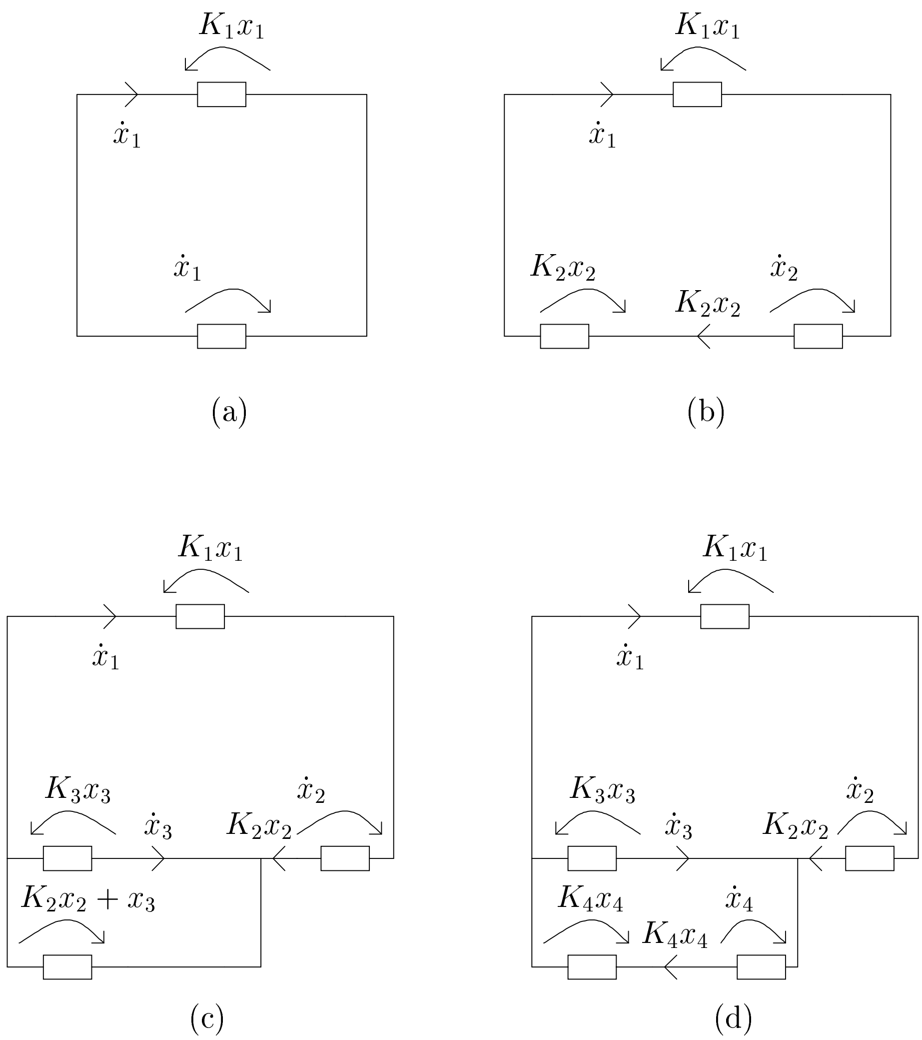

Figure 6: The circuits (a) $C(1)$, (b) $C(2)$, (c) $C(3)$, and (d) $C(4)$

- If n is odd then $C(n)$ is the circuit obtained by replacing the resistor in $C(n-1)$ with the parallel connection of a capacitor K_n and a 1Ω resistor.
- If n is even then $C(n)$ is the circuit obtained by replacing the resistor in $C(n-1)$ with the serial connection of an inductor K_n and a 1Ω resistor.

Fig. 6 depicts the circuits $C(1) - C(4)$.

To obtain a state-space description of the circuit's dynamics we also define state-space variables x_i. Whenever we add a capacitor (inductor) with parameter K_n we add a state-space variable x_n, where K_nx_n is the capacitor's voltage (inductor's current).

Lemma 1 *The circuit $C(n)$ defined above is described by the state-space equation: $\dot{\mathbf{x}} = A(n)\mathbf{x}$, where $\mathbf{x} = (x_1, ..., x_n)^T$ and $A(n)$ is the tridiagonal matrix:*

$$A(n) = \begin{pmatrix} 0 & K_2 & & & & & \\ -K_1 & 0 & K_3 & & & & \\ & -K_2 & 0 & K_4 & & & \\ & & \cdot & \cdot & \cdot & & \\ & & & \cdot & \cdot & \cdot & \\ & & & & -K_{n-2} & 0 & K_n \\ & & & & & -K_{n-1} & -K_n \end{pmatrix} \tag{31}$$

Proof. A recursive proof follows immediately from the recursive definition of the circuit $C(n)$. □

Let $P(n)(s) = det(sI - A(n))$ be the characteristic polynomial of $C(n)$. It turns out that the first column of the RH array of $P(n)(s)$ has a particularly simple form.

Lemma 2 *The first column of the RH array for $P(n)(s)$ is:*

$$1, K_n, K_n K_{n-1}, K_n K_{n-1} K_{n-2}, ..., K_n K_{n-1} K_{n-2} \cdots K_1 \tag{32}$$

Proof. Define the matrix $M(n) = diag(1, K_2, K_2K_3, ..., K_2K_3 \cdots K_N)$. Then it is easy to verify, using (31), that

$$M(n)A(n)M^{-1}(n) = \begin{pmatrix} 0 & 1 & & & & & \\ -K_1K_2 & 0 & 1 & & & & \\ & -K_2K_3 & 0 & 1 & & & \\ & & \cdot & \cdot & \cdot & & \\ & & & \cdot & \cdot & \cdot & \\ & & & & -K_{n-2}K_{n-1} & 0 & 1 \\ & & & & & -K_{n-1}K_n & -K_n \end{pmatrix}$$

That is, $M(n)A(n)M(n)^{-1}$ is a Schwarz matrix (Ogata, 1967, Ch. 8). Now Theorem 5 in (Parks, 1962) (see also (Ogata, 1967, Theorem 8-10)) implies that the first column of the RH array of the polynomial $det(sI - M(n)A(n)M^{-1}(n))$ is:

$$1, K_n, K_n K_{n-1}, K_n K_{n-1} K_{n-2}, ..., K_n K_{n-1} K_{n-2} ... K_1$$

and since $det(sI - M(n)A(n)M^{-1}(n)) = det\left(M(n)(sI - A(n))M^{-1}(n)\right) = det(M(n))det(sI - A(n))det(M^{-1}(n)) = det(sI - A(n)) = P(n)(s)$ this completes the proof. □

Note that Lemma 2 and the definition of the modified RH array imply that the first column in the modified RH array of $P(s)$ is just:

$$1, K_n, K_{n-1}, K_{n-2}, ..., K_1$$

In other words, constructing the RH for $P(n)(s)$ retrieves the values of the parameters in the circuit $C(n)$ that we started with.

Finally, let $D(s)$ be some monic characteristic polynomial $D(s) = s^n + d_{n-1}s^{n-1} + ... + d_0$, and let $\mathbf{z} = (1, z_1, .., z_n)$ be the first column in the modified RH array of $D(s)$. Consider the circuit $C(n)$ with $K_n = 1$,$K_{n-1} = z_1$,...,$K_1 = z_n$, and let $P(n)(s)$ be its characteristic polynomial. We know that the first column in the modified RH array of $P(s)$ is just: $1, K_n, K_{n-1}, K_{n-2}, ..., K_1$, that is, the vector $\mathbf{z}$. Hence, $D(s)$ and $P(n)(s)$ have the same modified RH array[6] and, therefore, $D(s) = P(n)(s)$. □

[6]Note that if two monic polynomials lead to RH arrays with an identical first column then they must be the same polynomial (Parks, 1962). It is easy to see that this holds for modified RH arrays as well.